Institute of Education

UCL

University College London
UCL Institute of Education
UCL Center for Applied Linguistics

Understanding Trans-semioticising Expressions on Chinese Social Media

A Thesis in Applied Linguistics by

Tiancheng Chen

Supervised by Professor Li Wei and Professor Ana Pellicer-Sanchez

Submitted in Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

July 2023

Declaration

I, Tiancheng Chen, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Word Count (excluding Reference List and Appendix): 95236 words

Understanding Trans-semioticising Expressions on Chinese Social Media

Abstract

Previous research has investigated Trans-semioticising expressions (TSEs) on social media from different perspectives. However, there is still a lot of work to be done on the topic, such as the categorisation of translanguaging expressions, the understanding processes of different types of TSEs, and the cognitive effects achieved during the understanding processes. This study aims to demonstrate the structural and composite diversity of TSEs that are encountered on Chinese social media and, more importantly, explore the cognitive processes of their comprehension.

The first part of data is a TSEs database constructed with posts from 8 Weibo Super-topics. Through Grounded Theory, the TSEs are categorised into different types. The second part of data is recordings of think-aloud tasks, coupled with follow-up interviews to gather the thoughts of participants when they understand different TSEs. The strategies and patterns used by the participants are analysed and summarised. Relevance Theory was adopted as a theoretical tool to help analyse and explain the participants' understanding paths and strategies used during the think-aloud tasks.

The study finds that the formation of TSEs falls into four macro-mechanisms of innovative lexical formation, resulting in four types and eight sub-types. In terms of the combinations of semiotic elements involved, five different semiotic systems are involved. TSEs are differentiated into explicit and implicit ones according to whether the formation mechanism and the semiotic elements were visible on the surface. Three different understanding paths and eight different understanding strategies are found. The participants' level of familiarity with the TSEs and the types of the TEs are found to be influential. The participants' understanding of TSEs is constrained by the presumption of optimal relevance when participants decode and integrate the meaning

of the TSEs. In addition, the extra effort expended on TSEs can generate some non-propositional effects, namely social associations and aesthetic associations.

Impact Statement

Focusing on Trans-semioticising expressions (TSEs), this study serves to generate the following impacts.

To start with, by systematically categorising TSEs and generalizing their rules and elements of formation and exploring the cognitive process of TSE comprehension, this study broadens the research scope of Translanguaging and dynamic lexicology. In the former case, by focusing on the phenomenon on the lexical level, it reveals how elements from different linguistic and non-linguistic systems work together to serve meaning-making purposes in the context of social media communication and provides a more comprehensive view of TSEs based on a large amount of data collected. More importantly, it complements previous research on Translanguaging by revealing the paths, strategies and routes of TSE comprehension. In the latter case, it introduces Translanguaging or Trans-semioticising into the scope of analysis, showing that creative lexical formation can transcend the boundary between different languages and semiotic systems. It testifies the impact of technological affordances of social media communication such as inputting technology and the popular use of emojis on creative lexical formation.

Secondly, this study contributes to our understanding and application of Relevance Theory and enriches the field of lexical pragmatic research. By investigating the processing of multi-lingual or multi-semiotic expressions, it expands the application scope of Relevance Theory in the study of lexico-pragmatic processing, as previous studies addressed lexical pragmatic processes in understanding mono-lingual terms. It testifies the applicability of Relevance Theory in the study of TSEs, as it finds that in the tackling the meaning of TSEs, the constraint of relevance functions both on the decoding and the integrating processes of TSE understanding. Also, the study shows that understanding TSEs not only involves relevance-constrained cognitive processes of narrowing and broadening, but also that of integration, which is also

relevance-constrained. It provides new evidence for the argument of Relevance Theory that utterance interpretation does not always start from decoding what is said and (if at all) end at inferring the implicature; rather, understanding utterance, including lexical items like TSEs in this study, is a bidirectional process. It finds that the association about sensitivity avoidance is a brand new type of non-propositional effect revealed in the present study. The new discovery is a natural outcome as online communication is inevitably subject to various restrictions. It turns out that using TSEs is a suitable way of coping with the restrictions.

Finally, the outcome of this study can be applied to many fields in daily life. For instance, by providing the understanding paths and strategies to people who have difficulties following social media communication, this study can help them understand all kinds of TSEs in possible ways. Potentially, this study can provide a tool for the elder generation and the younger generation to better communicate with each other. Moreover, the understanding paths and strategies may offer useful logic chains for Artificial Intelligence (AI) studies on understanding social media language. Furthermore, supported by the present study, related Internet censorship organisations can better track those illegal behaviors which are hidden beneath all kinds of TSEs.

Acknowledgments

Completing a thesis of this length and complexity could by no means be plain sailing. In the process, many people helped me in one way or another to remove all sorts of obstacles on my way to the final destination of voyage. Without their generous and timely help, the completion of the thesis would have been impossible.

First of all, I would like to express my sincere gratitude towards my principal supervisor, Professor Li Wei. It was a great fortune of mine to meet and become his student some four years ago. I could not have made it without his expert guidance and patient encouragement. In my first year, Professor Li led me into the hall of academic scholarship with his enthusiasm and wisdom. He illustrated with his own stories what a mature researcher should be like. When I failed my first upgrade viva, he offered me many useful suggestions instead of blaming me for my failure. He was very open-minded when I planned to give up the design of eye-tracking and adopt think-aloud tasks instead. His optimistic attitude, sense of humor, spirit of scholarship, and creative thinking influenced me all the time. When I was suffering the most difficult time during the second year amid the trouble of the Covid-19, he never gave up on me. He showed his great patience and warmth towards me and helped me out when I doubted whether I deserved to continue my doctoral student life. When I was behind schedule, I could feel his worry and concern. However, he always encouraged me in a gentle way and spared no effort to help out me. I sometimes thought that I might be far from being his best student, but I could still feel his high expectation on me. He was not only a perfect supervisor to me, but also a fatherly friend with love and support and a role model for me. With full gratitude towards him, I will resolve to continue my research career. The beginning might be tough but, with his guidance and expectation, I have confidence to forge ahead and succeed step by step.

I also want to thank my subsidiary supervisor, Ana Pllicer Sanchez. She is so nice to me throughout my PhD study. I planned to take eye-tracking as my research method at the very beginning. Ana, as a specialist in this area, helped me a lot to understand this high-tech method. However, when I was questioned about the feasibility of my original design and noticed that eye-tracking might not be able to answer my research questions, Ana did not leave me isolated and helpless. In turn, she provided great support to help me polish and complete the think-aloud design. She offered useful comments and advise on my naive plan with great patience. Without her professional guidance, I could not build up the crucial foundation of my thesis and passed the second upgrade viva. I would also to give my appreciation to my upgrade examiners. Their strict requirements and pertinent questions helped me avoid many potential problems in my research design. By thinking over their suggestions, I learned how to think critically and write convincingly. Without their advice and insight to my thesis, I could not have become a qualified PhD candidate at all.

During the years of my PhD study, my fellow PhD candidates, friends, and senior scholars helped me a lot. I am so gratitude to Dr. Yi Wang and Dr. Rowling Hu. As a newcomer who was not familiar with everything in PhD study in the UK, they supported me not only academically, but also mentally. When I had no clue to get

access to school information, Dr. Yi Wang generously offered her help without any complaints. Her academic achievement also made her a role model that I wished to follow. Dr. Rowling Hu and her husband Will came to my rescue in the most difficult time of my PhD study. During the time that I isolated myself from outside, they invited me to their Christmas dinner. They were the first local friends I made in the UK. They were also willing to discuss with me about my research, which provided different perspectives to my research design. My friends far in China were also important to me. They were always ready to share my negative emotions and accompany me when I was down. When I had difficulty recruiting participants, they spared their time from heavy work to support my study. I could complete my study without the assistance of Professor Zhoumin Yuan and Professor Wenyu Wang. They provided precious advice to my database design and think-aloud design with their professional research experience.

One special person that must be mentioned is my girlfriend, Weizhao Gong, who is also working on the PhD study. She enlightened my life when I was alone suffering all the difficulties in a small room in London. We overcame the toughest time during the pandemic. Every time I was busy with deadlines and loads of writing tasks, she took care of my life and encouraged me to keep on. We could discuss each other's research design and have a small panel at any time. Her talent of critical thinking and spirit of scholarship are exactly what I lack. I have learned a lot from her perspectives. We left our footprints together on many cities in the UK, which made the journey of PhD study no longer boring. I am so grateful for her encouragement, accompany, and love.

Finally, I want to present my deepest gratitude to my parents. On the one hand, the dream of becoming a successful researcher like my father has always been the biggest motivation to support my PhD study. He has played as a perfect role modal from the first day of my life. I could feel his influence and inspiration in his long sitting in front of the computer day in and day out, his patient comments and notes on my rough writing, and his carefully designed and typed messages to check my progress. In fact, I could also feel his concern, expectation, trust, and love to me behind all of these. He is the source of my interest in language, teacher of my academic writing, and lighthouse of my future career. He could tolerate my slow action and still offer his most to help me afterwards. In effect, no language could convey my gratitude towards him, not even Trans-semioticising. I would like to devote myself to the future career as the biggest acknowledgment to him. On the other hand, my completion of PhD study could not have been possible without my mother's devotion. I still remember the day she took her PhD examination in her later 30s. Her courage to challenge her limits motivated me when I was at school. She offered me the most needed mental support when I got lost. Her expertise in consulting skills was the most powerful tool to help me out and survive at the most difficult time. She has always played the role of my friend and is ready to hear my woes. With the generous support of my parents, I could study abroad.

This doctoral thesis is a gift not only for myself, but also for those mentioned above, as well as those who also cared about me during the years, including my

grandparents, my aunts and uncles, my brothers and sisters, my former teachers, and all the participants in my data collection process.

Table of Contents

List of Images

List of Tables

List of Abbreviations

TE Translanguaging Expression

TSE Trans-semioticising Expression

IPA Interpretative Phenomenological Analysis

LLTT Looking, Listening, Talking, and Thinking

RT Relevance Theory

PC Personal Computer

CNY Chinese Yuan

TXT Text

LLTSE Logographic Characters and Latin Alphabet Letters

LNTSE Logographic and Numeral Trans-semioticising Expression

LETSE Logographic and Emoji Trans-semioticising Expression

List of Figures

Chapter 1 Introduction

This introductory chapter provides a brief overview of the thesis. First, an introduction of the research problem is presented. It follows up by the statement of the research aims and significance of the study. The chapter ends with an overview of the thesis structure.

1.1 Statement of the problem

We are living in an age of digital social media. To get acquainted and communicate with the external world, we read and write digitally. One important aspect of our cyber life consists in browsing various sorts of social media on the Internet and conducting interactions with other Internet users. We do not just read messages posted by them but also post our comments. A close look into the digital content here and there on the Chinese social media may reveal an interesting phenomenon. Specifically, expressions made up of elements from different languages and even those from other semiotic systems like pictures and emojis are widely used to make meaning.

What exactly do these expressions on Chinese social media look like? Generally speaking, they are not in the traditional or prototypical Chinese form. For example, netizens use numbers (*2333*, meaning "big laugh"), English letters (*XSWL*, abbreviated from "笑死我了", "xiao si wo le", meaning "making me laugh to death"), both English and Chinese (*笑skr人*, meaning *"*making some feel very funny"), and so on. These expressions go definitely beyond the range of Chinese but are widely used by millions of Chinese netizens. How are they constructed? What do people think about them? How do social media users understand them?

The questions above trigger the present study. While gathering examples of these new items, I was introduced to the notion of Translanguaging (e.g., Li, 2016a, 2016b; Li *et al*, 2020; Li & Zhu, 2019; Zhang & Ren, 2022). The core belief of Translanguaging is to break traditional language boundaries and set up a space where

elements from different languages are used as resources to serve communication purposes for people. For example, case like 笑 *skr* 人 can be treated as a Translanguaging word which uses Chinese and English elements. However, the studies on the aforementioned expressions that appear on Chinese social media are not limited to one single language. For example, *7456* is a mixture of Chinese ("气死我了", meaning "I'm saddened to death") and numbers by means of homophony. Literally, the notion of Translanguaging is not inclusive enough to capture creative expressions like *7456* which go beyond one semiotic system.

Instead, I will introduce a new term, namely, Trans-semioticising expression (TSE for short). By definition, it refers to a lexical item (word or word group), explicitly or implicitly composed of elements from more than one linguistic or semiotic system such as languages, pictures, numbers, gifs, and hyperlinks. By the distinction between "explicitly" and "implicitly", I refer to the fact that some TSEs are composed of elements from more than one linguistic or semiotic system as we observe from their surface form, whereas some other TSEs apparently contain elements from only one linguistic or semiotic system but actually their sense-making or comprehension involves elements from more than one linguistic or semiotic system (for detailed explanation, refer to Section 4.3). In other words, the notion of Trans-semioticising expression is broader than the more familiar notion of Translanguaging expression (TE) (if it literally refers to an expression consisting of elements from two languages), because it also includes those containing one linguistic element plus some non-linguistic element(s), as well as "tranßcripting" (Li and Zhu, 2019, p. 153). In the existing literature, however, we do find some researchers occasionally use TE loosely to include those like the aforementioned ones that contain non-linguistic element(s) (e.g. Garcia & Li, 2004; Ren & Guo, 2022; Ren, Guo, & Li, 2022). In this sense, TE used in that way is the same as TSE. Nevertheless, to be precise, I prefer to use TSE throughout, understanding that there is no clear-cut boundary between languages on the one hand and between linguistic system and other

semiotic systems on the other.

One caveat must be provided here. In this dissertation, by "semiotic system", I refer to any system of symbols or signs that serves the purpose of meaning making. There are various sorts of semiotic system. Scholars may categorise the notion of semiotic system differently. For example, when studying translanguaging strategies of self-praise on Weibo, Ren and Guo (2022) distinguish between multimodal resource (e.g. photo, emoji, image macro, video), multilingual resource (including inter-linguistic resource such as Chinese + English, Chinese + Arabic numeral(s) and intra-linguistic resource such as Pinyin initialism and Pinyin), and multi-semiotic resource (such as hashtag and repeated punctuation marks). In this study, I view "semiotic system" as the broadest term, as understood in semiotics, which covers linguistic system and non-linguistic system. Yet, since my study focuses on TSEs, I shall categorise semiotic systems differently. Specifically, five different semiotic systems will be differentiated, namely logographic characters (including characters in Chinese and Japanese), Latin alphabet letters (including Chinese Pinyin, English letters), Arabic numerals, emojis, and other symbols (including mathematical symbols, @, punctuation marks, etc.). Photos, videos, and hashtags are generally not used in forming TSEs, as my study will show.

According to Christison and Murray (2020), there are five semiotic systems, which are audio, gestural, linguistic, spatial, and visual. Logographic characters, Latin alphabet letters, numerals, and mathematical symbols in my categorisation all belong to linguistic semiotic systems, while emojis are the combination of linguistic and visual systems. The reason I distinguish between logographic characters and Latin alphabet letters is that Chinese is a logographic language which uses characters to represent a word or morpheme. The written language and oral language in Chinese are separated from each other. One can understand the meaning of a character with having to read it aloud, and vice versa. With the development and publication of Pinyin in the 1950s, people can read a character based on the Pinyin without understanding its meaning. After decades of promotion and popularisation, people

learn both Chinese characters and Pinyin from the beginning. Nowadays, the typing system of computer and mobile devices is designed based on Pinyin system. In this study, Chinese characters and Pinyin are considered as two different semiotic systems as Chinese characters are logograms while Pinyin are phonograms. Therefore, the process of using Pinyin to input Chinese characters involves the transformation of two semiotic systems, which is actually in line with the definition of Trans-semioticising in this study. For example, when we say "霭" (means "mist" in English) and "矮" (means "short" in English) are homophones, we would provide the Pinyin spelling of each character, which is "ǎi". Also for characters with similar but not the same pronunciation, such as "矮" ("ǎi", which means short in English) and "爱" ("à", which means "love" in English), we also need to present the Pinyin of each character.

1.2 Aims and significance of the study

Based on a self-built database, this exploratory study aims to provide a multi-faceted categorisation of the TSEs people encounter on Chinese social media regarding the mechanisms of their formation, the types and combinations of semiotic elements involved, and the explicitness of their Trans-semioticising. Then, drawing on the think-aloud protocols of social media readers, it seeks to probe into the cognitive processes of TSE comprehension, a mysterious dark box that cannot directly be accessed from outside. In addition, it attempts to offer a Relevance-theoretical account of the comprehension process with special attention to various non-propositional effects resulting from the processing of the TSEs.

With these aims at hand, the present study may prove significant in various ways.

First, by fairly comprehensively categorising the TSEs people encounter on Chinese social media and generalizing their mechanisms of formation, this study will enlarge the research scope of lexicology from a multi-semiotic perspective and extend research on lexical-level Translanguaging from a cross-linguistic perspective to a multi-semiotic perspective. The contribution to the studies of word formation lies in introducing Trans-semioticising to the fields, which extends the scope of applicability

of traditional word formation theory and Translanguaging and breaks the wall between different languages and semiotic symbols, especially for logographic languages like Chinese, phonetic languages like English, and iconic symbols like emojis. Regarding research on Translanguaging, this study brings the focus of attention from the discourse level down to lexical level. Also, by investigating the formation of Chinese online TSEs, it reveals how elements from different semiotic systems work together to serve the communication and meaning-making purposes. The creativity of TSEs resides in the combination of the smallest units of languages and other semiotic symbols.

Second, this study will enrich Relevance Theory by demonstrating how the comprehension of the semantic meaning and the non-propositional effects in tackling TSEs is constrained by the presumption of relevance. The application of Relevance Theory enables the researcher to discuss why readers understand the TSEs in particular ways. In addition, this study expands the application of Relevance Theory to lexical pragmatics, as previous studies focus on the understanding of various underspecified lexical forms. However, in this study, Relevance Theory is used to explain the lexical processes of comprehension pertaining to TSEs, specifying semantic integration as a new lexical pragmatic process involved in understanding TSEs.

Third, from the perspective of methodology, this study distinguishes itself by employing think-aloud protocols in the study of people's understanding of TSEs. It involves participants in some tasks (reading or doing something that is designed for studies) and asks them to verbalise their thoughts while working on the tasks. This study, instead, narrows down the context of using think-aloud protocols and makes it a useful tool to showcase the readers' understanding processes that cannot easily be displayed directly. By applying think-aloud protocols, the present study is able to look into the understanding paths, strategies, and association mechanisms used by the readers.

Fourth, still methodologically, compared to the few studies that focus on

Translanguaging in online communication, this study provides a more macro and overall view towards the types of TSEs based on a great amount of data collected.

1.3 Organisation of the thesis

The thesis is composed of seven chapters.

In the first chapter, a brief introduction of the research problem of the study will be provided. It begins with the aims of the study, followed by the elucidation of the significance of the study, including the theoretical contributions made to the field of translanguaging, word formation, and Relevance Theory, and the methodological contribution to the study of people's understanding of words and expressions and the scope of application for think-aloud protocols. The chapter ends with the organisation of the thesis.

In Chapter 2, four strands of related literature will be reviewed. Firstly, previous studies on creative word formation is introduced based on Lipka's (2002) categorisation of innovative lexical formation, to provide a framework for my own categorisation of the formation mechanisms of the TSEs. Then I will review some previous studies on Translanguaging to lay a solid foundation for understanding TSEs conceptually. Section 2.3 turns to wordplay including semiotic play and multilingual wordplay to assist in my later discussion of the non-propositional effects of TSEs. Section 2.4 presents existing research on lexical pragmatic processing from the relevance theoretical perspective, serving as the background of my Relevance-theoretic approach to TSEs. Section 2.5 provides a synopsis of some relevant arguments of Relevance Theory (RT), which will serve as the theoretical framework for interpreting the comprehension of TSEs. Then, The last section summarises these strands of previous studies on the various issues and pinpoints the research gaps to be filled up. On the basis, the research questions for the present study will be proposed in order to fulfill the aims of the study.

Chapter 3 will describe the research methodology employed in this study. In the study, two types of data are collected from the Weibo super-topics and think-aloud tasks. In all eight super-topics about four different topics selected, 2670 different TSEs have been found. I also introduce how to analyse the TSE database, and how to categorise these expressions from different perspectives. Finally I touch upon the ethical issues that are related to this study at the end of this chapter.

Chapter 4 will categorise the TSEs identified on Chinese social media, particularly, Weibo super-topics, from three perspectives, namely the semiotic elements employed, word formation mechanisms involved, and explicitness of Trans-semioticising. A four-way categorisation is also proposed by following the word formation mechanisms raised by Lipka (2002). Nine combinations of TSEs are also identified based on the semiotic elements that are involved in the formation processes. Finally, implicit TSEs and explicit TSEs are differentiated, depending on the transparency of the semiotic elements used.

In Chapter 5, the details of understanding TSEs will be elaborated and discussed from three perspectives. Firstly, three understanding paths are identified and illustrated on the basis of the think-aloud protocols of 22 participants. Then, eight understanding strategies are summarised and illustrated with examples. Section 5.3 presents how the routes of understanding may vary according to the word formation types of TSEs and readers' familiarity levels.

Chapter 6 will discuss TSEs' understanding from the theoretical perspective of Relevance Theory. Section 6.1 uses examples to demonstrate how the readers' lexical semantic process is constrained by the presumption of optimal relevance. Section 6.2 focuses particularly on how integration as a lexical pragmatic process distinct from either broadening or narrowing is constrained by the presumption of relevance in the course of TSE comprehension. Section 6.3 discusses the non-propositional effects achieved at the cost of additional effort by the participants during the understanding processes. Based on the discussions in previous sections, as well as the participants'

think-aloud protocols, Section 6.4 generalises a Relevance-theoretic framework of TSE comprehension.

Chapter 7 summarises the major findings of the present study, discusses its theoretical, practical and methodological implications, points out some limitations of the study, and offers some directions for future research.

Chapter 2 Literature Review

This chapter provides a general survey of existing literature to mark the point of departure for the present study. Considering that Trans-semioticising expressions (TSEs) are creative in their formation, it will begin by reviewing the previous literature on various types of creative word formation in section 2.1. Then, Section 2.2 will present a review of related literature on Translanguaging, which is part and parcel of Trans-semioticising treated in this study. Considering that Trans-semioticising is to be discussed as a form of wordplay, Section 2.3 will review previous research that pertains to wordplay. In Section 2.4, a brief account of lexical processing mechanisms will be offered, as this study is also concerned with the lexical processing of Trans-semioticising expressions. To pave the way for the discussions of TSE comprehension in Chapter 6, Section 2.5 presents a sketch of Relevance Theory, particularly those theoretic arguments that help to account for the comprehension of TSEs. The chapter will conclude by proposing the research questions to be dealt with in the present study after summarising the research gaps emerging from the aforementioned venues of literature review.

2.1 Research on mechanisms of creative word formation

In this study, an expression is considered creative if it deviates from regular or conventional ways of word formation and meaning making as implicitly or tacitly shared by native speakers. Compared with regularly formed words or expressions, the innovative word or expression reflects the users' creativity in using linguistic or other semiotic resources to construct new lexical forms or make new meaning. For example, the use of "i" to replace "I" for first-person reference (as the American poet e.e. cummings did) is creative because it breaks the rule of capitalisation. Similarly, Translanguaging expressions (TEs) such as "扌丁 call" are also creative because their

formation transcends the boundary between languages. Thus, lexical creativity is different from lexical productivity, because the latter refers to the extent to which a certain mechanism of formation can be used to form new words. For example, "-er" as a suffix meaning "someone who does something" in English is more productive than "-ist" because the former is found to be part of significantly more nouns than the latter.

According to Lipka (2002, p. 109), there are four mechanisms for lexical innovation or creative lexical formation: a. morpho-semantic neologisms, such as affixation, composition, and phonetic motivation; b. semantic neologisms, or semantic transfer (ST), including metaphor and metonymy; c. morphological neologisms, in other words, reduction processes like clipping, blending, and acronymy; and d. external neologisms, such as loanwords.

2.1.1 Morpho-semantic neologisms

Morpho-semantic neologism falls into three main types, namely affixation, composition, and phonetic motivation (Lipka, 2002).

Affixation is the most frequently used method of forming new words in English. By adding a prefix or/and a suffix to another word, we give it a new meaning (O'Dell, 2016). Some words created by means of affixation are not creative because the morphological process conforms to the general rule or convention that native speakers share. For example, the Latin prefix "de-" meaning "the opposite of" or "remove" is generally used before verbs, adjectives or nouns, such as "decentralise", "defrost", "debug" and "derisk". The nouns used in the affixed words generally denote something negative. However, words like "defriend" is creative in the sense that the affix is applied to a noun that denotes something positive. Similarly, in English, the prefix "un-" is generally affixed to a verb. However, in the words "un-birthday" and "un-cola", it is attached to a noun. Therefore, it is a case of creative affixation.

Composition or word-compounding refers to the word-building practice where

new words are produced through combining two or more stems (Gontlšarova, 2013). Some classic examples of composition are “dining-room”, “blackbird”, “mother-in-law”, etc. From the semantic perspective, Antrushina *et al* (1999) divide composition into three types: 1) compounds whose meaning is the sum of their composite meaning (e.g. “classroom”, “bedroom”, “working-man”, “dining-room”); 2) compounds with the meaning of one or two of the components changed (e.g. “blackboard”, “blackbird”, “football”); 3) compounds with irreversibly lost meaning (e.g. “ladybird” is not a bird, but an insect; “tallboy” is not a boy but a piece of furniture). As in the case of affixation, some composition cases are creative whereas others are not. Unlike the examples given above which are regular compounding cases, “job-job” and “friend-friend” are creative ones because they involve duplicating a noun to convey the meaning of typicality. Thus, “job-job” means “a typical job”; “friend-friend” means “a typical friend”. In addition, in “Don’t you do-you-mind me!”, “do-you-mind-me” is also a creative compound verb because it is made by conjoining the words of a sentence rather than two or more semantically related but syntactically unrelated words.

Phonetic motivation is realised by building the relation between the meaning and phonetic form (Zheng, 2015). In Chinese, new words with phonetic motivations are created by means of homophonic word-formation, which refers to neologisms made by using homophones with similar pronunciation but different meanings. For example, some Chinese netizens use “神马” to mean “什么” or use “有木有” to refer to “有没有”. These homophonic forms are creatively made as they differ from regular homophonic forms such as “送钟” (meaning “giving a clock as a gift”, for “送终”, meaning “pay sb. last respects”) because the former are not full lexical items but results of malapropism, i.e., deliberate mispronunication.

2.1.2 Semantic neologism

Semantic neologism, also called semantic transfer, refers to the assigning of new meanings to existing word forms. Peprník (2006, p. 113) points out that people can give "a new or additional meaning to the existing lexical forms". In semantic neologisms, the original meaning can be replaced or co-exist along with the new meaning (Peprník, 2006). One situation for semantic transfer to take place is metaphor, as the meaning of certain words such as "fox" and "pigeon" is expanded so as to create a neologism (Aduda, 2013). Another situation is metonymy, as in the case of "hand" used to refer to "a person like a worker". Nunberg (2004, p. 344) defines semantic transfer as a process where speakers "use an expression that denotes one property as the name of another property". The semantic properties are the components of meanings of words. Some cases of semantic neologisms can be creative when there seems to be no semantic similarity or resemblance between the original meaning of the word and the newly derived meaning. For example, in Chinese, the expression "山寨" used to mean "cottages on mountainous areas", but now it can be used to mean "imitation or copy, or product of imitation or copy". We can hardly detect any semantic similarity between the new meaning and the original meaning of the word. another commonly heard example is "恐龙", which is used creatively to refer to those who look ugly rather than dinosaurs.

2.1.3 Morphological neologism

Morphological neologism refers primarily to the process of blending whereby parts of two words are mixed to form a new word. Blending has become increasingly popular and many linguists are devoted to the examination of the properties of blends (Lehrer, 2007, p. 115). According to Lipka (2002), blending is one of the morphological neologisms which reveals the creativity of word-formation.

Blends are defined as "underlying compounds which are composed of one word and part of another, or part of two (and occasionally three) other words. The word part

is called Splinter, which is a clipping but cannot appear alone as a word." (Algeo, 1977, p. 48). For instance, in "dramedy < drama + comedy" (Lehrer, 2007, p. 116), the two splinters "dram" and "edy" themselves are not existing meaning-making words. They cannot be used in sentences like "I like that dram player". In infotainment < information + entertainment (Lehrer, 2007, p. 116), "info" as a clipping for "information" can be used as an independent word, but "-tainment" is a suffix which needs to be attached to something else (e.g. "entertainment"). According to Lipka (2002), blending is one of the processes which typically reveal the creativity of word-formation. A most creative way of blending is based on the sharing of the same part of two words, as in "sexploitation", in which "ex" is part of both "sex" and "exploitation". Lehrer (2007) introduced five types of blend structures in his book. The first common type is a full word followed by a splinter, like *wintertainment* < *winter* + *entertainment*. The second and third type are blends consisting of two splinters with either 1) the beginning of one word followed by the end of another, such as *psychergy* < *psychic* + *energy*; or 2) both splinters being the beginning of words, such as *biopic* < *biographical* + *picture*. The fourth type of blend involves complete overlap of one or more phonemes, often of whole syllables. For example, *clandestiny* < *clandestine* + *destiny*, and *wheatables* < *wheat* + *eatable*. The fifth type is less common, which involves a discontinuous elements. A word or clipping is embedded in part of another sources word as an infix, like *entreporneur* < *entrepreneur* + *porn*(*ography*).

Another common type of morphological neologism is initialism or acronymy by means of clustering the initial letters or characters of a phrase or several phrases. Some cases can be creative, as they run counter to what people generally take for granted. For example, when "三高" is used to refer to 高血压、高血脂、高血糖, it is a regular contraction form. But when it is used to refer to 高学历、高身材、高薪金, it is creative because it goes against the general usage of the form to refer to some unhealthy physical conditions.

2.1.4 External neologism

External neologism, mainly in the form of loanword, categorised by Lehrer (2007), is another way of making new and even creative expressions. By loanword (or lexical borrowing), it refers to "a word that at some point in the history of a language entered its lexicon as a result of borrowing (or transfer, or copying)" (Haspelmath, 2009, p. 36). For example, in China, people sometimes use "PK" as a creative or novel way to mean "compete". Actually, the expression is not a Chinese expression but is an initialism (contracted from "play" and "killer") that comes from English.

2.2 Research on Translanguaging

2.2.1 Defining the notion of Translanguaging

By Translanguaging, it was originally used to refer to "a pedagogical practice in bilingual classrooms where the input (e.g. reading and listening) is in one language and the output (e.g. speaking and writing) in another language" (Li, 2011, p. 1222). More broadly, it refers to communicative practice in which the boundaries between traditional named languages such as English, French, Chinese, and Japanese are blurred and replaced by a semiotic system integrating lexical, morphological, and grammatical linguistic features to serve communication aims (Vogel & García, 2017). In other words, it is assumed that there are no independent languages that are separated from each other. Typically, it is realised by the combination of elements from two languages such as "期待 ing" ("expecting"). García and Li (2014) point out that broadly speaking, translanguaging involves the mixture of not only linguistic resources, but also multimodal, multisensory and multi-semiotic resources. Indeed, we may find innovative expressions, sentences or discourse that are also realised by the combination of linguistic element(s) and non-linguistic element(s) such as "I ❤ London". Translanguaging may take place at lexical, sentential and discoursal levels. "I ❤ London" is an example of sentence-level Translanguaging if it appears as part of a text, but an instance of discourse-level Translanguaging if it occurs alone as a

slogan. Of course, discourse-level Translanguaging can be much longer and more complex, as found in a classroom context. It is not a one-shot production but operates across turns or sequences of utterances.

Researchers tend to use Translanguaging as a descriptive label for language mixing practices such as in classroom, daily communication, and social media (Li, 2022). With Translanguaging welcomed by more and more researchers, this notion gradually evolves from a pedagogical strategy, a communicative strategy, to a practical theory of language.

2.2.2 Translanguaging as a pedagogical strategy

The broad discussion of Translanguaging phenomenon originates from Welsh bilingual education in the 1980s (Lewis *et al*, 2012). Cen Williams coined the term of "Trawsieithu" in Welsh, which was later translated into English as "Translanguaging". It was first constructed as a cross-curricular strategy for "the planned and systematic use of two languages for teaching and learning inside the same lesson" (Lewis *et al*, 2012, p. 3). Li (2018) discusses the pedagogic potential of Translanguaging and considers that it can be beneficial to language teaching and learning. In studies of multilingual classrooms, researchers begin using the term of "Translanguaging" to describe oral interactions and written texts that involve the mixed use of more than one language (Creese & Blackledge, 2010; Canagarajah, 2011; Cenoz & Santos, 2020; García, 2009; García & Kano, 2014; Setyaningrum, Setiawan & Anam, 2022). Translanguaging is considered to be a pedagogy that may support language learning and communication between bilinguals and multilinguals.

Recent developments in the field of Translanguaging, which spans from bilingual education to translingual and transmodal communication, focus mainly on language teaching and learning in bilingual and multilingual contexts.

2.2.3 Translanguaging as a creative strategy of communication

Translanguaging can be used not only as a language learning scaffold, but also as a creative strategy to make special meaning in communication. For example, Baker (2011, p. 288) defines Translanguaging as "the process of making meaning, shaping experiences, gaining understanding and knowledge through the use of two languages". From this point of view, studies on novel words from the perspective of Translanguaging blur the boundaries between languages and investigate on words that contain more than one language (e.g. Chen, 2014; Li, 2016a, 2016b; Li *et al*, 2020; Li & Zhu, 2019; Zhang & Ren, 2022).

The Internet has become an important space for people to communicate with each other, which arises great attention by the researchers who are interenting in online Translanguaging expressions and collecting data from the social media platforms (Chen, 2014; Li *et al*, 2020; Li & Zhu, 2019; Zhang & Ren, 2022). Compared to the language used in traditional face-to-face communication, the kind of language used by people on the Internet has developed distinctive features. As Crystal (2006) argues, the language used online is more like speaking than writing. To be more specific, when people communicate digitally, they are allowed to feedback and respond to others immediately and spontaneously, which is similar to face-to-face chatting (Baron, 2003). Under these circumstances, people develop the desire to use all possible resources they can utilise at hand to "conserve energy and space" (Baron, 2003). Correspondingly, semiotic symbols including acronyms (Rúa, 2007), emojis (Dürscheid & Siever, 2017; Ortner, 2013), and numbers (e.g. Qi & Zhang, 2020; Yang, 2007; Yuan, 2011; Zhang, 2017), which can be easily accessed, have been creatively used to form new expressions.

Li and Zhu (2019, p. 153) investigate a particular type of Translanguaging expressions, which they propose to call "tranßcripting". It is defined as "creating a script with elements from different writing systems, such as Chinese and English, or by mixing conventional language scripts with other symbols and signs including emoji" (ibid). They discuss three types of tranßcripting practices regarding Chinese and English, including Chinese + English (e.g. 无 fuck 可说), Chinese characters +

alphabetic letters (e.g. ) , and Chinese characters + numerals like 森 7. However, the notion of Tranßcripting is limited by the idea of “mixing scripts”, which to some extends focuses more on the writing form of the creative expressions. As we can tell from the examples raised by Li and Zhu (2019), they discuss mainly about Translanguging expressions that contain different semiotic elements in their written forms, which are actually referring to a particular type of the TSE that will be discussed in this study, namely, explicit TSE (definition will be provided later). Moreover, when classifying Tranßcripting expressions, Li and Zhu (2019) only identified three types of semiotic elements, which are Chinese character, alphabetic letters and numerals. While in this study, elements from more semiotic systems will be analyzed. Therefore, Tranßcripting cannot cover the target expressions that will be discussed in this study, as Trans-semioticsing is actually a larger term that involves more semiotic systems and more types of written forms.

Zhang and Ren (2022) investigate the use of a popular online expression “skr” on Micro-blog, considering its usage as a TE practice. By analysing the 1861 instances of “skr” used by Micro-blog users, they sort out the expressions containing “skr” by frequency and conclude eight types, namely phonetic substitution of Chinese, affective exclamation, adjective as appreciation, hip-hop element, intensification device, searchable hashtag, metonymy, and onomatopoeia. Zhang and Ren’s categorisation provides some types of TL expressions such as 欧了 that traditional word-formation that cannot account for.

Sultana (2014) finds that some English morphological changes in Bangladesh, such as the current continuous marker of the plural suffix -*s* mixed with Bangladeshi root words, produce new words. Bengali speaking young people add the English plural suffix -*s* to Bengali words (such as bonds’ friends).

Li, Tsang, Wong and Lok (2020) investigate the TE practices in *Kongish Daily*, a Facebook page which makes use of tranßcripts like traditional Chinese characters, Romanisation and made-up characters, simplified Chinese, pinyin, English, Hong Kong English, emoji, etc. Through analysing the formation and meaning making

process of the Tranßcripts, they conclude that the TE practices suggest the creativity and political criticality behind the news items.

As more and more research has been done, scholars come to the consensus that Translanguaging is not to be viewed as "an object or a linguistic structural phenomenon to describe and analyse but a practice and a process—a practice that involves dynamic and functionally integrated use of different languages and language varieties, but more importantly, a process of knowledge construction that goes beyond language(s)" (Li, 2018, p. 15).

2.2.4 Translanguaging as a practical theory of language

The concept of Translanguaging as proposed by Li (2011) is derived from the notion of Languaging (Becker, 1991), which supposes that 'there is no such thing as Language, only continual languaging, an activity of human beings in the world' (p. 34), and the argument that we should view language "as an accomplished fact, as a thing made and finished, but as in the process of being made" (Gasset, 1957, p. 242). On the new view, there are no so-called divides between the linguistic, the paralinguistic, and the extralinguistic dimensions of human communication. Rather, communicators integrate "their feeling, experience, history, memory, subjectivity, and culture" (Li, 2018, p. 17) in the use of language. The theory of Translanguaging thus offers a new angle of investigating expressions that contain elements from different semiotic systems (including different languages). These expressions break down the boundaries between languages and other semiotic systems as well from a rather micro level, which happens during the formation and understanding processes of these expressions.

Language users, particularly bilinguals or multilinguals, have a "Translanguaging Instinct" (Li, 2016b, 2018). In essence, Translanguaging is "a transformative, resemiotisation process, whereby language users display the best of their creativity and criticality" (2018, p. 22). To capture the process, Li (2011) comes up with Moment Analysis in his study of Translanguaging in communication between

bilinguals. He claims that this methodology is a paradigm change, “away from frequency and regularity oriented, pattern-seeking approaches to focus on spontaneous, impromptu, and momentary actions and performances of individuals” (p. 1224). The methodology focuses on how bilingual participants combine elements from different named languages and from other semiotic systems to create innovative expressions and the consequences of actions and re-actions by the communicators. The meaning and sense-making at a special moment during the communication (i.e. online and face to face) can be investigated through this method (Li, 2022). It is argued that the moment that is worth analysing should be creative while this moment should show the criticality of the communicators. Different strategies, tools, and techniques can be used to investigate the moment, such as observation (O’Neil & Roberts, 2019), Interpretative Phenomenological Analysis (IPA) (Smith & Osborn, 2015), and Looking, Listening, Talking, and Thinking (LLTT) (Li & Zhu, 2013).

Moment Analysis used in analysing Translanguaging phenomena is significant because it can transform the way researchers handle applied linguistics research (García *et al*, 2021). However, Moment Analysis still has its own shortages that are not a suitable method that can be taken in this study. First, Moment Analysis focuses more on the meaning making processes when bilinguals communicating with each other. Researchers devote efforts on how the meaning of a particular Translanguaging expression or sentence being made sense between the speaker and listener on a macro level. It cannot go to a deeper level on specific understanding processes. Secondly, tools like observation, IPA, and LLTT, can provide some general perceptions made by both the researchers and participants about when they notice the creative expressions and how they think the meaning of these expressions. However, while emphasizing on “moment”, these methods cannot offer a real-time data that happens exactly on that moment. These methods rely on either the researchers’ objective observation or on afterwards interviews. Thirdly, Moment Analysis also focuses on identifying the creativity and criticality of Translanguaging. The effects brought by using Translanguaging are in fact more complicated, which requires a theory that provides a

more all-round discussion about the cognitive effects. Last but not least, the understanding processes of different types of TSE cannot be revealed by Moment Analysis based on Think aloud protocols that being introduced to this study.

2.3 Research on creative word formation as wordplay

As has been pointed out (Tebaldi, 2019), Translanguaging often involves a kind of wordplay to achieve the effect of mockery. To get a deeper understanding of how TSEs on Chinese social media serve as a kind of wordplay to generate some special effects and how they differ from the kind of wordplay practice discussed in previous literature, it is necessary to review related studies.

2.3.1 Defining wordplay

Wordplay is a unique but exciting topic in the context of linguistic research. Usually, wordplay, as a topic of interdisciplinary study, focuses on how playing with words reveals the language used in a humorous form through manipulating lexical units from formal and semantic manipulation perspectives. A typical practice is the use of puns.

Wordplay can help achieve specific, practical purposes in different contexts. It can be used to enact different linguistic functions (e.g. Kabatek, 2015; Partington, 2009; Renner, 2015; Thaler, 2012). Its first and most crucial function is to produce a humorous effect and amuse the hearers during communication. For example, Monsefi and Mahadi (2016) explore how wordplay is employed in English online news headlines to attract readers. Besides, wordplay can be used to achieve an even broader range of other social functions from personal image construction (like admiration of the speaker's wisdom, creativity, etc.), to social relation-building (such as in-group/out-group identity establishment) (Lloyd, 2007). Moreover, it can serve some very practical functions. For example, In advertising slogans or discourse, wordplay is a means to enhance the memory of the speaker's words and a means of persuasion (Tanaka, 1992). Sometimes, teachers use wordplay in classroom instruction to increase students' motivation or attention (Tai & Li, 2020).

Recent research on wordplay is motivated by a fundamental change in the way language is understood; language is not considered a static and fixed entity anymore, but as a flexible and dynamic resource (see Blommaert, 2010). This conceptual change is due to the increasing trend of language globalisation and is characterised by a more transient or fluid social relationship (Baumann, 2000).

2.3.2 Multilingual wordplay in Translanguaging practices

Sometimes, wordplay may serve as a cross-linguistic bridge that transcends language boundaries and connects linguistic elements from different tongues, as found in Translanguaging practices. In their book, Knospe, Onysko, and Goth (2016) discuss multilingual wordplay in a great variety of contexts such as newspaper language, everyday communication, Internet forums, and literacy works. According to Dovchin (2016), multilingual wordplay happens in daily language practices, especially among the younger generations. Young people integrate a variety of multilingual resources into their language practice and create their own mixed forms, to have fun or be creative (Godin, 2006).

One typical multilingual wordplay strategy is to borrow words or expressions from different languages used in their multilingual community (Schoonen & Appel, 2005, p. 88). Doran (2004) introduces a popular language case called Verlan, which is used among the Parisian young generation. The study shows that this alternative language code becomes a symbol of marginalised young people who stand outside the traditional Parisian cultural norm. Leppänen *et al* (2009, p. 1099) provide another bilingual wordplay with mixed forms associated with digital communication. The overall way of Finnish online wordplay is to integrate various nouns and verbs used in English extreme sports into the grammar system. What is more, some English words are partially changed to fit the Finnish vocabulary custom (Leppänen *et al*, 2009, p. 1099). Dovchin (2016) conducts a study on hybrid forms in Mongolian Facebook posts through participant observations and causal group discussions and hangouts. These Facebook users are bilinguals who can use various linguistic resources,

including English and Russian. The data show that languages such as English and other additional languages, such as Russian, are creatively and interestingly mixed with the phonetic, lexical, and syntactic systems of Mongolian. Mongolian Facebook users tend to "create a new type of multilingual wordplay by which they reference locally relevant meanings and accommodate their linguistic practices" (2016, p. 97) with multilingual resources. Three types of creating multilingual wordplay are summarised by Dovchin (2016), which are creating through English names, through Facebook linguistic features, and Russian linguistic resources.

2.4 Research on lexical comprehension

In communication, many lexical items do not encode fully-fledged or clearly defined concepts, which may affect the comprehension of the utterances in which they occur. To understand the utterances properly, it is necessary to undertake the interpretation of the underspecified lexical items concerned. In the existing literature, two cognitive processes have been well documented. One is narrowing and the other broadening (Carston, 2002; Wilson, 2004; Sperber & Wilson, 2008; Wilson & Carston, 2006, 2007). These narrowing and broadening mechanisms apply when a lexically encoded concept triggers a pragmatic process whose outcome is either a narrower or broader lexical concept.

2.4.1 Lexical narrowing

Lexical narrowing refers to the process of "*using a word or a phrase to convey a more specific concept (with a narrower denotation) than the linguistically encoded meaning*" (Wilson& Carston, 2006, p. 14). Here are some examples of lexical narrowing:

(1) All politicians *drink*. (Wilson & Carston, 2006, p. 409)

Here, when understanding "drink", we need to narrow the scope of what is drunk so that we know that the speaker means drinking alcohol rather than any liquid.

(2) I have a *temperature*. (Wilson, 2004, p. 344)

Here, “temperature” does not refer to any degree of temperature but rather refers to a higher body temperature than expected in a healthy human being.

(3) a. The *fish* attacked the swimmer. (Moreno, 2007, p. 44)

b. The *fish* was nice but the potatoes were cold. (Moreno, 2007, p. 44)

In (3a), the “fish” refers to some dangerous fish instead of other fish as it attacks swimmers. In (3b), the “fish” here refers to fish that can be eaten other than those inedible.

In cases (1) and (2), the concepts communicated are understood to be more specific than their lexical meaning. In cases 3 (a) and (b), the concepts of “fish” are narrower than their lexical meaning as they are used to refer to a certain kind of fish.

2.4.2 Lexical broadening

Lexical broadening involves “the use of a word or phrase to convey a more general concept (with a broader denotation) than the linguistically encoded meaning” (Wilson, 2006, p. 15). Depending on the way an ad-hoc concept goes beyond the conventions of meaning, broadening cases can be categorised into four types, namely approximation, category extension, hyperbole and metaphor (Wilson, 2004; Sperber & Wilson, 2008; Wilson & Carston, 2007). Here is an illustration of various types of lexical broadening:

A. Approximation(examples from Wilson, 2004, p. 345)

(4) This coat cost *1,000 dollars* [meaning ‘about 1,000 dollars’]

(5) The stones form *a circle* [meaning ‘approximately a circle’]

(6) The injection will be *painless* [meaning ‘nearly painless’]

What is more, approximation varies in degree across different contexts. Taking the adjective “flat” for an example:

(7) *a. My garden is flat.*

b. My country is flat. (Wilson, 2004, p. 345)

Obviously, the degree of flatness in (7a) is much higher than that in (7b).

B. Category extension

It refers to the use of salient brand names like *Kleenex* or people's names and some common nouns. A detailed explanation on broadening is provided by Sperber and Wilson (2008, p. 91, 94), "category extension ... involves extending a word with a relatively precise sense to a range of items that clearly fall outside its linguistically specified denotation" by projecting "defining, or at least characteristic, properties of the encoded concept onto a broader category". Here are two examples from Sperber and Wilson (2007, p. 48; 2008, p. 93):

(8) *(Handing someone a tissue): Here's a Kleenex.* (Sperber & Wilson, 2008, p. 91)

(9) *Žižek is another Derrida.* (Sperber & Wilson, 2008, p. 93)

C. Hyperbole

Compared with approximation, it appears to involve more substantially broadening of the encoded lexical concept. For example, in the sentence of "You are a genius!" (Moreno, 2007, p. 48), "genius" is broadened to refer to someone cleverer than expected other than its lexical meaning which is to describe that someone's natural mental ability is very great.

D. Metaphor

It involves that type of broadening "on the basis of relatively peripheral or contingent properties" (Moreno, 2007, p. 48), such as using "butterfly" to compare to someone who is beautiful but so delicate as to be hurt like a butterfly.

Metaphor and hyperbole can be regarded as consisting in an obvious violation of the Quality Maxim of Grice's (1967, 1989) Cooperative Principle (the other three maxims are Quantity Maxim, Relation Maxim, and Manner Maxim) because the former uses metaphor to implicate a related simile or comparison and the latter uses hyperbole to implicate a weaker but related proposition (Grice, 1967, 1989). In order to unify the two main distinct processes, researchers turn to Relevance Theory to find some common ground underlying the processes, and come up with two main claims (Sperber & Wilson, 2012b). First of all, literalness hypothesis does not hold: in the process of pragmatic interpretation, the context informational available to both sides

of communication is usually used to adjust the meaning of words. On the other hand, the range of usage from approximation to metaphor (such as exaggeration and metaphor) is expanding. These usages all involve the same interpretation mechanism and can be interpreted in the same way.

2.5 Relevance Theory

Relevance Theory was proposed by Sperber and Wilson (1986, 1995) as a Post-Gricean model to explain how people communicate meaning in context. It is built around only one of Grice's (1967, 1989) four maxims under the Cooperative Principle, namely Relation Maxim, and treats relevance as a constant cognitive constraint, which guides the dynamic search and construction of an appropriate context in which to process utterances.

2.5.1 Defining relevance

In their theory, relevance is seen as a property of inputs (including utterances, thoughts, memories, actions, sounds, sights, smells and so on) to cognitive processes:

> *that human cognition has a goal: we pay attention only to information which seems to us relevant. To communicate is to claim someone's attention, and hence to imply that the information communicated is relevant. Thus, a single property - relevance - is seen as the key to human communication and cognition (Sperber & Wilson, 1986, back cover).*

Sperber and Wilson's definition of relevance as a property of utterances is based on a few simple assumptions:

a. Human cognition is relevance-oriented. In the procedure of communication, we always pay attention to information that seems relevant to us.
b. Every utterance has a variety of possible interpretations, all of which are compatible with the information that is being linguistically encoded.

c. Not all these interpretations occur to the hearer simultaneously; some take more effort to create.

d. Hearers are equipped with a single, very general criterion for evaluating interpretations as they occur to them.

e. This criterion (i.e. relevance) is sufficiently powerful to exclude all but, at most, a single interpretation, so that once having found an interpretation that satisfies it, no further interpretations will be produced.

The notion of relevance is further explicitly defined in terms of both the cognitive effects an utterance can yield in the context and the processing effort needed or expended to obtain these cognitive effects (Sperber & Wilson, 1986/1995a). Thus, an assumption is relevant in a context to the extent that its cognitive effects in this context are substantial, and an assumption is relevant in a context to the extent that the effort required to process it in this context is small. What makes an input, or more specifically an utterance, worth processing is that its cognitive effects are substantial enough and its processing cost is reasonably small.

According to RT, there are three main types of cognitive effects, namely:

A. Strengthening

The new information conveyed by the current utterance provides further evidence for, and therefore strengthens, old assumptions.

B. Contradiction

The new information conveyed by the current utterance may provide counter-evidence for some old assumption, which could potentially lead to abandoning it.

C. Contextual implication

The new information conveyed by the current utterance is combined with the old information to yield a contextual implication -- an implication which is derivable from neither the old nor the new information alone. This is when we derive a conversational implicature.

While the three types of cognitive effects are propositional, there are non-propositional cognitive effects as well (Yus, 2017, 2022), when metaphorical, hyperbolic, ironical and other figurative utterances are processed. These effects are usually some vivid associations or emphases that are triggered by the processing of the metaphors, hyperbole, or irony involved.

Regarding processing effort, RT defines it in terms of the number of inferential steps required to derive the cognitive effects, often depending on the accessibility/amount of related background assumptions necessary for the interpretation of the utterances.

Relevance is not only a qualitative notion but a quantitative concept as well. In the same context, different utterances, as inputs, may vary in their degree of relevance (in other words, not all of them are optimally relevant in the context). Suppose Peter knows he is sick and goes to see the doctor. Imagine that the doctor uses some of the following:

(1) a. You've got a cold.

b. You've got a cold and look feverish.

c. You've got a cold and China has a long history.

(1a) is relevant to Peter because it can strengthen his existing assumption that he is ill. (1b) is relevant in a different way because it not only strengthens his existing assumption that he has got a cold, but also adds that he may have a fever. To process the latter part for the cognitive effect, the hearer needs to expend additional processing effort. (1c) is an interesting case. While the first part is highly relevant, the second part is not relevant at all, in the given context. The processing of this will incur a loss of effort and disappoint the processor. Therefore, (1c) is generally rejected as an impossible utterance. As rational and efficient information processors, human beings tend to allocate attentional resources to information that seems relevant and rewarding.

2.5.2 Presumption of optimal relevance

As a cognitive pragmatic theory, RT is intended to uncover the cognitive processes that people follow in interpreting utterances and in communicating with each other. It distinguishes between two principles of relevance (cognitive and communicative principle) (Sperber & Wilson, 1986/1995a, p. 260):

The cognitive principle: human cognition tends to be geared to the maximisation of relevance;

The communicative principle: every act of ostensive communication communicates a presumption of its own optimal relevance.

While the former principle governs our general cognition that is not affected by immediate communicative pressure, the latter principle guides our communication in a specific context in which economy of effort for both speaker and hearer is normally expected.

According to Relevance Theory, relevance is determined by both cognitive effects of an utterance and the effort needed to process the information the utterance contains. In communication, we do not pursue as many cognitive effects as possible (i.e. maximal relevance from the perspective of cognitive effects) without considering the processing cost. Rather, we pursue optimal relevance, which is determined not by effects or effort alone, but by the balance between the two. This is what is meant by the presumption of optimal relevance. Thus, we generally expect that every utterance conveys the presumption of its optimal relevance.

Sperber and Wilson (2002, 2005) argue that human cognitive system consists of a lot of domain-specific capacities whose function can be activated when people need to deal with certain problems through a fast and efficient computation. As defined by Carston (2010, p. 2), the pragmatic capacity is a dedicated system: “its specific

domain is ostensive stimuli (verbal utterances and other acts of ostensive communication) and the comprehension procedure it employs is a fast and frugal heuristic". The procedure of comprehension can be illustrated as the following steps:

> *for any ostensive stimulus, assess hypotheses about its interpretation in order of their accessibility and accept the first one that meets current expectations of relevance (where occasion-specific expectations of relevance are regulated by a general presumption of 'optimal relevance' conveyed by all instances of ostensive communication)* (Carston, 2014, p. 3)

2.5.3 Context construction as relevance guided

In Relevance Theory, context is not pre-given or something out there. Rather, it is dynamically constructed for processing utterances. The construction process is guided by the presumption of optimal relevance.

According to RT, the context constructed for processing utterances is a set of assumptions coming from different sources. One type of assumption comes from the perception of the surrounding physical setting. Another type of assumption is derived from processing preceding utterance(s) and still stored in the working memory. The third type of assumption stems from the long-term memory. To access a certain portion of our long-term memory or background knowledge, we depend on the processing of words encoded in the current utterance. A word encodes the semantics of a concept which embodies an address in human memory system that offers access to information stored in the mental environment. The access to the information can be reached through three entries: logical, encyclopedic and lexical (Walaszewska, 2010). Logical entry is a small, finite and relatively stable entry, independent of speakers and times. Encyclopedic entry has the messages about the concept's extension and/or denotation of the concept including "*expert and lay assumptions, cultural beliefs and*

personal experiences stored in the form of propositional representations, assumption schemas, prototypes, scenarios or scripts and mental images" (Walaszewska, 2010, p. 316). This entry is usually open-ended and varies with speakers and times. Lexical entry has to do with the lexical properties like the phonetic structure and the grammatical properties of a word representing a concept (Carston, 2002, p. 321).

Thus, RT provides an account of how certain background information is activated through encoded concepts and retrieved to serve together with other assumptions as the context for processing the utterance. However, RT further claims that the activation and retrieval of certain background information is guided or constrained by the presumption of optimal relevance. As numerous pieces of background information can be stored under the encyclopedic entry of the same concept, only the very piece of information that can help to generate a substantial cognitive effect without incurring too much processing cost will be chosen.

2.5.4 Weak implicature and "non-propositional effects"

According to Relevance Theory, utterance interpretation is a two-phase process: a decoding phase and an inferential phase. The former provides input to the latter "in which a linguistically encoded logical form is contextually enriched and used to construct a hypothesis about the speaker's informative intention (Wilson & Sperber, 2012a).

The speaker's informative intention can be some literal meaning intended by the speaker and derived by the hearer through pure decoding, or some strong implicature (i.e., implicature in Gricean sense), plus some weak implicature(s) (Sperber & Wilson, 1986/1995a). Some weak implicatures, which bring about a wide array of weak cognitive effects, may be anticipated or intended by the speaker, and some may not. The hearer can actively decide the content of the weak cognitive effects, although the recovery of them is triggered by the speaker (Sperber & Wilson, 2012). The greater the room allowed to the hearer in processing an utterance for weak implicatures, the

wider the range of such cognitive effects. Sperber and Wilson (1995: Chapter 4, section 6; 2012b) identified the cognitive effects achieved by conveying weak implicatures as "poetic effects" or "non-propositional effects" (Sperber & Wilson, 1995, p. 228). The notion is similar to what we generally call "rhetorical effects", but is inclusive enough to include all kinds of unencodable cognitive effects ranging from aesthetic ones to social ones that are mentally entertained and processed (Chen, 2014). Yus (2017, 2022) also used the term "non-propositional effects", on the grounds that such effects are different from strong implicatures that are propositional in nature.

According to Relevance Theory, "[t]he production of genuinely relevant poetic effects can be a powerfully creative form of language use (creative on the part of both communicator and audience)" (Sperber & Wilson, 2012a, p. 118). There are many ways of communicating unencodable poetic effects or non-propositional effects. One of the commonest ways is the use of creative metaphors. Suppose a mother says "You're a piglet!" to her child. Apart from conveying the strong implicature that her child is a dirty child, the mother also communicates some weak implicatures, "in a way that makes it quite appropriate to talk of "poetic effects" (Sperber & Wilson, 2012a, p.121). She might intend to convey weak cognitive effects such as the impression of a naughty girl and the feeling of dissatisfaction. In turn, when hearing her mother's words, the daughter might also derive one or both of the weak cognitive effects and even some other poetic effect such as her mother's love for her.

2.6 Research questions

The literature review above started with research on creative word-formation mechanisms as methods of analysing the constructions of novel expressions. It then showed previous research on various aspects of translanguaging, which is most relevant to the present study, such as the origin, definition, functions and method of analysis. These studies go further than traditional word-formation studies as they not only look into the lexical level of TEs but also their characteristics and functions in

digital communication. Subsequently, I briefly reviewed previous studies on wordplay, with particular attention to multilingual or even multi-semiotic wordplay. It was shown that some scholars address multilingual wordplay from a Translanguaging perspective, a direction I shall also pursue in this study. Then, I surveyed existing pragmatic studies on the cognitive processes of lexical item comprehension, given that the present study is also primarily concerned with the comprehension of lexical items. I demonstrated that the cognitive processes fall into two major categories (narrowing and broadening), which could be further categorised into sub-processes.

Admittedly, the previous research along different fronts has laid the groundwork for the present study in many ways. For example, the discussion on creative word formation and lexical narrowing and broadening may build the foundation of analysing the TSEs' formation and comprehension processes, as these mechanisms can help to explain how TSEs are formed and how the readers understand the TSEs in a certain way. My study will show that the process of broadening also applies to the interpretations of some TSEs. For example, *多亿点* involves hyperbole; both  *精* and *立flag* involve metaphorisation. Thus, broadening occurs in the interpretation of these TSEs. But, I have found a lot of space that previous research has left for further exploration. Specifically, the review of related literature above has suggested the following research gaps for the present study.

First, while existing literature on creative word formation tends to focus on monolingual lexical items, few scholars have fairly systematically directed their attention to lexical items whose formation involves more than one language or even more than one semiotic resource. Although sporadic studies have touched upon the formation of particular TSEs, little research has systematically probed into how creative TEs or TSEs are formed with regard to the various word formation mechanisms involved and the various semiotic resources employed. Whether these mechanisms also operate and how in the formation of TSEs that go beyond a single language or, more broadly, Trans-semioticising expressions that go beyond a single language and even go across two or more semiotic systems is left to be studied. Nor

do we have any idea of the extent to which various mechanisms and semiotic resources for creative word formation are applied in forming TSEs on Chinese social media. It is not clear whether some mechanisms and combinations of semiotic resources are favored over others. In addition, although some semiotic word formations are introduced and discussed, the origins of these words are not clearly explained. Moreover, apart from word formation mechanisms, whether there is any other way to categorise TSE is also worth investigating through authentic data collected. Some other studies investigate the Translanguaging elements separately from the perspectives like acronyms, numerals, emojis that are used in creating TSEs on Chinese social media. It is to be noted that while the study provides a rough categorisation of tranßcripting expressions, there are still some other situations left behind. For one thing, TSEs used on Chinese social media may contain more than one language, and combine Chinese characters in different ways. For another, since Tranßcripting allows communicators to make use of more than one element to express their intentions, TSEs created with more types of semiotic symbols should also be discussed. I will focus in current research on how people create TSEs, especially words, using different sorts of semiotic symbols in digital communication.

Second, while previous lexical pragmatics mainly focuses on the interpretation of underspecified lexical items involving only one linguistic/semiotic system, scholars have not explored TSEs that involve more than one linguistic/semiotic system. It is worthy of investigation whether narrowing and broadening previously discovered in lexical interpretation also apply to the comprehension of TSEs and whether any additional new process is involved and nevertheless is guided by the same principle of optimal relevance.

Third, while the previous research on wordplay has some implications for the present study, especially on how people interpret TSEs and why these expressions sound so special as to attract readers' attention, there is still much to be desired. Most relevant studies about wordplay, as mentioned above, provide a linguistic view of this phenomenon in different language contexts. In effect, with the fast development of

online communication, digital interactions between people go beyond the linguistic boundary. Chinese netizens use all kinds of Translanguaging resources to fulfill communicative purposes, including elements from diverse languages, pictures, gifs, short videos, etc. Correspondingly, wordplay is evolving into semiotic-play when people use Translanguaging resources to communicate creatively or critically. As we shall demonstrate, TSEs on Chinese social media usually contain acts of playing with different languages and semiotic resources to achieve certain communicative purposes such as humor, mockery, and entertainment. Moreover, whether the use of TSEs can serve the same functions or may bring about other special effects and how the special effects arise have not been explored from a theoretical perspective or empirically attested.

Fourth, methodologically, using Moment Analysis to analyse Translanguaging practices is a good approach but still has its limitations. Moment Analysis along with strategies like observation and IPA can offer accounts towards the creative and critical Translanguaging practice at a moment. However, this way of investigation is mainly based on the researcher's perspectives. For example, it cannot reveal what readers actually think when they are comprehending the TSEs they encounter. Interviews with the participants can provide some retrospective data on a certain moment during a conversation which may happen days ago or longer. Some data of the participants' thoughts may be ignored. Therefore, an introspective way to collect data is needed to be introduced to Translanguaging studies to answer questions like what the readers are thinking when they meet these novel expressions and how they make sense of these expressions. In addition, the limited number of studies that do exist on TSEs by Chinese social media users tend to focus on a small number of restricted types of TSEs. My thesis will offer a more comprehensive treatment of the creation of TSEs and develop new typologies of TSEs on the basis of a large TSE database. Furthermore, most existing studies tend to analyse the language phenomena in isolation without considering how they make sense with Chinese characters together in the posts concerned.

To fill in the above research gaps, this study proposes to address the following research questions.

A. How are the Trans-semioticising expressions (TSEs) Chinese readers encounter in Weibo posts created?

a. What creative formation mechanisms are applied in the TSEs encountered? What is the distribution of each type of mechanism?

b. What kinds of semiotic elements are combined in the TSEs encountered? What is the distribution of each type of combination?

c. Are all semiotic elements explicit in the TSEs encountered?

B. How do Chinese readers of the Weibo posts understand the different types of TSEs they encounter in the text of the posts?

a. What paths are taken by the readers in understanding the TSEs?

b. What strategies are used by the Chinese readers in understanding the TSEs along each path?

c. Does the readers' choice of paths and strategies vary according to the mechanism of creative formation applied and their level of familiarity?

d. What associative mechanisms do the readers use in their implementation of the strategies?

C. Why do Chinese readers of the Weibo posts understand the TSEs the way they do?

a. What directs the readers' search for an appropriate interpretation in understanding the TSEs they encounter?

b. What lexical pragmatic processes operate in their course of understanding the TSEs?

c. What additional effects do they obtain for the extra effort invested in the lexical pragmatic processes?

A database of TSEs will be constructed to address the first research question. To answer the second question, I will combine think-aloud protocols with interviews. To answer the third research question, I will resort to Relevance Theory (RT for short)

(Carston & Powell, 2006; Sperber & Wilson, 1986/1995a, 1995b; 1997, 2012a, 2012b; Wilson, 2004; Wilson & Carston, 2006, 2007) to explain the comprehension process. I will choose this particular theoretical framework as the theoretical framework for the interpretation of TSEs for the following two reasons:

To help account for how the participants interpret the TSEs, a brief review of Relevance Theory was offered in Section 2.5. The theory is a cognitive pragmatic theory that explores the cognitive process and mechanism of utterance interpretation (as well as the interpretation of lexical items involved), with particular attention to how various background assumptions are activated to participate in the interpretation process. The present study explores the interpretation of TSEs, with an interest in how the readers/viewers of the novel TSEs draw on various background assumptions instead of just decoding. Thus, RT is appropriate to my research purpose.

Relevance Theory has been extensively applied to the study of how some lexical items that are semantically underspecified in utterances are interpreted through processes such as narrowing and broadening, giving rise to the sub-field of lexical pragmatics. Since TSEs as constructions are also semantically underspecified or opaque in many ways, this study will follow previous lexical pragmatics research (Wilson, 2004; Wilson & Carston, 2006, 2007) by adopting RT as the theoretical framework. Meanwhile, it will seek to find out new process(es), apart from narrowing and broadening, that are involved in the interpretation of TSEs guided by the principle of relevance.

This study focuses on the most up-to-date language phenomenon which is widely used among the public but is seldom investigated by researchers. By answering the first research question, we may hope to extend the study of Translanguaging by revealing how various non-linguistic semiotic elements are explicitly or implicitly integrated with or without linguistic elements to make meaning. By addressing the second research question, we may hope to enrich the study of Translanguaging by uncovering the paths, strategies and influencing factors of comprehension as reported by the social media users themselves. By tackling the third research question, we may

hope to substantiate Relevance-based explanation of lexical comprehension by specifying some comprehension-facilitating cognitive associations involved.

Furthermore, the present thesis will also expand the existing studies on TSEs to include an investigation of the interpretation process of novel expressions. In this regard, this study will turn to Relevance Theory to find possible answers. As an application of RT, lexical pragmatics has provided a well-developed research system for analysis on words and how words get their contextualised meaning as understood by readers or hearers. However, in this study, TSEs go beyond the boundary of a single word as some TSEs are performed in the form of short phrases. What is more, the process of TSE interpretation involves multi-semiotic semantic integration, which requires consideration of how people make flexible use of the semantic meaning of words, background information and other semiotic elements' meaning. Considering that innovative items appearing in online reading or digital communication make use of all types of semiotic resources to serve the purpose of communication, it is meaningful and reasonable to take advantage of Relevance Theory as a theoretical approach to investigate the process whereby people process the TSEs under study.

Chapter 3 Methodology

This chapter presents a detailed account of the methodology employed in the present study. Section One explains how the different types of data used for answering the research questions are collected, including the data source, data description and the procedure of data collection. Section Two explains how the different types of data are analysed, including the methods of analysis for each type of data, the coding schemes involved. Section Three states some ethical considerations involved in the collection, transcription and analysis of the data.

3.1 Data collection

To answer the research questions proposed in this study, the researcher collected three types of data.

3.1.1 Constructing a TSEs database

In order to address the first research question, this study constructed a TSEs database by collecting data from Weibo (微博, Microblog). Weibo is a Chinese microblog website, run by Sina Corporation in 2019. It is one of the biggest social media platforms in China. According to the Weibo 2020 Annual Report on User Development, there were 511 million monthly active users and 224 million daily active users on September 2020. Weibo users are mainly young people as the majority of Weibo users are those born between 1990 and 2000 (48%) and after 2000 (30%). Among all these users, female users account for 54.6%, and the percentage of female users rises to 61.6% in post-2000 generation.

Weibo has become an all-round media for people to acquire different types of information including social, cultural, financial, and political news. One of the functions of Weibo is Super Topics, which refers to 超话 in Chinese. “Super” means 超级, while “topic” means 话题. The term of 超话 is the ancronym of the phrase 超级话题. The data are the microblogs posted by Microblog users selected from

twenty “Super Topics” covering topics like celebrities, sports, public issues, etc. Super Topic is a function provided by Weibo, in the form of a chatting platform where people can post their thoughts or comments under the same topic through putting “#+ Topic Name#” in the front and back of their blogs, like “# Translanguaging#”. Thus, blogs with the same tag can be sorted out and seen under the same Super Topic, which makes up a giant group with people sharing the same interest communicating with each other. As a Super Topic contains millions of Weibo posts about the same topic, the size of the topic can be described as “super”.

The website where I depended on to collect the data is Micro-hotspot (微热点), a public website tool for data collecting and processing. Micro hotspots (micro public opinion) are based on Chinese Internet big data and official data from Sina Weibo, focusing on topics such as Internet information, business monitoring and socialised big data scenario applications. Micro-hotspots use distributed network technology to complete the collection of information-related data sources on the Internet while supplementing the entire network data according to the user’s predetermined monitoring keywords. From micro-hotspots, we can collect information on relevant websites, forums, microblogs, public accounts, blogs, and other Internet open platforms quickly through big data processing technologies such as Chinese intelligent word segmentation and natural language processing.

I used Micro hotspots to collect Micro-blog posts from 10 Super-topics on May 2021. For each Super-topic, I collected 1000 Weibo posts. Since the activity of different super-topic differs from each other, it takes different length of time to collect the same amount of Weibo posts. Therefore, I set the goal of collecting 1000 Weibo posts without any time restriction. The ten super-topics selected were Zhang Yixing Super-topic, Chuangzaoying Super-topic, Cai Xukun Super-topic, Wang Yibo Super-topic, Richang Super-topic, Wangzherongyao Super-topic, Yingxionglianmeng Super-topic, Hanfu Super-topic, Meiriyishan Super-topic, and Qinggan Super-topic. However, when collecting the data, the activity of the Qinggan Super-topic and Hanfu Super-topic was relatively low. As a result, for these two super-topics, after five hours

of collecting Weibo posts, only 253 and 364 pieces of Weibo posts were collected, which was far fewer than other super-topics. Therefore, these two super-topics were given up and only the remaining eight super-topics' data were used in the following analysis. Table 3.1 below showed the basic information about the selected eight super-topics.

Table 3.1 A sketch of the selected super-topics

Yingxionglianmeng Super-topic 英雄联盟超话	Meiriyishan Super-topic 每日一善超话	Wangzherongyao Super-topic 王者荣耀超话	Richang Super-topic 日常超话
Game	Daily	Game	Daily
Under this super-topic, people who played the game of Wangzherongyao would communicate with each other and talk about their gaming experience.	Under this super-topic, people would like to share their kind acts and express their kindness.	Under this super-topic. People who played the game of League of Legends would communicate with each other and talk about their gaming experience.	In this super-topic, people would like to share their daily stories and feelings with each other.
1000	1000	1000	1000
2146	152	1058	241

Name of the Super-topics	Zhang Yixing Super-topic \张艺兴超话	Chuangzaoying Super-topic 创造营超话	Cai Xukun Super-topic 蔡徐坤超话	Wang Yibo Super-topic 王一博超话
Topic type	Idol	TV show	Idol	Idol
Content introduction	In this super-topic, fans of Zhang Yixing gathered together to talk about and share information about their idol.	In this super-topic, people who watched a pick-up show would post Weibo and communicate with each other about the show and the candidates.	In this super-topic, fans of Cai Xukun gathered together to talk about and share information about their idol.	Under this super-topic, fans of Wang Yibo gathered together to talk about and share information about their idol.
Number of Weibo posts collected	1000	1000	1000	1000
Number of TSEs identified (without deduplication)	371	168	741	1280

In the database, TSEs of different types appeared for 6157 times without duplication and 2670 different TSEs were manually identified, based on the definition of TSE provided in Chapter 1. Any lexical-level expression that explicitly or implicitly contained elements from two or more semiotic systems (i.e., logographic characters,

Latin alphabet letters, Arabic numerals, emojis, and other symbols) was recognised as a TSE. By lexical level, we encompassed word-like TSEs (like *niubility* and *ungeilivable*) and word-group TSEs (like *打 call* and *种花家*).

To deal with the data collected from Micro-hotspots, I first manually coded the TSEs in the data based on the definition of TSEs. Then the data were put into the software AntConc for concordance analysis so that I was able to locate different types of TSEs during data statistics for further analysis. Table 3.2 shows how I marked the TSEs that appeared in the data.

Table 3.2 Marks of TSEs in the database

Forms of Marks	**Meaning of marks**	**Examples**
<TSE=CN/EN>	Chinese + English	*笑skr人*
<TSE=CN/NU>	Chinese + Numerals	*520*
<TSE=CN/EM>	Chinese + emojis	关🐷
<TSE=CN/JP>	Chinese + Japanese	*桥豆麻袋*
<TSE=EN/NU>	English + Numerals	*V10*
<TSE=EN>	Only English	*CP*
<TSE=CN>	Only Chinese (including pinyin)	*TMD*
<TSE=CN/EN/NU>	Chinese + English + Numerals	*U1s1*
<TSE=CN/PU>	Chinese + Punctuation	*+关注*
Etc.		

The form of marks follows the model of <TSE=abbreviation of semiotic system 1/ abbreviation of semiotic system 2/ abbreviation of semiotic system 3…>.

In the data, there is no original data that contains similar expressions like this form, so that this mark can be successfully sorted in this data analysis software. Table 3.3 shows the distribution of each type of TSEs:

Table 3.3 The distribution of each type of TSEs

Types	Number
<TL=CN/EM/PU>	4
<TL=CN/EM>	351
<TL=CN/EN/EM>	1
<TSE=CN/EN/JP>	1
<TSE=CN/EN/NU>	6
<TSE=CN/EN>	1068
<TSE=CN/JP>	134
<TSE=CN/KR>	1
<TSE=CN/NU/PU>	1
<TSE=CN/NU>	37
<TSE=CN/PU>	10
<TSE=CN/TH>	1
<TSE=CN>	1980
<TSE=EM/PU>	3
<TSE=EM>	8
<TSE=EN/EM>	8
<TSE=EN/NU>	145
<TSE=EN/UN>	2
<TSE=EN>	2352
<TSE=JP>	1
<TSE=NU/EM>	4
<TSE=NU/EN>	2
<TSE=NU>	2
Total	6157

The data were stored as a txt document which was further analysed by means of corpus software. In the TXT document, each TSE was presented with the context where it appeared and TSEs with emojis were recorded in the form of Unicode. Unicode is an information technology standard for the consistent encoding, representation, and handling of text expressed in most of the world's writing system (Unicode Standard, 2019).

3.1.2 Think-aloud protocols design

In order to address the second research question, that is, how readers of the Weibo posts understand the TSEs used, I designed and used a think-aloud task.

Think-aloud method was prominently advocated by Ericsson and Simon (1993). Its purpose is to make observable at least some proportion of the information processing that takes place during a given task (Ericsson and Simon, 1993). It requires participants to engage in some tasks (reading or doing something that is designed for studies) and constantly verbalise their thoughts while working on the tasks. Researchers assume that the source of the think-aloud output or protocols comes from the information currently active in one's short-term memory. By qualitatively coding the think-aloud protocols, researchers should be able to develop a fairly accurate understanding of the cognitive processes one has just undergone during the task performance. Compared to methods like interview or eye-tracking, think-aloud protocols can provide more detailed, real-time, and direct information about people's thought without bringing the researcher's subjective perceptions to the data.

Think-aloud method is a non-directive technique. The only probe used after initial instruction is that when participants apparently stop verbalizing for some time, they are simply reminded to continue thinking aloud. The think-aloud task is normally conducted on small samples of between 10 and 30 participants, who are usually students (see Carmel, Crawford & Chen, 1992; Crampton, 1992; Darken & Sibert, 1996; Hill & Hannafin, 1997). The products of think-aloud interviews are generally

coded and analysed qualitatively (Carmel *et al*, 1992), although some researchers analyse them quantitatively as well (e.g., Hill & Hannifin, 1997). Former studies (e.g., Jourdenais, Ota, Stauffer, Boyson & Doughty, 1995; Yayli, 2010) usually use think-aloud protocols as a tool to analyse the participants' thoughts during reading or writing tasks, using a product, etc. In my research, I used think-aloud qualitatively as an empirical method to reveal how Chinese readers of Weibo posts understand the TSEs used in the posts and what factors have a close bearing on the way and result of their understanding.

3.1.2.1 Pilot Study

Before the formal think-aloud tasks, I planned a pilot study to verify the feasibility of the plan and test the materials and online communicating tools.

Participants

To ensure the validity of the study, I did a pilot think-aloud study on 5 participants (with 2 males and 3 females), who were not included in the participants that would take part in the final study. These participants were Chinese undergraduate and postgraduate students from a Chinese university aged from 20-25. The recruitment did not take into consideration the students' majors, grades, and so on. What mattered was that all of them use social media on a regular or even daily basis. When I recruited the participants, I asked the volunteers whether they were good talkers as they would talk a lot alone during think-aloud task. Only those who were adequately communicative were finally invited as the participants. All the 5 participants volunteered to anonymously take part in the pilot study and all their personal information was carefully protected. All the 5 participants were rewarded with one-year membership of Weibo for participating in the project at my cost, which values 118 CNY for one year. Considering that all the participants' first language was Chinese, to ensure that they could best express themselves during the think-aloud task without any language difficulties, the pilot study was done in Chinese and recorded with a professional recorder.

Materials

In this study, based on my typological analysis that had been completed by drawing on Lipka (2002) before the pilot study, TSEs were classified into 4 main groups (morpho-semantic, semantic, morphological, and external), as shown in Table 3.4.

Table 3.4 TSEs used in the pilot study

TSE types		**Test items**
morpho-semantic TSEs	affixation	*ungeilivable, niubility*
	composition	*打 call, Book 思议*
	phonetic motivation	 ，放 p
semantic TSEs		*华粉,* 
morphological TSEs		*sb*, *cpdd*
external TSEs		*买米当卡，牛逼普拉斯*

Preparation for the pilot study

Before the pilot study, I gave the participants instructions about the think-aloud method in Chinese one by one, followed up with a training on how to participate in a think-aloud task, including a training video clip and a trial task. The training material was made into a video, which contained two sections. In the first section, I orally introduced the concept and definition of think-aloud in Chinese with literal explanation shown on the screen (in the format of PowerPoints). The second section was a video recorded example of a participant who was doing a think-aloud task on the understanding of a Chinese expression. By watching the instructional video, the participants could have a direct idea of what think-aloud was and what they needed to do during think-aloud tasks. Then a Q&A section was provided for each participant to check if they still felt confused on any details. Then I went on with some specific requirements about the upcoming task, as detailed below.

In order to help the 5 participants to get familiar with the think-aloud task, I then invited them to conduct a trial task after they watched the training video. To avoid practice effect, I used 2 Chinese expressions that were once very popular among Chinese young people, namely “喜大普奔” (*xi da pu ben*) and “奥利给” (*ao li gei*). These two Chinese expressions are not Trans-semioticising expressions. “喜大普奔” (*xi da pu ben*) is a contracted form of four Chinese idioms, namely “喜闻乐见”, “大快人心”, “普天同庆” and “奔走相告”, meaning that people should share the great news and celebrate it together. When it is formed, the first character of each idiom is picked out and then combined as a new word. People endow a stronger and richer meaning to this expression as it represents four different but similar idioms. “喜闻乐见” literally means “happy to hear and see (something happens)”. “大快人心” literally means “something makes people feel extremely happy”. “普天同庆” means “people who are under the same sky celebrate something together”. “奔走相告” means “people rush about telling others good news”. All these idioms vividly illustrate people’s reactions when something awesome happens. So when adding these four idioms together as an acronym, “喜大普奔” can express a “four-times” stronger joyful feeling. “奥利给”, however, is not a Chinese expression that can be found in Chinese dictionary. Unlike “喜大普奔”, which is formed based on other existing Chinese idioms, “奥利给” is a recently created modal particle used to cheer people up. It was first used by a Chinese streamer who loved to spread positive energy to his audience. Every time when he finished saying an inspiring sentence, he would add an “奥利给” to end his speech. For example, “One shall never give up when life beats you! Fight back and overcome all the barriers! Come on, 奥利给”. The streamer always used a very powerful and loud voice to speak out “奥利给”, which left people

with a deep impression. “喜大普奔” is an expression that follows a certain word-formation rule (clipping), while “奥利给” can only be understood through the context. These two Chinese expressions can be understood based on their formation or context; however, none of them belongs to TSE as they do not contain other semiotic elements in their forms and they do not need other semiotic elements to help understand their meaning. Therefore, “喜大普奔” and “奥利给” can help the participants to get familiar with think-aloud. In the process of understanding TSEs, these two expressions should not influence the participants’ understanding.

I encouraged the participants to keep talking and say as much as they could about how they understood the expressions. Specifically, I told the 5 participants to talk continuously about how they understood the sample TSEs during the task. I informed them that I would remind them to keep talking. Whenever they fell silent for up to 15s, I would only say “请继续讲” (Keep talking please) to break silence. I would not give the participants any hints that might help them understand the target Translanguaging expression. Any type of result was a meaningful outcome, which would be considered as a possible situation people might come across. The think-aloud task for each TSE expression would stop when the participant signaled they had nothing more to say (raising their hands up and saying they were done). Then I would move to the next TSE. The hypothesis of this task is that people may employ different ways to understand different types of TSEs. As Relevance Theory can only provide general steps of understanding lexical items, this method can further figure out the nuanced differences in understanding different TSEs.

Then the pilot study started. I asked the participants to report on how they understood the various TSEs presented to them one after another.

The TSEs used for the pilot study were not presented to the 3 participants in isolation but rather occurred in the context of the original Weibo posts. Here are 4 examples of the posts in which the test items occurred (The personal information of the users has been pixelated for ethical consideration).

Example 1 *666*

Image 1 Example of 666

[**Translation**: It's not going well this afternoon. Haven't won a single game. In this game I used the champion of Zed. I got the highest kill. When my teammates were all terminated, I managed to kill Jax (full health) with only one quarter of my health. Two of my teammates said "**666**" and the other said "nb". However, it's still useless. We still lost the game. All my teammates' KDAs are negative and they contributed little damage to the game. What was left to us was 持て余す鼓動に (for the heartbeat that I cannot hold).]

[**Annotation:** The 666 here conveys praise. The author performed well in a League of Legend game.]

Example 2 *气skr人*

Image 2 Example of *气skr人*

[**Translation**: # Eliminated# What a pity for the team to be eliminated! They did well in the two games. However, G's team lost due to the misjudgment in last episode. If they are eliminated in this episode, it can really ***气skr人***(make me piss off). #This is Slam Dunk#]

[**Annotation:** The TSE of *气skr人* in the example can be rephrased as "气死个人", meaning "making someone piss off". "skr" inside the TSE is interpreted as "死个" instead of "是个" or

other words because the interpretation is consistent with the speaker's intended meaning for the whole utterance (i.e., the author believed that it was utterly unfair for the team to be eliminated). Thus, the interpretation of the TSE in this way will convey not only an emphasis (which "气死个人" can also achieve), but also some additional effect, because the processing of "skr" will call forth the fun story of Wu Yifan.]

Example 3 *xswl*

Image 3 Example of *xswl*

[**Translation**: The whole homepages (of the Micro-blog) are worried about the weather, traffic and temperature of Beijing hahahahahaha, ***xswl***, which makes me feel so excited too.]

[**Annotation:** In this case, *xswl* is interpreted as the acronym of the Chinese expression of "笑死我了", which means something is so funny and is killing the hearer. This interpretation is in line with the context, unlike other potential interpretations such as "学生未来" or "雪山无路". The use of *xswl* is more effective than its direct equivalent "笑死我了" in that it saves typing effort and looks exotic as well.]

Example 4 *欧克*

Image 4 Example of *欧克*

[**Translation**: I drew my loved boy in spare time. I've been a fan of him for seven years. Fangfangzi is the most handsome guy. Do you think the painting is ***欧克*** (OK)?]

[**Annotation**: In this example, the TSE expression of *欧克* is written in the form of Chinese characters but actually its interpretation depends on one's knowledge of the English word "OK" as well as its pronunciation. The innovativeness of the TSE conveys additional special effects because it builds up a homophonic effect.]

After the pilot study, two of the test items were replaced. Specifically, *Book 思议* was substituted with 笑 yue, because from the think-aloud protocols, the TSE also belonged to the category of phonetic motivation, which has already had two test items. Besides, one of the two test items for phonetic motivation, that is, "放 p" was improper and replaced with  , because it was reported that the expression did not sound elegant and made it awkward to talk about it. Thus, for the main study, I used the test items shown in Table 3.5.

Table 3.5 Test items used in the pilot study

TSE types		**Test items**
Morpho-semantic TSEs	Affixation	ungeilivable, niubility, zhuangbility, unbengable, undingable, ikun
	Composition	打 call, 冲  , 卧　, 带 jio 布, 鱼香肉丝, book 思议, 兔粉
	Phonetic motivation	 了,  , 无 1 无靠, 456, 271, 9 敏
Semantic TSEs		司马, , 丹砂, xxs, 几九, 集美, 花粉
Morphological TSEs		种花家, ban 位买房, 可拷, 可真刑, 阿婆主, cpdd
External TSEs		酸 q, 兔瓦斯, 谢特, 077, 艾派德, 乌拉, 苏卡不列, 屁屁踢, 欧巴, 大丈夫, 法克鱿, 妈惹法克, 蟹不肉, 买米当卡, 撒浪嘿, 安可, 哈拉少
Complex TSEs		桃浦, u1s1, 笑 s, 西八, 粉丝, 芜湖 nese

When conducting the pilot think-aloud task, I followed the following stages.

Stage 1: Preparation

I confirmed a meeting time with each of the 5 participants. For personal reasons, not all of the participants participated in the task on the same day. It took me 2 days to have all the participants complete the task. At most, 2 participants participated on the same day; 3 of the participants completed the task on separate days.

I sent a consent form and information sheet to the participant through email. I explained the general purpose of my research and what I expected them to do. Also, I introduced the procedure of think-aloud task they were invited to attend, and answered the participant's questions in understanding the task. I double checked that the participants really understood what they were expected to do. After they had agreed to participate in the think-aloud task, I asked each of them to sign on the consent form (The participants could use e-signature to sign the consent forms) and send it back to me through email.

Then, I instructed them to download and install Tencent Meeting on their PC if they had not installed such software, and made sure the participants could use the screen sharing function.

Due to network issues, Zoom is unstable in China. Using it cannot guarantee the completion and quality of the recording. In addition, Chinese participants are unfamiliar with Microsoft Teams. They need to buy the software in order to use its full functions. During the pandemic, the most popular online meeting and teaching software in Chinese is Tencent Meeting, provided for free by Tencent supporters. Chinese students are generally so familiar with this software that they will not feel anxious while using it.

Meanwhile, I prepared the PowerPoint to show the Weibo posts, each with one test item to be used for the think-aloud task. When arranging the order of the presentation, I deliberately separated those items of the same type to avoid possible repetition effect. The first slide of the PowerPoint was a title page. Then there were ten slides of blank pages, which could prevent the participants from unconsciously seeing the TSEs before the task started. Since the researcher might begin to comprehend the TSE at first sight, the result of their think-aloud data might be

affected. For each of the 20 TSEs, there were two slides. The first slide showed the TSE without any context and also provided three-point-scale options for the participants to assess their familiarity with the TSE. The other slide then provided one screenshot of the Weibo post which contained the target TSE. In this way, we could provide the participants with an online reading environment close to their natural Weibo surfing environment. Considering that different computer systems may have different ways of encoding emoji, which may lead to the failure to show Emoji-related TSEs, all emojis that appeared in the PowerPoint were shown in pictures.

Also, I prepared 20 sheets containing a think-aloud completion table (see Table 3.6), with which I would be able to mark the completed TSEs with different level of familiarity.

Table 3.6 Completion table for different levels of familiarity TSEs of each participant

Level of familiarity / TSE Types		High familiarity	Low familiarity	Non-familiarity
morpho-semantic TSEs	affixation			
	composition			
	phonetic motivation			
semantic TSEs				
morphological TSEs				
external TSEs				
Complex TSEs				

In this pilot study, the participants' familiarity with Trans-semioticising expressions fell into three levels, which are 1) high familiarity (which means that they have met the TSE before and know its exact meaning); 2) low familiarity (which

means that they have met the TSE before but do not know or are not sure about its meaning; 3) non-familiarity (which means that they have never met the TSE before and thus know nothing about its meaning).

I took the participants' levels of familiarity with the TSEs because they are an important variable for consideration in previous research on lexical processing. Former studies usually focus on the quantitative relationship between familiarity and lexical cognition. Researchers (Balota, Pilotti, & Corteze, 2001; Chaffin, Morris, & Seely, 2001; Connine, Mullenix, Shernoff & Yelen, 1990; Ferraro & Sturgill, 1998; Gordon, 1985; Juhasz & Rayner, 2003; Whalen & Zziga, 1994; Williams & Morris, 2004) use lexical decision tasks and eye-tracking techniques to examine how readers process words during L1 reading. All the studies achieve a similar finding that if a reader is familiar with the word encountered in the text, it can be processed much faster than unfamiliar words. In other words, the level of word familiarity can be directly used to predict the time needed to process the word (Dolgunsöz, 2018). The two studies in the L1 context are important because the current study has many similarities to them. Chaffin *et al* (2001) record the reader's eye movements when reading a pair of sentences containing target words from one of the three subjective familiarity conditions: high familiarity, low familiarity, or novelty. Novel words are pseudo-words and can be pronounced, but readers do not know them at all. Their results show that readers spend more time on novel words than on highly familiar or unfamiliar words. However, the consensus still exists. There is no significant difference in the initial processing time (first fixation and duration of fixation), in a similar study by William and Morris (2004). In the study, two eye-tracking experiments scrutinised the effects of word familiarity on word processing and text comprehension during silent L1 reading. The text stimuli were composed of words with different familiarity and frequency, and multiple-choice tests were implemented to assess comprehension ability. The results show that there is no significant difference in the time spent on high-frequency words and low-frequency words.

Compared with familiar words, readers spend more initial processing time on new words.

Before the task started on the scheduled day, I checked and made sure that the Internet connection on both sides was stable. In addition, both the participant and I were arranged in a quiet and undisturbed environment to ensure the success of recording and the recording quality. The participants could attend the think-aloud task either at home or in any other quiet place where they felt relaxed. They needed to open up the Tencent meeting software on their PC and got ready for taking the task.

I contacted the participants through platforms like WeChat, QQ, email, etc., and used Tencent Meeting for Think-aloud tasks. In addition, I helped the participants deal with all technique problems such as Internet connection, PC software download, etc. These jobs were done before the day of Think-aloud tasks.

Stage 2: Task performance

The think-aloud tasks were conducted following the procedures designed below.

Table 3.7 Think-aloud task procedure

Think-aloud task	**Explanations**
1. The researcher sends the consent form and information sheet to the participant through email. Then the researcher introduces the procedure of Think-aloud protocols, plays the prepared guidance video, and answers the participant's questions. The researcher has to make sure that the participant understands what he or she is required to do. Finally, the researcher asks the participant to sign on the form if he or she is willing to continue to join in the task. 2. When the researcher receives the signed consent form, the researcher will send the PowerPoint used for the Think-aloud task to the participant through email and let the participant start screen sharing. Then, with the permission of the researcher, the participant	1. The researcher needs to contact the participants through platforms like Wechat, QQ, email, etc., and use Tencent Meeting for Think-aloud tasks. In addition, the researcher needs to help the participants deal with all technique problems such as Internet connection, PC software download, etc. These jobs should be done before the day of Think-aloud tasks. 2. The participants can use e-signature to sign the consent forms. In addition, a consent form, information sheet, and PowerPoint will be sent to the participants through email. 3. The participant does not need to turn on his or her webcam. 4. The researcher needs to emphasise the definition of "having not met the TSE before"

opens the PowerPoint.

3. The researcher starts screen recording and informs the participant: "I have opened the screen recording. Then, we will begin the Think-aloud task. Please play the PowerPoint and click your mouse ten times."
4. The next slide will show a new (test) TSE. First, the researcher identifies the type of TSE. Then the researcher lets the participant assess the level of familiarity with the TSE: "Before the Think-aloud task, which level of familiarity best suits your knowledge of this TSE? Please choose from the three options: A. having met the TSE before and knowing its meaning; B. having met the TSE before but knowing nothing or am not sure about its meaning; C. having not met the TSE before and knowing nothing about its meaning."
5. The researcher puts an √ in the corresponding box in the completion table based on the TSE's type and the participant's familiarity. Then, the researcher lets the participant move to the next slide of PowerPoint: "Please click your mouse once and move to the next slide."
6. The researcher informs the participant that the think-aloud task of this TSE begins: "Now you can read the Weibo. When you start reading, please speak all thoughts that appear in your mind aloud. You should follow your reading habits when you surf on Weibo. If you need to use any tool, please let me know and show me how you use the tool. You cannot use search engines to search for its meaning. If you cannot figure out its meaning based on the information Weibo provides and your daily reading habit is to search its meaning when you cannot understand, tell me that you would like to search its meaning. Then the think-aloud task for this TSE will stop. Please try your best to figure out its meaning, but also feel free to quit. When you think you have finished your reading, no matter whether you have figured out the and "knowing its meaning" to the participants. "Having not met the TSE before" means that the participants have not met the TSE in another context before this TSE Think-aloud task. "Not sure about the TSE's meaning" means that before seeing the shown TSE, the participants do not know its meaning. Sometimes some TSEs' meanings can be easily figured out without any context at first sight so that the participants may consider they know these TSEs' meanings.
5. When the TSE is shown on the slide, the researcher will know the type of this TSE. The √ should be put in the corresponding box as soon as the participants give their assessments on the familiarity of the TSE showing that a certain type and level of familiarity of TSE has already been tested.
6. Considering that this Think-aloud task of this study is conducted online, the participants' attention is focused on the PowerPoint slides while reading the Weibo posts and doing think-aloud. Therefore, we cannot use light or other visual signals to remind the participants to continue speaking their minds aloud when they go silent. The researcher can use language to remind and encourage the participants as this study focuses on the comprehension process instead of their language. The researcher's words will not influence the way participants talk. The participants may not help to asking questions to confirm whether they properly understand the TSEs. The researcher should not provide any information related to the comprehension. No hints should be given to help the participants understand the TSEs. To help the participants better understand what they are expected to do and get used to the way they talk during the think-aloud tasks, the researcher uses the first three TSEs as samples provided for the participants. The researcher should not put √ in the completion table when the participants are doing the first three TSEs.
7. As the study focuses on the thinking

TSE's meaning, please let me know by saying 'I'm done.' When I notice that you are not speaking your thoughts aloud, I will remind you. You do not need to feel stressed when I remind you, because it is just an encouragement that reminds you to speak aloud."

7. When the participant becomes quiet and does not speak, the researcher will directly use sentences to remind the participant to speak his or her mind aloud: "Don't forget to speak your thoughts aloud. Just speak whatever that appears in your mind."
8. When the participant informs the researcher that the reading procedure is done or he or she can no longer figure out the TSE's meaning, the researcher will stop the think-aloud task of this TSE and start doing a follow-up interview on the participant: "Okay, thank you for your cooperation. I will ask you some questions about the silences during your think-aloud task and double-check some unclear expressions you just mentioned."
9. After the follow-up interview, the researcher lets the participant move to the next slide of TSE: "Now we are done with this TSE. Please click your mouse once and move to the next TSE."
10. The researcher will repeat the step 4-9 until all boxes in the completion table have been ticked.

a) Situation 1: When a new type of TSE appears, the researcher repeats steps 4-9;

b) Situation 2: When a TSE's type has already appeared, the researcher still needs to let the participant assess his or her familiarity with the TSE. If the participant chooses a new type of the same familiarity, the researcher enables the participant to click the mouse once, conduct steps 5-9, and put an √ in the corresponding box in the completion table. If the participant has done another TSE (same type and same level of familiarity), the researcher lets the participant click the mouse

procedures through which the participants comprehend the TSEs, direct Chinese reminders will not influence them or provide additional information.

8. Only when the participants have finished comprehending the first three TSEs can the researcher tell them these three TSEs are for practice. And the formal think-aloud task starts from the fourth TSE.
9. Step 9 illustrates how to accomplish the 15 different types and levels of familiarity of TSEs orderly and efficiently. There are mainly two different situations based on whether the participants have encountered the same type of TSE before the new TSE. In situation 1, when the participants have not met the type of the newly shown TSE, the researcher will return to steps 4-9. For example, when the TSE shown on the slide is a Phonogram TSE, and the participant has not shared his or her comprehension procedures aloud, the researcher will let the participant assess its familiarity (step 4) and start the think-aloud task this TSE (step 5-9). In situation 2, when the participants have encountered the same type of TSE before in the Think-aloud task, the researcher will let the participants assess his or her familiarity with the TSE first. If the participants have done a TSE with the same level of familiarity, the researcher will let the participants share how they understood the TSE when they first met it. If the participants have not done a TSE with the same level of familiarity, the researcher will return to steps 5-9. For example, suppose that the participant first accomplished a Type 1 TSE with A level of familiarity. If the next TSE is type 2, the researcher will repeat steps 4-9. If the next TSE is still type 1, the researcher needs to let the participant assess its familiarity. If the level of familiarity is also A level, the researcher will let the participant share his or her former experience of comprehending this TSE. If the participant considers the TSE as B level, the

once and say: "Please introduce how you figured out the meaning of this TSE when you first met it?" After that, the researcher lets the participant click the mouse once and move to the next TSE, returning to step 4. 11. When the 15 boxes are all marked with √ and the participant has finished the follow-up interview of the 15th TSE, the participant's think-aloud task is accomplished. The researcher says the participant: "Thank you for your cooperation. We have finished today's think-aloud task." 12. The researcher stops the screen recording, closes Tencent Meeting, and saves the record to an encrypted PC. All records will be deleted when the researcher finishes the thesis.	researcher will repeat from step 5 to step 9. 10. For those TSEs which are marked as "having met the TSE before and know its meaning, the researcher also needs to ask the participants to share their comprehending experience. For one thing, the processes of understanding familiar TSEs' are simple but important. It is also a way that the participants take when they meet the TSEs while reading. For another, it can test whether the participants truly know the TSEs' meaning. Sometimes the participants may think they know its meaning at first glance or the same TSE can be used with a different meaning in different contexts. When this situation happens, the researcher should accept the participant's familiarity with this TSE, as they believe they know the meaning (even if it is not the "correct" meaning) based on their former comprehending experience. 11. When the researcher does think-aloud tasks with the participants, he needs to record or mark every silence position where the participants fail to express clearly. Then, every time the participants finish the think-aloud task of the particular TSE, the researcher should conduct a follow-up interview on these marked places with the participants before they forget what they were thinking. 12. None

The participant has asked to turn on the cameras so that the researcher could have better control over the experimental procedure.

When the researcher received the signed consent form, the researcher sent the PowerPoint used for the think-aloud task to the participant through email and let the participant start screen sharing. Then, with the permission of the researcher, the participant opened the PowerPoint.

The researcher started screen recording and informed the participant: “I have opened the screen recording. Now, we begin the think-aloud task. Please play the PowerPoint and click your mouse ten times.”

The next slide showed a new (test) TSE. First, the researcher identified the type of TSE. Then the researcher let the participant assess the level of familiarity of the TSE: “Before the think-aloud task, which level of familiarity best suits your knowledge of this TSE? Please choose from the three options: A. having met the TSE before and knowing its meaning; B. having met the TSE before but knowing nothing or is not sure about its meaning; C. having not met the TSE before and knowing nothing about its meaning.”

The researcher put a √ in the corresponding box in the completion table based on TSE type and the participant's familiarity. Then, the researcher let the participant move to the next slide of PowerPoint: “Please click your mouse once and move to the next slide.”

> “Now you can read the Weibo. When you start reading, please speak out all thoughts that appear in your mind aloud. You should follow your reading habits when you surf on Weibo. If you need to use any tool, please let me know and show me how you use the tool. You cannot use search engines to search for its meaning. If you cannot figure out its meaning based on the information Weibo provides and your daily reading habit is to search its meaning when you cannot understand, tell me that you would like to search its meaning. Then the think-aloud task for this TSE will stop. Please try your best to figure out its meaning, but also feel free to quit. When you think you have finished your reading, no matter whether you have figured out the meaning of the TSE, please let me know by saying “I’m done.” When I notice that you are not speaking your thoughts aloud, I will remind you. You do not need to feel stressed when I remind you. Because it is just a reminder for you to speak aloud.”

When the participant became quiet and did not speak, the researcher directly used sentences to remind the participant to speak his or her mind aloud: “Don’t forget

speaking your thoughts aloud. Just speak up whatever that appears in your mind."

When the participant informed the researcher that the reading procedure was done or he or she could no longer figure out the meaning of the TSE, the researcher stopped the think-aloud task of this TSE and started doing a follow-up interview on the participant: "Okay, thank you for your cooperation. I will ask you some questions about the silences during your think-aloud task and double-check some unclear expressions you just mentioned."

After the follow-up interview, the researcher let the participant move to the next slide of TSE: "Now we are done with this TSE. Please click your mouse once and move to the next TSE."

The researcher repeated the steps until all boxes in the completion table had been ticked.

Two situations happened during the procedure.

a) Situation 1: When a new type of TSE appeared, the researcher repeated steps 4-9;

b) Situation 2: When a type of TSE had already appeared, the researcher still let the participant assess his or her familiarity with the TSE. If the participant chose a new type of familiarity, the researcher enabled the participant to click the mouse once and begin the understanding, and put an √ in the corresponding box in the completion table. If the participant had done another TSE (same type and same level of familiarity), the researcher let the participant click the mouse once and say: "Please introduce how you figured out the meaning of this TSE when you first met it?" After that, the researcher let the participant click the mouse once and move to the next TSE.

When the 21 boxes were all marked with √ and the participant had finished the follow-up interview of the last TSE, the participant's think-aloud task was accomplished. The researcher said to the participant: "Thank you for your cooperation. We have finished today's think-aloud task."

The researcher stopped the screen recording, closed Tencent Meeting, and saved the record to an encrypted PC. All records would be deleted when the researcher finishes the thesis.

The researcher needed to clarify the definition of "having not met the TSE before" and "knowing its meaning" to the participants. "Having not met the TSE before" meant that the participants had not met the TSE in another context before this TSE Think-aloud task. "Not knowing the meaning of the TSE" meant that before seeing the shown TSE, the participants did not know its meaning. Sometimes some TSEs' meanings could be figured out at first sight without any context so that the participants may believe they know these TSEs' meanings.

When the TSE was shown on the slide, the researcher would know its type. The √ should be put in the corresponding box as soon as the participants gave their assessments on the familiarity of the TSE showing that a certain type and level of familiarity of TSE has already been tested.

Since the think-aloud task of this study was conducted online, the participants' attention was focused on the PowerPoint slides while reading the Weibo posts and doing the think-aloud task. For this reason, I did not use light or other visual signals to remind the participants to continue speaking their minds aloud when they went silent. The researcher used language to remind and encourage the participants as this study focused on the comprehension process instead of their language. The researcher's words would not influence the way participants talk. The participants may not help asking questions to confirm whether they properly understand the TSEs. The researcher should not provide any information related to the comprehension. No hints should be given to help the participants understand the TSEs. To help the participants better understand what they were expected to do and get used to the way they talked during the think-aloud tasks, the researcher used the first three TSEs as samples for the participants. The researcher did not put √ in the completion table until the participants were doing the first three TSEs.

Stage 3: After the task

After the study, all the files of the recordings were stored in an encrypted computer and UCL research log e-disk.

3.1.2.2 Main study

Participants

I invited 22 volunteers before the formal tasks. However, one participant quit due to his personal reasons. Thus, 21 Chinese undergraduate and graduate students aged 20-25 participated in this study. I recruited the participants from my former and current classmates, schoolmates, friends, and relatives. Instead of recruiting strangers from social media, inviting familiar people to do the tasks had some advantages. First, I was familiar with the participants' social media usage and their communicative ability. All the invited participants were Weibo users for at least five years and were able to clearly express themselves orally. Secondly, familiar people would not feel very nervous when doing think-aloud tasks along with me. They would feel more comfortable to speak up their minds and be more talkative. Thirdly, familiar participants would be more cooperative during the study. They were willing to participate in the study and arrange time for the study with more flexible time slots. Last, inviting familiar participants would improve the efficiency of recruitment comparing with randomly waiting for responses from social media and checking whether they were available. The recruitment did not take into consideration the participants' major, grade and so on. It also kept a balance in the number of male and female participants (with 10 male and 11 female students). They were told that they would be rewarded with one year membership of Weibo for participating in the project at my cost, which was worth 118 CNY for one year.

Materials

The materials used in the main study were exactly the same as those used in the pilot study (refer to Table 3.2). The TSEs appeared in the Weibo screenshots and were displayed to the participants digitally.

For each type of TSEs, I prepared at least six different TSEs for the participants to understand. For each participant, they would be presented with 21 different TSEs from the table. For each group of TSEs, only three of them were used for each participant.

Think-aloud tasks procedures

After I completed the pilot study, I found many drawbacks and problems that needed to be fixed in the main study.

The first issue was that using PowerPoint slides and letting the participants operate the mouse to move to the next TSE brought many troubles during the pilot study. The participants sometimes double clicked the mouse so they would move to the next TSE slide, which would influence their understanding of the target TSE displayed. Also, the slides' form sometimes changed when participants opened on their own computers. Moreover, not all participants had access to the computer. Some of the participants would like to use phones to participate in the study.

The second issue was that when the participants evaluated the level of their familiarity with the TSEs, the understanding processes were actually happening because some TSEs required analysis of their formation, which were not strictly dependent on the context. As a result, the understanding processes were significantly affected when the participants formally started the understanding. Therefore, I excluded the step of checking the level of familiarity before the tasks. Participants would directly understand the TSEs as soon as they saw the Weibo posts and the TSEs.

The third issue was also related to the way of displaying TSEs and familiarity check. Although I designed a check table to count the accomplish status, I found it not feasible during the think-aloud tasks. The check table would influence the coherence of the whole process of doing the tasks with the participants. Sometimes there was not enough time to put a mark on the check table as the participants spent very little time on the TSE and I had to pass to the next TSE immediately.

It was also noticed that if the participants were told that they could use tools to help them understand the TSEs, participants' decisions might be affected by this tip. Participants got an extra encouragement to use this method of understanding the TSEs. It would be better to leave it open and not to mention this before the tasks. However, if the participants asked whether they could use other tools like search engine or typing systems, I could allow them to do it but needed to ask them to explain how they used the tools.

To address these problems, I redesigned the way of displaying TSEs and changed the processes of conducting the tasks.

First of all, during the understanding processes, the participant might come across TSEs that they had already known before the task, which might lead to a very short understanding process. In this study, all kinds of situations were acceptable. even though the participant knew the meaning of the selected TSE, it was not a problem as this situation could reflect a more natural and real outcome that might appear in the participant's daily life. There was still a little chance that the participant knew all the 21 TSEs presented to him or her. But by recruiting 21 participants, the chances would be reduced a lot. In this way, the familiar-level checking process could be excluded.

Secondly, since all the TSEs were randomly selected from Weibo, the Weibo writers might provide different levels of background information or context to support the comprehension of the TSEs. In some extreme cases, the participants might come across TSEs with no context at all. For example, in the case of *cpdd*, many Weibo posts contained nothing but the four letters, so that the participants had no contextual support. In example 4 listed above, after using *xswl*, the writer wrote many *哈* s, which is an onomatopoeia of laughing. In this case, the participant could get some related information from the context. In some other examples, the writer may directly provide both the TSEs and their meaning at the same time. For instance, a writer mentioned both *cxk* and *蔡徐坤* (cxk) in her Weibo. *Cxk* is the pinyin initialism of

蔡徐坤, which became an obvious hint for the participants to understand the TSE. As we can see, in different cases, the participants may have different levels of contextual information while understanding the TSEs' meaning, which may influence their understanding processes to certain degrees. According to Relevance Theory, when processing an utterance, the reader will make use of assumptions derived from the surrounding environment and co-text to find the most relevant interpretation. So if the context information is limited, the readers may get into trouble when they need help from the context. However, it is very difficult to evaluate the exact level of context support provided to each target TSE. Firstly, different readers may not rely on the context as much to understand the TSEs. That is to say, even though a context may be enough for Reader A, it may not be enough for Reader B. Secondly, it is very difficult to find enough original TSEs with relatively the same level of contextual information. Thirdly, some contextual information is time-sensitive and social-sensitive. For example, if a Weibo that contains a TSE is posted many months after a social event happens, it may become very difficult for a reader to relate the social background to the Weibo. Therefore, in this study, I planned to minimise the impact of the level of contextual information rather than take it as a variable in the task.

To minimise the effect of the two factors mentioned above (whether the participant knows the TSE before and how much contextual information is available), I designed a small programme to automatically and randomly select target TSEs for the participants to work with through Python. As mentioned, I collected at least six different TSEs for the six sub-types with their own Weibo context. In all, 55 TSEs were selected and stored in a .txt file. Each TSE was numbered and labeled with its type name. The Weibo screenshots were saved in another folder, and were numbered accordingly. When I started the programme, it would randomly pick one TSE from the 120 TSEs, and retrieved the corresponding screenshot from the folder. At the same time, the programme would read the information of TSE's number and type mark. If a number had been read, the programme would not pick the TSE with this number again in the rest of the task. Meanwhile, each type of TSE could only be selected for three

times. Once a type was picked for three times, the programme would not select TSE with that type mark. Then, on the screen, the screenshot of the Weibo would be presented to the participant. When the participant finished the understanding task of this TSE, I would click the “next” button, and the programme would continue to randomly pick a new TSE which met two conditions, 1) its number had not been chosen in previous rounds; 2) its type mark had not been chosen for more than three times. The programme would stop automatically when each type of TSEs had been picked for three times. Figure 3.1 is the flow chart of the TSE selection programme:

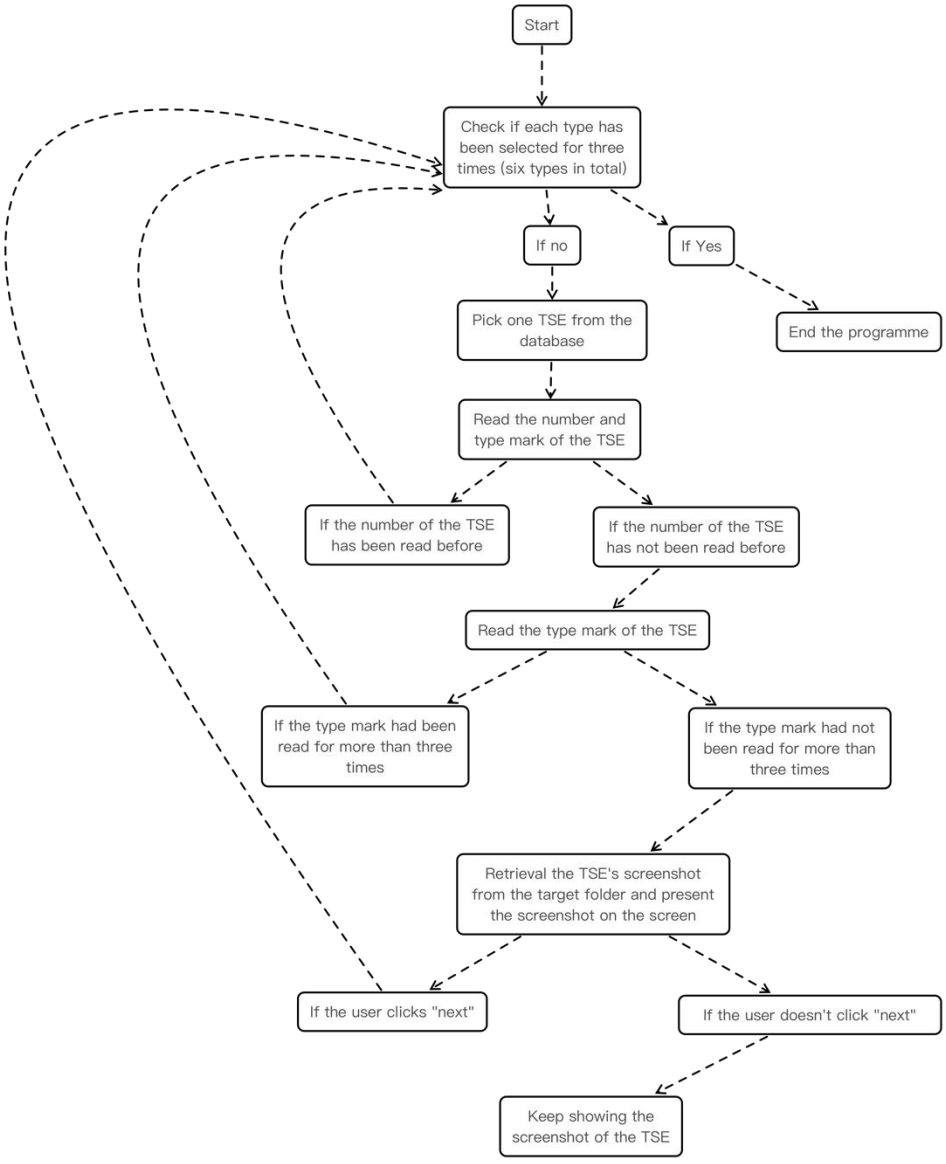


Figure 3.1 The flow chart of the TSE selection programme

Here is a screenshot of how the software looked like during the tasks:

Image 5 Screenshot of Random TSE Selection Software

When conducting the think-aloud task, I followed the procedures below.

Stage 1: Preparation

I confirmed a meeting time with each of the 21 participants. For personal reasons, not all of the participants participated in the task on the same day. It took me 2 days to have all the participants complete the task. 10 participants participated on the first day; the other 11 participants completed the task on another day.

I sent a consent form and information sheet to the participant through email. I explained the general purpose of my research and what I expected them to do. Also, I introduced the procedure of Think-aloud task they were invited to attend, and answered the participant's questions in understanding the task. I double checked that the participants really understood what they were expected to do. After they had agreed to participate in the think-aloud task, I asked each of them to sign on the consent form (The participants could use e-signature to sign the consent forms) and send it back to me through email.

Then, I instructed them to download and install Tencent Meeting on their PC if they had not installed such software, and made sure the participant could use the screen sharing function.

Meanwhile, I prepared the TSE Selection Programme to show the Weibo posts to the participants. Before the task started on the scheduled day, I checked and made sure that the Internet connection on both sides was stable. In addition, both the participant and I were in a quiet and undisturbed environment to ensure the success of recording and the recording quality. The participants could attend the think-aloud task either at home or in any other quiet place where they felt relaxed. They needed to open up the Tencent meeting software on their PC and got ready for taking the task.

I contacted the participants through platforms like WeChat, QQ, email, etc., and used Tencent Meeting for the Think-aloud tasks. In addition, I helped the participants deal with all technique problems such as Internet connection, PC software download, etc. These jobs had been done before the day of Think-aloud tasks.

Stage 2: Task performance

The participant was asked to turn on the cameras so that the researcher could have better control over the experimental procedure.

When the researcher received the signed consent form, the researcher sent the PowerPoint used for the Think-aloud task to the participant through email and let the participant start screen sharing. Then, the researcher started the TSE Selection Programme.

The researcher started screen recording and informed the participant: "I have opened the screen recording. Then, we will begin the Think-aloud task. "

When the Weibo post with the TSE was shown on the screen, the researcher asked the participant to start the think-aloud task. The researcher informed the participant that the Think-aloud task of this TSE began: “Now you can read the Weibo. When you start reading, please speak all thoughts that appear in your mind aloud. You should follow your reading habits when you surf on Weibo. If you believe that you have complete the understanding or want to stop the understanding, please say “I have completed”. Please try your best to figure out its meaning, but also feel free to quit. When I notice that you are not speaking your thoughts aloud, I will remind you. You do not need to feel stressed when I remind you by say *Keep on*

talking. It is just a reminder for you to speak aloud."

When the participant became quiet and did not speak, I directly used sentences to remind the participant to speak up his or her mind: "Don't forget speaking your thoughts aloud. Just speak whatever that appears in your mind."

When the participant informed the researcher that the reading procedure was done or he or she could no longer figure out the meaning of the TSE, I stopped the think-aloud task of this TSE and started doing a follow-up interview on the participant: "Okay, thank you for your cooperation. (What were you thinking when you were silent?) Let us move to the next TSE."

Then I continued the think-aloud tasks by clicking the mouse until the programme stopped. I told the participants: "Thank you for your cooperation. We have finished today's think-aloud task."

I stopped the screen recording, closed Tencent Meeting, and saved the record to an encrypted PC. All records would be saved for ten years unless the participants informed me to delete their records.

Since this think-aloud task of this study was conducted online, the participants' attention was focused on the screen while reading the Weibo posts and doing think-aloud, I could not use light or other visual signals to remind the participants to continue speaking their minds aloud when they went silent. I used language to remind and encourage the participants as this study focused on the understanding processes instead of their language. My words would not influence the way participants talked. The participants might not help asking questions to confirm whether they properly understand the TSEs. I did not provide any information related to the comprehension. No hints were given to help the participants understand the TSEs.

Stage 3: After the task

I answered the participants' questions about the meaning of the TSEs they failed to understand.

3.2 Data analysis

The data of the study underwent both quantitative and qualitative analyses.

For the former, frequency calculation for the different types of TSEs in terms of the mechanism of creative formation, combinations of semiotic elements, and the association mechanisms involved in the think-aloud protocols were done. In addition, a frequency count was done for the different paths and strategies used in understanding different types of TSEs.

For the latter, the data collected through think-aloud protocols were initially done qualitatively to identify the understanding strategies and paths. The researcher followed grounded theory to generalise the paths and strategies in understanding the TSEs. Grounded theory originates from a sociological research program launched by Glaser and Strauss in 1960's (Corbin & Strauss, 1990). They have crafted a method that enables researchers to generate systematically a substantive theory grounded in empirical data. Glaser and Strauss (1967) argue that grounded theory is a qualitative method can help grab a theory, which can fit the data and work in the real world. In grounded theory, data analysis begins with basic description of the data and then moves to conceptual ordering and ends with theorizing (Patton, 2002). Actually, researchers may not aim or be able to construct a theory in real studies. It can only be used to help researchers to clarify and specify their ideas (Thornberg *et al*, 2015). Grounded theory welcomes different types of data collection such as qualitative interviews, field observation, focus groups, etc. (Thornberg *et al*, 2015). In this study, think-aloud is also a valid method of collecting data. Grounded theory requires researchers to gather and analyse data simultaneously throughout the research process (Charmaz, 2006; Corbin & Strauss, 2008; Glaser, 1979, 1998). When a tentative theoretical categories have been defined, researchers will engage in theoretical sampling, which involves "seeking data to enable them to fill out the analytic properties of the defined categories" (Thornberg *et al*, 2015, p. 407). Then when the first data is identified, researchers will start the process of coding, including initial coding and focused coding. During the process of coding and analysing, "memos" are needed to help researchers memorise their analytical, conceptual or theoretical notes.

Finally, researchers sort out their ideas and findings based on the memos and integrate categories into grounded theory. In this study, I followed the routine of grounded theory to help investigate the understanding process of readers towards TSEs based on the think-aloud data. Different paths, strategies and associative mechanism taken by the participants were marked and coded. Draft understanding flow charts were drawn along with memos to support my analysis and the integration of the final understanding patterns. The processes of analyzing think aloud data is strictly followed by the routines of Grounded Theory.

In the analysis of the cognitive constraint on the comprehension process of TSEs and the non-propositional effect of the TSEs, I combined theoretical analysis and contextual analysis with think-aloud protocols from the participants. For example, for a TSE used in the context, if the participants reported that the words or emojis made them laugh, I would claim that this TSE has an effect of entertainment or humor. Finally, by comparing the data collected from the think-aloud protocol, I assessed the applicability of the account offered by Relevance Theory through a theoretical analysis, which is roughly and provisionally described as follows:

When people read the TSEs on the Internet, they will take it as an ostensive stimulus and reach out to related encyclopedic information stored in their mental lexicon through either logical, encyclopedic or lexical entry (Sperber & Wilson, 1986/1995a). For people who have already known the TSE expression, the meaning of each part of the expression has been stored in their memory and can be easily retrieved. They will activate related background information triggered by the component parts of the expression. By integrating the meaning of each component (different TSEs involve different elements and modes of integration), they will derive an interpretation that accords with the expectation of relevance. Since the integration requires some extra effort, they will be rewarded with some extra cognitive effects like humor, irony, etc. For those who meet the expression for the first time, they may further need to go through logic or lexical entries. When they find that it takes too much effort to understand the expression, they may either choose to skip the

expression or they go on to process it by investing a huge amount of cognitive effort. In the latter case, they may succeed in understanding the meaning and secure more added cognitive effects, or fail.

3.3 Ethic considerations

This research project followed the research guidelines endorsed by the UCL Institute of Education Research Ethics Committee. Informed consent forms were provided in writing (in Chinese) to the voluntary participants before undertaking the research. I explained all aspects of my project in terms of research aims, my role as the researcher, and data collection and storage. I also made sure that they knew they could withdraw from the project at any time without justifying their decision. Participants in the data would be kept in anonymity to guarantee confidentiality. The Weibo posts selected for think aloud tasks were all authentic and posted on Weibo by users whose name were encrypted. The content of these Weibo posts passed the Weibo censorship check and double-checked before they were used for this study, such that no sensitive or illegal content appeared in the selected materials. As for Weibo censorship check, it refers to the automatic censorship check system provided by Weibo platform. All contents that are illegal or sensitive cannot be posted publically. For example, if someone posts a Weibo that contains violent images, the Weibo post will be deleted and cannot be read by anyone. Generated data would comply with current GDPR. Ethical approval was received in June 8th with the approval from IOE.

Chapter 4 Categorising Trans-Semioticising Expressions in Weibo Posts

In this chapter, I will discuss the categorisation of Trans-semioticising expressions that readers encounter in Weibo Posts from mainly three perspectives, which are 1) the mechanisms of creative word formation applied; 2) the semiotic elements contained; 3) the explicitness of Trans-semioticising expressions. The distributional information of each categorisation will be offered to better understand the formation features of TSEs, particularly those not shared by mono-lingual neologisms and bilingual TEs.

4.1 Categorisation based on the mechanisms of creative formation applied

My data shows that the formation of TSEs also follows the various mechanisms of creative lexical formation (Lipka, 2002). Altogether four main mechanisms (including eight sub-processes) are found in forming TSEs, as shown in Table 4.1.

Table 4.1 Mechanisms of creative TSE formation (N=2669) (Borrowed from Lipka, 2002)

Mechanisms	**Sub-mechanisms**	**Illustrations**	**Numbers and frequencies**

Morpho-semantic process	Affixation	*niubility*	10
	Composition/ compounding	*立 flag*	985
	Phonetic motivation (homophones)	*酸 q*	2256
Semantic process	Metaphor	None	0
	Metonymy		463
Morphological process	Clipping	*up 主*	302
	blending	*笑 skr 人*	394
	Acronymy	*c 位*	105
External process	loanwords	*Ky*	1615
Total			2669

In the following sections, I will demonstrate how different types of TSEs are formed though the four formation mechanisms.

4.1.1 TSEs through morpho-semantic mechanism

TSEs created through morpho-semantic mechanism can be divided into three sub-types, namely those through affixation, composition, and phonetic motivation respectively.

TSEs through affixation

When people use affixation as a method to create TSE, they add a prefix or/and a suffix to another expression to give it a new meaning. The formation process of the TSE involves more than one semiotic element as the components of the TSE.

First of all, let us have a look at *niubility*, which is a typical example of affixation TSEs used by Chinese bilinguals who can speak both Chinese and English. *Niubility* is the Trans-semioticising counterpart of *牛逼*, whose pinyin spelling is *niu bi*. The expression is usually used as an adjective to describe something that is very good or

someone who does something very well. The character *牛* refers to "cow" in English and the character *逼* is a homophone of the Chinese equivalent of "vagina" in English. The formation of *niubility* consists in adding a suffix to the pinyin spelling of *牛*. In English, the suffix of "-ity" is usually used when transforming adjectives that end in "-able" to nouns, such as "able" to "ability" and "noble" to "nobility". The pinyin spelling of *niubi* coincidentally ends in "bi". By adding "-ity" (actually -ility because the meaning of "able" is implied) to *niubi*, people believe that the created expression of *niubility* looks like an English word. However, this TSE does not inherit the conventional rule of transforming an adjective to a noun when adding the suffix of "-ility". Rather, people still use *niubility* as an adjective which has the same meaning as *牛逼*.

Another interesting example of morpho-semantic TSE that are formed through affixation is *ungeilivable* (or *ungelivable*). *Ungeilivable* is the Trans-semioticising equivalent of *不给力* (bu gei li), which means "not being helpful enough" or "not being awesome" in English. The Chinese word *给力* (gei li) comes from Minnan dialect, *激力* (ji li). In Minnan dialect, the character *激* (ji) is pronounced similarly to *给* (gei) in Mandarin. The word *给力* (gei li) first became popular due to a Chinese dubbing of a Japanese anime. Then it was widely used during the 2010 World Cup. The formation of *ungeilivable* includes three steps. The first step is to transform the Chinese characters into pinyin spelling. The second step is to add the prefix, "un-", to the pinyin spelling. The third step is to add the suffix of "-able" to the end of *ungeili*. Interestingly, the "v" letter is inserted between "ungeili" and "-able" possibly because the word "give" (给) has the letter in it, or because the pronunciation of *ungeilivable* is close to the English word "unbelievable". The Chinese word *不给力* is an adjective both before and after its Translanguaging transformation.

TSEs through composition

My data indicates that compounding is also a very productive way of creating TSEs. For example, *立 flag* (literally meaning "setting a goal" and usually referring to "someone setting a goal but failing to achieve it") is a TSE that is composed of two words, one from Chinese and the other from English. So *A4 腰* (A4 waist) is characterised by a tiny midriff which can be fully covered by a vertical sheet of A4 paper (Wang & Su, 2018). When people combine Chinese characters and emojis together, they may also create compound TSEs. For example, in the case of 🐇*粉*, which refers to the fans of Twice (a K-pop girl group in Korea), its original expression is *兔粉* (tu fen). The Chinese character *兔* (tu) comes from the Korean pronunciation of Twice, which sounds like "tu wai si". The first syllable "tu" shares the same pronunciation as all the Chinese characters that are pronounced as "tu". And of all the characters that are pronounced as "tu", *兔* (tu), which means rabbit, can best describe fans' love to the girl group as they are as cute as rabbits. Finally, people use the emoji of rabbit to replace the Chinese character *兔* (tu). The Chinese character *粉* (fen) means fan in English. When someone says "I am the *粉* (fen) of somebody", it means this person is the fan of somebody. *粉* (fen) comes from the Chinese word *粉丝* (fen si), which is the Trans-semioticising equivalent of fans as they share similar pronunciation. Therefore, in the combination of 🐇*粉*, the emoji of rabbit and the Chinese character, *粉* (fen), are combined together to form a new expression. Before the combination, 🐇 refers to Twice and *粉* (fen) refers to fans. After the combination, both 兔子 and *粉* (fen) keep their original meaning. So 🐇*粉* (🐇 fen) belongs to the first type of composition whose meaning is the sum of their composite meaning.

Book 思议 (book si yi), however, belongs to the second type with the meaning of one component changed. *Book 思议* (book si yi) is the Trans-semioticising equivalent of *不可思议* (bu ke si yi), which means "unimaginable" in English. *思议* (si yi) means "imagine" or "think" in English, which is often used in the Chinese idiom of *不可思议* (bu ke si yi). "Book" is an English word, which has a similar pronunciation to that of the Chinese characters *不可* (bu ke). The composition of *book* and *思议* (si yi) forms a TSE. In this composition, the meaning of "book" changes while the meaning of *思议* (si yi) remains intact.

The third type of composition with irreversibly lost meaning can also be found in the formation of TSEs. Let us take *打 call* as an example. *打 call* (da call) literally means "making a phone call" in English. *打* (da) is a verb which refers to the act of dialing in Chinese. However, as a Translanguaging expression, *打 call* means "cheer up for somebody". *打 call* comes from Japanese WOTA-gei (ヲタ芸) culture, where fans use light sticks and dance to support their idols during their concerts. This expression has become popular in China since 2017 and is gradually used to replace *加油* (jia you), which means "come on" or "cheer up". In this case, the meaning of *打 call* is irreversibly changed from making a phone call to "cheer up".

TSEs through phonetic motivation

In my data, many TSEs are created based on similar pronunciation[1], which brings

[1] Homophones are common in Chinese. Many Chinese characters share the same or similar pronunciation. Before the Pinyin was developed, the discussion of homophones in Chinese was based on people's experience and the storage of Chinese characters. Pinyin as a verbalised version of oral Chinese can intuitively reflect the pronunciation of characters and make it easier to locate as many homophones as possible. Also with the support of typing system, when the conversation or writing activity takes place on technical devices, people can directly get access to all homophones that share the same or similar pronunciation by simply typing the Pinyin. For example, when we type "ai", the typing system will present all the Chinese characters that are pronounced as "ai"

different semiotic elements into the process of formation. Here are some examples of TSEs that are built on phonetic similarity.

First of all, many logographic TSEs are created based on similar pronunciation. For example, *妈惹法克* (ma re fa ke) is a TSE that only has four Chinese characters but refers to an English phrase. As *妈惹法克* (ma re fa ke) is pronounced similarly to "mother fucker" ([ˈmʌðə ˈfʌkə]) in English, people create a Translanguaging expression with only one type of semiotic resource. When reading the TSE, people can think about the English pronunciation through the relation between the meaning and phonetic form.

Secondly, TSEs that contain more than one type of semiotic resource also use homophonic skills during their formation. In the TSE of *带 jio 布* (dai jio bu), both Chinese and Latin letters are used. This TSE is a Trans-semioticising equivalent of a Japanese expression, 大丈夫 (だいじょうぶ), which is pronounced as "dai jouu bu). The TSE's pronunciation is very close to Chinese pinyin, "dai jiao bu". So people find Chinese characters to present the pronunciation of *dai jiao bu*. Among many potential options, *带脚布* (dai jiao bu) becomes one of chosen character combinations with its interesting literal meaning, which is "wearing a feet cloth". Here, *jio* replaces *脚* (jiao), which means foot in English as 1) the pronunciation of *jio* is close to Japanese and 2) in some Chinese dialects, *脚* (jiao) can be pronounced as *jio*, which sounds

---

in the list for users to choose. Here is the screenshot of the typing system when "ai" is typed.

cute and interesting. Finally, the homophonic TSE from Japanese to the combination of Chinese characters and English letters is formed.

TSEs that involve Arabic numbers are also found formed due to phonetic motivation. In English, people have created innovative expressions that use the pronunciation of numbers to replace words for part of a word in sentences. For example, in *4u*, the number *4* refers to "for", so that *4u* means "for you" in English. *B4* refers to "before" as *4* is pronounced in the same way as "fore" in the word "before". Similar situations also happen in the Chinese context, while people use numbers to find suitable homophones other than their English pronunciation. Take a look at Table 4.2.

Table 4.2 Numerals and corresponding pinyin spellings

1	yi (yī, yí, yǐ, yì), yao (yāo, yáo, yǎo, yào)
2	er (ér, ěr, èr)
3	san (sān, sǎn, sàn)
4	si (sī, sǐ, sì)
5	wu (wū, wú, wǔ, wù)
6	liu (liū, liú, liǔ, liù)
7	qi (qī, qí, qǐ, qì)
8	ba (bā, bá, bǎ, bà)
9	jiu (jiū, jiú, jiǔ, jiù)
0	ling (līng, líng, lǐng, lìng), 洞 (dòng), 圈 (quān)

For each number from 0 to 9, we can find corresponding characters with the same pinyin spelling but different tones. For example, the pinyin of the number *1* is *yī*. It can be used to replace all characters whose pinyin is yī, yí, yǐ, or yì. Moreover, since *1* can also be pronounced as *yāo*, it is also capable of replacing characters that are pronounced as yāo, yáo, yǎo, and, yào.

Let us take a look at some examples. *555* is the Trans-semioticising equivalent of

呜呜呜 (wu wu wu), which refers to the sound of sobbing. The number *5* is pronounced as *wǔ* in pinyin while *呜* is pronounced as *wū* with the difference lying on the tone. *666* is made up of three *6*s, which is pronounced in the same way as *溜溜溜* (liu liu liu). The Chinese character *溜* (liu) is an oral expression from the northern part of China, which means "being good at something" in English. For example, when we say someone plays a video game very *溜* (liu), we mean that this person is very good at this game. The more *溜* (liu) or *6* people use, the higher degree of awesomeness the expression can express.

In the above two examples, numbers performing as alternatives of Chinese characters in original Chinese expressions brings no extra meaning to the TSEs. However, in some cases, numbers are used as homphones of Chinese characters but also add new meaning to the created expressions. For example, in *无1无靠* (wu 1 wu kao), the number *1* replaces the Chinese character of *依* (yi). *无依无靠* (wu yi wu kao) is a Chinese idiom which means "having no one to depend on" in English. However, in this TSE, *1* refers to people who are tops among gays. By changing *依* (yi) to *1*, people create a new meaning of this idiom, which means there is no top to rely on. People usually use this TSE when they find many bottoms on social media but few tops.

As shown in Table 4.3, besides considering numerals as Arabic numbers, there are also numerals that are used to represent numbered musical notations.

Table 4.3 Numerals and musical notations

1	2	3	4	5	6	7
do	re	mi	fa	so	la	si

These notes also have homophones in Chinese. *67* is a typical case of this type.

67 refers to two notes which are *6* (la) and *7* (si). The pronunciation of *la si* sounds close to *la xi* in pinyin. Therefore, people sometimes use *67* to replace *拉稀* (la xi) in Weibo, which means diarrhea in English. *456*, which is pronounced as *fa so la* according to the musical notation form, is the Trans-semioticising equivalent of *发骚啦* (fa sao la). The pronunciation of *5* (so) is similar to that of *骚* (sao) in Chinese, which means "totty" in English. *4* (fa) has the same pronunciation as *发* (fa) and *6* (la) is pronounced in the same way as *啦* (la). The whole expression of *456* refers to a status of someone whose behavior is very totty.

4.1.2 TSEs through semantic neologism

One example of giving new meaning to an existing word via semantic neologism or semantic transfer is *花粉* (hua fen). Literally, *花粉* (hua fen) means pollen in English. However, in Weibo, people assign a new meaning to this word. *粉* (fen) is the Trans-semioticising equivalent of *fans* in English. As mentioned before, someone's *粉* (fen) means "the fans of someone". *花* (hua) is the homophone of *华* (hua), which is the first character of *华为* (hua wei), a famous technology company from China. Therefore, *华粉* (hua fen) is the acronym of *华为的粉丝* (hua wei de fen si), which means "the fans of Huawei". As *华粉* (hua fen) shares similar pronunciation with *花粉* (hua fen), Huawei's fans choose the latter as the name of themselves.

Emojis can also be given new meanings to become semantic transferred TSEs. Although the emoji official website has given precise definition on each emoji, people still use emojis flexibly and creatively. In Weibo, many emojis are given meanings beyond their original designs. Let us start with  which is an emoji that represents a

cup of tea. Due to its color, people usually consider it as green tea. In Chinese culture, green tea is related to purity, elegance, and transparency. However, green tea has been given another meaning as a metaphor of girls who look pure and harmless but scheming and sinister since a notorious feast in 2013. Some girls who joined the activity looked fancy but were actually looking for sugar daddies. Hence, 🍵 bears a new meaning extended from its original meaning.

Emojis sometimes achieve new meanings by homophones. For example, ⚽⚽ is formed by two football emojis but its meaning has nothing to do with football. In effect, ⚽⚽ is the Trans-semioticising equivalent of *求求* (qiu qiu), which means "begging" or "pleading" in English. *求* (qiu) shares the same pronunciation as *球* (qiu), which means ball in English. Thus, *求求* (qiu qiu) is replaced by *球球* (qiu qiu) in advance. Then people use the emoji of football (as it usually comes first when people type through the keyboard) to express the characters of *球* (qiu). Finally, ⚽⚽ gets a brand new meaning transferred from *求求* (qiu qiu).

4.1.3 TSEs through morphological mechanism

In my data, many TSEs employ blending. For example, in Chinese, we often use 大咖 to refer to someone who is very important or successful in his or her field. However, in the TSE low 咖, only the splinter "咖" is used, together with another adjective, which refers to the losers who have very bad taste or act like an under-educated rural person. Here are three more examples of TSEs that are formed through morphological mechanism.

The first TSE is *cpdd*. *Cpdd* is composed of two splinters, which are *cp* and *dd*. *Cp* is the acronym of "character pairing" from English. However, the concept of "character pairing" is from Japanese Doujin works, which are amateur manga

publications, featuring either a cast of original characters and plot or characters from another manga or an anime. Now in the Chinese context, the meaning of *cp* is expanded to couples or any other online relationships between two people. *Dd* is the pinyin initial of 滴滴 (di di), which is the notification sound of QQ, an instant messaging social media in China. By combining the two separate splinters from different language background, *cpdd* is formed as an innovative TSE, which means "if you want to find someone to pair with you, please send me message".

阿婆主 (a po zhu) is another TSE containing one splinter which comes from Japanese and a Chinese character. In Japanese, the expression うｐ主 is used to describe people who post or upload videos on social media. The first splinter is うｐ, which is the Translanguaging expression of "up", which is an abbreviation of "upload". The character 主 means the person who conducts the action, which performs as the "er" in "uploader". In the Japanese expression, the *u* in "up" is replaced by Japanese Hiragana character う, whose Romanji is *u*. Therefore, うｐ主 itself is a Japanese-English Translanguaging expression using Japanese Hiragana character to express English letter that shares similar pronunciation. When people bring this expression to Chinese social media, うｐ主 experiences further transformation. Chinese characters, 阿婆 (a po), which sound close to うｐ, are used, which mean "grandmother" or "the older women" in some Chinese dialects. In Japanese, 主 is pronounced as ぬし (nu shi), but in Chinese, it is pronounced as *zhu*.

The third TSE expression is 笑 *skr* 人 (xiào sǐ gè rén, meaning "something is very funny so that the person laughs to death"). Zhang and Ren (2020) did a study on *skr*, finding that people created a lot of new blends with the splinter *skr,* such as 气 *skr* 人 (qì sǐ gè rén, meaning "something is very annoying so that the person feel

extremely angry"), *s 谁在唱 kr* (shì shéi zài chàng gē ér, meaning who is singing a song), and *skr 狠人* (shì gè hěn rén, meaning someone is very hard on himself or someone is very brave), etc. Some of these examples (*气 skr 人* & *笑 skr 人*) follow the third type of Lehrer's classification, which is blending. However, the case of s 谁在唱 kr is different as the splinter is discontinuous and divided into two parts with a phrase (or a short sentence) embedded. Moreover, the meaning of this splinter in this case is changed as *s* here refers to "是" (shì, meaning is) and kr refers to "歌儿" (gē ér, meaning "song"), which together means "someone who is singing a song". Also, *skr 狠人* is in the opposite form of type one with an entire word following a splinter. In this case, *skr* means "是个" (shì gè, meaning "is a") rather than "死个" (sǐ gè, meaning "a person is dead").

4.1.4 TSEs through External mechanism

Many TSEs in my data are external neologisms as they involve foreign languages during the process of formation. Here are some types of TSEs that are formed by borrowing. The external mechanism refers to TSEs that involve elements that come from external languages, which in other words means that languages that are not Chinese have been used to form the TSEs.

A. Borrowing pronunciation and meaning partly from other languages

Some TSEs involve borrowing meaning from other languages. For example, in the TSE *花粉*, "粉" actually refers to "fans", which comes from English. Also in the TSE *low 咖* mentioned above, "咖" is also a phonological loanword originating from "cast" in English. "Low" is an English adjective, however, whose meaning has undergone changes in Chinese online communication, referring to "being bad-tasted, rural and under-educated".

In some other TSEs, people may borrow words from other languages and mix these words with Chinese words to create TSEs. For instance, *牛逼普拉斯* (niu bi pu la si) is a TSE formed by *牛逼* (niu bi) and *普拉斯* (pu la si). *牛逼* (niu bi), as introduced in *niubility*, means "awesome" in English. The second part is *普拉斯* (pu la si) originates from "plus" in English due to the similar pronunciation they share. Many smart phone companies have put forward different types of phones. For those phones that have large screens, they are often named with a "plus" as a suffix, such as iPhone 8 Plus. Therefore, people consider that if something is the "plus" version, it must be stronger, bigger and better. So *牛逼 plus* (niu bi plus) means that something is "extraordinarily awesome" as it is a "plus" version of *牛逼* (niu bi).

B. Borrowing pronunciation and meaning wholly from other languages

Some TSEs only have Chinese characters but are actually borrowed from other languages. For instance, *买米当卡* (mai mi dang ka) is a TSE that comes from Thai. In Thai, ไม่มีตังค์ค่ะ (Mai mee tang ka) means "I don't have money". As the pronunciation of *买米当卡* (mai mi dang ka) is similar to the Thai phrase, people use this TSE to replace *我没有钱* (wo mei you qian) in Chinese, which also means "I don't have money". A similar case is *欧巴* (ou ba). This is the Trans-semioticising equivalent of the Korean word *오빠* (oppa), which means "elder brother". *오빠* (oppa) is usually used by women to address "men who are older than them, especially when they have close relationship". As South Korean love TV series are very popular in China, the addressing of *오빠* (oppa) leaves a deep expression to the audience. People then use Chinese characters, *欧巴* (ou ba), that have similar pronunciations to express the Korean word.

C. Borrowing logographic elements from other languages

Some TSEs, however, only borrow logographic elements from other languages to

form new expressions. For instance, 王の蔑视 (wang de mie shi) is a Chinese-Japanese mixed TSE that means "the king's contempt" in English. In this TSE, the second character の (no) is a Japanese character that illustrates a possessive relationship, like "of" or "'s" in English. The TSE 王の蔑视 (wang de mie shi) is the Trans-semioticising equivalent of 王的蔑视 (wang de mie shi) with the 的 (de) replaced by の (no). On the one hand, the meaning of 的 (de) is similar to の (no) in Japanese. On the other hand, in Chinese calligraphy, the graph of 的 (de) also looks like の (no). After borrowing Japanese character to form this expression, the meaning of 王的蔑视 (wang de mie shi) does not change but may achieve special pragmatic effects in certain context (refer to 6.4 for more details).

Another example is *i* 了. This TSE is composed of one English letter and one Chinese character, which is used as a declaration that the speaker has fallen in love with something. People may be asking since both English alphabet and pinyin contain the letter *i*, why the *i* is from English. It is decided by the meaning of this TSE. *i* 了 is the Trans-semioticising equivalent of 爱了 (ai le), so *i* should be pronounced in the same way as the Chinese character 爱 (ai). In pinyin, *i* is pronounced similarly to [i:], while in English it is pronounced as [ai]. In this TSE, only the pronunciation of *i* is borrowed to replace 爱 (ai), bringing no extra meaning to itself.

The third example, "maybe", is in the form of an English word but it is given a new meaning from Chinese. It is the Translanguaging expression of 霉逼 (mei bi), which means "a person who is very unlucky". 霉 (mei) literally means "mould", which is used as a metaphor to denote bad luck in Chinese. 逼 (bi) is an oral vulgar expression. The pronunciation of 霉逼 (mei bi) is similar to that of the English word *maybe* [ˈmeibi], so people use the English word as an alternative to the Chinese vulgar

expression with its English meaning abandoned.

As mentioned above, for each of the mechanisms of creative word formation mentioned by Lipka (2002), there can be corresponding TSEs that adopt the same mechanism. However, it is to be noted that although Trans-semioticising expressions found in my data use the same set of mechanisms, their formations are more complex in five significant ways. First, they employ different linguistic or semiotic elements, as in the case of "打 call" and "无 1 无靠" (for more details, refer to section 4.2). Second, they may sometimes apply more than one type of formation mechanisms. For example, *阿婆主* (a po zhu) falls into the categories of both morphological and external TSE at the same time. It is formed by two splinters, which makes it a morphological TSE. But when we look into its formation process further, we can see that this TSE also borrows Japanese word, so that it is also an external TSE. Third, some TSEs apply the mechanisms implicitly as in the case of "买米当卡" for "I don't have money" and "maybe" for "霉逼" (for more details, refer to section 4.3). Fourth, some TSEs undergo a semantic change or loss, as in the case of "打 call" and "low 咖". Fifth, some TSEs involve complex processes, as in the case of "ungeilivable", in which "geili" and "give" are both incorporated.

Another noteworthy point is that homophony plays an important role in the formation of many TSEs using other mechanisms such as semantic transfer and external neologism. Homophonic association is a bridge that provides opportunities for different semiotic (language) systems to transfer from each other. TSEs (e.g. *无 1 无靠*, *牛逼普拉斯*, *阿婆主*) whose formation involve more than one language all use homophones to replace original characters or letters during their formation.

4.2 Categorisation based on the semiotic elements contained

TSEs in my data may contain two or more types of semiotic element, suggesting the

varying complexities of such creative lexical items.

4.2.1 TSEs with two semiotic elements

Table 4.4 presents the different combinations of two linguistic semiotic elements of TSEs that are found in my data:

Table 4.4 Combinations of two linguistic or semiotic elements in TSEs (N=1611)

Combinations of linguistic or semiotic elements	**Illustrations**	**Numbers & Frequencies**
Logographic characters + Latin alphabet letters	笑 skr 人	1068 (66.3%)
Logographic characters + Numerals	遍地飘 0	37(2.2%)
Logographic characters + emojis	牛🍺	351(21.8%)
Latin alphabet letters + Numerals	y1s1	145(9.0%)
Logographic characters + other symbols	+关注	10(0.7%)
Total		1611(100%)

Five combinational types of TSEs consisting of two semiotic types of elements are found in my data, involving five different semiotic systems, which are logographic characters, Latin alphabet letters, numerals, emojis, and other symbols (mostly mathematical symbols). In addition, there is a further type of TSEs that contain more than two semiotic types of elements, suggesting potential complications of Trans-semioticising. From Table 4.4, it is evident that the TSEs that are encountered on Chinese social media are made up of Logographic characters + Latin alphabet letters, which are mostly Chinese characters plus English letters. This reflects the impact of English learning on Chinese netizens.

Here is a caveat. When naming which language a certain logogram or letter belongs to, a processing action has been conducted, which will be discussed in the chapter of understanding of TSEs. Logographic characters include Chinese and Japanese characters, which are simplified Chinese characters and Japanese Kanji (including adopted traditional Chinese characters and syllabic kana). In Weibo, the main language used by the netizens is Chinese. Thus, most of the Logographic characters are Chinese characters. However, due to the similarity between Chinese and Japanese characters, we cannot directly distinguish whether a character is Chinese or Japanese. Therefore, we simply categorise all the characters used into logographic symbols. Similarly, pinyin letters and English letters both are based on Latin alphabets with two main differences: 1) pinyin has four tones to mark on the vowel letters and 2) it does not have the letter "v". However, when people use pinyin in Weibo, the signs of tones are usually omitted. For instance, "wǒ" can be written as "wo". Therefore, to avoid judging which language system a certain TSE belongs to, I categorise English and pinyin letters into Latin alphabet system. Numerals and emojis are two independent symbolic systems although they are used together with logographic symbols. The last group of semiotic symbols is "other symbols" including symbols like "+", "@", "=", etc.

a. Logographic characters + Latin alphabet letters

TSEs composed of logographic characters and Latin alphabet letters (LLTSE) are widely used in Weibo posts. In this type of TSE, the logographic characters are usually Chinese characters while the Latin alphabet letters can be either English letters or pinyin letters. Here are some examples of LLTSEs found in my data.

In TSEs that contain English letters, 笑 *skr* 人 is a representative example. 笑 *skr* 人 (xiao skr ren) is formed with Chinese characters, 笑 and 人, and English letters, *skr*. This expression means that something is very funny so that it makes people laugh fiercely to death. It is an exaggerated way of expressing the degree of fun, similar to the English social media expression, *lmao* (laugh my ass off). In this

TSE, 笑 is the Chinese character that refers to "laugh" and 人 literally means "people". 笑 *skr* 人 can be regarded as the Trans-semioticising equivalent of 笑死个人. The two Chinese characters 死个 are pronounced as "si ge", which sound similar to *skr* in English. *Skr* is an onomatopoeia that is usually used in Hip-Hop songs mimicking the sound made by tires when the car is drifting. In 笑 *skr* 人, *skr* loses its original English meaning but only provides its pronunciation during the formation of this TSE.

The same formation pattern can also be found in some other LLTSEs. Let us take 牛 *B* as an example. This TSE contains one Chinese character, 牛, and one English letter, *B*. 牛 *B* is the Trans-semioticising equivalent of a classic Chinese vulgar word, 牛逼 (niu bi), which literally means the virginal (逼) of a cow (牛) and is usually used to describe something that is awesome. In Chinese, the word 逼 is one common but impolite expression of female's reproductive organ. It is pronounced as *bi*, which sounds very similar to the English letter *B*. Therefore, people use *B* to replace the impolite character 逼, by borrowing the pronunciation of *B* while neglecting its meaning of being an alphabet letter. It must be mentioned that both *B* and *b* can be used in this expression as capitalisation does not influence its pronunciation.

However, there are also some LLTSEs that borrow both English pronunciation and meaning in the course of their formation. In the TSE of *p* 图 (p tu), *p* is the initial letter of "Photoshop", which is a famous photo editing software. Before all kinds of selfie editing applications become popular, Photoshop is the most widely used tool to make photos look fancier, so people use Photoshop as a verb to describe the process of using apps to beautify photos. Then the long word Photoshop was simplified to a single initial letter, *p*, to refer to such behavior. 图 is the Chinese character that means "photo" or "picture". 图 usually refers to photos, especially selfies, shot by

cellphones. *P* and *图* together illustrate the act of using apps to beautify photos. In this TSE, the meaning of *p* is from English word and the pronunciation of the letter *p* is also used.

Sometimes, the meaning of the English part in TSEs gets expanded in the Chinese context. For instance, in the expression of *low 咖* (low ka), part of the English meaning of "low" is reserved. *Low 咖* is composed by one English word, "low", and one Chinese character, *咖*. In this expression, "low" is used to describe someone who has poor moral character, bad fashion taste, or is corny. Its meaning in the Chinese language context has been extended. *咖* is a personal pronoun, originating from Min Nan dialect. When we say someone is a *大咖* (da ka), we consider this person is an important or successful figure in his or her career.

Pinyin letters can also be the phonographic element of LLTSEs. There are mainly four types of pinyin-related LLTSEs found in my Weibo data. The first type is the use full of pinyin spelling in a TSE. For instance, *强 jian*. It is a TSE that means "rape" in English. It contains one Chinese character and one pinyin expression. The Chinese character *强* (qiang) means "force someone to do something" and the pinyin expression is the pinyin of *奸* (jian) without tone marker, which means "having sex unwillingly". By shifting the Chinese word *强奸* (qiang jian) into a Translanguaging expression, *强 jian* (qiang jian), the new way of expression changes its form but keeps its original meaning.

The second type of pinyin-involved TSEs are those that use pinyin to express some rare characters, or some oral expressions used in dialects that cannot be input through the keyboard. One of the popular cases is *笑 yue 了* (xiao yue le). *笑* (xiao) means "laugh" and *了* (le) is a Chinese character that helps to express the condition where something has already been done. *Yue* is the pinyin of *哕* (yue) without tone

marker. The character 哟 (yue) is a rare character that often used in Shandong dialect. 哟 (yue) in English refers to the action of vomiting but nothing goes out. This expression means that the speaker laughs too hard so that he or she wants to vomit.

The third type of LLTSE does not involve a full pinyin spelling. For example, the TSE 代*y* (dai y) is formed by one Chinese character 代 (dai) and one pinyin letter *y*. 代 (dai) in Chinese means "to do something for others". *y* is the pinyin initial of the Chinese character 孕 (yun), which means "pregnant" in English. Therefore, 代孕 (dai yun) means "surrogacy". Surrogacy is an illegal behavior that is not permitted in China. People use this TSE when discussing or commenting on a Chinese actress's surrogacy affair. By using the letter *y* to replace 孕 (yun), 代孕 is transformed to 代*y*, which can still convey the same meaning but in another form.

The fourth type is using an English letter to replace the Chinese character which has similar pronunciation. Let us take 放 *p* as an example. 放 *p* (fang p) is the Trans-semioticising equivalent of 放屁 (fang pi), which literally means fart in English. However, within communication, 放*p* can also refer to "bullshit" in English, expressing other's opinion is nonsense. 放 (fang) is a verb that is used to describe the action of producing a fart. *p*, is not only the initial letter of the pinyin of 屁 (pi), but also is the homophone of the Chinese character 屁 (pi).

b. Logographic characters + Numerals

Logographic characters along with numerals can also form Trans-semioticising expressions (LNTSE). In my data, the most common type of LNTSE is Chinese characters combined with numerals. Here is an example that illustrates this type of TSE. In the expression, 遍地飘 *0* (bian di piao 0), there are three Chinese characters and one numeral, *0*. This TSE originates from a Chinese idiom, 遍地飘零 (bian di

piao ling), which can be used to describe a depressive scene in autumn where leaves are all over the ground. However, the numeral, *0*, brings a new meaning to this expression. The Chinese character *零* (ling) means “zero”, which is also equal to *0* in Arabic numerals. The numeral *0* is now used to refer to a gay who performs as a Bottom in the Chinese language context. The culture of queer sex, Tops generally like to be the more dominant person during sex, while Bottoms follow their lead. Therefore, the expression of *遍地飘 0* means there are so many Bottoms (gay) on the road (on social media), implying that there are limited Tops (gay) that can be seen. In this case, people use a different meaning of the numeral and give the traditional idiom a new meaning.

c. Logographic characters + emojis

In my data, Chinese characters can be used together with emojis to form a TSE. There are three types of formation of logographic character and emoji TSEs (LETSE).

The first type of LETSE is using the literal meaning of an emoji to replace a certain Chinese character in an expression. For instance, in 理 (he li), one character emoji and one Chinese character form the Translanguaging expression. *合理* (he li) in Chinese literally means something is “in line with certain rules or reasons”, so when people consider a point of view is *合理* (he li), it means that "this point of view “makes sense or is reasonable”. In the emoji system, there is an emoji ( ) that employs the same calligraphy with the Chinese character *合* (he). This emoji comes from the Japanese language, 合「ごう」, which can be used to refer to 合資会社 (ごうしがいしゃ), which means “limited partnership”. As Chinese and Japanese share some similar characters due to the history of these two languages, Chinese people can find some Japanese character emojis useful in the Chinese context. In this case, Chinese people consider the meaning of as its Chinese meaning.

Some LETSEs use the Chinese pronunciation of the emojis’ meaning to create

new meaning. For example, in the TSE of 你🐴死了 (ni 🐴 si le), three Chinese characters and one emoji form the expression. In this TSE, the emoji is a horse, whose corresponding Chinese character is 马 (mǎ). 马 (mǎ) also share a similar pronunciation with 妈 (mā), which means "mother". The Chinese character 你 (ni) means "your" and 死了 (si le) means "dead". Therefore, 你🐴死了 (ni 🐴 si le) refers to a curse expression which means "your mother is dead". The user of the TSE first finds a homophone of the key or most sensitive character in Chinese. Then he or she uses the corresponding emoji of the homophonic expression to replace that character.

The third type of LETSE uses the original meaning of emojis to replace the Chinese characters. In the case of 打💉 (da💉), one Chinese character and one emoji form the expression. 打 (da) here is a verb meaning "to give". The emoji is a medical needle. So 打💉 (da 💉) means "to give an injection". This type of TSE directly uses the meaning that emoji conveys to replace the character.

d. Latin alphabet letters + Numerals

Latin alphabet letters can form TSEs together with numerals in Weibo data. Since the Latin alphabet letters include English letters and pinyin letters, I will discuss each type separately.

For TSEs formed with English letters and numerals, let us take *p1* as an example. *p1* contains English letter *p* and numeral *1*, which means "photo No.1". *p* is the initial letter of *photo* while *1* is a serial number. Following this way of formation, people also use *p2*, *p3*, etc. to mark the name of the photos they share on Weibo.

Pinyin letters are often used together with numerals in Weibo. In the TSE of *y1s1*, there are two pinyin letters, *y* and *s*, and two numerals, *1*. *y1s1* is the Translanguaging expression of 有一说一 (you yi shuo yi). 有 (you) means "have" and 说 (shuo) means "say" or "talk". 一 (yi) means "one". Therefore, the literal meaning of 有一

说一 (you yi shuo yi) is “if we have one thing to talk about, we should talk about this thing”. By saying this, the speaker implies that he or she will talk about something and stick to it rather than offering compliments. Then its meaning is extended to “to be honest" or “to tell you the truth”. Note that the formation of *y1s1* involves the pinyin initial letters (*y* and *s*) of *有* (you) and *说* (shuo) and the numeral equivalent of the Chinese character 一 (yi), which is *1*.

e. Logographic characters + other symbols

The last type of semiotic combination in creating TSE has Chinese (logographic characters) characters and other symbols such as punctuation marks and mathematical symbols. For instance, people can add guillemets to certain Chinese expressions. In the case of «*真香*», the two Chinese characters are put between the two guillemets. In French writing, Guillemets are usually used to indicate the names of books or movies. However, in Chinese, “真香” (zhen xiang) literally means something smells or tastes very good. However, in this case, the TSE is related to a famous story that happened during a Chinese reality show. In that show, a city boy was sent to a small village to experience life there. The boy did not want to eat the food made by the host at first because he thought everything was dirty and unhealthy in the village. He claimed that he would rather die than eat the food. However, when he felt too hungry to survive, he finally gave up his promise. When he was pigging out on the food, he said satisfactorily to the camera, “真香!”. When this TV show was broadcast, the boy’s reaction left a deep impression on the audience. Gradually, people use «*真香*» to describe similar situations where people cannot realise the advantages until they give a trial.

People also use mathematical symbols with Chinese characters to form TSEs. Let us take +*关注* (+ guan zhu) as an example. In this TSE, there are one mathematical symbol and two Chinese characters. *关注* (guan zhu) refers to the action of

following someone's social media account. +, which means "plus" or "add", refers to the Chinese character 加 (jia). The Chinese expression 加关注 (jia guan zhu) means to click the follow button and follow the account. On Weibo, the button of following uses the symbol of +, so that people are used to use 加 (jia) as the verb to describe the action of adding.

Above are the semiotic types of TSEs found from the Weibo data. We can see diverse semiotic systems are actively used to form TSEs. These TSEs have more than one type of symbols in their formation and can be directly identified without any difficulties.

4.2.2 TSEs with multiple semiotic elements

In my data, while two types of semiotic symbols are used during their formation **i**n most cases of TSEs, there are a limited number of TSEs that involve more than two semiotic elements in their formation. Here is an example. Like *y1s1*, the TSE *u1s1* also contains two explicit semiotic elements, which are Latin letters and numerals. In addition, *u1s1* has the same meaning with *y1s1*, which means "to be honest". *u* refers to the Chinese character 有 (you). *s* refers to 说 (shuo). Two numerals *1* refers to 一 (yi). The difference of *u1s1* and *y1s1* lies in the letters *u* and *y*. In *y1s1*, the letter *y* is the initial pinyin letter of 有 (you). However, unlike *y1s1*, *u1s1* involves the use of the English letter *u* to replace the Chinese character because *u* and 有 (you) share similar pronunciation. In this case, when using *u* to express 有 (you), the pronunciation of the English letter is involved along with pinyin letter and numerals. Thus, *u1s1* is a multiple-element TSE. This type of TSE is special as it can be identified as it apparently has two semiotic symbols in its literal formation, but it involves another semiotic systems during its actual formation.

Another example of multiple-symbol TSEs is *077*. *077* is a nickname created by Chinese NBA fans for Luka Dončić, a Slovenian NBA player. *077* is the Trans-semioticising equivalent of *东契奇* /dong qi qi/, the Chinese translation of Dončić. *77* (七七) is pronounced as /qi qi/, which is the same as *契奇* /qi qi/. People use the homophones to replace the original Chinese characters. Using *0* to represent *东* is also related to homophonic mechanism. Here, *0* is pronounced as /ling/, which is not the same as its usual pronunciation /ling/. In fact, it is because the numeral *0* looks like a round hole. Hole in Chinese is written as *洞* /dong/. Then since *洞* and *东* are homophones, people use *0* to replace the character *东*. Thus, to understand this TSE, readers need to make use of four types of semiotic elements, namely Chinese, English, numerals, and pictures through homophony and other kinds of association (for more details, refer to Section 5.3).

4.3 Categorisation based on the explicitness of Trans-semioticising

My data shows that not all TSEs explicitly combine two or more linguistic or semiotic elements in their formation. Thus, explicit TSEs and implicit TSEs can be distinguished according to the explicitness of Trans-semioticising.

Explicit TSEs cover all kinds of TSEs that explicitly have more than one type of symbolic symbols. In the previous section, the TSEs I mentioned are almost exclusively explicit TSEs (but *u1s1* is somewhat different, as it is both explicit and implicit).

In contrast, in implicit TSEs, there is only one type of semiotic element in its surface form and manner of formation, like a regular word or phrase. Thus, the identification of implicit TSEs is more complicated than that of explicit ones.

To identify implicit TSEs, we must involve a certain degree of interpretation on the TSEs. This process of interpretation differs from figuring out what does a TSE

exactly mean in the context. To identify whether a one-semiotic-system expression is a TSE or not, we just need to use our knowledge to count the number of involved semiotic systems. Based on this standard, Table 4.5 presents the different types of implicit TSEs that are found in the data.

Table 4.5 Types of implicit TSEs

Types of Implicit TSEs	**Sub-types of Implicit TSEs**	**Illustrations**
Logographic characters	Chinese + English	*粉丝*
	Chinese + Japanese	*八格牙路*
	Chinese + Korean	*欧巴*
	Chinese + pinyin (Chinese Homophone)	*卧槽*
	Chinese + other symbols	*艾特*
Latin alphabet letters	Chinese + English	*uu*
	Chinese + pinyin (initials)	*dddd*
	Chinese + pinyin (full spelling)	*zhuan*
	English + pinyin	*vx*
	Japanese	*ky*
Numerals	Chinese + numerals	*520*
Emojis	Chinese + Emojis	
	Chinese + Emojis + Numerals	*077*

There are four main types and 13 subtypes of implicit TSEs found in my data. As we can see from the examples, even though some TSEs share the same semiotic system, they require background knowledge about different semiotic resources to understand. In the sub-type of classification, I separate Chinese characters and Chinese pinyin into two different systems for three reasons: 1) Chinese characters belong to logographic system while pinyin belongs to phonographic system; 2) pinyin

cannot directly match specific Chinese character as one pinyin string can refer to many Chinese characters; 3) pinyin is not a must in Chinese as many people can still speak and write Chinese without learning pinyin. Take "dddd" as an example. If a person whose mother language is Chinese does not know pinyin, he or she cannot identify that "dddd" is the pinyin initial acronym for "懂得都懂/ dǒng de dōu dǒng". The type of Chinese + pinyin TSE works purely on homophony. The existence of pinyin enables us to explain how Chinese homophones function. Chinese + English/Japanese/Korean types are similar as they all use Chinese characters which have similar pronunciation to its original pronunciation in Chinese. For instance, "粉丝/ fěn sī" sounds like "fans" in English. In the following part, I will discuss each type and sub-type of TSEs in detail.

A. Logographic characters

Logographic character TSEs found in Weibo are all in the form of Chinese characters. The first three sub-types are all Translanguaging practices that use Chinese characters to express meaning by finding cross-language homophones. Let us start from Chinese + English TSEs.

a. Chinese + English TSEs

There are two different types of Chinese + English TSEs that are found in the data. The first type is entirely based on using Chinese-English homophones to express the meaning. Here is an example. 爱豆 (ai dou) means "idol", referring to those young and good-looking stars who can sing while dancing. The concept of idol industry came from K-Pop and became overwhelming in China since the 2010s. The pronunciation of idol is ['aidl], which sounds like "ai dou" in Chinese. Then young people use the character 爱 (ai), which means "love" or "admiration", to express the pronunciation of "i", as the character 爱 (ai) can show the fans' love for their idols. As for the pronunciation of "dol", people use 豆 (dou) as an alternative. In Chinese, the character 豆 (dou) can be used to describe a person who is small and cute.

Finally, *爱豆* (ai dou) becomes a TSE for "idol".

However, not all Chinese + English TSEs are formed with careful choice of characters. In fact, for one English word, there may be more than one TSE version. For the most famous English vulgar expression, "fuck you", there are at least three different corresponding TSEs, which are *法克鱿* (fa ke you), *法克油* (fa ke you), and *法克尤* (fa ke you). All these three expressions have the same Chinese pronunciation but vary on the third character. *法克* (fa ke) is pronounced like "fuck [fʌk]", but no evidence shows why people choose these two characters instead of others. One possible reason may be *法* (fa) and *克* (ke) are more frequently used comparing with other characters that are also pronounced as "fa" and "ke". It may also be influenced by the typing software as it can automatically recommend characters based on the pinyin people input. Similar situations also happen when choosing the homophonic character of "you". Both *油* (you) and *尤* (you) can be a choice as they have the same pronunciation. But no matter which Chinese characters are chosen, they need to be pronounced similarly to the English phrase. For *法克鱿* (fa ke you), there is a background story that explains the choice of *鱿* (you). *鱿* (you) refers to "squid" in English. In early 2009, people voted for top 10 Internet Mythical Beast in online forums. "Fuck you" as an international bad language got its position in social media. In order to associate the TSE of "fuck you" with an animal, people chose *鱿* (you), squid, to express "you". After that, *法克鱿* (fa ke you), which literally means "fuck squid", became a widely accepted TSE version of "fuck you".

Another type of Chinese + English TSEs involves some special translation techniques during the formation of the expression]. *油管* (you guan) is a typical example of this case. *油管* (you guan) refers to YouTube but not all syllables of

"YouTube" are transferred to Chinese homophones. Only "You" is replaced by 油 (you), which is similar to what happens in 法克油 (fa ke you). However, when it comes to "Tube", people translate it literally as 管 (guan). Finally, 油管 (you guan) becomes the TSE of "YouTube", although the Chinese word 油管 (you guan) does have its own meaning in Chinese, which means oil pipeline.

b. Chinese + Japanese TSEs

Implicit TSEs that only have Chinese characters in their forms sometimes involve Japanese knowledge during the process of formation. Let us have a look on several examples.

The first example is 搜嘎 (sou ga). In this TSE, both 搜 (sou) and 嘎 (ga) are Chinese characters, which do not mean anything in the Chinese context. Indeed, 搜嘎 (sou ga) is the homophone of a Japanese expression, そっか (souka), which means "I see" or "I get it". People choose two characters from Chinese characters that are pronounced like "sou" and "ka".

Another example is different from 搜 嘎 (sou ga), as it combines both homophones of Japanese characters and Chinese characters. The TSE of 霓虹国 (ni hong guo) means "the country of Japanese". In Japanese, "Japan" is pronounced as にほん (ni hon), so that Chinese people use a homophonic word to refer to the name of Japan, which is 霓虹 (ni hong). 霓虹 (ni hong) in Chinese refers to neon lights, which are very beautiful and in line with people's impression towards Japan. The character 国 (guo), means "country". In all, 霓虹国 (ni hong guo) becomes the name for Japan in China.

c. Chinese + Korean TSEs

Implicit TSEs that are formed with Chinese and Korean are also found in my data. The Translanguaging expression of "I love you" in Korean is used by Chinese Weibo

netizens. *撒浪嘿哟* (sa lang hei yo) has a similar pronunciation to that of the Korean expression, 사랑해요 (salangheyo). When people watch Korean TV series, they are deeply moved by the expression of 사랑해요 (salangheyo). Apart from *撒浪嘿哟* (sa lang hei yo), people also use *欧巴* (ou ba) to refer to boys who are tall and handsome. *欧巴* (ou ba) is the homophone of 오빠 (oppa), which means "brother" in Korean.

In the data, besides these TSEs that involve English, Japanese, and Korean, there are also some TSEs that involve some other languages. For example, *买米当卡* (mai mi dang ka), the homophone of ไม่มีตังค์ค่ะ, means "I don't have money"; in Thai *苏卡不列* (su ka bu lie), the homophone of Сука блядь, means "bitches" or "whore" in Russian; *逮虾户* (dai xia hu), the homophone of déjà vu, refers to "the feeling of having already dreamed something that is currently being experienced" in French.

d. Chinese character + pinyin letter TSEs

After introducing the implicit TSEs that involve languages from other countries, let us have a look at how pinyin performs an important role in Chinese homophones and helps to form TSEs.

TSEs that involve pinyin during their formation can also be seen as using Chinese homophones to replace the Chinese characters. Let us take *集美*(ji mei) as an example. It has two Chinese characters that are homophones of *姐妹* (jie mei), which means "sisters" in English. Apparently, it is not a TSE, as it does not seem to involve elements from two more semiotic systems. However, by my definition, it does contain elements from two semiotic systems, one being the system of logographic characters and the other being the system of Latin alphabet letters. Without the participation of the two systems, the formation and comprehension of the TSE would be impossible. There is another reason why expressions like *集美* (other similar

cases in my data include *蒜泥狠，可拷，可真刑*, and *司马*) generally have related anecdotal origins and perform special functions like mockery and criticality. For example, the novel pronunciation of *姐妹* (jie mei) as *集美* (ji mei), does not come from any Chinese dialect. It was first used by a Douyin influencer, Guo Laoshi. She often talked in a very strange tone, which helped her become popular in 2020. *集美* (ji mei) is one of the famous productions of her accent. When her fans commented on her videos, they used Chinese homophones to refer to the accent. In this case, Guo Laoshi used *ji mei* to express *姐妹* (jie mei), so her fans chose the characters that are pronounced as *ji mei*. Coincidentally, the word *集美* (ji mei) is the name of a university in Fujian Province.

In other TSEs that involve pinyin during their formation, Chinese dialects perform important roles. For instance, *丢雷老母* (diu lei lao mu), which means "fuck your mother", is a vulgar expression used in Cantonese. In Cantonese, the character, *丢* (diu), means "fuck" and *老母* (lao mu) means "mother". The character, *雷* (lei), is the homophone of the character of *你* (ni), which means "you" or "your". The pronunciation of *你* (ni) in Cantonese is *néi* (in different areas, the pronunciation may change a little). For people who do not speak Cantonese, they will choose the character *雷* (léi) when they try to denote this pronunciation. In this TSE, pinyin performs as a medium that transfers Cantonese oral pronunciations to Chinese Mandarin characters.

The third type of Chinese TSE which involve pinyin is not formed based on Chinese dialects. People create this type of TSEs because they want to use other characters to replace original characters for different reasons. Let us take *割割* (ge ge) as an example. *割割* (ge ge) is the Trans-semioticising equivalent of *哥哥* (ge ge), which means "elder brother". The Chinese character *哥* (ge) has the same

pronunciation as the character 割 (ge), however, only 哥 (ge) can refer to the meaning of "elder brother". 割 (ge) means "cut" in Chinese. In this case, it refers to the action of "cutting leeks" specifically. In Chinese, "cutting leeks" is a metaphorical expression implying the harvest of capital to the consumers. As leeks can be harvested for many times, comparing the consumers to leeks means that the capital can make money from consumers again and again. This economic activity also happens in the entertainment industry, as idols and their companies make money from their fans repeatedly. By replacing 哥 (ge) with 割 (ge),割割 (ge ge) is usually used by haters to address idols who make money from crazy fans.

e. Chinese + other symbols TSE

Some TSEs that only have Chinese characters may involve other symbols during their formation. In the case of 艾特 (ai te), which refers to the symbol of @, is an interesting example of this kind of TSE. The symbol @, in English, is called "at sign" or "at mark". When people read the symbol @, they usually say "at". In Weibo, when using the symbol @ before a username, the person with that username will be notified. Therefore, @ becomes a verb to express the action of using @ to mention someone else. Since the pronunciation of 艾特 (ai te) sounds similar to that of @, 艾特 (ai te) becomes the Trans-semioticising equivalent of the symbol @.

B. Latin alphabet letters

In Weibo, people not only use logographic (Chinese) TSEs, but also use TSEs that only have Latin alphabet letters. This type of TSE may also be composed of elements from different semiotic systems, including Chinese, English, pinyin, Japanese, etc. The following section will introduce each sub-type of Latin alphabet letters TSEs through examples.

a. Chinese + English TSEs

For TSEs that only have letters, their formation may involve Chinese and English. For example, *uu* is a TSE that is formed by two English letters, *u*. In English, though,

uu is not a meaningful word. In fact, *uu* shares a similar pronunciation with the Chinese word 友友 (you you), which is a new expression of 朋友 (peng you), meaning "friend" in English. Compared with 朋友 (peng you), 友友 (you you) is a more intimate way of addressing others, which was first used by elder people and online sellers. After this TSE went viral, many people used it to address other netizens friendly. This TSE contains English letters but does not involve any actual English word.

Nbcs, on the contrary, is an English related TSE, which does involve English words. *Nbcs* is the acronym of an English short phrase, "nobody cares". *Nbcs* is composed of the initial letter of "nobody" and "cares", the letter "b" of "body", and the letter "s" in "cares". The formation of *nbcs* is not a common way of English acronyms, such as *ASAP* ("as soon as possible"), *IDK* ("I don't know"), *GOAT* ("greatest of all times"), etc. These acronyms usually do not involve the consonant letter "s" in the combination. In effect, the formation of this TSE is related to the pronunciation of Chinese. In Chinese, the letter "s" can be expressed by Chinese characters that are pronounced as "si". Therefore, although "s" is a consonant in English, it is pronounced as a full syllable, which makes it an initial letter of the syllable "si". In this way, the phrase "nobody cares" can be shorten into *nbcs*.

Vx, which means 微信 (wei xin), is formed by one pinyin letter and one English letter. The letter *v* has a similar pronunciation as the character 微 (wei), so that people use the letter *v* to replace the Chinese character, whereas *x* is the initial letter of the Chinese character 信 (xin). The formation of this TSE is a combination of Chinese pinyin initial and a homophonic English letter.

b. Latin letters + pinyin TSEs

Many TSEs that are formed with letters are based on Chinese pinyin system. Despite the fact that some researchers realise pinyin as a unique language system that differs from English, some other types of TSEs are created thanks to the special features of mobile devices such as the cell phone pinyin typing systems. People's

typing habit differs as some use Sudoku typing system while others use full keyboard typing system or even Wubi (五笔), a specially designed typing system based on Chinese handwriting logic. For example, let us use *小学生* to illustrate the effect of typing on TSE formation. If someone wants to type *小学生* on his phone or computer, he or she can either type the full pinyin of *小学生* (i.e., xiaoxuesheng) or type the pinyin initials of *小学生* (i.e., xxs).

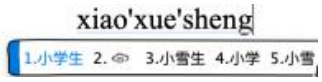


Image 6 Screenshot 1 of the typing software of *xxs*

When typing the full pinyin of the word, the typing system will provide all possible words that share the same pinyin combination on the waiting list. The more the user types, the more precisely the typing system will provide the waiting list.

Image 7 Screenshot 2 of the typing software of *xxs*

As shown in the typing system's screenshot, the waiting list provides four Chinese words whose initials are *xxs*. Apart from *小学生*, *小小思* /**x**iao **x**iao **s**i/, *休息室* /**x**iu **x**i **s**hi/, and *学习上* /**x**ue **x**i **s**hang/ are all words whose characters' pinyin initials are *xxs*. If the user does not want to choose any of these words and only keep the pinyin letters, *xxs*, he or she can press the "Enter" button on the keyboard, while otherwise he or she needs to use the number buttons of the keyboard to select corresponding words or use the mouse to click on the target word. Comparing with spending additional effort on pressing number button or using the mouse to click, pressing "Enter" button is relatively faster.

Generally speaking, there are two types of Latin letter TSEs, which are TSEs based on pinyin initials and TSEs based on fully spelt pinyin.

The first kind of TSEs are formed with several unrelated letters, which are actually the initial letters of the pinyin of corresponding Chinese characters. For

example, in the TSE of *dddd*, there are four "d"s. *dddd* is the pinyin initial acronym of *懂得都懂* (dong de dou dong), which means "if you get it, you get it". As shown in *懂得都懂*'s pinyin, the first letter of each character's pinyin is "d", which together form the TSE, *dddd*. Here are some other examples that follow the same path of formation. *Cnm* refers to the initial acronym of *操你妈* (*cao ni ma*), which means "fuck you" in English. *Xyxf* refers to the initial acronym of *血雨腥风* (xue yu xing feng), which describes scene of a field where a fierce battle has just been fought.

The second type of TSEs are formed with fully spelled pinyin letters. The TSE *zhuan* is the fully spelt pinyin of *转* (zhuan), which is the first character of a Chinese word *转发* (zhuan fa), meaning "forward a post" in English.

I also find TSEs with the form of Latin letters but involve Japanese in their formation. One of this kind of TSEs is *ky*. *Ky* is the Japanese acronym of 空気読めない (kuuki yomenai). *K and y* are the initial letters of *kuuki* and *yomenai*. Literally, *ky* means that one cannot read the atmosphere. It can be extended to mean saying something that is not appropriate in certain conditions. Chinese Internet users borrow this expression from Japanese social media. In this case, the Latin letters are from Japanese, which is called Romanji.

C. Numerals

In the Weibo data, some TSEs only have numerals in their formation. There are four types of numeral TSEs that are found in the data.

The first type is using numerals as homophones to Chinese characters. In *520*, the pinyin of *520* is "wu er ling". People use *520* to replace the Chinese characters *我爱你* (wo ai ni), which means "I love you". The first character *我* (wo) has the same initial *w* with *5* (wu), and the tone of these characters are all the third tone. The numeral *2* (er) is pronounced like the character *爱* (ai) in some Chinese dialects.

Also, *0* (ling) sounds similar to *您* (nin), which means *你* (ni) in Chinese. Therefore, *520* (wu er ling) becomes the Translanguaging expression of *我爱你* (wo ai ni).

The second type of numeral-only TSE is related to the code of an emoji. The TSE, *2333* (er san san san), means "laughing". The origin of this TSE is connected with an emoji that was used in Tieba, an online discussion forum like Reddit. The No. 233 emoji was a smile face. When people wanted to use that emoji, they just needed to type *233*, then the system will automatically provide the emoji for people to choose. Gradually, people used *233* directly to express laughing. The more *3's* people put behind *2,* the bigger they are laughing.

The third type of TSE that only has numerals requires music knowledge to understand. The TSE, *67,* is formed by two numerals that each represent the "la" and "si" in numbered musical notation. The pronunciation of "la si" sounds like the characters of *拉稀* (la xi), which means "have got a run" or "suffering from diarrhea". Based on numbered music notation, the seven notations of "do", "re", "mi", "fa", "so", "la", "si", and "do" can represent many different Chinese characters that have similar pronunciations.

The last group of TSEs that only have numerals in their formation are created based on their meanings as numbers. *996* (jiu jiu liu) is a classic example of this type of TSE. *996* represents a working schedule, where the employees go to work at 9 a.m. in the morning and go off work at 9 p.m. in the evening for 6 days a week. People depict this kind of life as *996* life.

D. Emojis

I found several TSEs that only use emojis to express their meaning in two ways. The first way is to use the homophones of emoji's meaning to form TSEs. For instance, in 🐮🍺, only two emojis are used, which means *牛逼* (niu bi). The emoji of a cattle refers to *牛* (niu) in Chinese, while the beer emoji refers to *啤酒* (pi jiu).

In this way, the emojis of 🐮🍺 can be transferred as 牛啤 (niu pi). In some Chinese dialects, the word 牛逼 (niu bi), which means "awesome", can also be called as 牛批 (niu pi). Both 逼 (bi) and 批 (pi) are ways of expressing vagina. The other way to use emoji is to use the emoji's original meaning to express something. For example, 🌬🐂, is formed by an emoji that means "吹气 blowing air", and the emoji of a cattle refers to 牛 (niu). 🌬🐂 then can express 吹牛 (chui niu), which means "boast" in English.

4.4 Summary

In this chapter, I tried to demonstrate the compositional diversity and complexity of TSEs that readers encounter in posts on Chinese social media. To achieve this goal, I examined the various word formation mechanisms through which TSEs in my data are created, the various combinations of linguistic or semiotic elements that the TSEs contained, and the distinction between explicit and implicit TSEs.

Compared to previous research on the creation of mono-lingual neologisms, my study showed that although the four major types of word formation mechanisms (Lipka, 2002) also apply to the creative formation of TSEs, differences exist. For example, although we have quite a few cases of affixational TSEs, the roots and the affixes come from different languages, as in *niubility* and *ungeilivability*. Also, with regard to the mechanism of compounding or composition, instead of combining two or more stems from the same language (Gontšarova, 2013), TSEs through composition combine stem-like elements from different linguistic or semiotic systems. Indeed, the analysis showed that all the TSEs in my study contain elements from different linguistic or semiotic element, either explicitly or implicitly. With respect to external neologism, former studies on loan-blends (e.g. Haspelmath, 2009; Lipka, 2002) provide crucial enlightenment on the analysis of formation of some TSEs.

However, they cannot explain all of them. For instance, cases like *zqsg* and *nb* are in the forms which go beyond the scope of loan-blends as they only borrow the English letters but these letters mean nothing in English. Moreover, traditional research on loan-blends is usually restricted to written or spoken languages in daily communication. TSEs, however, are derived from the Internet. For example, in the Translanguaging expression  *山大* (yā lí shān dà, meaning "feel very stressful"), the emojis of duck (鸭 yā) and pear (梨 lí) in Chinese are pronounced similarly to *压力* (yā lì) with the tone of *梨* and *力* being different, which means "pressure" in English. The two emojis together function as a noun in place of *压力*, while *山大* is a metaphor referring to the pressure that is as heavy as a mountain. However, in traditional Chinese, *压力山大* is not a word or an idiom but a phrase. The reason why people create this word is an imitation of the pronunciation of "Alexander" in English. Moreover, the Translanguaging expression of *欧克* belongs to loan-shift as the meaning of OK is borrowed but it also borrows the pronunciation by using Chinese characters that sound similarly. As for TSEs, they also share the characteristics of such complex loanwords. Sometimes Chinese netizens borrow foreign language elements directly into their communication system, but the meaning and usage of the word have shifted and changed in actual use. In other words, some TSEs may simply borrow the components from foreign words and combine them with Chinese words, which superficially match the type of loan-blends.

Compared to previous research on TEs (e.g., Zhang & Ren, 2022; Li, Tsang, Wong & Lok, 2020), my analysis of TSEs on Chinese social media turned out the following new findings.

First, unlike previous research that focuses on the functions of TEs, this study uncovered the mechanisms of their formation. It showed that unlike TEs that are generally created using one single mechanism, some TSEs may involve more than one formation mechanism.

Second, unlike TEs that are generally made up of two linguistic elements as discussed in most previous research, TSEs in my study were found sometimes to contain more than two linguistic or semiotic elements.

Third, apart from explicit TSEs, this study also revealed TSEs that are implicit in various ways. Those TEs that are treated in previous literature clearly involve the participation of two languages (often English plus Chinese). Yet, for some TSEs, we cannot identify them from their surface formation, simply by drawing on the four types of creative word formation mechanisms (Lipka, 2002). Only when we associate their pronunciation/form and meaning with an item from another language or semiotic system, can we discover their nature of being Trans-semiotic. In my data, implicit TSEs have four main types and thirteen sub-types, namely logographic characters TSE, Latin alphabet letters TSE, numerals TSE, and emojis TSE. Under each main type, combinations of different semiotic symbols form diverse interesting TSEs that cannot be directly identified.

All the findings above suggested that TSEs are much more complex and varied than both single-language creative expressions and TEs that we have long been familiar with. They demonstrate some unique ways of creating novel TSEs used by Chinese social media users. Thus, my study has expanded the scope of the existing models of innovative word formation.

Thus, this chapter showed that the TSEs consist of different semiotic elements, creatively integrating these elements into novel expressions which go beyond the traditional definition of word or phrase in one single language. To account for the diversity and creativity of TSE formation, two important reasons are not negligible. One is that the users of the social media who create or use the TSEs are more or less bilingual or even multilingual. At least, they are familiar with and could use a small amount of words from other languages than their mother tongue. The other reason has to do with the manner of typing on the computer, laptop, or the mobile phone.

Note that there are still some other ways to input Chinese characters on digital devices. For example, some people may use Wubi input system, which allows people

to type Characters based on their strokes. On smart phones, people can also directly write characters on the screen, which is similar to traditional hand writing. What is more, audio input is available to people who cannot write or find it inconvenient to write. With the development of AI speech recognition, the accuracy of audio input is greatly improved. However, in comparison to Pinyin input system, these methods have their own drawbacks such as lower input proficiency and lower input accuracy. Audio input is also restricted when it is not convenient to speak. According to *Chinese Third Party Mobile Input Industry Development Research Report* in 2022 by IiMeida Research, Pinyin input method is the most commonly used, which accounts for 75.7% of mobile phone users. Therefore, Pinyin input can be considered as the major way of input on Chinese social media. Using Pinyin to type Chinese characters, people actually give up the traditional way of writing. People do not necessarily remember how to write a character when typing Pinyin. Instead, they only need to recognise the characters and their meaning. The process of thinking of a Chinese character, typing the Pinyin of the character, and selecting the character from the list form a new way of Chinese input that makes Pinyin an inevitable tool. The realisation of Chinese homophones in digital context relies on Pinyin input if the writer or reader is a Pinyin input users. With the support of Pinyin input system, using homophones becomes relatively easier than traditional writing as all homophones are provided on the screen for users to choose. Pinyin also helps illustrate the realisation of homophones across languages and other semiotic systems. For example, “三克” (“sān kè”) , is used as the homophone of “thank”, which is pronounced as /θæŋk/. People can directly input a Pinyin spelling that is similar to the pronunciation of other languages or symbols of other semiotic systems and choose the Chinese characters provided to complete the process of establishing homophones across semiotic systems.

It has to be mentioned that besides saving efforts, people may also deliberately use initials instead of original word due to diverse motivations such as avoiding mentioning the exact name of the person they want to gossip about, or avoiding typing sensitive words that are rude or abandoned on social media, etc. We will return to

discuss these effects in Chapter 6.

Chapter 5 Paths, Strategies and Routes of Understanding TSEs on Social Media

In last chapter, I categorised TSEs that readers encounter in Weibo posts from the perspectives of their formation mechanisms, the semiotic elements within the expressions, and the explicitness of Trans-semioticising. The categorisation provides an overall description of the formal and compositional characteristics of TSEs. In this chapter, I will report, on the basis of the think-aloud protocols and the follow-up interview, how Chinese readers understand different types of TSEs as they encounter them on social media, including the paths, strategies and routes of comprehension used, as well as possible variations involved. Some mechanisms of association that operated in the participants' course of comprehension will also be illustrated.

5.1 Paths of TSE Understanding

My analysis of the think-aloud protocols reveals three major paths along which different participants processed the semantic meaning of the TSEs.

5.1.1 Path A Direct access

Some participants met and knew the TSEs tested prior to the think-aloud task. In other words, they had stored the meaning of the TSEs in their mind. When they were asked to understand the TSEs that they knew, they would directly extract the meanings of the TSEs from their memory as in the case of processing formulaic expressions or chunks, and then output the meaning. This path is relatively simple and smooth. Here are three extracts from Participants A, C and T respectively.

Ex. 1

Background information: The TSE, 集美 /ji mei/, means "sisters" in English. It is the Trans-semioticising equivalent of the Chinese word 姐妹 /jie mei/ meaning "sisters" or, more broadly, "female friends". 集美's pronunciation is similar to that of 姐妹 in some dialects, which can be seen as homophones to each other. However, the creation of using 集美 to replace 姐妹 is not related to any Chinese dialect. This expression became popular due to a web anchor on Douyin. Douyin and TikTok are basically the same created by the same company, while

Douyin is provided only for Chinese users. The Douyin anchor was called Teacher Guo, whose video and live streaming became viral due to her humorous jokes, life attitude, and her special pronunciation. She created her own tone of speaking Chinese, which left millions of people a deep impression. Her fans mimicked her accent, which later became a trend among young people. Among all the strange pronunciations of Chinese, *集美* was one of the most famous.

Image 8 Example of *集美*

[Translation: Goodnight *集美 s*. (Which one looks good? Pick one for me! I will see (your comments) tomorrow morning.]

Transcript extract of Participant A's think-aloud protocol:

集美们应该是姐，姐妹们，我知道

集美 s should be sis, sisters. I knew it.

As participant A mentioned, she knew the TSE before. Therefore, when she saw the TSE, she referred to her knowledge of this TSE in her mind and output the meaning of the TSE.

Ex. 2

Background information: The TSE *花粉* means the fans of Huawei. This is a TSE equivalent of *华为的粉丝* (the fans of Huawei). First, the TSE is combined by character initials of the phrase *华为的粉丝*. The phrase is a nominal phrase formed by one adjective and one noun. In Chinese, *华为的* means "Huawei's". The character *的* is used as a symbol to reflect a possessive relationship, meaning "belonging to Huawei". *粉丝* is a homophone of *fans*. *粉丝* pronounced as /fen si/ in Chinese, while *fans* is pronounced as /fens/ in English. Literally, *粉丝* means "vermicelli", which is a type of transparent noodle made from mung beans. However, *粉丝* is entitled to the new meaning of *fans* with the development of social media. It is difficult to trace back when *粉丝* got the new meaning. Perhaps using *粉丝* to refer to *fans* was created even earlier. At least in a TV series called *Xianren Madajie* (Mrs. Idler) in 2000, the expression of *粉丝* was used to refer to *fans*. It became even more popular in 2006, when draft shows like *Super Girls* and *Happy Boys* swept the country. The culture of fandom set up its foundation back then. *华为* (Huawei) is a Chinese technology company, which sells all types of new products, such as cell phones, laptop computers, and routers. Huawei earned a lot of market portion in the global cell phone competition, which made it the pride of Chinese customers and received millions of fans.]

Image 9 Example of *花粉*

[**Translation:** I can bet, all *花粉* students must have bought Huawei AX3 series Wi-Fi6 router in the past few years?
This time AX6 comes with its new product. Wanna take one?]

Transcript extract of Participant C's think-aloud protocol:
Participant C：*花粉*很通用的，应该是华为的粉丝。
花粉 is very popular. It should mean "the fans of *华为*".

As the participant mentioned, the TSE was popular, which means that he had seen this expression used by other people many times. He then directly output that this TSE meant "the fans of Huawei". The whole understanding process was quickly finished by the participant, as reflected by his assertive one-sentence comment. Here is another example.

Ex. 3

Background information: Compared with Professor Xue, other guests' performance in the TV show was mediocre. The Weibo user writes a Weibo post to complain about a TV show which does not provide enough footage about a guest called Professor Xue. Professor Xue was an economic researcher at Beijing University and attended some TV shows such as I CAN BB (a debate show) and Super Brain (a show that focuses on brain games). He won many fans because of his logic and wise comment on these TV shows. However, the Weibo user finds that in the new season, Professor Xue gets less chance to talk and to be edited into the show. Yet, other guests who are considered to be "ungeilivable" are exposed more often to the camera.

Image 10 Example of *ungeilivable*

[**Translation**: Professor Xue got so little scene this season... Please save the other guests' scene for the professor... Anyway, all the other guests were *ungeilivable*.]

Transcript extract of Participant T's think-aloud protocol:

ok 这个和那个 undingable 很像，之前好像也看到过，大概也是同样的意思。那 ungeilivable 是不太给力的意思。

Ok, this is similar to undingable. I used to see it before. It probably has the same meaning. So *ungeilivable* means not being *给力*.

As Participant T mentioned in the shadowed line, he had seen this TSE before, and he supposed that it probably conveyed the same meaning as the "dingable", which he had seen before. Then he output the meaning of "ungeilivable" without any difficulty.

Thus, the direct path is relatively smooth and fast, because the participant only needed to reach out to his or her mental lexicon to retrieve the meaning of that expression that had been stored in their brain. In some cases, even though the participants did not comment metapragmatically like "I knew the meaning before", they did directly come up with the meaning of the TSEs and asked the researcher to move on to the next TSE.

5.1.2 Path B Access via external assistance

Occasionally, the participants might seek external assistance by using search engines like Google and Baidu, or typing on the keyboard in order to understand the meaning of the TSE because they did not know the TSEs and did not bother to devote efforts to the understanding or after they failed despite trying different strategies. Here is an extract from Participant G's think-aloud protocol.

Ex. 4

Background information: *xxs* is the Chinese pinyin initial combination of the word *小学生* /xiao xue sheng/, which means "primary school student" in English. Using Chinese pinyin initials to replace corresponding characters is often used by people on social media. This path to create TSE is owed to the mechanism of typing system of simplified Chinese and the pinyin system. When typing a Chinese character on the cell phone or computer, one can just type the pinyin initial of the character and the typing system will automatically provide possible characters which share the same initial being typed in the waiting list for users to choose. It is the same when typing a phrase that is composed of a few characters.

Image 11 Example of *xxs*

[**Translation**: A high school student lives like a *xxs*.]

Transcript extract of Participant G's think-aloud protocol:

一个高中生活的跟个小学一个高中生活得跟个小学生一样, 这个应该也是比较常见的一个缩写，但是最初我也是不明白是什么意思，我来查下百度百科。

A high school student lives like a primary school student. This should be a relatively common abbreviation. But I don't know what it means, let me look it up in Baidu Baike.

Here, Participant G opted to look for external assistance to help herself understand the meaning of *xxs*. Unlike academic reading, reading on Weibo is more of an entertainment action that allows the reader to read freely without any pressure to know every single word. Some people may give up understanding while some can use the tools at hand to find the meaning. The advantage of reading online is that readers can have easy access to the Internet to search for information at any time with a few clicks or touches. It is not necessary for readers like Participant G to devote too much effort to the understanding process. This path may save time and increase the reading experience as readers will not be stuck by difficult or novel expressions.

Alternatively, the participants would type the letters to find an answer as mentioned by Participant Q.

Ex.5

Background information: Refer to Ex. 4 above for details.

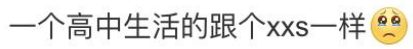


Image 11 Example of *xxs*

[**Translation**: A high school student lives like a *xxs*.]

Transcript extract of Participant Q's think-aloud protocol:

参与者 Q: 一个高中生的，一个高中生活的跟个叉叉 s。这应该又是个什么汉语拼音的缩略。我想一个高中，一个高中生活的跟个叉叉 s 是个什么简拼，叉叉 s 我可以用手机打一下拼音吗?

研究者: 可以，你自己试一试，可以。

参与者 Q 打字中

参与者 Q: 我我打一下看看。小学生了，可以被汉语拼音打出来，首字母打出来就显示小学生。应该是一个高中生活的小学生一样。

研究者：你是输入的“xxs”是吧?

参与者 Q: 对。

Participant Q: A high school student. A high school student lives like a *叉叉 s* /cha cha s/. This again should be an acronym of a certain word's Chinese pinyin. Let me think. A high school student lives like a *叉叉 s*, which should be the acronym of what. *叉叉 s*. Can I use my phone to

type the pinyin?
Researcher: Yes, you can try. Yes.
Participant Q typing
Participant Q: Let me type it and have a look. (It's) 小学生. It can be typed in Chinese pinyin. Typing the initials, it shows 小学生. It should be "a high school student lives like a primary school student.
Researcher: You just typed "xxs", right?
Participant Q: Yes.

As the transcript shows, Participant Q chose to seek external assistance when he found it difficult to transfer *xxs* to get its meaning. 叉叉 *s* referred to *xxs*. The shape of *x* looked like a cross mark, which is written as 叉 /cha/ in China. When realizing that *xxs* was an acronym, Participant Q asked the researcher whether he could use his phone to type the pinyin. With the consent of the researcher, he used the typing system on his phone and quickly output his understanding that *xxs* meant "primary school student". The researcher then checked what he typed on his phone, which was just the three letters. Participant Q touched upon another external tool that could help understand TSEs in the form of acronym. He then used this external tool. Using the typing system to help to understand acronym TSEs might be helpful as the typing system and Weibo were always on the same device or easy to access when doing online reading. Using the mechanical association function of typing system could be a wise strategy to deal with TSEs formed by the initials of Chinese characters' pinyin.

5.1.3 Path C Access via using comprehension strategy(ies)

Unlike the two paths above, Path C refers to participants' head-on approach to tackling unfamiliar TSEs by deriving the meaning, without seeking assistance from others or search engines. The whole understanding process was completed by the participant alone, using the information provided by the Weibo post and his own background knowledge. Here is an extract from Participant T's think-aloud protocol:

Ex. 6
Background information: TSE of  is the TSE equivalent of the Chinese word 约 (yue). In Chinese, 约 is a character used to ask whether others want to meet and have fun. The emoji of the

moon replaces the Chinese character 约 because they have the same pronunciation /yue/ in Chinese. In this case, people use the single character to refer to the moon. However, in modern Chinese, people seldom use a single character 月 to express the moon. In fact, 月亮 (yue liang) is often used word to express the notion of the moon. This may bring some trouble for readers to understand the meaning of this emoji in this TSE. The character, 约 is pronounced with the first tone of /yue/, which should be /yuē/. 月 should be pronounced as /yuè/ with the fourth tone. However, in using emojis to express Chinese characters with a similar pronunciation, people can tolerate minor deviations in their pronunciation, such as different tones or front/back nasal, since the amount of emojis is relatively inadequate for thousands of Chinese characters.]

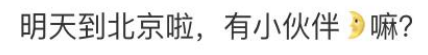


Image 12 Example of 嘛

[**Translation**: Arriving at Beijing tomorrow, any buddy wants to ?]

Transcript extract of Participant T's think-aloud protocol:

明天到北京有小伙伴月亮吗？我是这样理解，这个我是第一次看的，首先第一反应，我看到的这个符号是个月亮的符号，我在想是不是和夜晚和深夜比较有关系，但是好像发现用深夜、夜晚这种思维往里边代入有点不太通顺，后来发现应该就是这个表示约的意思，因为月亮和约是读音上很像的，因为这句话他说明天到北京了，有小伙伴什么吗？那么大家经常会说有没有人约一下或者是聚一聚这种，用约来表示这种聚会的状态，那么月亮应该就是一个谐音，也是一种谐音的方式，去代表约这个意思，可以这样。

Arriving at Beijing tomorrow, any buddy moon? I understand it in this way. It is the first time I've met this (expression). First, at first glance, I see this emoji, which is a symbol of the moon. I'm thinking whether it has a stronger relationship with evening and night. However, (I realise that) it seems that the thought of midnight and evening could not make sense in the context. Then I find that the moon emoji should represent the meaning of 约 as it sounded like 月 in 月亮 from the perspective of pronunciation. Because he said that he would arrive at Beijing tomorrow, are there any little friends (to do) what? Then people usually say anyone wants to meet or gather around. Using 约 may express the status of holding a party. So the moon's emoji should be a homophone. It's also a way to use homophone to convey the meaning of 约. It can work in this way.

In this example, the participant used four different comprehension strategies (to be detailed below) to find access to the meaning of 吗. For example, each time he had a guess of the meaning, he referred to the context to verify whether he was right.

My data analysis shows that the participants might make use of a variety of strategies in order to arrive at a meaningful interpretation, to which I turn in the next section.

Sometimes, the participants who took Path 3 might not succeed in comprehending the TSEs and finally gave up although they had tried their best to

understand the TSE using different strategies and associating many possible explanations to the TSE. Here is an example of Participant R who gave up understanding the TSE after he had tried different ways.

Ex. 7

Background information: *xxs* is the Chinese pinyin initial combination of the word *小学生* /xiao xue sheng/, which means primary school student in English. Using Chinese pinyin initials to replace corresponding characters is often used by people on social media. This path to creating TSE is attributed to the mechanism of typing system of simplified Chinese and the pinyin system. When typing a Chinese character on a cell phone or computer, one can only type the pinyin initial of the character and the typing system will automatically provide possible characters which share the same initial being typed in the waiting list for user to choose. It is the same when typing a word that is composed of few characters. In this example, let us use *小学生* to illustrate the mechanism. If someone wants to type *小学生* on his phone or computer, he or she can either type the full pinyin of *小学生* or type the pinyin initials of *小学生*.

Image 6 (Screenshot 1 of the typing software of *xxs*)

When typing the full pinyin of the word, the typing system will provide all possible words that share the same pinyin combination in the waiting list. The more the user types, the more precisely the typing system will provide the waiting list.

Image 7 (Screenshot 2 of the typing software of *xxs*)

As shown in the typing system's screenshot, the waiting list provides four Chinese words whose initials are *xxs*. Apart from *小学生*, *小小思* /**x**iao **x**iao **s**i/, *休息室* /**x**iu **x**i **s**hi/, and *学习上* /**x**ue **x**i **s**hang/ are all words whose characters' pinyin initials are *xxs*. If the user does not want to choose any of these provided words and only keeps the pinyin letters, *xxs*, he or she can press the "Enter" button on the keyboard, while otherwise he or she needs to use the number buttons of keyboard to select corresponding words or use mouse to click on the target word. Compared to spending additional effort on pressing the number button or using mouse to click, pressing "Enter" button is relatively faster. It has to be mentioned that besides saving efforts, people may also deliberately use initials instead of original words due to diverse motivations such as avoiding mentioning the exact name of the person they want to gossip about, or avoiding typing sensitive words that are rude or abandoned on social media, etc.

Image 11 Example of *xxs*

[**Translation**: A high school student lives like a *xxs*.]

Transcript extract of Participant R's think-aloud protocol:

跟个 ss，xxs 一样。一个高中生活的。一个高中，生活的跟个，我操，跟个什么一样。跟个小傻瓜，小。就是我一开始思维肯定是想去拼一个词，跟个潇潇潇洒一样，跟个小傻逼一样。不对。一个高中生活的跟个，肯定感觉是一个贬义的词，一个高中生活肯定是，然后最后配一个 emoji，这个肯定是一个不好的词，大概类似那种跟个监狱。但是这又不是监狱，这是

三个字母。三个字母的话就是跟个,一个高中，高中能生活的跟个什么？跟个，监管所，但这也不是监管所的谐音，这是一个 xxs 的谐音。跟个新，谢，受，跟个，跟一个高中生活的，跟个新鲜神，新鲜血，操想不到诶，一个高中生活的跟个，一个高中高中一般生活的就很枯燥，现在刷题对吧？可以往枯燥这种地方想。我感觉就有点像。监管所那种什么监狱一个高中生活的跟个。这个确实不知道，能过吗？

Like a “ss”, “xxs”. A high school student lives. A high school student, living like. Damn. Like a what. Like a *小傻瓜* /xiao sha gua/, *小* /xiao/. So firstly my mind must be thinking to spell a word. Like a *小小潇洒* /xiao xiao xiao sa/. Like a *小傻逼* /xiao sha bi/. No. A high school student living like a. It must be a derogatory word. High school life must be, followed by an emoji at the end. This must be a passive word. Maybe like that in prison. But it’s not prison (*监狱* /jian yu/). It has three letters. If (it has) three letters, like a high school. How will high school life gonna to be? Like a supervision department (*监管所* /jian guan suo/). But it is not the homophone of *监管所*. It is a homophone of “xxs”. Like a *新*, *谢*, *受*. Like a, like a high school, high school life. Like a *新鲜神* /xin xian shen/, *新鲜血* /xin xian xue/. Damn, I cannot think about its meaning. A high school lives like a. A high school, a life like in high school, which is very boring. So it’s about taking quizzes, right? (Maybe I) can think about something boring. I think it’s a little bit like a supervision department or something. A high school lives like a. I really don’t know about this. Can I skip it?

In this example, Participant R finally gave up understanding the TSE after he spent 2 minutes trying to understand the meaning. In the transcript, he quickly realised that the TSE *xxs* was the initials of certain Chinese characters. So he kept associating the three letters with different Chinese characters or words. In all, he tried six different Chinese characters combinations, such as *小傻瓜* /xiao sha gua/, *小小潇洒* /xiao xiao xiao sa/, *小傻逼* /xiao sha bi/, *监狱* /jian yu/, *监管所* /jian guan suo/, *新鲜神* /xin xian shen/, and *新鲜血* /xin xian xue/. Every time he thought about a new Chinese word, he would go back to the context and check the initials to make sure it could match the contextual meaning and the initial letters. He also read the implied information of the passive attitude conveyed in the text. However, he made a mistake in fragmentation. The original sentence should be understood as “a high school student lives like a…”. However, the participant read the sentence as “a high school live like…”. This was very tricky because *高中生* meant high school student and *活* meant to live, while *高中* meant high school and *生活* meant life. Both *高中生+活* and *高中+生活* could make sense in the sentence, especially when the participant

did not know the meaning of *xxs*. If the participant tried to consider *高中生活* as the combination of *高中生+活*, he would have a higher chance to associate *xxs* with *小学生* /xiao xue sheng/, which meant primary school student, opposite to high school student. When the participant could not understand the TSE any more, he gave up understanding the TSE's exact meaning. However, he did know that *xxs* was used in the Weibo to talk about some passive things. This result was still meaningful as in daily reading activities a relatively vague or general understanding could be enough in certain circumstances.

5.2 Strategies of TSE understanding

In this study, the participants were invited to think aloud about their comprehension process while reading Weibo posts containing different TSEs. By combing through the transcripts of the participants' think-aloud records, coupled with follow-up interviews, I found that the participants may adopt various comprehension strategies to tackle the semantic meaning of the TSEs. By comprehension strategies, I refer to strategies that the participants adopt to decode and infer the meaning of the TSEs, which contribute to the final understanding of the TSEs. When the participants took Path A, they just used a recalling strategy, simply mapping the meaning of the TSEs they remarked onto the encountered TSEs. When they took Path B, they would turn to others or use some search engines to set up the problem by themselves. When the participants selected Path C, they might use one or two of the following strategies. The following part will provide details about each strategy in my test adopted by the participants for the purpose of understanding the TSEs they encountered in Weibo posts.

5.2.1 A. Resorting to the text

Resorting to the text refers to reading aloud the Weibo post's whole text word by

word in order to assist understanding. This strategy was usually taken when the reader was not familiar with the target TSE and thus had trouble figuring out its meaning. In this case, when he failed to understand it, he began reading the whole text containing the TSE for help. Consider the case of Participant A understanding the TSE "ungeilivable" in Ex. 7:

Ex. 7

Background information: Refer to Ex. 3 of this chapter for details.

薛教授这一季的镜头，也太太太少了吧……求多切几个嘉宾的镜头补给教授吧……反正这一季的嘉宾个个ungeilivable的…

Image 10 Example of *ungeilivable*

[**Translation**: Professor Xue got so little footage this season... Please save the other guests' scene for the professor... Anyway, all the other guests were *ungeilivable*.]

Transcript extract of Participant A's think-aloud protocol:

Un---gei---li---vable。薛教授这一季的镜头也太太太少了吧，求多切几个嘉宾的镜头补给教授。反正这一季的嘉宾个个 ungeilivable 的。应该就是不给力。我猜的，没见过。

Un---gei---li---vable. [1.0] Professor Xue got soooo little footage this season... Please spare the other guests' scene to Professor... Anyway, all the other guests were *ungeilivable*. It should mean "not being good enough". I guessed the meaning. Haven't seen it before.

As we can see from the participant's think-aloud protocol, the participant read the TSE very slowly syllable by syllable and paused for around one second. Then, she began to read the whole Weibo text of the post and arrived at an interpretation ("不给力"). In this process, we could say that the participant had been trying to figure out the meaning of the TSE directly from reading the text aloud syllable by syllable. However, she failed to understand the TSE which she had not seen before ("没见过"). Based on the reading of the whole text, she guessed ("我猜的") at the meaning of the TSE. Although she could have benefited from other factors to make the guessing, her resort to the meaning of the whole text through reading played a definite role in

this case. After all, her output of the meaning of the TSE immediately after the completion of the text reading process.

It must be mentioned that reading the full text was not always a strategy participants resorted to for understanding TSEs. Sometimes, a participant might just read the whole text before starting to analyse the TSE involved. This occurred when the target TSE appeared at the end of a post. In this case, the reader might read from the beginning of the text to the target TSE and then start further analysis. For instance, in the Weibo of *zhuangbility*, there were 93 characters before the target TSE appeared.

Ex. 8

Background information: *zhuangbility* is the Translanguaging expression of *装逼* /zhuang bi/, which means "someone being a poser". It is a combination of Chinese pinyin and English suffix.

> 看到刘的采访 我想到了某东的数学诺贝尔奖 还蛮喜欢某东的 但是这个事件之后就彻底脱粉了 路人粉的粉哈哈哈哈哈之前他发的微博我一条都看不懂 就觉得可以写出来的人好厉害 但现在只觉得好zhuangbility

Image 13 Example of *zhuangbility*

[**Translation:** When I saw Liu's interview, I thought about Dong's Mathematics Nobel Prize. (I used to) like Dong before, while after this affair I did not fan him at all. The fan of "passerby fan" lol. I could not understand a single Weibo post of his, thinking that the person who could write these must be very extraordinary. But now I could only feel very *zhuangbility*.]

Transcript extract of Participant Q's think-aloud protocol

看到刘的采访，我想到某东的数学诺贝尔奖喜欢某东西，但是这个事件之后彻底脱粉，路人粉的粉，他之前他发的微博一条都看不懂，就觉得可以写出来的人好厉害，现在只觉得好装。zhuangbility 就是装逼的意思，这个感觉也在之前忘了是什么词，但是也看到有人用类似的这种造词写过，所以也是挺好理解的应该。

Translation: When I saw Liu's interview, I thought about Dong's Mathematics Nobel Prize. (I used to) like Dong before, but after this affair I did not fan him at all. The fan of "passerby fan". I could not understand a single Weibo post of his, thinking that the person who could write these must be very extraordinary. But now I could only feel very *zhuangbility*. *Zhuangbility* means *装逼*. I had this feeling of this meaning before. I forget what was that word. But I have seen some people using this type of created words before. So it is relatively easy to understand. It should be.

In this case, Participant Q read almost the whole text except the *哈哈哈*. He did not provide his understanding of this TSE until he finished reading the text. According to his think-aloud transcript, he was very sure about the meaning of the TSE, as he used

another understanding strategy (referring to processing history, see below) to analyse the structure of this TSE. Although this TSE was not difficult for the reader to understand as he mentioned that "this is relatively easy to understand", he still read the whole text before trying to understand its meaning. Thus, we could take his reading of the full text as a deliberate or motivated strategy for understanding the TSE.

In the same vein, when a participant had the habit of reading the full text, we could not regard such reading as a comprehension strategy. Among all the 20 participants, 3 participants appeared to have this habit of reading the whole text containing the TSE in the course of completing the think-aloud task. It has to be mentioned that even for those who tended to read the whole text before understanding TSEs, they might have skipped reading the text as they were very familiar with the meaning of some TSEs. For those TSEs they were very familiar with, they directly spoke out the meaning of the TSEs without devoting extra effort to reading the text. Their reading of the whole text might have resulted from their improper understanding of the design of the think-aloud task. They might misunderstand that they were required to share their thoughts of understanding the meaning in great detail.

5.2.2 B. Judging from the topic

The second strategy used by the participants for understanding TSEs on social media might be called judging from the topic of the Weibo post. Unlike the first strategy, topic-based judgment concerns primarily what the Weibo post is about, rather than the specific content of the whole Weibo text. The reader got an idea about the topic of the text, which would support the process of understanding the meaning of the TSE. Let us take Participant B's understanding process of *ban 位买房* (to buy a house at the banning position) as an example.

Ex. 9

Background information: To buy a house at the banning position means a champion is often banned during game ranking (Wangzherongyao in this context, which is a mobile moba

(multiplayer online battle arena) game similar to League of Legend). In the game, for each match in rank (usually higher than Diamond rank), players from two sides can select four or six champions that they want to ban for this match. The banned champions' profiles will be shown on the top left and right. For those over powerful champions, they will be banned for most of the matches. Figuratively speaking, these champions can always be seen on the top left and right corner. The players metaphorically describe this interesting situation as these champions buy a permanent house at the top left and right corner and settle down forever, as shown in the yellow squares on the top left and right in the screenshot below.

Image 14 Screenshot of the game user interface of Wangzherongyao

八位Ban买房的英雄，你都会玩哪个？

Image 15 Example of *ban 位买房*

[**Translation:** Regarding the eight champions who bought houses at the banning position, which ones can you play? #wangzherongyao##reai jing xinchun#]

Transcript extract of Participant B's think-aloud protocol

8 位 ban 买房的英雄，你都会玩哪一个？底下应该这也是王者荣耀，然后王者荣耀就是说比赛的时候应该要 ban 掉，就是互相挑选掉，觉得比较强的把这样让对面英雄没法选。

eight champions who bought houses at the banning position, which ones can you play? The picture below should also be about Wangzherongyao. And in Wangzherongyao, (the champions) will be banned during each match. In other words, (players) will pick the champions that need to be banned if they think the champions are overpowering so that opponents cannot pick them.

When the participant started to read the Weibo post, he saw the picture the author posted along with the post. In the picture, the poster listed several champions' ban/pick ratio in ranking. There was also a tag mentioning that it is a Weibo post

related to the topic of Wangzherongyao. The participant realised that this Weibo was talking about something about Wangzherongyao. Therefore, he said that "The picture below should also be about Wangzherongyao" ("底下应该这也是王者荣耀"), which was an act of thinking about the Weibo post's topic. Then the reader talked about the meaning of *ban* in this TSE. Although he had not played the game for a long time, he believed that the TSE was a new expression that related to the topic of the game. The reader used the strategy of topic-based judgment to narrow down the scope of the information he needed to understand the TSE, which stimulated his further understanding and activated his knowledge of Wangzherongyao.

In the following example, Participant F also judged from the topic to find information when understanding the TSE.

Ex. 10

Background information: In this context, *456* refers to an idol from a Chinese girls group. Literally, *456* refers to "someone, being in heat or being oestrous." It is the Trans-semioticising equivalent of *发骚啦* /fa sao la/ in Chinese. The TSE is based on homophony between Chinese and the pronunciation of music notes, while it only contains numerals in its word formation. In numbered musical notation, 4 (fa) is pronounced as /fa:/, 5 (sol) is pronounced as /sɒl/, and 6 (la) is pronounced as /la/. Both 4 and 6 are pronounced almost in the same way as 发 and 啦. While 骚 and 5 are homophones which share the same consonant and similar vowels. *发骚* in the Chinese context is a rude sexist word that is usually used to describe a woman who acts like she is in heat. *啦* is a particle which functions like *了* to express an action being done. *发骚啦* together means a woman is in the state of being in heat. Using *456* to express the meaning of *发骚啦* was first widely known to the public during an affair of a Chinese female idol from a pop group. She was accused of cheating with a music producer who had a girlfriend. The producer's girlfriend posted many screenshots of the chat record between the idol and the producer. Both fans and non-fans were attracted by this chat record, especially by the expression of *456*. In the chat record, the idol mentioned she was *456* to the producer. Hence, people use this TSE to refer to the idol.

Weibo screenshot of the TSE

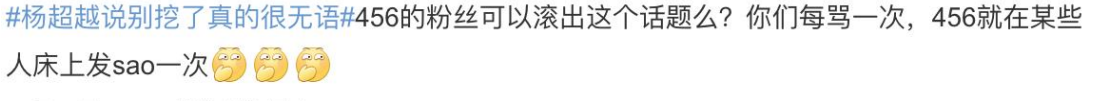


Image 16 Example of *456*

[**Translation**: #Yang Chaoyue said please don't dig, feeling very speechless# *456*'s fans can you please leave this topic? Every time you curse, *456* will be in heat on someone's bed.]

Transcript extract of Participant F's think-aloud protocol

456 我看到这个我不太理解，我觉得可能是饭圈词，然后我想的是可能是偶练里面是杨超越的名字是不是 *456*，我看一下。饭圈文化我不太懂。我看一下 1234456 也不是，嗯哼。我这个就不太懂了，就不知道。

456. I don't really understand when I saw this. I think it might be a word from fandom. And I think maybe it refers to Yang Chaoyue from Idol Trainees. Whether her name was *456*, let me think. I don't really know about fandom. Let me see, 1 2 3 4 4 5 6. No. uh-huh. I don't quite understand. Yeah, I don't know.

In this case, the participant claimed that she did not know the meaning of *456* at first. Then she referred to her background knowledge about Yang Chaoyue, who was a team mate of the "*456* idol". Both of these two idols debuted in a pick-up show called *Idol Trainee*. Fandom is a subcultural circle where people have their favorite idols and do certain activities to support their idols. So the participant's knowledge about Yang Chaoyue helped her judge the topic. Although she does not "really know about fandom", with the help of knowing the topic, she realised that she did not have enough knowledge to support further understanding of this TSE. So she gave up understanding this TSE quickly. In a daily reading context, people sometimes may not want to continue devoting efforts to figuring out TSEs that they find difficult to understand. This can also be considered as a meaningful result of understanding TSE, while it ends up with failure.

5.2.3 C. Decoding the TSE

Decoding the TSE is a strategy that was most widely used by the participants. From the data, participants decoded the TSEs from four different perspectives, namely, C1. analysing TSE's pronunciation; C2. analysing the internal structure of the TSE; C3. analysing the shape of TSE characters; C4. Transferring the semiotic system.

5.2.3.1 C1. Analysing TSE's pronunciation

analysing the TSEs' pronunciation refers to a strategy through which the participants get access to the meaning of TSEs from their pronunciation. Some TSEs were formed by homophones. By analysing the TSEs' pronunciation, participants could get information for the next steps of understanding.

Ex. 11

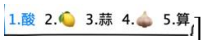
Background information: *酸 q* is the Trans-semioticising equivalent of "thank you" in English. *酸 q* and "thank you" share a similar pronunciation. The former is pronounced as /suān kju:/, while the latter is pronounced as /θæŋk ju:/. Due to the influence of the Chinese accent, some people may not pronounce /θ/ correctly. /s/ is often used to replace /θ/ in Chinese oral English. So /θæŋk/ is pronounced like /sæŋk/. However, in this TSE, /sæŋk/ further changed into another version of pronunciation, adding an /u/ after the consonant /s/. This variant requires additional background knowledge to understand. On the May of 2021, a TikToker, called Liu Tao or Teacher Liu uploaded an English introduction about his hometown, Guilin. Liu's view became viral due to his impressive appearance and his special English pronunciation. His Guilin accent added sauce to the word "thank you", which sounded like /suan kju:/. Since then, many Chinese netizens liked to use /suān kju:/ to replace /θæŋk ju:/. Many Chinese characters that are pronounced /suān/ or similarly to /suān/ were selected to represent this pronunciation, such as *酸* /suān/, *蒜* /suàn/, and *栓* /shuān/. The emoji of lemon was also used to replace *酸* as the form of *q*, while *酸* literally means "sour" in English. When typing *酸*, the emoji of lemon will automatically appear on the waiting list for users to select. As shown in the screenshot. []

Image 17 Screenshot of the typing software of *酸 q*

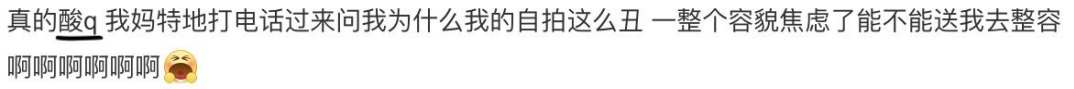


Image 18 Example of *酸 q*

[**Translation**: Really *酸 q*. My mom specially called me to ask why my selfie looked so ugly. I felt overwhelmed by appearance anxiety. Can (anyone) send me to do plastic surgery? Ahhhhhhhhh.]

Transcript extract of Participant P's think-aloud protocol

<E2>因为其实一开始不是有栓 q，<C2>就是木字旁一个全的栓，然后英文字母的 q。栓 q 这个词一开始因为平时之前看的有一个主播会经常这样用，<D>然后大概知道这个是因为结合他的语境，他一般是在比如说谢谢人家的那种有给他打钱给他打送礼物的时候，他会说酸 q，比如所以就知道这大概知道栓 q 是 thank you 的意思。<C1>然后在这里再看到酸 q 的话，因为也是音相同的，所以能理解应该是 thank you 就是谢谢的意思，<E1>但这又扯到另外一个梗，就是我真的会谢。这个词其实一般他当然不是正面表示感谢的意思，他就觉得说我真的谢谢你有一种这种反讽的意思在里面，<D>能够结合上下这里讲一段语境就可以理解，这确实是一个在讽刺了一次。

<E2>Because there was *栓 q*, <C2>which is composed of a *木* and a *全*, along with the English letter Q. As for the word *栓 q*, I saw a streamer often using this expression. <D>Then I understood its meaning basically based on its context. It was usually used when (the streamer) wanted to thank the audience who donated gifts. Therefore, I know that *栓 q* meant 'thank you". <C1>Then when I see the *酸 q* here, I recognise that they are also homophones. <E1>So I can understand that it should mean thank you as well, while it might relate to another meme, which is "I really would thank". This expression actually is not often used to express positive meaning of gratitude. <D>It is used to ironically say "I really thank you", which could be understood according to the context. In this case, it is indeed sarcasm.

When participant P tried to understand this TSE, she analysed the pronunciation of the TSE by establishing correspondence between *栓* and *酸* because they have the same pronunciation in Chinese. Then she figured out the meaning of *酸 q*, thinking it should convey the same meaning as *"*Thank you*"*.

5.2.3.2 C2. Analysing the internal structure of the TSE

analysing the internal structure of the TSE is a strategy used, when the readers analyse the TSEs part by part. It often occurs when the TSEs are formed by more than one semiotic element and the TSEs could be easily decomposed into different parts. In my data, this strategy was used when the reader found it difficult to figure its meaning out as a whole. The readers would decompose the TSE into several parts and analyse the TSE part by part. For example, when a TSE contained both Roman letters and Chinese characters, the readers would split the TSE into a letter part and a character part before further analysis. Here we can also take *ban 位买房* as an example.

Ex. 12

Background information: *niubility* means "extraordinarily good" in English. It is the Trans-semioticising equivalent of *牛逼* /niu bi/. Literally, *牛逼* means ox's vagina. People usually use this word as an adjective to describe someone or something that is extremely good. This expression is relatively impolite and often used in oral communication. As we can see, *牛逼*'s pinyin ends up with a "bi", which is also contained in the English suffix "bility". Therefore, people add the suffix to the end of *牛逼*'s pinyin, *niubi*, to form "niubibility". However, two "bi"s do not look right and sound strange, so people eliminate one "bi" in the expression. *niubility* is often used as an adjective in the Chinese context, while it looks similar to "ability" and "capacity", both of which are nouns. Chinese users neglect the English grammar logic when creating this expression.

Image 19 Example of *niubility*

[**Translation**: In the past two years, our diplomacy has received great results. The international environment is getting better and better, along with more foreign friends and increasing respect when travelling abroad. Nowadays, foreigners will say *niubility* when they talk about Chinese people. This should be attributed to the successfully established wolf warrior image, which is amiable, adorable, and respectable. Keep fighting, diplomatic star group!]

Transcript extract of Participant G's think-aloud protocol:

niubility 这读起来应该就是就直接用了，斜音就是中文"牛皮"的意思吧。 是因为之前有别的词什么 ability，然后什么有这种后缀的，然后可能是中文觉得牛逼两个字可能不是很文雅，然后就用到了国外的这种谐音的后缀，然后就创造了这么一个新的词来表达这个意思。

Niubility reads like should be used directly. It's the homophone of *牛皮* in Chinese. It is because there were some other words that (have) "ability" and this kind of suffix. And maybe because (the writer) thought *牛逼* was not very polite in Chinese, (the writer) used the foreign (language)'s homophone and suffix. Then this new word was created to express the meaning.

In this example, the participant's understanding process could be divided into two phases, which correspond to the two parts of the TSE that had been decomposed by the reader. The first phase was about the suffix. The reader mentioned that the suffix of *niubility* was related to *ability*. Then in the second phase, she analysed *niubi* and sought to know why pinyin instead of its original Chinese characters was used.

5.2.3.3 C3. Analysing the shape of TSE characters

Some TSEs are formed based on the shape of Chinese characters. Chinese characters are logograms that convey their meaning regardless of their pronunciations. As logograms, the shape of characters is available to be changed to meet certain goals of communication. When understanding such types of TSEs, participants would analyse the shape of TSE characters to get more information to support their understanding. Let us take a look at Ex 13.

Ex. 13

Background information: *几凡* (fán fán) is the Trans-semioticising equivalent of the word *凡凡* (fán fán). The only difference between *几* and *凡* is the position of the dot. In the character, *凡*, the dot is in the middle of the character *几*, while in *几* the dot is on the top of *几*. Both these two characters are pronounced as /fán/. *凡* is a common character that is usually used by people in the words like *平凡* (píng fán), which means "ordinary", and *凡人* (fán rén), which means "ordinary people". The meaning of *凡* is about being ordinary. However, the character of

几 is not a commonly used character in Chinese. 几 in Chinese equals 凡, which is an old, eliminated version of 凡. People no longer use this character in modern Chinese, although it can still be typed and found in typing system. Chinese characters have experienced many stages of evolution during the past century from traditional Chinese to simplified Chinese. Some characters like 凡 may have few different old versions existing but never being used today. People start to use 几 again due to an affair of a former Canadian-Chinese pop star called Kris Wu, whose Chinese name is 吴亦凡 (wú yì fán). Fans used to call him 凡凡 to show a close relationship. However, after being accused of raping underage girls, he lost his fame overnight. Many ironic memes including some TSEs were created to ridicule him. This TSE was created during that period of time. Both 凡凡 and 几几 were used to refer to Kris Wu. In Chinese, there was an old saying to describe someone who has a crush to have sex, which is having the worm of sperm climbing on someone's head. Therefore, by moving the dot from the middle to the head, the TSE 几几 vividly illustrates how Kris Wu was like when he committed seduction.]

Image 20 Example of 几几

[**Translation**: Take a look. Who is the real 几几?]

Transcript extract of Participant M's think-aloud protocol:

我知道凡凡他有这个，嗯，因为他名字叫吴亦凡，我有见过这个字，但我不知道这个字念什么。我看过这个字，然后因为凡它中间不是有个点，然后这个点就经常大家会可能用中国的汉字各种改造。怎么改造，我现在具体也忘了，但是它放在头上我不知道什么意思，但是我知道他是凡凡的意思，但是肯定就是讽刺他这两个字不一样，讽刺他的某一些特质，但具体什么特质我也看不太出来。

I know that 凡凡 has this (nickname). Yeah. Because his name is Wu Yifan. I have seen this character before but I don't know how to say it. I have seen this character. And because it has a dot in the middle, does it? So this dot usually is modified by people on different Chinese characters. I forgot how to modify it in detail now. But I have no idea why it is put on the head. While I do know it refers to fanfan. Anyway, it must be used to satirise him and some of his propensities. But I cannot tell what propensities (this TSE satirises).

As we can see from the shadowed lines, the reader noticed the change of the position of the dot in 凡. She mentioned that she saw this character before and knew what 几 meant in this context. So she did not have trouble in understanding the literal meaning of this TSE, by relating 几 to 凡. However, she devoted effort to figuring out why the TSE's user or creator moved the dot's position. She sensed that there should be an

ironic effect when using the character 几. We can also find that when understanding this type of TSE, the readers do not necessarily need to know or recognise the new characters before. The only thing they need to consider is what the original version of these characters is like. The phenomenon is a typical example that illustrates the gap between Chinese characters and their pronunciation. People can read without knowing how the character is pronounced. TSEs with the characters' shape or font change make use of this characteristic of Chinese characters.

5.2.3.4 C4. Transferring the semiotic system

Transferring the semiotic system is a strategy that the participants used when there were emojis in the TSEs or other semiotic elements that could not be directly understood.

Ex. 14

Background information: Refer to Ex 6 of this chapter for details.

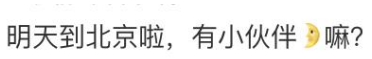


Image 12 Example of 嘛

[**Translation**: Arriving in Beijing tomorrow, any buddy wants to ?]

Transcript extract of Participant T's think-aloud protocol:

<A>明天到北京有小伙伴月亮吗？我是这样理解，这个我是第一次看的，首先<C3>第一反应我看到的这个符号是个月亮的符号，我在想是不是和夜晚和深夜比较有关系，<D>但是好像发现用深夜夜晚这种思维往里边代入有点不太通顺，后来发现应该就是这个表示约的意思，<C1>因为月亮和约是读音上很像的，因为这句话他说明天到北京了，有小伙伴什么吗？<E1>那么大家经常会说有没有人约一下或者是聚一聚这种用约来表示这种聚会的状态，那么月亮应该就是一个谐，也是一种谐音的方式，去代表约这个意思，可以这样。

<A>Arriving at Beijing tomorrow, any buddy 月 /yue/ (moon)? I understand it in this way. It is my first time met this (expression). <C3>First, at first glance, I saw this emoji, which was a symbol of moon. I was thinking whether it had a stronger relationship with evening and night. <D>However, (I realised that) it seemed that the thought of midnight and evening could not make sense in the context. Then I found that the moon emoji should represent the meaning of 约 <C1>as it sounded alike 月 in 月亮 from the perspective of pronunciation. <D>Because he said that he would arrive at Beijing tomorrow, are there any little friends (to do) what(to do) what? <E1>Then people usually say anyone wants to meet or gather around. Using 约 to express the status of holding a party. So the moon's emoji should be a homophone. It's also a way to use homophone to refer to the meaning of 约. It can work in this way.

This transcript shows how participant T used the strategy of transferring the semiotic meaning when he understood the meaning of the emoji of moon. When he read the TSE, he directly read the emoji's meaning *月* /yue/ transferring the emoji from a picture to a corresponding Chinese word. Then, based on this understanding, he began his first trial, thinking that this TSE was talking about something related to evening or night. On the basis of the homophonic association between "月" and "约", he finally managed to derive the meaning of the emoji.

5.2.4 D.Referring to the context

Referring to the context is the fourth strategy which the participants used. It refers to summarizing the meaning or implication of the context to assist in understanding the meaning of the TSEs. Readers usually use the context to verify their preliminary understanding of the TSE or use the context to activate possible association of relevant words or knowledge. Unlike the use of the first strategy, when the reader refers to the context, he or she does not need to read the text character by character. In the data, this strategy is reflected as directly mentioning information provided or implied in the context. Reading the text can be seen as a preparation for this strategy. As for the second strategy, determining the topic, the reader may simply determine the topic based on a few words or an affiliated photo. They can also use the context to support the determination of the topic of certain Weibo posts. These three strategies are closely related in the process of understanding the TSE's meaning though they are different from each other. The following example will show how participants used this strategy to assist in their understanding of the given TSEs in Weibo.

Ex. 15

Background information: In this Weibo, the TSE *种花家* refers to China. It is a new nickname created by Chinese netizens to kindly call their homeland. *种花家* literally means "a family who plants flowers for living", which has nothing to do with "China". Indeed, *种花* is pronounced as /zhòng huā/, which shares a similar pronunciation with *中华* /zhōng huá/. The latter is the traditional name of the Chines nation. The character *家* still means "family group". Due to the rise of online live streams and online shopping live, many streamers like to call their customers and fans *家人*, which means "family members" in Chinese. *XX 家* becomes a popular saying

referring to possessive nouns such as some brands' or some streamers'. *中华家* together means "a family of the Chinese nation". The word *种花*, which means "planting flowers", also gives others a feeling of being elegant, which can express people's nice expectations toward their country.

Image 21 Example of *种花家*

[**Translation**: Whose family's medal standing ranks the first beautifully? Oh it turns out to be our 种花家 .]

Transcript extract of Participant D's think-aloud protocol:
"种花家"不太理解，"种花家"是比较厉害的那一种。因为他说奖牌榜第一，所以我觉得是种花家是对他表示一个非常厉害非常赞美的一种意思，所以我觉得可能是粉丝那一类的。
种花家, I don't really understand. *种花家* is a type of being relatively good. Because he said ranking the first in the medal standings. I think *种花家* means very good or sincerely praising someone. So I think it may be (related to) fans or something.

As we can see from the transcript, the participant did not know the exact meaning of this TSE at first. When she was invited to share how she figured out the meaning, she quoted a sentence from the Weibo post, saying that "ranking the first in the medal standings". By combining this information with the target TSE, the reader used the contextual information to help her understand the meaning of TSE. Although she finally failed to get the right meaning of this TSE, she did make use of the context concerned in her process of comprehension.

Ex. 16
Background information: *丹砂* /dan sha/ means "solo kill" in games. It is the Trans-semioticising equivalent of *单杀* /dan sha/. *单* means "solo" and *杀* means "kill". *丹砂* and *单杀* are homophones that are pronounced exactly in the same way. In Weibo context, the writer used *丹砂* to replace *单杀*, which might be due to the consideration of the sensitive character of *杀* (kill). This picture in the Weibo is a screenshot of Wangzherongyao's game. In the screenshot, two champions are battling with each other.

Image 22 Example of *丹砂*

[Translation: Anyway, just solo kill! #Wangzherongyao# Commentator Pingzi fan group's video]

Transcript extract of Participant A's think-aloud protocol:

丹砂没见过，但是我猜联系底下这个图就是单杀，就是单独的单，然后杀，自杀的杀。

I haven't seen *丹砂* before, but I guess it means solo kill if considering the picture below. It is the *单* of *单独* (solo). And *杀* is the *杀* in *自杀* (suicide).

As the shadowed line reflects, the participant used information not only from the text of the Weibo but also from the context including the picture in the Weibo. The picture did not directly provide any information about the meaning of the TSE while it contained some information implied. From the picture, she saw two champions in the game fighting. She might know some information about this game, which triggered her identification of the Chinese characters that were replaced by *丹砂*.

5.2.5 E. Activating background knowledge

Activating background knowledge is a common strategy that participants used when the context and the TSEs themselves could not provide enough information for them to identify the accurate meaning of the target TSEs. Even when the participants were given selected TSEs to understand, the TSEs and the Weibo posts were all real materials picked from Weibo. Background information about the TSEs or Weibo content could be used as a resource when participants or readers happened to know something about them.

This strategy sounds similar to strategy B, determining the topic. However, there lies great difference between them. Activating the background knowledge from related topics is talking about readers referring to information from other topics

differing from the topic of the presented Weibo post. For example, if a Weibo post is talking about an actor being rude to his fans, background knowledge about related topics can be the actor's other affairs or other actors or actresses' stories. These are different topics. In contrast, determining the topic of the Weibo focuses more on how the reader used the information from the Weibo post to identify the main topic of the Weibo, regardless of other related topics. Let us use an example to illustrate how participants used this strategy.

Ex. 17

Background information: Refer to Ex. 11 of this chapter for details.

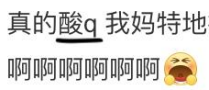


Image 18 Example of *酸 q*

[**Translation**: Really *酸 q*. My mom specially called me to ask why my selfie looks so ugly. I felt overwhelmed by appearance anxiety. Can (anyone) send me to do plastic surgery. Ahhhhhhhhh.]

Transcript extract of Participant C's think-aloud protocol:

这个是 thank you 的意思。因为一方面是它本来音就很近，然后前两天翻抖音的时候看到那个是广西还是桂林的大叔，什么栓 q 跟它差不多一个意思。

This means thank you. Because for one thing, it has a similar pronunciation (with thank you). For another, I saw this recently on TikTok. That Guangxi or Guilin's uncle. That *栓 q* shares a similar meaning with it.

As we can see from the understanding process of Participant C, he referred to his previous experience to find supportive information to verify the meaning of this TSE. He had the knowledge reserve of topics related to *酸 q* as well as the background story mentioned in the annotation about this TSE. Therefore, it became easier for him to build connections between *酸 q* and *栓 q*, and further noticed the similarity on the pronunciation level. Understanding TSEs on social media like Weibo is a dynamic and complex behavior. Most TSEs are grounded in the social media, which requires the readers to restore more information and knowledge in their mind and flexibly activate the related knowledge to support their understanding of new TSEs.

Ex. 18

Background information: This TSE is the Trans-semioticising equivalent of *看麻了* in Chinese. *看麻了* literally means "feeling numb when someone sees something". People can use it to express the status of being speechless and shocked. *看🐴了* literally means it is time "to watch a horse". It is composed of two Chinese characters and one emoji. *看* /kàn/ means "to look at", "to read", or "to see something" in Chinese. *了* /le/ is a character that functions as a symbol of the perfect tense in Chinese grammar. It describes the final state when something has been done. In this TSE, *了* expresses that the reader has readily felt numb after he or she read the news. The emoji in this TSE is a horse which is running. Horse in Chinese is pronounced as /mǎ/, which is written as *马*. This character is a homophone of *麻* /má/. They share the same consonant and vowel but have different tones. People transfer the character *麻* to *马* and then use the emoji of *马*, horse, to replace a Chinese character with the same meaning. When typing "ma" in pinyin typing system, *马*, 🐴, and *麻* can all be listed on the waiting list (as shown in the screenshot) . It is uncertain why people use the emoji to replace the original Chinese character.

Image 23 Example of *看🐴了*

[**Translation**: This transfer made me *看🐴了*. To my surprise, James Harden went to the 76ers.]

Transcript extract of Participant B's think-aloud protocol:

这转会给我看🐴了，他居然去了 76 人，这个麻四川话其实，这个我其实很早之前研究过，因为我有一次和我有个同事是，四川乐山的，然后我跟他喝酒的时候，他说他喝麻了整个人，然后我就说我就觉得很好笑，我就问他为什么是麻，他说四川话喝醉了，或者说整个人就不好了，他就都可以用麻来替代，所以这个就是"麻"的意思。

This transfer made me feel🐴. To my surprise, he went to the 76ers. This *麻* is Sichuan dialect in fact. I studied this a long time ago. Because when I had a drink with my colleague who was from Leshan, Sichuan, he said that he drank till *麻*. Then I thought it was very funny. So I asked him what *麻* meant. He said that in Sichuan dialect, when someone was drunk or did not feel well, he or she could use *麻* to express the feeling. So this is the meaning of *麻*.

When the participant came across this TSE, he first read the text, and then transferred the emoji to the Chinese character of *麻*. The participant referred to a story about how he knew the meaning of *麻*. He did not consider the topic of basketball or the transfer of James Harden when he understood this TSE. Instead, he referred to another topic to seek information to help him understand this TSE. This strategy successfully helped him understand this TSE without the help of current context information.

Ex. 19

Background information: *cpdd* means "If you want to find an online couple, please send me a message". This Weibo post only contains the TSE without any other text. *@Wangzherongyao* is a hashtag, which shows that this Weibo is posted under the Super-topic of Wangzherongyao.

Super-topic is a function embedded in Weibo. It provides an online forum for people who share the same hobbies to communicate. When adding the hashtag of *Wangzherongyao*, everyone who follows this Super-topic can see this Weibo post. This TSE needs to be analysed by dividing it into two parts, which are *cp* and *dd*. *Cp* is the shortened form of "couple", which takes the two initials of the two syllables of ***c**ou**p**le*, /'**k**ʌ**p**l/. Using *cp* to mean "couple: originated from Japanese doujin or fan-fiction culture. Doujin culture was about fans creating their own fiction for the characters that appeared in their beloved cartoons, comics, games, or movies. In Japanese, it was called カップリング, which in English was "coupling" meaning "matching couples", especially between people of the same gender. After the Japanese culture flowed into Chinese social media, the meaning of *cp* expanded to a wider range. *Cp* now could be used to refer to all types of couple relationships. *dd* is a TSE expression, which means sending messages. *dd* in English is pronounced as /di: di:/. The sound of /di:/ sounds similar to the electronic beep of the message notification sound of some online instant communication media. In China, only QQ uses the sound of /di:/ to notify the users of the coming messages. Therefore, the double *d* can vividly mimic the sound of coming messages of QQ. Using *dd* to represent sending messages is a metonymy created by the users of this TSE. However, understanding the meanings of these two parts of the TSEs is not enough for understanding the TSE's meaning as a whole. It is because this TSE is actually an extremely shortened sentence with subjects, objects, verbs, and conjunctions all omitted. The full

sentence is *if you want to build a **cp** relationship, please **dd** me (send me a message)*. One can hardly understand this TSE without knowing all of the information above.

Image 24 Example of *cpdd*

[**Translation**: cpdd@Wangzherongyao]

Transcript extract of Participant A's think-aloud protocol:

Participant A：感觉是那种 05 后比较爱用的词，我认识就是处 CP，然后就滴滴。滴滴应该就是找滴滴我，就找我聊天应该是吧？因为之前我也不太懂，我主要是 CP，然后不懂 *dd*，后来听他们说应该是这个意思。

Participant A: I feel like people born after 2005 like to use this word. I know it. It means to build a *cp* relationship and then *didi*. *didi* should mean "come to find", *didi* me, which means "come to me and talk to me". Because I did not get it before. I mainly understand what *cp* means but don't understand *dd*. I've heard it from others afterwards.

When understanding the TSE, participant A took two paths to figure out the meaning. For the first two letters, *cp*, she referred to her mental lexicon to get the meaning, while for *dd*, she activated her memory of the meaning she got from others. This example showed that the reader could take even different paths if he or she found it necessary or convenient.

5.2.6 F. Associating the meaning of the TSE

Associating the TSE's meaning is a strategy used when the readers extend their

thought to find other possible corresponding characters or words based on the TSE or components of the TSEs, which usually are emojis, numerals, and homophones. This strategy in this study is limited to the character or word association level other than topics or word formation level. Characters or words are not isolated units in languages. Readers may think of other characters or words that have connections to the original ones when understanding the meaning of the TSEs. These connections can be based on their meanings, background, pronunciations, and even shapes. This strategy was usually applied after the analysis of the target TSEs. The participants used the information gained from strategy C (analysing the TSEs) to start association with possible characters or words in order to find an interpretation that made sense in the context. Here is an example of how participant D used this strategy.

Ex. 20
Background information: Refer to Ex 18 above for details.

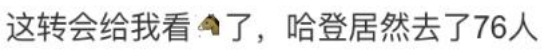


Image 23 Example of *看 了*
[**Translation**: This transfer made me *看了*. To my surprise, James Harden went to the 76ers.]
Transcript extract of Participant D's think-aloud protocol:
可能是和我今天理解的可能是一样的吧，看马了，这个麻和中文的麻痹的麻应该是一样的，所以我觉得是麻痹的麻。
Maybe it is the same with what I understand today. *看马了*. This *麻* should be the same with the *麻* in *麻痹* (numbness) in Chinese. So I think it is the *麻* of *麻痹* (numbness).

In this case, participant D first transferred the emoji to the corresponding Chinese character *马* by reading it aloud. While she did not mention how she built a connection between *麻* and *马*, she used the association strategy, thinking of what Chinese words contained the character *麻*. By using her knowledge of the Chinese language, she believed that the word *麻痹* could best suit the context. *麻* has many different meanings in Chinese so the participant needed to figure out which meaning was the suitable meaning in this TSE. Using *麻* to represent the meaning of *麻痹* is

one of the different meanings of *麻*. In this study, to find a whole word based on a single character can be considered as a process of association. Although this participant successfully identified the "correct" answer with one trial, there were still many other potential choices for her to consider, including other Chinese words that had *麻*, such as *酥麻*, which means "itchy"; *椒麻*, which means "spicy", etc.

The second example is how Participant E associated a TSE that contained an English element.

Ex. 21

Background information: The TSE, *阿婆主*, is an implicit TSE which only contains Chinese characters on the surface. However, it requires knowledge of English or Japanese to understand its meaning. *阿婆主* means "video uploader" on social media platforms, such as Bilibili, YouTube, and Weibo. The first two characters *阿婆* literally mean "maternal grandmother" in English. The pronunciation of *阿婆* /ā pó/ is close to that of the English word "up", so people use these two characters to replace "up". The word "up" is the shortened form of "upload", which is a necessary process of post videos on social media. Using "up" to replace "uploader" first became popular in Japan. *アップロード者* is the Japanese expression of *阿婆主*. Users used "up" to replace "upload" for convenience. The character *主* /zhǔ/ means "person" in English. In Chinese, the character *主* has a meaning, referring to a person who has power. For example, *家主* means the person who has the greatest power in the family. *国主* means the king of a nation. *阿婆主* then is transferred from *up 主* to its final version.]

深夜看吃播到底是什么臭毛病
人生清单里加一条
能不能成为一个吃播阿婆主

Image 25 Example of *阿婆主*

[**Translation**: What is wrong with me to watch eating broadcasts (mukbang) late at night?
One more piece added to my bucket list,
Whether (I) can become an eating broadcast *阿婆主*.]

Transcript extract of Participant E's think-aloud protocol:

深夜看吃播到底是什么臭毛病，人生清单里加一条，能不能成为一个吃播阿婆主阿婆。阿婆主应该是叫什么 b 站上面的语言，阿婆主网站上面传视频的博主，嗯哼。阿婆主，但是好像不是谐音梗。Let me think，婆肯定是 post 的谐音，我猜，但是阿婆主阿婆我现在在猜的是它的来源，但是我感觉尽力了猜不到，但应该是这个意思。

What is wrong with me to watch eating broadcasts (mukbang) late at night.

One more piece added to my bucket list, whether (I) can become an eating broadcast *阿婆主*. *阿婆*. *阿婆主* should be called as language on Bilibili. *阿婆主* (is) the vlogger who uploads videos on websites. Uh huh. *阿婆主*, but it seems not a homophonic meme. Let me think. *婆* must be the homophone of post, I guess. But *阿婆主*, *阿婆*, now I am guessing its origin. While I think I

have tried my best, I cannot guess it. But it should mean this.

In this case, the participant associated the pronunciation of 婆 with the English word *post*. As mentioned in annotation, 阿婆 is pronounced as /ā pó/. The participant did not take the character 阿 into consideration when trying to associate the meaning. He somehow focused on the single character 婆. The participant had the meaning of this TSE stored in his background knowledge in advance, which provided extra assistance to the understanding of 婆. In Chinese, *upload* and *post* have relatively close meanings in this specific context. Both of them can be used to describe the action of putting content on social media. Therefore, based on the pronunciation /pó/ and the meaning of *post*, he thought the word *post* should be the full version of *po*. It was a coincidence that the participant's association still made sense for this TSE.

As we can see from these two examples, the participants used the strategy of association to seek more possible background knowledge when they found information from the TSEs themselves and the context was limited for the purpose of understanding. The effect of using strategy may be determined by what the readers can think of at that moment and how much related background information they have stored in their mind about the topics, which means the readers may come up with both useful and useless information or even nothing. Then they need to make judgment according to the TSEs' formation and the context.

5.2.7 G. Referring to processing history

Referring to the processing history is an understanding strategy that readers may use when they find this TSE's way of formation is similar to that of another that they have processed before. As discussed in last chapter, TSEs can be categorised into several different types based on their formation characteristics. If a reader has met one type of TSEs before, he or she may refer to the processing history and use this experience to assist in understanding the new TSE that is believed to be the same type. When

designing the think-aloud task, I included at least six different TSEs in each word formation type. Therefore, the participants could meet TSEs from the same type three times. Some participants mentioned that they just dealt with a similar TSE, while some did not. In some cases, the participants directly used their own processing experience (not gained in this think-aloud task) to understand target TSEs. Here is an example of one participant using this strategy.

Ex. 22

Background information: *Ungeilivable* is the Trans-semioticising equivalent of *不给力的* /bù gěi lì de/. *给力* is a Chinese word originating from Minnan dialect, which is a dialect of Southern Fujian. In Minnan dialect, the word *够力* /gòu lì/ literally means "being powerful enough". It can also be used to describe someone who is good or supportive enough in doing something. The character *不* means *no*, which denotes a negative meaning. The English prefix *un-* can convey a similar meaning like *不* in this expression. The suffix *-able* is usually used at the end of some adjectives such as "unbelievable", "capable" and "sustainable". Correspondingly, *不给力的* is an adjective for which the creator of this TSE adds *-able* to convey a similar effect. Particularly, due to the pronunciation similarity between *geili* and *belie-*, the TSE adds an additional *v* to form the TSE. Finally, *不给力* is transformed into *ungeilivable*. It has to be mentioned that there do exist some other versions of *不给力*, for example, *ungeiliable*, which does not has the letter *v*. For Chinese readers, both versions can convey the same meaning of *不给力*.]

Image 10 Example of *ungeilivable*

[**Translation**: Professor Xue got so little footage this season... Please save the other guests' scene for the professor... Anyway, all the other guests were *ungeilivable*.]

Transcript extract of Participant T's think-aloud protocol:

这个给力， ok 这个和那个 undingable 很像，之前好像也看到过，大概也是同样的意思，就是你结合这个 un，什么 ble 的形式，中间的这个词就是替换成了一个中文的拼音叫比较给力了。那 un 给了不太给力的意思，好了，结束。

This *给力*. Okay, this (TSE) is very similar to *undingable*. I seem to have seen it before. Probably (they have) the same meaning. So you can consider the form of this *un-* and what *-ble*. The word in the middle is replaced by a Chinese (word's) pinyin, which is called relatively *给力*. That *un-* provides a meaning of "not being *给力*". Okay. Over.

The participant mentioned two periods of processing history when he met the TSE. First, he said that it was similar to "*undingable"*. This participant was given the TSE, *undingable*, to understand earlier in the think-aloud task. He failed to figure out the exact meaning of "*undingable"* but he analysed the internal structure of "*undingable"*. He noticed the pre- and suffix structure of the TSE. Thus, when he met "*ungeilivable"*, he identified the similarity between these two TSEs. He also claimed that he saw this TSE before. His previous understanding experience helped him fully understand its meaning. When he was able to deconstruct this TSE, he only needed to figure out the meaning of "*geili"*. Some readers might not have to refer to former understanding experience to understand this TSE, as basic English vocabulary knowledge could be sufficient to analyse this TSE. However, Participant T did mention that he thought of his previous experience.

The second example is how the participant referred to similar expressions' structures to understand the TSE.

Ex. 23

Background information: Refer to Ex.2 of this chapter for details.

我敢打赌，各位花粉同学，过去这几年里肯定买过华为AX3系列的Wi-Fi6路由器吧?

这回AX6的新品来了，要不要整一个?

Image 9 Example of *花粉*

[**Translation:** I can bet, every *花粉* student must have bought Huawei AX3 series Wi-Fi6 router in the past few years?
This time AX6 comes with its new product. Wanna take one?]

Transcript extract of Participant C's think-aloud protocol:

Participant C：花粉很通用的，应该是华为的粉丝。

花粉 is very universal. It should mean the fans of *华为*.

Interviewer：你最开始是咋猜到的？

How did you figure out its meaning at the beginning?

Participant C：也是音差不多，比如说各个品牌它的，或者是各个就跟 idol 一样，我会起一个后援的名字，可能会取它其中一个字的谐音或者是近似的一个名字，作为他的 fan 的名字。

It's also about the phonetic similarity. For instance, each brands' (fans). Or it's like every idol. I will give (their fans) a name to (their) fan clubs. (For the names), (fans) may give it (the fan club) a name using the homophone of one of (the idol's) name's characters or using an alternative character that looks similar to his fans' name.

In this case, the participant referred to her former processing history of fan groups' names to help to understand the TSE, which was a name of the fans of a Chinese technology company. She mentioned the mechanism through which fans used homophones or similar characters to form fan groups' names. These names might not all be TSEs, although they might share similar formation processes. Holding this understanding or processing history in mind, the participant did not have much difficulty in understanding the meaning of the TSE.

5.2.8 H. Guessing

Guessing was sometimes used by the participants when they could not clearly find evidence to support their understanding. They made their guess based on some vague feelings about the TSEs.

Ex. 24

Background information: Refer to Ex. 9 of this chapter for details.

Image 15 Example of *ban 位买房*

[**Translation**: Eight champions who bought houses at the banning position, which ones can you play? #wangzherongyao##reai jing xinchun#]

Transcript extract of Participant N's think-aloud protocol:

8 位 ban 买房的英雄，这个 ban 我是知道的，因为我虽然不玩王者荣耀。但是我玩英雄联盟。因为英雄联盟，所以我就知道它是一个怎么说，就知道它的具体实际含义是什么。然后买房我其实稍微钝了一点，但是这也很好理解，虽然我那个时候刚刚还没有看清图片，但是我那时候猜的意思应该就是指 ban 率很高，然后再看下面这个图片，也确实这个图片这个英雄都是 ban 率，从上往下，那也验证了我的猜测是正确。买房的意思就是一直被困在那个地方就出不来。

8 champions who bought houses at banning position. I knew the meaning of this *ban*. Although I don't play Wangzherongyao, I play League of Legends. Because of LOL, I knew what it meant. I knew the exact meaning of this. However, I hesitated when I saw *买房* but actually it is very easy to understand. Although I didn't take a close look at the photo just now, I had a guess that it should mean (these champions) have a high ratio of banning. Then I saw the photo below. It was a photo sharing these champions' banning ratio, from high to low, which proved my guess was right. *买房* means being stuck at that place and cannot get out from there.

Participant N successfully understood the meaning of this TSE through the strategy of guessing, apart from analysing the internal structure of the TSE. In the extract, the sentences underlined are the participant's understanding of *ban*, while the shadowed part is about how he understood *买房*. He divided this TSE into two parts, which were Chinese characters and English letters. Then with his background knowledge about MOBA (multiplayer online battle arena) gaming, he did not have trouble understanding the meaning of *ban*. Then for the second part of Chinese characters, he mentioned that he hesitated about its meaning. So he made a guess, and turned back to the Weibo to find more information to support his guess. Finally he understood the metaphorical meaning of *买房* and output a correct meaning of this TSE.

The frequency of each strategy used by the participants is shown in Table 5.1. Table 5.1 clearly indicates that the participants resorted to the strategy of analysing the TSEs most often, followed up by that of activating their background knowledge. Also quite often, they depended on the context and the text for assistance. Guessing was used the least often, suggesting that the participants were very active to find their way to understanding the TSEs.

Table 5.1 Comprehension strategies used by the participants

Comprehension strategies	**Frequency**
A. Resorting to the text	195
B. Judging from the topic	20
C. analysing the TSE C1. analysing the TSE's pronunciation; C2. Splitting the TSE; C3. Transferring the TSE's semiotic meaning; C4. analysing the TSE characters' shape.	296
D. Referring to the context	124
E. Activating background knowledge	272
F. Associating the meaning of the TSE	85
G. Referring to processing history	40
H. Guessing	26

It is to be noted that in the actual process of comprehension, the participants often used one or some of (not all) these strategies to understand the TSEs. For instance, in Ex.8, Participant Q used both resorting to the text and referring to processing history. In Example (14), apart from transferring the semiotic systems, he used other strategies as well. One was that of referring to the context, to check if the meaning of moon made sense. When realizing that the first guess did not make sense, he found another direction from the perspective of phonetic similarity. He mentioned that 约 "sounded alike 月 in 月亮 from the perspective of pronunciation". Then he referred to the context. Based on the missing context, "to do what", he thought that the emoji of the moon, which is pronounced in the same way as 约 (to date), should represent a verb of doing certain activities. So he used the fourth strategy of activating background knowledge of related topics, saying that "people usually say anyone wants to meet or gather around. "Using 约 to express the status of holding a party". Finally, he was determined that the emoji of the moon referred to the Chinese character of moon, which was replaced by its homophone, 约. Here is another example.

Ex. 25

Background information: *undingable* means “cannot tolerate something” in English. In this context, the writer used this expression to complain that he or she did not want to go to school anymore and could not insist on going to school. It is the Trans-semioticising equivalent of a Chinese expression called *顶不住了* /ding bu zhu le/. To understand this TSE, we need to divide the TSE into three parts, which are “un-”, “ding”, and “-able”. The first and the third parts are English prefix and suffix, while the second part is the pinyin of a Chinese character, *顶*. *顶* /ding/ has many meanings in Chinese, such as “using head or shoulders to hold or support something”, “being extremely good”, and “to insist on doing something even if it is very difficult”. In this TSE, *顶* uses the third meaning. The original expression *顶不住了* means “cannot keep insisting on doing something”. *不* is often used to express a negative meaning, *住* is used to express a certain status, and *了* is a particle which is used to mean that something has already happened. *不住了* together illustrate a negative status that has already or is about to happen. Considering the meaning of *不住了*, people find it suitable to use the prefix, “un-”, and the suffix, “-able”, to replace *不住了* in the original expression as they can convey relatively similar meaning. However, maybe because *un 顶 able* does not look like a “pure” English word, people change *顶* to its pinyin form, which is also composed of Latin letters. So finally, *undingable* is created.

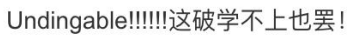


Image 26 Example of *undingable*

[**Translation**: *Undingable*!!!!!! I would rather not go to the damn school to study!]

Transcript extract of Participant T’s think-aloud protocol:

<F>这个是之前就知道 ungeilivable，这个我是第一次看到。但是<A>我读下来之后差不多也能理解他的意思了，Undingable 应该就是”顶不住了”的意思。之所以能够猜到这个意思，<C2>首先拼接的英文单词前面有个 un，后边有一个 able，它这个就是表示“无能为力的“这种感觉。<E1>那么中间的这个词这个字它这个叫 ding 它不是一个英文单词，那就联想到了是不是中文的拼音，<E2>那么就是有”顶“的意思在里边，<D>后来结合后边这段话就破学不上也罢，<E1>因为现在很多年轻人喜欢说顶不住了，或者难顶这种，undingable 应该就是”顶不住“了的意思。

This is because I knew ungeilivable before. It was the first time I saw this. But I could almost understand its meaning after I read through (the Weibo). *Undingable* should mean *顶不住了*. (Here are) the reasons why I could guess the meaning. First, the spliced English word had an “un-” in the front and an “-able” at the end so it should illustrate a feeling of ”powerless“. Then the word in the middle. This word was called *ding*, which was not an English word. Then I thought maybe it was a pinyin of Chinese. So it had the meaning of *ding* in it. Afterwards, I considered the text of “would rather not go to the damn school to study”. Since many young people nowadays liked to say *顶不住了* or *难顶*, *undingable* should mean *顶不住了* (“cannot sustain something”).

In this case, the participant used five different strategies in all to understand this TSE. The first strategy used by Participant T was referring to understanding process. At the beginning, he mentioned that he knew the meaning of another TSE called *ungeilivable*,

which implied that he could use his processing experience to support his understanding of *undingable*. *Undingable* and *ungeilivable* have a similar structure, which use the English prefix, "un-" and suffix, "-able", and the pinyin of one Chinese word's letters to form the expressions. The second strategy used was resorting to the text. Then, although he did not read the text aloud, he shared his act of reading the Weibo text by saying "I read through (the Weibo)", which helped him a lot to understand the TSE as he "could almost understand its meaning". After providing his understanding of the meaning of the TSE, he shared how he understood *undingable*. Participant T first used the strategy of splitting the TSE. He analysed the TSE part by part. He mentioned that the TSE was spliced and he analysed the prefix and suffix. He also inferred from the affixes that "un-" and "-able" together denoted powerlessness. Then the participant focused on the meaning of the middle part. Here, he used his background knowledge to identify which type of semiotic system "ding" belonged to. In the process, he might reach out to his English knowledge but came up with no corresponding English word called "ding". So he gave up the idea that the word "ding" was an English word. Then he used his Chinese language knowledge, which supported his assumption of "ding" being the pinyin of a Chinese character. When he was sure that "ding" was the pinyin of a certain Chinese word, he used the fourth strategy, which was associating the meaning of "ding". He might search with the relevant Chinese characters with the pronunciation of "ding" in his mind. He mentioned that "so it had the meaning of ding in it", which showed that he was thinking of a Chinese character pronounced as "ding". However, there were many Chinese characters having the pronunciation. Therefore, to decide which character was the correct one, the participant used the fifth strategy, referring to the context, to narrow down the range of possible characters. He referred to the text of "would rather not go to the damn school to study". Even though he tried to find clues from the context, it was not enough for him to locate the answer. So he used his background knowledge again by referring to his knowledge stored about relevant topics. This strategy helped him identify the right choice, as he knew that "many young people

nowadays liked to say '顶不住了' or '难顶'". In the expressions "难顶" /nan ding/ and "顶不住了" /ding bu zhu le/, the character of "顶" both meant someone could not bear the pressure of some challenges. Moreover, these two Chinese expressions both conveyed the feeling of "powerlessness" as the participant mentioned. The characters of "难" and "不住了" both conveyed negative attitudes similar to the affixes of "un-" and "-able". Finally, the participant matched the middle word "ding" with the Chinese character "顶" in "难顶" and output his understanding of this TSE as "undingable should mean '顶不住了' (cannot sustain something)".

The two examples above well illustrate how the participants use multiple comprehension strategies when dealing with relatively complex TSEs. As we can see from the two examples, strategies like resorting to the text and judging from the topic are usually used at the beginning of the understanding process. Then the reader will use strategies from analysing the TSEs to find clues that can provide information to direct them to the possible meaning of the TSEs. As soon as they believe that they find a possible interpretation, they will refer to the context to check if the understanding makes sense in the context. When the TSEs themselves and the context cannot provide enough useful information, they may activate their background information (strategy E) to get more information. The strategy of referring to process history can help them quickly form the right route to figure out the TSEs' meaning.

Decoding the TSEs is a strategy that is usually used by the readers to find clues and get ready for further comprehension procedures. The process of decoding the TSE only focuses on the TSE itself from the four perspectives. No associations with other words or expressions will be made in this strategy. For instance, when analysing the TSE *笑s*, the reader can analyse its pronunciation, split the TSE into *笑* and *s*, and so on. However, when the reader starts associating *笑s* with other expressions or words like *笑啥* (means what are you laughing at, which can also be written as *笑s*), the reader is actually using another strategy, associating the TSE's meaning.

5.3 Routes of TSE understanding

My data suggests that the choice of routes (i.e., the sequential order in which strategies of comprehension are applied) varied among individuals, regardless of the formation type but in line with the participants' level of familiarity with the TSEs. We introduce the notion of "route" to signal the dynamic choice and sequence of strategy choice so as to further reveal the complexities of TSE comprehension.

5.3.1 Types of TSEs and understanding routes

In this section, I will show how the comprehension of each type of TSE varied among participants from the perspective of the route choice. Here, we focus on the situation of Path 3, given that Path 1 and Path 2 are relatively straightforward.

5.3.1.1 Affixation

When dealing with affixation TSEs, the participants chose different paths and strategies. Let us have a look at the paths and strategies of understanding the TSE of *ungeilivable* by four participants who came across this TSE.

Ex. 26

Background information: Refer to Ex. 3 and Ex. 21 of this chapter for details.

Image 10 Example of *ungeilivable*

[**Translation**: Professor Xue got so little scene this season... Please save the other guests' footage for the professor... Anyway, all the other guests were *ungeilivable*.]

Transcript extract of Participant M's think-aloud protocol:

我能猜出来是什么意思。Ungeilivable。可能就是模仿英语当中的 unbelievable，然后改了把

belie 改成”给力“的意思吧。我看内容我会联想到，对，我刚才是拼了一下，我还以为是什么 get，但是应该就是给力，就是“不给力“的意思，然后加上还行。
I can guess what it means. Ungeilivable may mimic the word “unbelievable” in English. Then I change the meaning of “belie” to “geili”. I will associate the meaning when I see the content. Yeah I spelled it a little bit just now, thinking that it might be “get” or something. But it should be “geili”, meaning that “not being geili” with the rest part added.

Transcript extract of Participant T’s think-aloud protocol:
ok 这个和那个 undingable 很像，之前好像也看到过，大概也是同样的意思，就是你结合这个 un，什么 ble 的形式，中间的这个词就是替换成了一个中文的拼音叫比较给力了。那 ungeilivable”不太给力”的意思。
Ok, this one is very similar to that “undingable”. It seems that I have seen it before. Might be the same meaning. So you can consider the form of “un-” and “-ble”. The word in the middle is just replaced by a Chinese pinyin, called “geili”. So “ungeilivable” means “not being geili”.

Transcript extract of Participant R’s think-aloud protocol:
是就有一个词叫 unbelievable 很难相信的，这边以前可以念了之后发现这是一个掺杂的中文在里面的恶搞的单词，意思是“给力”，在中文里面就是一个很常见的网络词，所以这边就是“不给力”的。
Yes. There was a word called “unbelievable”, very difficult to believe. So here I could read it aloud and realise that it was a funny word with Chinese word mixed in it. It means that *给力* in Chinese is a very common online word. So here it refers to “not being geili”.

Transcript extract of Participant U’s think-aloud protocol:
薛教授这一季的薪酬也太少了，求多切几个嘉宾的镜头，给教授，反正这一季的嘉宾各个 ungeilivable, ungeilivable，它这个也是中文和英文的一个结合，“给力”是中文，大家都知道，“给力”可能是觉得他操作得很好，让人看着很爽这种意思，然后 geilivable 就是把它变成了一个形容词，然后在前面加上前缀，un-反义，也就是说这一届的嘉宾个个都是可能是观众觉得不怎么讨好的，或者觉得带着看着不够劲的，也不给劲的，然后是这个意思，对。
Professor Xue was paid too little this season. Please reduce other guests’ cut and save them to Professor. Anyway all the guests of this season a “ungeilivable”. This is also a combination of Chinese and English. *给力* is a Chinese word, which is known to many people. *给力* might mean that someone is good at doing something, which is a feast to eyes. Then “geilivable” changed to an adjective. Then a prefix is added to its front. “un-” expresses an opposite meaning. To sum up, the guests of this season are not good enough for the audience or their performance is not exciting and satisfying. So this should mean this. Yeah.

From these extracts, we could find that the four participants opted to understand the TSE of “ungeilivable” through different paths and strategies. For the first participant M, she began her understanding by associating this TSE with a similar word in English. However, she did not directly mention how she made the connection between

these two words, that is, from what perspectives she found “ungeilivable” and “unbelievable” similar. Then she claimed that “ungeilivable” was changed from “unbelievable” by replacing “belie” with “geili”. Therefore, it could be inferred from the lines that she analysed the TSE’s structure of “ungeilivable” ahead of making association, which she did not “think aloud”. Then she referred to the context to check whether the meaning of *给力* could make sense in the Weibo post. She tried other possibilities like “get” before she was sure that “geili” referred to *给力*. So for Participant M, the route of understanding the TSE was <C2> to <E> to <D>.

For Participant T, he came across the TSE of “undingable” in the think-aloud task before he was presented with “ungeilivable”. So when he saw the TSE, he referred to his previous experience of understanding similar type of TSEs. In effect, here the same situation as that of Participant M occurred, which was to know that “ungeilivable” was similar to “undingable”. The participant had to analyse the structure of the TSE before he got the conclusion. Therefore, we could suppose that Participant T also analysed the TSE’s structure before referring to processing history. Then he further explained how he understood the structure of the TSE by splitting the TSE into different parts as he mentioned that the TSE was in the form of “un- + -able”. Then he said that “the word in the middle was just replaced by a Chinese pinyin, called ‘geili’”. From this sentence, we can know that he was analysing the second part of the TSE after he split the TSE. He used another strategy in this step, which was transferring the semiotic meaning of the component in the TSE. He was aware that “geili” was the pinyin of *给力*, which showed that he successfully transferred pinyin to Chinese characters. Then when he got the meaning of the middle part, he concluded that “ungeilivable” meant “being not geili”. So the comprehension route of Participant T was <C2> to <G> to <C3>. He did not refer to the context at all when he understood the meaning of the TSE.

Participant R took a similar path from Participant M but with some differences in the strategies used. Like the two participants above, she did not speak aloud the

understanding process of analysing the TSE's structure before she connected the TSE with "unbelievable". After she analysed the structure of "ungeilivable", she introduced the meaning of "unbelievable", which implied that she referred to her background knowledge of related topics or words in this step. Then, Participant R used a strategy not used by the others to help her understand the meaning of the middle part of the TSE. She read the middle part aloud and realised that this TSE was a Chinese and English mixed expression, which showed that by reading the TSE, she noticed that the part of "geili" stood for the Chinese word. The pronunciation of "geili" for her was more like a Chinese pronunciation rather than an English word. So she associated the pronunciation of "geili" with a possible Chinese expression and finally came up with *给力*. She did not devote extra effort to understanding the meaning of *给力* as she mentioned that this word was very popular on Chinese social media, which implied that she knew the meaning of *给力*. So here she referred to her mental lexicon directly to get the meaning of this Chinese expression. Finally, she replaced "believe" in the meaning of "ungeilivable" by "geili" and got the meaning of "difficult to be geili". In other words, "ungeililvable" meant "not being geili enough". Thus, the comprehension route for Participant R was <C2> to <E> to <C> to <F>. Similar to Participant M, Participant R did not use the affixation mechanism to get the meaning. Instead, she made use of the similar English word she found to infer the meaning of "ungeilivable".

In contrast, the last participant strictly followed the word formation mechanism of affixation to understand "ungeilivable". Participant U first read through the Weibo text, which was a strategy not used by other participants above. Then, when she read "ungeilivable", she read it twice slowly. She stated that this TSE was a combination of Chinese and English. Therefore, we might infer that during the time she read it slowly twice, she was thinking about the TSE's features. During her first trial at understanding this TSE, she realised that the TSE was composed of two different languages. Then she split the TSE into two parts, which according to her were a

Chinese part and an English part. She started with transferring “geili” to the Chinese expression *给力*, and explained its meaning. Here she suggested that *给力* was known to many people, including herself. So she referred to her mental lexicon to get the meaning of *给力*. Then based on “geili”, she added the affixes to it. By adding “vable” to it, she mentioned that “geilivable” was changed to an adjective”. Then she analysed that “a prefix was added to its front”. Her knowledge of English supported her understanding that the prefix could be used to “express an opposite meaning”. So she concluded that “ungeilivable” should mean “not being geili”. We can see that she showed her strong background knowledge of both the Chinese and English grammar. Although the affixes were not English words, she could realise that they were from English grammar or word formation system. The understanding route of Participant U was <A> to <E> to <C2> to <E>.

All the routes of the four participants shared a similar strategy, which was the strategy of analysing the internal structure of the TSE. For the participants, to decode a TSE with affixes, it was necessary to decompose the TSE into two parts. One part contained the core meaning of the TSE as the meaning of the TSE was derived from the meaning conveyed from the core part. The other part contained affixes, which some participants considered as English, while others took to be just affixes that could change the core meaning. Participants T and U used the affixes to change the meaning of “geili” step by step and finally arrived at a meaning that they were satisfied with. In contrast, Participants M and R replaced the core meaning part of “unbelievable” with “geili” and then inferred the meaning of the TSE.

The examples of the four participants showed that despite adopting different routes, readers could still reach a similar and relatively “correct” understanding of the TSE. Based on the discussion above, Figure 5.1 illustrates all possible routes and strategies readers could take when understanding *ungeilivable*:

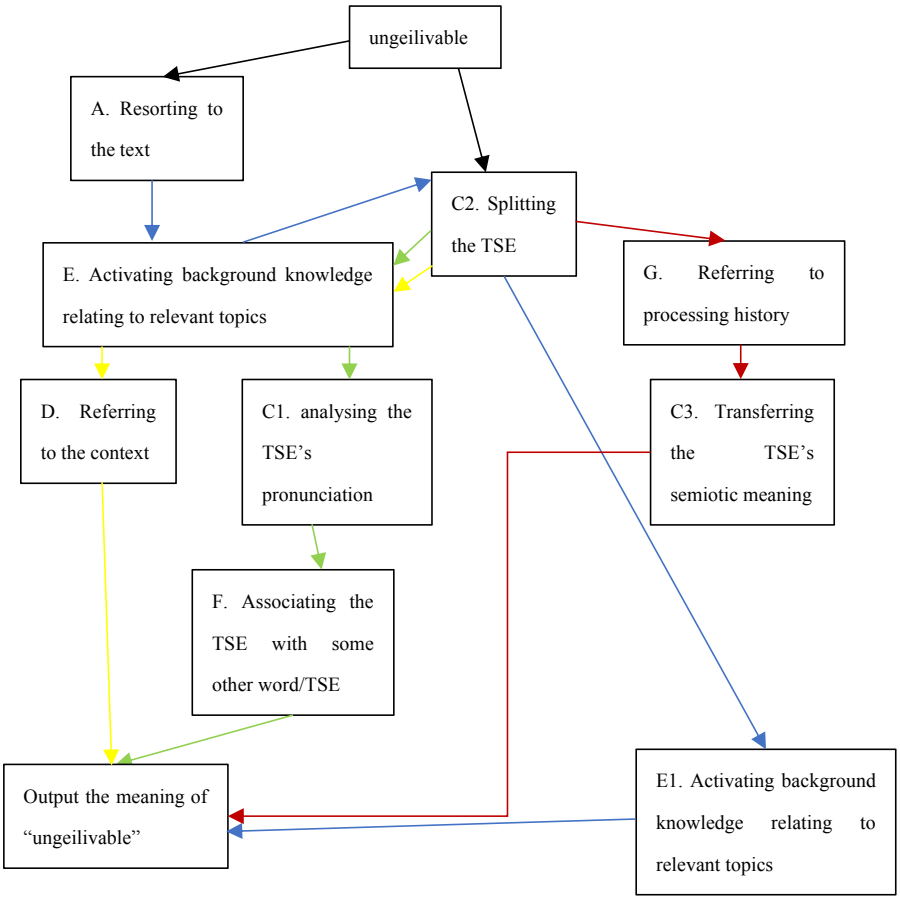


Figure 5.1 Routes for understanding *ungeilivable*

In this flow chart, four different routes were illustrated, which were 1) <C2> to <E> to <D>; 2) <C2> to <G> to <C3>; 3) <C2> to <E> to <C1> to <F> >; and 4) <A> to <E> to <C2> to <E>. All the four routes began with encounter the TSE, "ungeilivable". For Route 1 (yellow line), the participant started understanding by splitting the TSE into two parts. Then she activated her background knowledge of "unbelievable" and the Chinese word 给力, which helped her get the core meaning of this TSE. Then she referred to the context to check her understanding before finally outputting the meaning. For Route 2 (red line), the participant first split the TSE into two parts and then referred to his processing history of similar expressions to find a possible destination of understand this particular TSE. Knowing that the key to understanding the TSE was to figure out the meaning of "geili", he transferred the pinyin to Chinese characters, 给力. Finally, he output the meaning of "ungeilivable"

successfully. For Route 3 (green line), the participant also began with splitting the TSE. Then the pronunciation of "geili" reminded her that it might be a Chinese word's pinyin. Therefore, she associated the pinyin to possible characters. When she thought about the word *给力*, she directly extracted the meaning of *给力* from her mind. Finally, she made her way to the meaning of "ungeilivable". The last path (blue line) illustrated Participant U's understanding process. She began her path from reading the Weibo text. Then she activated her background knowledge of the English word "unbelievable". Referring to "unbelievable", she split the TSE into two parts, and understood the TSE part by part. For the middle part, she directly extracted her knowledge of *给力*, while for the affixes, she activated her knowledge relating to English affixes to help her change the meaning of *给力* gradually until she output the meaning of "ungeilivable".

The understanding routes of "ungeilivable" above involved the use of different strategies to serve readers' need. Participants might choose different strategies based on the information they got in the last step. In this example, the word *给力* was relatively very popular on Chinese social media. None of the participants who met "ungeilivable" in their think-aloud tasks had any trouble understanding the meaning of *给力*, which reduced the difficulty of understanding "ungeilivable" as they could easily get the meaning of "geili" by referring to their knowledge of *给力*. The next example of "undingable" would be more challenging for the participants as "ding" was not as commonly used as "geili". Let us have a look at how different participants tackled "undingable" in Ex. 27.

Ex. 27

Background information: Refer to Ex. 24 of this chapter for details.

Image 26 Example of *undingable*

[**Translation:** *Undingable*!!!!!! I would rather not go to the damn school to study!]

Transcript extract of Participant A's think-aloud protocol:

<A><C2>un---ding—able。这破学不算也罢，<B><G>应该是说上网课。钉钉。我猜的。
<A><C2>Un---ding---able. I would rather not go to the damn school. <B><F>It should be talking about online classes. *钉钉*. I guess.

Transcript extract of Participant H's think-aloud protocol:
<A><C2>Un---ding---able，这破学不少也罢，<C2>un---ding---a---ble，unding 是不？<E>这是个英文单词的，undingable，<D>破学，<F>他跟绷不住了也挺像，<D>他应该撑不住了，破学也不上也罢，就顶不住了。
<A><C2>Un---ding---able. I would rather not go to the damn school. <C2>un---ding---a---ble, unding, right? <E>This should be an English word. Undingable. <D>Damn school. <F>It looked like *绷不住了*. <D>He should fail to keep on doing something. I would rather not go to the damn school. It should be "cannot continue or keep on doing something".

Transcript extract of Participant S's think-aloud protocol:
对。<A>undingable 这破学校不上也罢，<D>消极情绪。<C2>然后前后缀这个样子，然后中间加个 ding，就是顶不住了，<C1>钉和顶一样的。
Yeah. <A>Undingable. I would rather not go to the damn school. <D>Negative moods. <C2>And the prefix and suffix. With a "ding" in the middle. So it should be "cannot *顶* anymore". <C1>"ding" was the same with *顶*.

Transcript extract of Participant T's think-aloud protocol:
这个是之前就知道 ungeilivable，这个我是第一次看到。但是<A>我读下来之后差不多也能理它的意思了，Undingable 应该就是顶不住了的意思。之所以能够猜到这个意思，<C2>首先拼接的英文单词前面有个 un，后边有一个 able，它这个就是表示无能为力的这种感觉。那么中间的这个词这个字它这个叫 ding 它不是一个英文单词，那就联想到了是不是中文的拼音，那么就是有"顶"的意思在里边，后来结合后边这段话就破学不上也罢，因为现在很多年轻人喜欢说顶不住了，或者难顶这种，undingable 应该就是"顶不住了"的意思。
This is because I knew ungeilivable before. It is the first time I have seen this. But I could almost understand its meaning after I read through (the Weibo). *Undingable* should mean *顶不住了*. (Here are) the reasons why I could guess the meaning. <C2>First, the spliced English word had a "un" in the front and an "able" at the end so it should mean "a feeling of powerless". Then the word in the middle was called *ding*, which is not an English word. Then I think maybe it is a pinyin of Chinese. <F>So it has the meaning of *ding* in it. Afterwards, I consider the text of "would rather not go to the damn school to study". Since many young people nowadays like to say *顶不住了* or *难顶*, *undingable* should mean *顶不住了* ("cannot tolerate something").

From these extracts, we could find that four participants opted to understand the TSE of "undingable" through different routes. For the Participant A, the understanding path was rather simple comparing with the other participants. She began understanding the TSE by resorting to the text. While reading the text, she read the TSE slowly and syllable by syllable. This behavior implied that she was trying to

analyse the TSE by splitting the TSE into several parts. Then she noticed that the middle part of this TSE was “ding”, according to which, she associated “ding” with *钉钉*. She also judged the topic of this TSE at this step as this Weibo was talking about learning and school. Therefore, she thought of an online classroom application, *钉钉*, which was widely used during the pandemic. Then she stopped the understanding process with a guess of “ding” referring to *钉钉*. Participant A’s understanding path was: <A> to <C2> to <E> to <B>.

For the second participant, she also started the understanding with strategies A and C2 by reading through the text and analysing the TSE at the same. Then she activated her background English knowledge and believed that this expression was like an English word. However, when she read the TSE again, she managed to associate it with another TSE, “unbengable”, which she met before. She then went back to the context and noticed that the writer was complaining about the school as she read the key word *破学*. Therefore, she realised that the writer was saying that she could not tolerate the school life any more. Based on this analysis, she got the meaning of “undingable” and output the original expression *顶不住了*. Participant H’s understanding route was <A> to <C2> to <E> to <D> to <F> to <D>.

For Participant S, she began understanding this TSE by resorting to the text as well. Before she analysed the TSE and split the TSE, she considered the context, which concerned a negative attitude. Then she mentioned that the TSE was formed with affixes and the middle part “ding”. Finally, she realised that the pronunciation of “ding” and that of *顶* were the same and successfully output the meaning of “undingable”. The route of understanding of Participant S was <A> to <D> to <C2> to <C1>.

The last participant directly referred to his former processing experience concerning “ungeilivable”. Based on this experience, he also went through two steps of resorting to the text and splitting the TSE. He first analysed the prefix “un-” and

suffix “-able” and picked “ding” out for further discussion. He activated his background knowledge of relevant topics as he considered that “ding” was not an English word but a Chinese pinyin. Then he realised that there should be a Chinese word that was pronounced as “ding” and had a meaning similar to *顶*. Afterwards, he “considered the text of ‘would rather not go to the damn school to study’” and activated his knowledge of how young people usually talked. Finally, he output his understanding of the meaning of “undingable”. His route of understanding was <G> to <A> to <C2> to <E> to <F> to <D> to <E>.

When understanding “undingable”, which was a TSE that was formed by affixes and a Chinese pinyin, all the participants above met this expression for the first time. They all started understanding the TSE from reading the text and splitting the TSE. After that, they would like to take different routes to further decode the TSE. Most of them activated their background information when they tried to understand the meaning of “ding”. Figure 5.2 is intended to illustrate how the different participants dealt with *undingable* along different routes.

In the flow chart, Participant A’s route of understanding is drawn in red lines, following the order of <A> to <C2> to <E> to <B>. Participant H’s route of understanding is drawn in green lines, following the order of <A> to <C2> to <E> to <D> to <F> to <D>. For Participant S, the route of understanding <A> to <D> to <C2> to <C1> is shown in blue lines. The last participant’s route of understanding follows the order of black lines as <G> to <A> to <C2> to <E> to <F> to <D> to <E>. The flow charts suggest that when understanding TSE formed by affixation, splitting the TSE was necessary for all the participants who were unfamiliar with the TSE. By the way, they all needed to read the text before they understood the TSE. After they split the TSE, the routes went on in different directions, which meant that

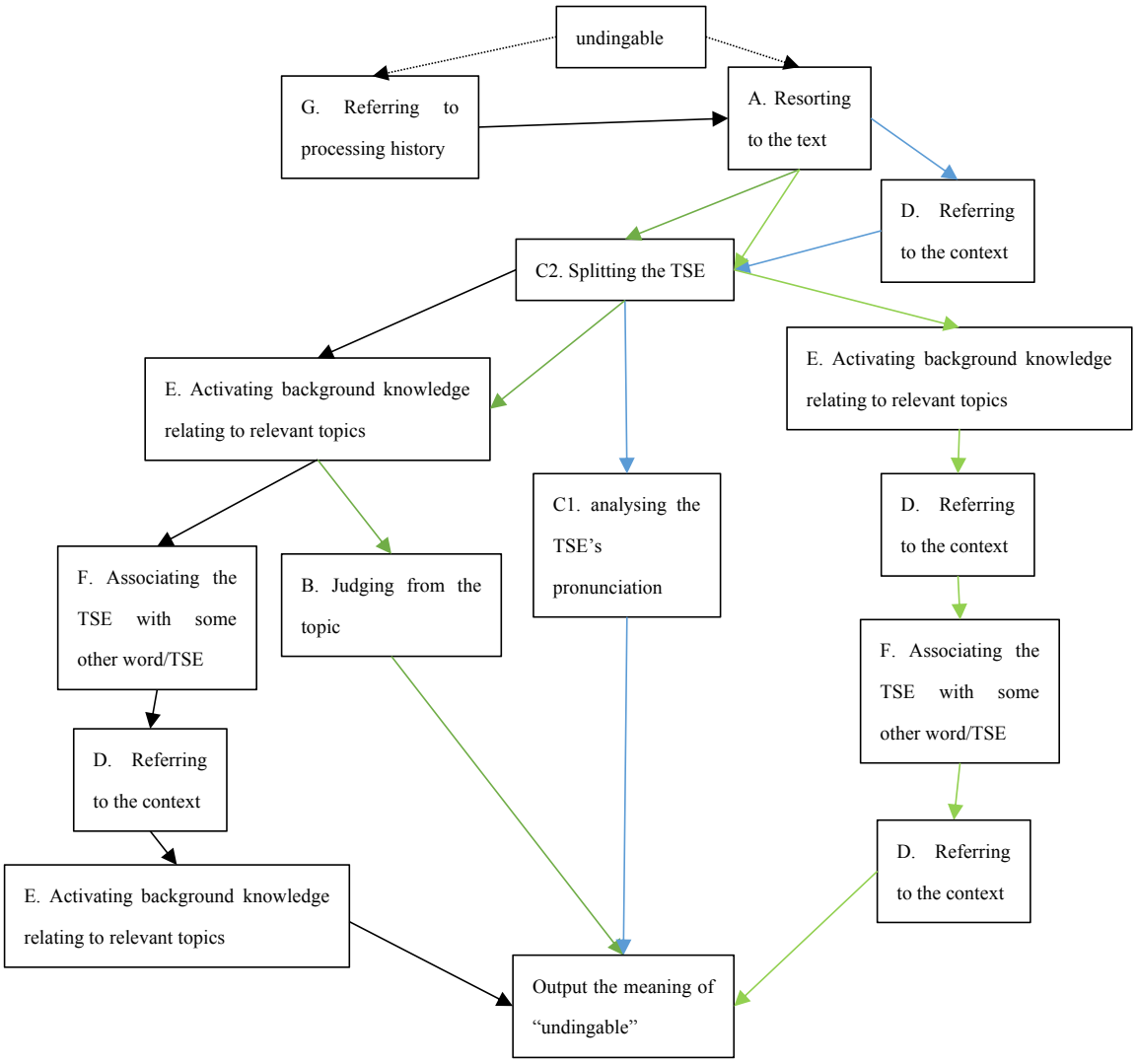


Figure 5.2 Routes for understanding *undingable*

they applied different strategies to support their decoding demand. Some participants activated their background knowledge to get additional information not provided in this Weibo, while some worked out the meaning through the pronunciation and the context. For Participant T, he was the only one who mentioned that he had the experience of decoding another TSE which had a similar structure with "undingable". In fact, the other participants might also have a similar experience, although they did not mention it. In other words, if they had processing history to refer to, their route of understanding might change a little bit, which would be more like Participant T's path. However, since they did not mention whether they referred to their process history,

here we could only draw the flow chart according to the content they spoke out.

For Participant A, she actually failed to get the "right" understanding of "undingable". Based on her understanding path, she might believe that "undingable" meant "do not want to use *钉钉* to study online at school". She did not provide her final understanding of this expression. This situation could happen in real life from time to time as readers might not need to or want to find out the exact meaning of the TSE while reading online. In effect, for the content that readers were not interested in, they would not bother to read at all. When doing the think-aloud tasks all the participants were required to understand all the TSEs and Weibo posts provided. In their real online reading experience, they might not choose to read relative Weibo posts at all. Since Participant A was satisfied with her guessing, her route of understanding would still contribute to the study, especially to the analysis of the route of understanding that could happen while reading online.

5.3.1.2 Phonetic motivation

Through the analysis of the data, the researcher noticed that participants' route of understanding shared some similarities when they were understanding TSEs that was formed based on phonetic motivation. In this section, I will discuss the route of understanding the TSEs that involve homophones, which are *456* and *9 敏*.

Ex. 28

Background information: Refer to Ex. 10 of this chapter for details.

#杨超越说别挖了真的很无语#456的粉丝可以滚出这个话题么？你们每骂一次，456就在某些人床上发sao一次🤭🤭🤭

Image 16 Example of *456*

[**Translation**: #Yang Chaoyue said please don't dig, feeling very speechless# *456*'s fans can you please leave this topic? Every time you curse, *456* will be in heat on someone's bed.]

Transcript extract of Participant B's think-aloud protocol

<A>456 的粉丝可以滚出这个话题吗？我知道。因为杨超越说别挖了，真的很无语。<E>首先这个是杨超越，所以跟杨超越在吵架的大概率就是跟他一期的，所以孟美琪或者是哪个女的。另一个然后孟美琪 456 我知道梗，因为孟美琪当时出这个事情的时候，她跟她做音乐的男朋友还不知道是不是男朋友，反正微信里有这句话，<C1>就是 456456 在音乐里就是 fa so，

fa so la，<E>他们那种玩音乐里面的那种喜欢用的这种婊话。但是对，所以它就是指孟美岐的粉丝可不可以滚出？
<A>456's fans can you please get out of this topic? I knew it. Because Yang Chaoyue said don't dig, very speechless. <E>First, this was Yang Chaoyue. So the one who was arguing with Yang Chaoyue probably debuted from the same grade with her. So (it had to be) Meng Meiqi or some other girl. Then I knew the meme about Meng Meiqi and 456. Because when the affair of Meng Meiqi happened, she and her boyfriend who was a musician, well, not sure whether he was her boyfriend. Anyway, there was a sentence in their WeChat chatting record, which was 456. <C1>456 in music was fa so, fa so la. <E>People who did music work would like to use such kind of bitchy talk. But yeah. So it should refer to whether Meng Meiqi's fans could get out.

Transcript extract of Participant F's think-aloud protocol
456 我看到这个我不太理解，<B>我觉得可能是饭圈词，<E>然后我想的是可能是偶练里面，是杨超越的名字是不是 *456*，我看一下。饭圈文化我不太懂。我看一下<C3>1234456 也不是，嗯哼。我这个就不太懂了，就不知道。
456. I don't really understand it when I saw this. <B>I think it might be a word from fandom. <E>And I think maybe it refers to Yang Chaoyue from Idol Trainees. Whether her name was *456*, let me think. I don't really know about fandom. <C3>Let me see, 1 2 3 4 4 5 6. No. uh-huh. I don't quite understand. Yeah, I don't know.

Transcript extract of Participant Q's think-aloud protocol
<A>杨超越说别挖了，真的很无语，456 的粉丝，<F>456 应该是个明星的谐，明星名字的谐音，<E>456 我真不知道这个人是谁，<E>应该因为微博上会带话题，不要人身攻击什么的，可能有清网活动，所以他们不敢在人的大名上，那 456 应该是某个人的代称，我是这么讲，但我真不知道这个人是谁。
<A>Yang Chaoyue said please don't dig, really speechless. 456's fans. <F>456 should be a homophonic name of an idol. <E>456, I really don't know who is this person. <E>Maybe because the topic will be added to the Weibo post and do not do personal attack. Maybe there was a Qingwang activity, so they did not dare to put the idol's real name on it. Then 456 should refer to someone. I would say this but I really don't know who this guy is.

This TSE was created based on homophonic mechanism based on the connection between the pronunciation of the music notes and Chinese characters. However, among all phonetic TSEs identified in the data, *456* was the only TSE that used the music notes and Chinese characters' homophonic information. Therefore, many participants might not have related background information or processing experience to understand this TSE. Participants needed to have music knowledge to think of the possibility that 4, 5, and 6 could be read as "fa", "so", and "la". Otherwise, they had to know enough background information about Meng Meiqi and the story of *456*. As a result, only Participant B successfully decoded this TSE, while both Participants F

and Q failed to understand this TSE to some degree.

Participant B started processing by resorting to the text. When he read the name of Yang Chaoyue, he recalled his background information of this person and related knowledge about Yang Chaoyue. Based on his knowledge about this topic, he considered that *456* referred to someone who attended the same debut show, which might be Meng Meiqi or someone else. Then he activated other information through the clue of Meng Meiqi. He claimed that he knew the story of Meng Meiqi and her boyfriend. In that story, he collected the information that they used the expression *456* to convey sexual innuendo messages. He decoded this expression from the perspective of music notes as 4, 5, 6 could represent fa, so, la. He then added that people who played with music would like to use such expressions. Finally, he output his understanding that in this context, *456* was used to refer to Meng Meiqi. His route of understanding was <A> to <E> to <C1> to <E>.

For Participant F, the processing began with judging the topic of this Weibo. When she saw the name of Yang Chaoyue, she realised that this Weibo was about fandom. Then she activated her background knowledge about Yang Chaoyue, who was thought to have attended the Idol Trainees programme. After the participant figured out the topic, she guessed that *456* might be used to refer to Yang Chaoyue in this context. However, she was not sure about this hypothesis due to her limited knowledge about fandom. At this point, Participant F tried to activate her background knowledge, while there was little information that could help her further understand this TSE. So she changed her strategy by attempting to transfer the meaning of *456* as numbers. She counted from 1 to 6 but could not associate any other plausible meaning with this expression. So at last, she gave up understanding the TSE. Participant F's route of understanding was <B> to <E> to <C3>.

Participant Q, the last participant also failed to totally understand this TSE. She did come up with an idea of using phonetic strategies to process this TSE; however, she did not realise that it was the music notes' pronunciation that played an important role in understanding the meaning. She began her understanding of *456* by reading

through the text. She noticed the grammar structure of "*456*'s fans", which implied that *456* had to be a celebrity who had fans. Therefore, she considered that *456* might be a homophone of a celebrity's name according to her background knowledge. Since there was no other information in this context that support Participant Q to understand this TSE from the perspective of homophones, her background knowledge might give her an idea that celebrities or idols might have names given by their fans that were created based on homophones. However, she could not further associate *456* with a particular celebrity's name. Then she activated her background knowledge about a related topic, which was about why TSEs like *456* was used on Weibo. She mentioned that it was because of the censorship policy and the avoidance of being attacked by other fans that the writer used *456* to refer to someone instead of calling the real name. Finally, based on the information she collected and analysed, she concluded that *456* was a noun but she was not sure who was the person behind *456*. Her route of understanding was <A> to <F> to <E>.

When understanding *456*, the three participants all activated their background knowledge of related topics to get more information beyond the Weibo text. However, due to the difficulty of the TSE, only one participant reached the "correct" meaning. The understanding routes and strategies of the participants could be illustrated in Figure 5.3.

The three participants' routes of understanding showed that while phonetic motivation was often used in the formation of TSEs, the corresponding homophones were seldom used when there were no direct hints provided, or the corresponding homophones required professional knowledge to support the association. Apart from the three participants, no other participants who met *456* in the think-aloud tasks successfully understood this TSE unless they knew the meaning before, which meant that they would take the direct access path to understand the meaning instead of devoting efforts to analyse the TSE. This example also showed that when the participants were unfamiliar with the provided TSEs, reading the texts and activating their background knowledge were two useful strategies that were often taken by the

participants. Here is another example.

Figure 5.3 Routes for understanding *456*

Ex. 29

Background information: *9 敏* means "oh my god" in this context. Literally, this TSE is the homophone of *救命* in Chinese, which means "to save my life". The Arabic numeral "9" refers to the character *九*, which is pronounced similarly to *救* /jiu/. The character *敏* /min/ is pronounced similarly to *命* /ming/, while the latter is a back nasal. Using *9 敏* to express the meaning of *救命* can be tracked back to a TikTok celebrity called *郭老师* (Teacher Guo). Her special accent of speaking Mandarin went viral on TikTok. Many people followed her way of speaking on both oral and written occasions. To bring an accent from pronunciation to written

form, using homophones became an effective method. *9 敏* was one of the Guo-language expressions that went popular. Shen Mengyao is a female idol from SNH48, a girl group in China. Public performance is provided by the girl group regularly, which offers a stage to the idols to show their talents and an opportunity for fans to directly watch their idols' live performance.

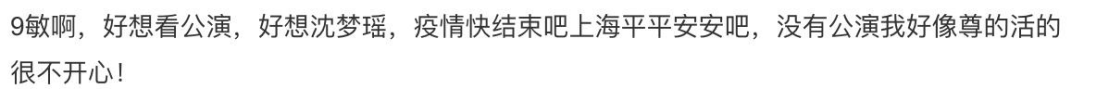


Image 27 Example of *9 敏*

[**Translation**: *9 敏*, I really want to watch the public performance. I really miss Shen Mengyao. The pandemic please end as soon as possible! Hope everything is good with Shanghai. It seems that I cannot live happily without the public performance.]

Transcript extract of Participant K's think-aloud protocol

<E>这个也是郭老师的徒弟，救命。

<E>This is also the student of Teacher Guo. 救命.

Transcript extract of Participant L's think-aloud protocol

9 敏，<C2> <C1>9 就是救的谐音，然后敏也是命的谐音，然后 *9 敏*就是救命的意思。

9 敏, <C2> <C1> "9" is the homophone of *救* (save) and *敏* is the homophone of *命* (life). So *9 敏* means *救命* (save life)。

Transcript extract of Participant P's think-aloud protocol

这个应该就是救命的意思，然后不知道这个是不是有，<E>因为以前在广东那一块生活过，<C1>感觉这样直接读起来有一点带广东口音的感觉。

This should be the meaning of *救命*. And I'm not sure whether this has. <E>Because I used to live in Guangdong Province. <C1>I feel that this expression reads like a little bit like Cantonese.

Transcript extract of Participant Q's think-aloud protocol

这个应该是救命的意思，这个也是之前在微博上看到过，<D>这之前第一次看的时候也不太明白，后来用的看的多了，然后发现大部分人写的就是救命语境。全都是很想急切的干一个事儿，或者就是想赶紧逃离现状这种语境下写的，然后后来才反应过来是救命的意思。

This should be the meaning of *救命* (save life). I also saw this expression on Weibo. <D>I didn't understand this TSE at my first encounter. But I met it many times afterwards. Then I found that most people used this expression in the context where *救命* should be used. All were eager to do something or wanted to escape from the reality. Then I realised that this expression meant *救命* ("save life").

Among all 13 participants who met this TSE in their think-aloud tasks, these four participants were representative in terms of understanding this TSE. For Participant K, this TSE was very familiar. When he saw this expression, he directly activated his background knowledge, mentioning that the writer was a fan of Teacher Guo. This showed that Participant K knew that this TSE originated from Teacher Guo's TikTok

and thus he knew the meaning of this expression. Then, he just output the meaning of *9 敏* without any delay. In this step, he might refer to his mental lexicon to get the meaning of *9 敏*, whose meaning had already been stored in his mind. So his route of understanding path was relatively very simple: <E>.

Participant L did not provide information about whether she had met the TSE before. She just started with reading the TSE aloud and then she began the understanding process. She split this TSE into two parts and analysed each element step by step. She analysed the phonetic characteristic of the two elements and pointed out that "'9' is the homophone of *救* and *敏* is the homophone of *命*". After each element was transferred to a corresponding Chinese character, she output her understanding of this TSE. Her route of understanding was <C2> to <C1>.

Participant P guessed at the meaning of this TSE at first. Then he analysed the pronunciation of this expression by combining his personal experience in Guangdong Province. He believed that the numeral "9" could be pronounced as /gau2/ in Cantonese, which sounded similar to *救* in Cantonese. Then he successfully understood the TSE. His way of analysing homophones was different from other participants. In fact, both in Mandarin and Cantonese, "9" or *九* are pronounced similarly to the character *救*. In this example, the understanding of homophonic TSE was not limited by dialect. Mandarin was just a path of establishing a relationship between two characters that sounded similar. For Participant P, the route of understanding was <E> to <C1>.

For the last participant, he took a totally different route to understand the TSE. Although *9 敏* was a typical phonetic motivated TSE, Participant Q did not realise that he could take this path. When doing the think-aloud tasks, he already knew the meaning of this expression. So he shared how he understood this TSE before. When he first met this TSE, he failed to figure out the meaning of *9 敏*. It was until he saw this TSE several times, he built a relationship between *9 敏* and *救命* through analysing the context. He found that for situations that *救命* could be used people used *9 敏* instead. So he finally understood this TSE. His route was <D>.

For the phonetically motivated TSE, *9 敏*, Participant L and P understood the TSE through its pronunciation by transferring "9" to the character *救* and *敏* to *命*. Participant K used his background knowledge and mental lexicon to directly output the meaning of this TSE. However, for Participant Q, *9 敏* did not seem to be an obvious homophonic expression. He had to turn to the context to analyse the meaning. Figure 5.4 illustrates how the different participants understood the TSE through different routes.

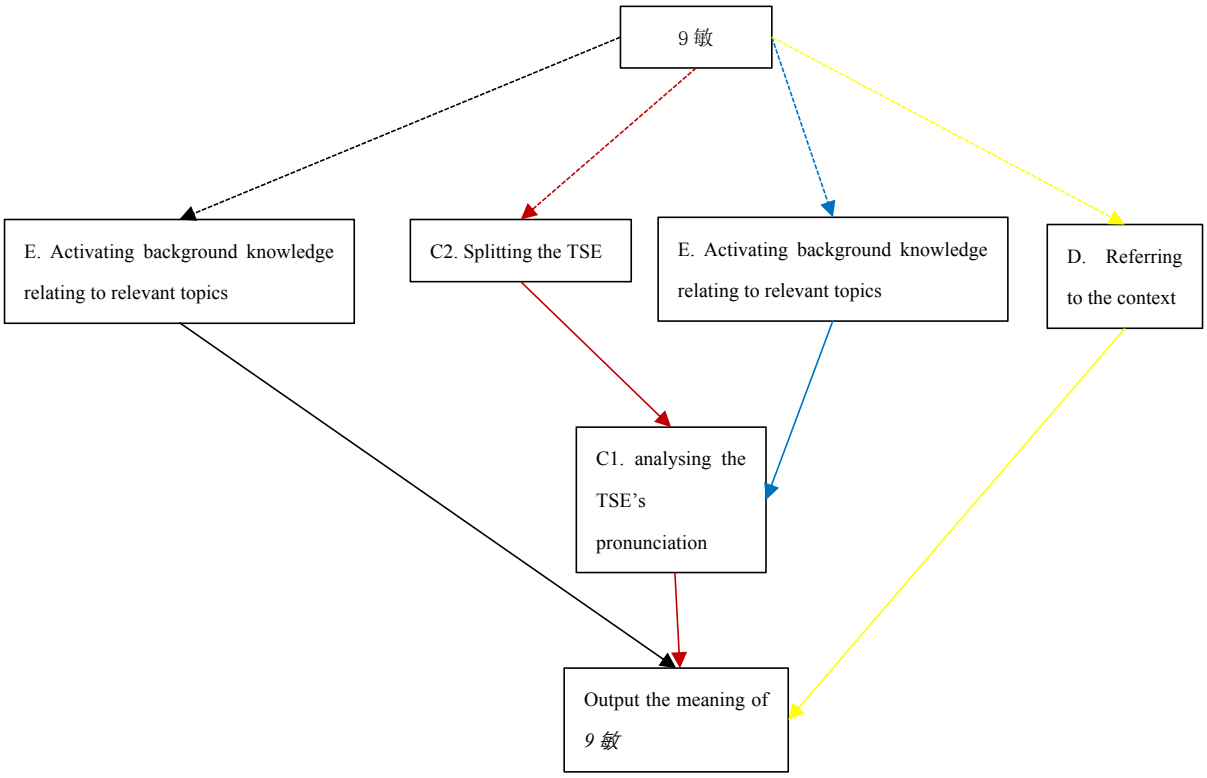


Figure 5.4 Routes for understanding *9 敏*

Comparing with affixation TSEs, the understanding routes of phonetically motivated TSEs were relatively simpler. Participants could easily identify the homophones and then found corresponding characters. The example of *9 敏* also showed that the readers' main dialect might be a factor that influenced their understanding routes Participants L and P were born and grew in Guangzhou, a southern city in China, while Participant Q was born and grew in Beijing, a northern city in China. In the northern parts of China, people distinguished the front and back nasals explicitly. In

contrast, southerners could not distinguish these two types of pronunciation clearly. Therefore, in this example, participants from the southern part of China would consider *敏* /min/ and *命* /ming/ to be homophones, while for people from the northern part, the similarity between these two characters was relatively low.

456 and *9 敏* showed two different types of phonetically motivated TSEs, as the former used rare homophones between music notes and Chinese characters, while the latter was relatively common homophones between Chinese characters and numbers. In fact, each number has a corresponding character in Chinese, such as "1" to 一, "2" to 二, and "9" to 九. Therefore, when reading numbers in Arabic form, Chinese readers could easily transfer the Arabic numbers to corresponding pinyin pronunciation. The above examples also proved that the routes of understanding for certain type of TSEs, like phonetically motivated TSEs, were not strictly limited by their types. TSEs formed by homophones could also be understood by using strategies such as referring to the context or seeking external assistance with no phonetic analysis involved in the understanding process.

5.3.1.3 C. External TSEs

TSEs that involved other languages were categorised into the type of external TSEs in previous chapters. The data of participants' think-aloud tasks revealed that readers might take different paths of understanding when they came across this type of TSEs. Foreign languages as semiotic elements in the TSEs could be a factor that affected participants' decoding process. The following two examples of *酸q* and *大丈夫* will show how participants understood these external TSEs.

Ex. 30

Background information: Refer to Ex. 11 of this chapter for details.

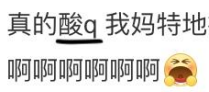


Image 18 Example of *酸q*

[**Translation**: Really *酸q*. My mom specially called me to ask why my selfie looks so ugly. I felt overwhelmed by appearance anxiety. Can (anyone) send me to do plastic surgery. Ahhhhhhhhh.]

Transcript extract of Participant A's think-aloud protocol

应该是栓 q<E>. 好像是我看最近好像比较火的一个乡村教师教英文的，好像他就说 Thank you，应该是。Thank you。对，他说的不太标准。

Should be *栓 q*.<E> It seems to be a relatively popular (expression). A rural English teacher. He seems to be saying "thank you", which should be. Thank you. Yes. He did not pronounce it properly.

Transcript extract of Participant K's think-aloud protocol

<E>最近比较火栓 q 是吧？Yeah，Thank you。太早了，已经比较火的 thank you。<C1>对，也是个谐音的。酸 q 也算对。带口音类型带口音的读英文，对。Thank You。

<E>(It's) kind of popular recently. *酸 q*, right? Yeah, thank you. It's been a while since the expression of Thank you went popular. Yes. <C1>It is also a homophone. *栓 q* can also be considered correct. It's a type of reading English with an accent. Yeah, thank you.

Transcript extract of Participant P's think-aloud protocol

<F>因为其实一开始不是有栓 q，<C2>就是木字旁一个全的栓，然后英文字母的 q。栓 q 这个词一开始因为平时之前看的有一个主播会经常这样用，<D>然后大概知道这个是因为结合他的语境，他一般是在比如说谢谢人家的那种有给他打钱给他打送礼物的时候，他会说酸 q，所以就知道这大概知道栓 q 是 thank you 的意思。<C1>然后在这里再看到酸 q 的话，因为也是音相同的，所以能理解应该是 thank you 就是谢谢的意思，<E>但这又扯到另外一个梗，就是我真的会谢。这个词其实一般他当然不是正面表示感谢的意思，他就觉得说我真的谢谢你有一种反讽的意思在里面，<D>能够结合上下这里讲一段语境就可以理解，这确实是讽刺了一次。

<F>Because there was *栓 q*, <C2>which was composed of a *木* and a *全*, along with the English letter Q. As for the word *栓 q*, I saw a streamer often use this expression. <D>Then I understood its meaning basically based on its context. It was usually used when (the streamer) wanted to thank the audience who donated gifts. Therefore, I knew that *栓 q* meant thank you. <C1>Then when I saw the *酸 q* here, I recognised that they were also homophones. <E>So I could understand that it should mean thank you as well, while it might relate to another meme, which was "I really would thank". This expression was actually not often used to express the positive meaning of gratitude. <D>It was used to ironically say that "I really thank you", which could be understood according to the context. In this case, it was indeed sarcasm.

The TSE, *酸 q*, originated from the English phrase "thank you". It was composed of one Chinese character, *酸*, and one English letter "q". Both parts were homophones of "thank you". The Chinese character *酸* refers to the syllable of "than" /ˈθæŋ/, while "q" refers to /kju:/. As mentioned in the background introduction, *酸 q* or *栓 q* has become a social media meme due to the usage of Chinese TikToker. Therefore, the three participants all activated their background knowledge to support their understanding, although some other participants also analysed the TSE from different

perspectives. Let us start from Participant A.

Participant A was familiar with this TSE. When she saw this TSE, she directly reached her background knowledge about the TikToker who spoke "thank you" in the way of *酸 q*. She also mentioned that the TikToker did not pronounce the phrase properly, which showed that she knew the right pronunciation of "thank you", while she could still understand the TikToker's way of speaking "thank you". Her route of understanding *酸 q* was just <E>.

For Participant K, the comprehension process also began with activating background knowledge about the related topic, which was Teacher Liu Tao pronouncing "thank you" as *栓 q*. However, in Participant K's background knowledge, Teacher Liu Tao said *栓 q* instead of *酸 q*. Therefore, further acts needed to be conducted, which was to build a connection between *栓 q* and *酸 q*. He mentioned that *酸 q* could also be considered correct as they were homophones. At this step, he analysed the TSE's pronunciation. Thus, his route of comprehension was <E> to <C1>.

Participant P processed this TSE along a more complicated route. He began understanding this TSE by talking about the expression of *栓 q*, which implied that he activated his background knowledge before he started to split the expression as he also mentioned that he saw a streamer using *栓 q*. He thought of *栓 q* when he saw the TSE of *酸 q*. He split the TSE into two parts, which were the Chinese character *栓* and the English letter "q". Then he took another strategy to understand this TSE. He referred to the context of how others use *栓 q*, and figured out that *栓 q* was often used to replace "thank you" or *谢谢* in communication. Then he analysed the pronunciation of *酸 q* and *栓 q* and believed that these two expressions were homophones. He did not stop when he got the meaning of *酸 q*. Instead, he activated his background knowledge about how people used "thank you" on social media. He mentioned that people would usually use *栓 q* or "I would really thank you" when people were unsatisfied. Therefore, he argued that *酸 q*, similar to *栓 q,* was used to express an ironic attitude in this Weibo, considering the context. Finally, he reached

a deeper understanding of this TSE including both its literal meaning and contextual meaning. Thus, Participant P's route of understanding was <F> to <C2> to <D> to <C1> to <E> to <D>.

All the participants successfully decoded the TSE and managed to figure out its meaning through different paths. Participant A mainly used their background knowledge of *栓 q* to support their understanding. Participants K and P both realised that there were slight differences between *栓 q* and *酸 q*. Therefore, they further analysed the pronunciation of these two expressions and claimed that they were homophones and could be used to express the same meaning. Although Participant A noticed that *酸 q* was not exactly the same as her knowledge about *栓 q*, she just considered it as an improperly pronounced version. Participant K managed to figure out the literal meaning of *酸 q* as "thank you", but did not bring this meaning into the context to check if there were further meaning conveyed. Participant P, however, activated another piece of background knowledge stored about "thank you" and realised that *酸 q* here also expressed an ironic tone in this Weibo. The more strategies used when understanding this TSE, the more information the participants might achieve about it. The understanding routes and strategies of these three participants could be shown in Figure 5.5:

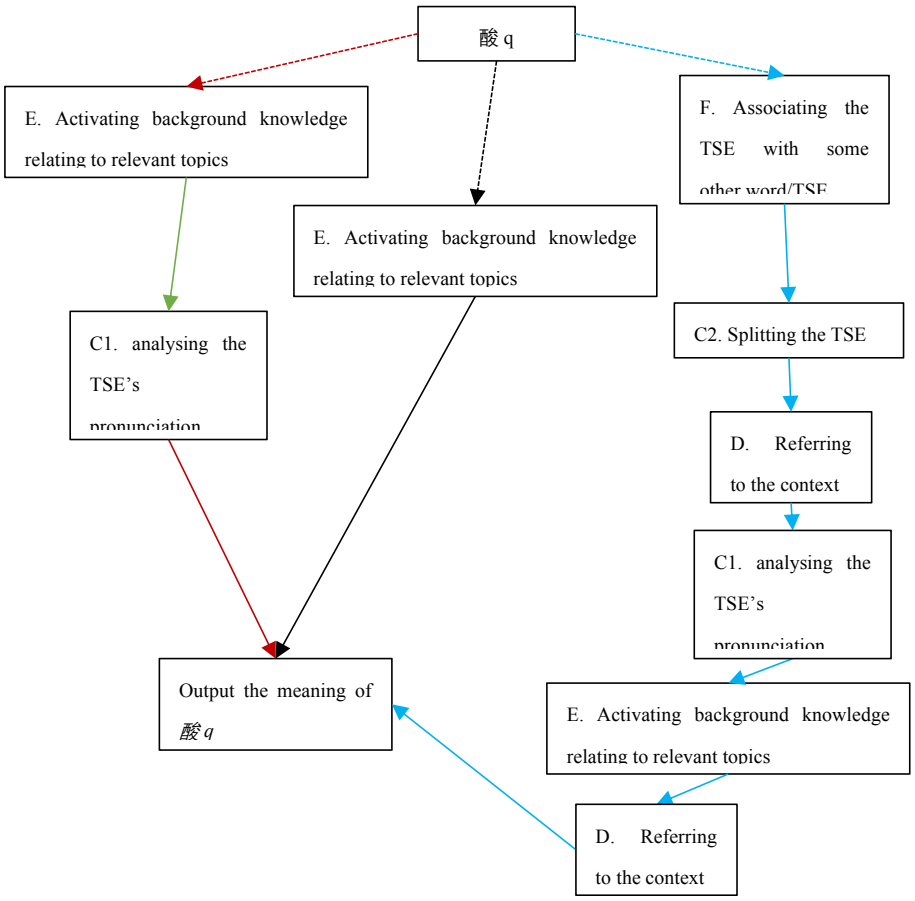


Figure 5.5 Routes for understanding *酸 q*

In the flow chart, the black lines represent the understanding route of Participant A, who only referred to her background knowledge to get the meaning of the TSE. Red lines showed the understanding route of Participant K, who used background knowledge and analysed the pronunciation of the TSE. Blue lines were relatively complex involving six steps before outputting the meaning of the TSE. He not only investigated the literal meaning of *酸 q* but also discussed its contextual meaning from the perspective of its pragmatic effect. The English phrase "thank you" was a daily expression which was known to many Chinese. Therefore, even though it was an external TSE, its English meaning did not function as a barrier when the participants understood it. Moreover, when the expression was popular enough for the readers, they might not need to refer to its meaning in the original language. Instead, they would rather activate their background knowledge to directly locate the meaning and

conduct further processes based on the meaning. The next example, *大丈夫*, on the contrary, brought difficulties to participants who did not know its Japanese meaning at all. Let us have a look at examples 31 and 32.

Ex. 31

Background information: *大丈夫* means "does not matter" in English. This TSE originated from a Japanese word *大丈夫* (だいじょうぶ) (dai jio bu). This Japanese word *Hiragana* is written literally in the same way as a Chinese word, *大丈夫*, which means "great males". In fact, it was very difficult to tell the difference between the Japanese *大丈夫* and Chinese *大丈夫*. Both Japanese typing system and Chinese typing system support the user to type the three characters. Chinese netizens who know Japanese bring the Japanese meaning to *大丈夫* in the Chinese context. In this example, the TSE *大丈夫* can be replaced by *没关系*, which means "does not matter".

Image 28 Example of *大丈夫*

[**Translation**: Luoluo is so cool at that frame. But is it really *大丈夫* with so much dust?]

Transcript extract of Participant R's think-aloud protocol

<A>不过那么多粉尘真的大丈夫吗？不过真的大丈夫吗？大丈夫真的。大丈夫吗？真的？<F>我觉得意思就是那种真的好，真的像个男人这种感觉，<E>因为大丈夫丈夫代表男人，真的是一个男人能做出来的事情，<D>或者是这种做法真的像一个男人一样吗？真的大丈夫可以这样子，没问题。

<A>But is it really *大丈夫* with so much dust? But is it really *大丈夫*? *大丈夫* really. *大丈夫*? Really? <F>I think this means "being really really like a man". <E>Because *大丈夫*, *丈夫* represents male. Really, it's a thing that a man could do. <D>Or was it a thing that a man could do? A real *大丈夫* could be like this. No problem (with this meaning).

Ex. 32

Background information: *带 jio 布* is another form of the expression of *大丈夫*, which also means "does not matter" in English. Unlike *大丈夫*, *带 jio 布* /dai jio bu/ is formed based on the

pronunciation of the Japanese word *大丈夫* /だいじょうぶ/ (dai jio bu). The characters *带* and *布*'s pinyin spelling are exactly the same as their Japanese Romanji spelling. Both *带* (Chinese character) --*大* (Japanese character) and *布* (Chinese character)--*夫* (Japanese character) are pronounced similarly to each other. As for "jio", there is no Chinese character that is pronounced as /jio/ so in the TSE, "jio" is used directly. What is more, "jio" is the Sichuan dialect pronunciation of the character *脚*, which means "foot" in English. Using "jio" to replace *脚* is a popular trend on social media, which is considered to be more cute and often used to refer to cats, dogs, and babies' small and soft foot. Therefore, *jio 布* means *脚布*, which means the cloth used for wiping feet and *带jio 布* literally means "to bring the feet wiping cloth". This literal meaning may affect readers' understanding when they are not familiar with this expression or have no Japanese knowledge.

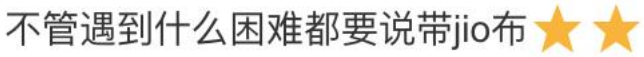


Image 29 Example of *带jio 布*

[**Translation**: No matter what difficulty one meets, one should say *带jio 布*.]

Transcript extract of Participant B's think-aloud protocol

<A>不管遇到什么困难都要说带 jio 布，<C>带 jio 布星星。卧槽，现在微博上真的是这么多骚话吗？完全我不理解，因为<D>不论遇到什么困难都要说带 jio 我不知道，我真说不出来。

<A>No matter what difficulty one meets, one should say *带jio 布*. <C>*带jio*, star star. Wow, why are there so many strange expressions on Weibo? I totally don't understand. <D>Because no matter what difficulty one encounters, one should say *带 jio*. I don't know. I really cannot say anything about it.

<A> to <C> to <D>

Transcript extract of Participant F's think-aloud protocol

下一个带 jio 布，<E>日语里面的 dai jio bu de su，就是大丈夫，<C1>然后他想用中文来把 dai jio bu 然后打出来，<C2>但是因为 jio，然后中文里面没有发音，然后他就用 jio 来代表 Jio 的音，然后带和布中文和日语里面都有这个音，然后带中在一个。

Next, *带jio 布*. <E>The *大丈夫* represents "dai jio bu de su" in Japanese. <C1>Then he wanted to use Chinese to type "dai jio bu". <C2>However, as there was no character in Chinese that could be pronounced as "jio", the he use "jio" to represent the pronunciation of "jio". Then both Chinese and Japanese have characters that are pronounced as *带* and *布*, with "jio" added in the middle.

<E> to <C1> to <C2>

Transcript extract of Participant J's think-aloud protocol

<A>不管遇到什么困难，都要说带 jio 布。<C1>难道是带脚布，听起来像脚布，<D>但是困难跟脚布有什么关系？带脚布，难的地方说带我猜就是带脚布，嗯嗯嗯，<C2>或者说带 jio 布它或者带脚布这三个词是一个意思，带 jio 布这三个字，<E>不要说应该可能是一个日文或者说韩文这种，带 jio，应该是个日文，他有是什么加油还是什么意思，他说不放弃。对，我基本上就是猜这个意思。就 ok。<D>不管遇到什么没问题。

<A>No matter what difficulty one meats,, one should say *带jio 布*. <C1>Isn't it *带脚布* (bring feet cloth)? It sounds like a feet towel. <D>But are there any connections between difficulties and feet towel? *带脚布*, need to say *带脚* when meeting difficulty. I guess it is the meaning of "bringing a feet towel". Emmmm. <C2>Or the three words in *带 jio 布* convey a meaning

together. The three characters of *带 jio 布*. <E>Maybe it could be a Japanese or Korean word. *带 jio* should be a Japanese word. It should mean "cheer up" or something else. It says that "don't give up". Yeah, I basically believe that it means this. Okay. <D>No matter what difficulty one meets, no problem.
<A> to <C1> to <D> to <C2> to <E> to <D>

Transcript extract of Participant S's think-aloud protocol
<A>都要说带 jio 布，就是不管遇到什么困难，都要说带 jio 布，<E>带 jio 布意思肯定就是日语，就没关系的意思。但其实一开始我刚刚读这个句子的时候，<D>我还以为是有人摆烂的，就不管遇到什么困难，都是他带 jio 布不要这个样子。因为我忘记了带 jio 布的意思是对我一开始用了一个错误的这个意思，我还觉得这个句子成立，然后现在我意识到他好像是意思就是我想起那个意思是没关系。<C1>因为我觉得他这个也有问题，带 jio 布应该带教，不就带教不，我知道那个词语读法，所以一开始没有完全联系上。
<A>All need to say *带jio 布*. Whatever difficulty one meets, one needs to say *带jio 布*. <E>The meaning of *带 jio 布* has to be "does not matter" in Japanese. <D>But actually when I first read this sentence, I thought it meant that someone gave up when meeting difficulty, because I forgot the meaning of *带 jio 布*. But I should consider this wrong meaning could also make sense in this sentence. But now I realise that it should be the meaning that I did not think of. <C1>I think this expression has its own problem. It should use the character *教* instead of *带 jio 布*. I knew the pronunciation of that word, so I failed to connect them with each other at first.
<A> to <E> to <D> to <C1>

In examples 31 and 32, two TSEs originating from the Japanese expression *大丈夫* were used in the Weibo contexts in different forms. Participants met both of these TSEs during the think-aloud tasks. Both *大丈夫* and *带 jio 布* were external TSEs that required the readers to harness corresponding knowledge of the Japanese vocabulary's meaning and even its pronunciation to understand them. The five participants took different paths to understand the TSE, while not all of them successfully figured out the meaning of *大丈夫* in the end. Whether they knew the meaning and pronunciation of *大丈夫* played an important role in the process.

The first participant came across the expression of *大丈夫*, which could be either considered as a pure Japanese expression or three Chinese characters that were used to represent the Japanese word. Participant R activated his Chinese vocabulary knowledge about this TSE and read it as *大丈夫* /da zhang fu/, which was the pronunciation of *大丈夫* in Chinese. In other words, Participant R did not realise that this TSE was an external expression, which influenced his further understanding process. He repeatedly read the text when the Weibo post was presented to him. Then

he activated his background knowledge of the meaning of 大丈夫 /da zhang fu/ in Chinese, which referred to being a great and responsible male or acting like a tough guy. Based on this information, he referred to the context to verify whether the meaning he believed could make sense. He transferred the expression of 大丈夫 into the meaning of "doing what a man should do" and thought that dancing in dust could be considered as an action that men would like to do. Therefore, he output his understanding of 大丈夫 as its literal Chinese meaning. His route of understanding was <A> to <F> to <E> to <D>.

Participants B, F, J, and S had to go through a more complex route. The TSE, *带 jio 布*, involved two languages and two types of semiotic elements in the formation process. For Participant B, he began by reading the Weibo text and tried to analyse the pronunciation of *带jio 布* by reading it slowly. When realizing that he did know the TSE, he read the text again and attempted to get more information from the context. However, he could not get enough information to support his understanding. Therefore, he gave up understanding the TSE. His route of understanding was: <A> to <C> to <D>.

Participant F did not directly output the exact understanding of this TSE. However, she first activated her background knowledge about Japanese, claiming that *带jio 布* referred to the Japanese word 大丈夫, which implied that she actually knew the Japanese word. She also explained how the TSE was transferred from its Japanese form by splitting this TSE into different parts. Her route of understanding was: <E> to <C1> to <C2>.

Participant J took a longer path to reach the meaning of *带jio 布* compared with other participants. He began his understanding by resorting to the text. Then he analysed the pronunciation of *带jio 布* and transferred "jio" to *脚*. He first believed that the TSE was talking about feet towel. So he brought this understanding to the context to verify the meaning. When realizing that the meaning of feet towel could not make sense, he reconsidered the TSE and began to think about it as a whole. He activated his foreign language knowledge and believed that this was a Japanese or

Korean expression. Then he read the TSE again and claimed that it had a Japanese pronunciation. He tried to reach his mental lexicon about Japanese and output that *带 jio 布* might mean "cheer up". After that, he went back to the context and figured out that this TSE might mean "don't give up". His route of understanding was: <A> to <C1> to <D> to <C2> to <E> to <D>.

Participant S began by reading the text when she encountered this Weibo. Then she activated her background knowledge about Japanese to look up the meaning of *大丈夫*. She then mentioned that according to the context, she at first misunderstood the meaning of this TSE as she did not build a connection between the pronunciation *带 jio 布* and that of *大丈夫* as she claimed that the pronunciation of "jio" was not correct so that she did not think that *带 jio 布* referred to the Japanese expression. Her route of understanding was: <A> to <E> to <D> to <C1>.

Four of the five participants began the comprehension process by resorting to the text. They generally used the strategy of activating background knowledge, referring to the context, and analysing the expression's pronunciation when processing the TSE. Not all of the participants could finally get the meaning of the TSE. In other words, as an external TSE, participants might not be able to further understand the TSE if their Japanese knowledge was limited. Figure 5.6 illustrates the different routes and strategies in understanding:

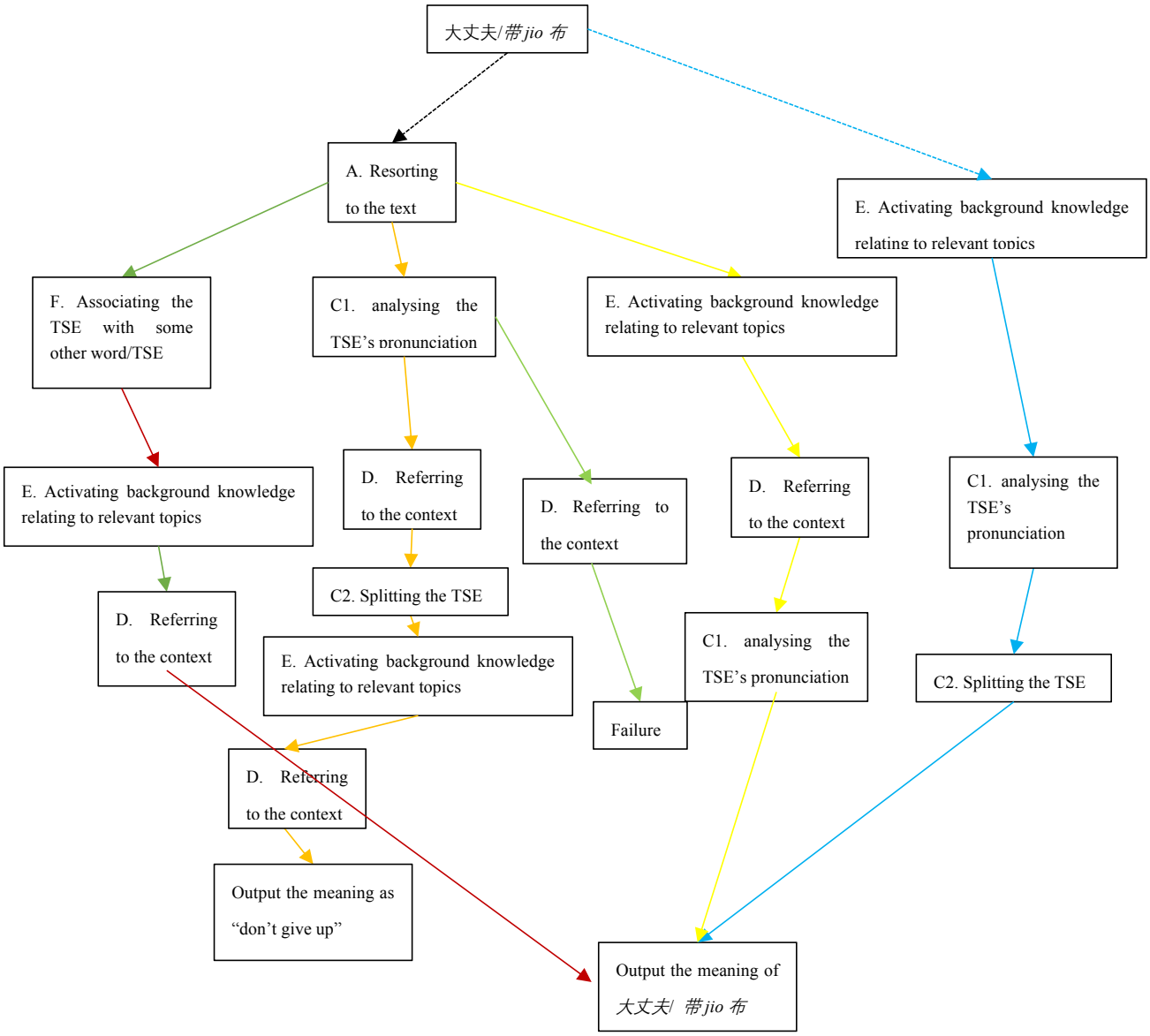


Figure 5.6 Routes for understanding *大丈夫/带jio 布*

Four understanding paths began with strategy A, resorting to the text. The red lines stood for Participant R. Participant B followed the green lines of understanding paths. For Participant J, the orange lines illustrated the paths from reading the text to considering the context and finally outputting his understanding of the meaning of the TSE. Participant S's understanding paths were drawn in yellow lines, while for Participant F, who was the only participant who skipped reading the text, her understanding route was shown in blue lines. Participant B failed to output a plausible meaning of the TSE even though he tried to use different strategies to achieve related information. The biggest barrier that stopped him from getting further progress was

that he had no background knowledge stored in his mind about Japanese or the direct meaning of the TSE. Therefore, when he read the TSE aloud, he could not build a connection between the pronunciation of 大丈夫 and its meaning. A similar situation also happened to Participant J. As we could see from Figure 5.6, he tried two different directions when understanding the TSE. When he did not realise that the TSE might involve another language in understanding, he analysed the pronunciation and attempted to use his Chinese background knowledge to seek a meaningful explanation. After noticing that the meaning of foot towel did not make sense, he shifted his mind and reconsidered the TSE from the perspective of external expression. He used his limited background knowledge about the foreign languages of Japanese and Korean to find a relatively more relevant meaning. He mentioned the meaning of "cheer up", which was called "頑張って (がんばって)" in Japanese. This expression was sometimes used on Chinese social media. Participant J was thinking of all possible Japanese words that he knew in his mind. Finally, with the help of the context, he concluded the meaning of "not giving up", which could express a similar attitude but a different meaning in this context. He then output this meaning as an outcome and ended his comprehension process. On the contrary, Participant R, F, and S all had the Japanese knowledge about the pronunciation and meaning of the word 大丈夫, which greatly supported their understanding as long as they could analyse the pronunciation of *带jio 布* and realised the similarity between *带jio 布* and 大丈夫. Therefore, for external TSEs' understanding processes, whether the reader knew the meaning and pronunciation of the original foreign language word or expression was very important to the success of the final come. Without such background knowledge, the understanding processes might be blocked or become quite difficult.

5.3.1.4 D. Semantic TSEs

Ex. 33

Background information: Refer to Ex. 13 of this chapter for details.

Image 20 Example of 几几

[**Translation**: Take a look. Who is the real 几几?]

Transcript extract of Participant T's think-aloud protocol:

<C4>这个的话这个字首先是不是一个汉字我也不确定，但是这个就表示"凡凡"的意思了。<D>因为下面的几张照片对应的是艺人吴亦凡，<E>吴亦凡有个外号叫"凡凡"，<A>哪个是真实的？凡凡这个应该是之所以写成这个样子，我看看有没有什么特别的意思在，没有想到什么特殊的含义了。

<C4>For this one, I am not sure whether this character is a Chinese character. But this should represent 凡凡. <D>Because the photos below are 吴亦凡's, who was a pop star. <E>He had a nickname called 凡凡. <A>Which one is real? Why 凡凡 was written like this? Let me see if there were any special meanings. Well I could not come up with any ideas.

<C4> to <D> to <E> to <A>

Transcript extract of Participant U's think-aloud protocol:

<A>看看哪个才是真实的凡凡。这个字其实我并不认识它叫什么，它是什么怎么读具体<E>我在网上看过，但也没去查证过，<D>但是配的图很明显看得出来是吴亦凡对，然后大家应该会用大家都是用去代指他，并且这两个字也是跟他个人的一种网上大家所知道的一些八卦娱乐新闻，<E>他的个人情感方面的一些问题，<C4>然后导致大家把它的原有的在中间的点放在了上面。对，所以我猜读凡凡，但是我真的不是很确定他叫什么，他应该是代指吴亦凡对。

<A>Check which one was the real 凡凡. I actually don't know the characters and could not pronounce them. <E>I have seen this on social media before but I did not check how to read it. <D>While the attached photos were definitely 吴亦凡. Yeah. <E>People should use this expression to refer to him. <C4>And these two characters were related to his personal affairs which were well known by people about his personal relationship issues. Because of this, people moved the two dots from the middle to the top. So I guess it should be pronounced in the same way as 凡凡. But I am really not sure how to call this expression. It should be used to refer to 吴亦凡. Yeah.

<A> to <E> to <D> to <E> to <C4>

Transcript extract of Participant M's think-aloud protocol:

<E>我知道凡凡他有这个，嗯，因为他名字叫吴亦凡，我有见过这个字，但我不知道这个字念什么。<C4>我看过这个字，然后因为凡它中间不是有个点，然后这个点就经常大家会可能用中国的汉字各种改造。怎么改造，我现在具体也忘了，但是它放在头上我不知道什么意思，但是我知道他是"凡凡"的意思，但是肯定就是讽刺他这两个字不一样，讽刺他的某一些特质，但具体什么特质我也看不太出来。

<E>I know that 凡凡 has this (nick name). Yeah. Because his name is Wu Yinfan. I have seen

this character before but I don't know how to say it. <C4>I have seen this character. And because it has a dot in the middle, does it? So this dot usually is modified by people on different Chinese characters. I forgot how to modify it in detail now. But I have no idea why it is put on the head. I do know it refers to "fanfan". Anyway, it must be used to satirise him and some of his characteristics. But I cannot tell what characteristics (this TSE satirises).
<E> to <C4>

For this TSE, 𠃌𠃌, all the three participants successfully output their understanding of these two characters although none of them knew how to correctly pronounce them. They all analysed the TSE's characters' pattern when they came across 𠃌𠃌. Participant T started with discussing the character of 𠃌 at the beginning. He claimed that he was not sure whether it could be considered as a Chinese character. Then to understand this character, he turned to the context to seek an answer. The Weibo contained not only a text but also pictures. Therefore, readers could get more information from the pictures. Participant T made use of the attached photos and recognised that the person shown in the pictures was Wu Yifan. He then activated his background knowledge about the person and realised that the character 𠃌 looked similar to 凡 in Wu Yifan's Chinese name. Finally, he read the text to verify whether calling 𠃌 as 凡 could make sense and output his understanding. Thus, Participant T's route of understanding was: <C4> to <D> to <E> to <A>.

For Participant U, the understanding paths started from reading the text. She mentioned that she saw this character before on social media, which proved that she activated her background knowledge about this expression to help her decode this TSE. Like Participant T, Participant U also did not know the pronunciation of these characters. Therefore, she went back to the context to find more information. She noticed that the person in the photo below was Wu Yifan. So she believed that 𠃌𠃌 should be used as a noun to refer to Wu Yifan, whose name was written as 吴亦凡. She did not finish her understanding when she knew that 𠃌𠃌 meant 凡凡. She tried to use her background knowledge to explain the pattern changes from 凡 to 𠃌. She believed that it was because of his personal affair that drove people to move the

position of the dot. The understanding route of Participant U was: <A> to <E> to <D> to <E> to <C4>.

Participant M knew this TSE before she participated in the think-aloud tasks. Therefore, when she saw this Weibo, she activated her background knowledge about the TSE. However, although she knew the meaning of the TSE, she could not read aloud as she did not recognise the Chinese character. Then she also attempted to analyse why the character 凡 was modified to 𠃌. She believed that people were using the movement of the dot of 凡 to ironically criticise Wu Yifan for the crime he committed. The understanding route for Participant M was: <E> to <C4>.

Semantic TSEs like 𠃌𠃌 did not involve other languages during the process of comprehension. The understanding process of 𠃌𠃌 required the readers to treat Chinese characters like a picture and analyse the change of the details of the picture. The meaning of the Chinese characters could sometimes be accessed according to the context or their similar characters. In this example, 𠃌 and 凡 looked similar to each other and the context provided a picture of a person whose name was 吴亦凡. However, due to the gap between Chinese characters and their pronunciation, even though all participants' first language was Chinese, they could not pronounce the character of 𠃌 without the help of pinyin. For online reading activities, it was not necessary for the readers to know the pronunciation of the TSEs when the meaning of the TSE could be obtained, especially for semantic TSEs. Figure 5.7 illustrates the understanding routes and strategies of the three participants:

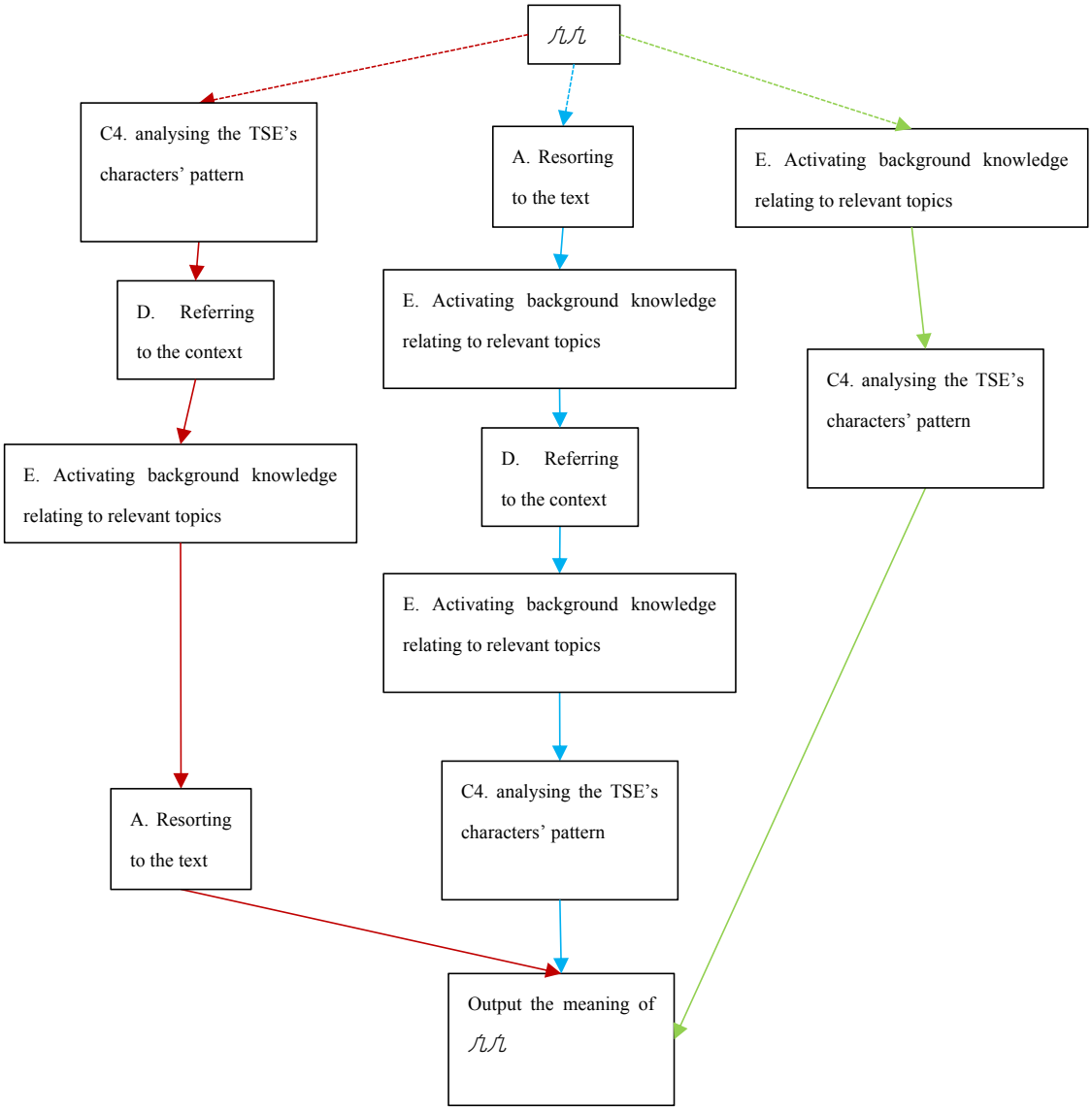


Figure 5.7 Routes for understanding 几几

The red lines showed the understanding routes of Participant T. The blue lines were the understanding routes of Participant U and the green lines drew the routes of Participant M. As shown in the flow charts, the strategy C4 (analysing the TSE's characters' pattern) appeared in all the participants' understanding process, which was a unique move that only semantic TSEs like 几几 would have. The context also supported the participants a lot. As in this particular example, the photo of Wu Yifan was an important piece of information that could assist the participants to build a connection between 几 and 凡. It also needed to be mentioned that the participants

all had relevant background knowledge about the person who appeared in the Weibo. When a TSE did not use homophonic strategies or other strategies, participants might need enough background information to support their understanding.

5.3.2 Routes for understanding TSEs of different familiarity on levels

It was also found that readers' familiarity with TSEs played an important role in the understanding processes. Readers with different levels of familiarity would take different routes and strategies, which was shown clearly in the understanding flow charts of each participant. Three levels of familiarity were distinguished in this study, which were A. When the participants knew the exact meaning of the TSE before the task; B. when the participants had some knowledge about the TSE before; and C. When the participants had no knowledge about the TSE. The knowledge about the TSE included both the TSE's meaning and sometimes pronunciation. The following section will introduce how the familiarity of TSEs influenced the paths of TSE understanding.

5.3.2.1 When the participants knew the exact meaning of the TSE before the task

When the participants knew the meaning of the provided TSEs, they would take the path of **direct access** to provide the meaning of TSEs. The understanding path would be very simple as participants only needed to refer to their mental lexicon to get the meaning of the TSE and then output the meaning.

Ex. 34

Background information: Refer to Ex. 1 of this chapter for details.

Image 8 Example of *集美*

[**Translation**: Goodnight *集美 s*. (Which one looks good? Pick one for me! I will see (your comments) tomorrow morning.]

Transcript extract of Participant A's think-aloud protocol:

集美们应该是姐，姐妹们，我知道
集美 s should be sis, sisters. I knew it.

In this example, Participant A directly output the meaning of this TSE and emphasised that she knew the meaning, without providing any other understanding processes. She might directly reach out to her mental lexicon to get the meaning and output the meaning quickly.

5.3.2.2 When the participants had some knowledge about the TSE before

The second level of familiarity refers to the situation in which participants did not know the exact meaning of the TSE. There were two different situations when the participants only had some information about the TSE.

When the participants came across a TSE that was believed to share the same origin with another TSE, understanding the meaning of the new TSE would not be very challenging even though the participants did not know its exact meaning. For example, the TSE, *酸 q*, originated from the phrase "thank you" in English, which had evolved into different TSEs on Chinese social media, such as *栓 q*, *蒜 q*, and *酸 q*. These different TSEs could all be used to express "thank you". Participants might be able to identify the original phrase when trying to understand the TSE. Let us have a look at Ex. 35.

Ex. 35
Background information: Refer to Ex. 11 of this chapter for details.

Image 18 Example of *酸 q*
[**Translation**: Really *酸 q*. My mom specially called me to ask why my selfie looks so ugly. I felt overwhelmed by appearance anxiety. Can (anyone) send me to do plastic surgery? Ahhhhhhhhh.]
Transcript extract of Participant C's think-aloud protocol:
这个是 thank you 的意思。因为一方面是它本来音就很近，然后前两天翻抖音的时候看到那个是广西还是桂林的大叔，什么栓 q 跟它差不多一个意思。
This means thank you. Because for one thing, it has a similar pronunciation (with "thank you").

For another, I saw this recently on TikTok. That Guangxi or Guilin's uncle. That *栓 q* shares a similar meaning with it.

In this example, Participant C mentioned that *酸q* was similar to *栓q*, which she met a few days before. Her knowledge of *栓 q*, which meant "thank you", helped her understand the meaning of *酸 q*. She used two strategies to identify the meaning of this TSE. One was analysing the pronunciation and the other one was comparing it with *栓 q*. The two strategies together assisted the participant in providing an appropriate meaning for herself.

When the participants knew how to understand this type of TSE based on their experience of understanding similar TSEs, the participants would take different paths. Sometimes, the participants' previous experience in understanding similar TSEs played a significant role in the understanding processes. In other words, the participants had mastered certain methods or key points of understanding similar TSEs. In this situation, although the participants did not know the meaning of the provided TSEs, they could use their experience of understanding other TSEs that were similar to the provided TSEs to support their understanding. The understanding experience could help the participants from different perspectives, including the experience of understanding TSEs with similar word formation type, the experience of associating the semiotic elements with related characters, and the experience of identifying homophones. These experiences could help the participants get more chances to understand the meanings of the TSEs. However, these experiences could not guarantee the final success of understanding the meaning of the TSEs. Here are some examples of how the participants used their previous experience to support their understanding of new TSEs.

Ex.36

Background information: *zhuangbility* is the Translanguaging expression of *装逼* /zhuang bi/, which means "someone being a poser". This TSE is a combination of Chinese pinyin and English suffix.

看到刘的采访 我想到了某东的数学诺贝尔奖 还蛮喜欢某东的 但是这个事件之后就彻底脱粉了 路人粉的粉哈哈哈哈哈之前他发的微博我一条都看不懂 就觉得可以写出来的人好厉害 但现在只觉得好zhuangbility

Image 13 Example of *zhuangbility*

[**Translation:** When I saw Liu's interview, I thought about Dong's Mathematics Nobel Prize. (I used to) like Dong before, while after this affair I did not fan him any more. The fan of "passerby fan" lol. I could not understand a single Weibo he posted, thinking that the person who could write these must be very extraordinary. But now I could only feel very *zhuangbility*.]

Transcript extract of Participant E's think-aloud protocol

zhuangbility，那就和之前那个 niubility 是差不多的，只不过之前那个是在夸别人牛逼，这个是在说你装逼，但它是动词变成了名词。

Zhuangbility, then, is similar to "niubility". However, the former one was talking about someone being *牛逼*, while this one was talking about *装逼*. It changed from a noun to a verb.

In this example, Participant E realised that the TSE *zhuangbility* was similar to *niubility*. He mentioned the changes of the middle word, which implied that he followed his understanding experience of analysing the middle word to figure out the meaning of the TSE. His familiarity with the formation of this type of TSE helped him successfully understand the meaning of the TSE.

Ex. 37

Background information: Refer to Ex. 18 of this chapter for details.

这转会给我看🐴了，哈登居然去了76人

Image 23 Example of *看　了*

[**Translation**: This transfer made me *看🐴了*, To my surprise, James Harden went to the 76ers.]

Transcript extract of Participant M's think-aloud protocol:

看🐴了，这个🐴它有很多种意思，它其实就是 mā má mǎ mà，反正我想的是，应该是，能碰得上，比如说这是看麻了，应该是看妈了，看马了，看骂了，应该都不是，所以我会猜就是看麻了，但是我应该是知道这个意思，主要是靠意会，中文很多都是靠意会。加上这种就那种 Emoji 的表情，我会猜它是看麻了。

看🐴了, this 🐴 has many different meanings. It actually should be mā má mǎ mà. Anyway, I think it should be able to match, such as *看麻*/má/*了*, *看妈*/mā/*了*, *看马*/mǎ/*了*, and *看骂*/mà/*了*. These should not be. So I would guess it must be *看麻了*. But I should know the meaning. Basically Chinese is based on sense. Plus, many expressions use the emoji. I would guess it is *看麻了*.

In this example, the participant mentioned that many expressions contained emojis in Chinese. So she had the experience of dealing with this type of expressions before.

When she started to understand the meaning of this TSE, she first said that the emoji of horse had many different meanings, which showed that she knew that when understanding TSEs with emojis, she needed to transfer the emoji to other corresponding characters. Then she tried four different tones of /ma/ in order to find a suitable character that could make sense in the context. The participant showed her awareness of using the pinyin system to match certain characters. It was her former experience that supported her to use the above strategies fluently without missing directions.

Ex. 38

Background information: *司 马* /sī mǎ/ means "mother died" in English. It is the Trans-semioticising equivalent of the Chinese word *死妈* /sǐ mā/. The character *死* means "death" and *妈* means "mother". This expression is a vulgar word that curses other's mother to die. The word *司马* is a surname in Chinese, which is used as a homophone of *死妈*.

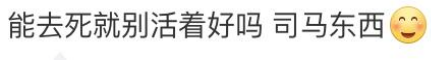


Image 30 Example of *司马*

[**Translation**: If you can die, please don't stay alive, okay? *司马* thing.]

Transcript extract of Participant M's think-aloud protocol:

这个是死妈？你想骂一个人，然后就会骂得很狠，然后就会说死妈了。对，这个很难听，然后我第一次看到这个词应该也是在微博，因为微博上面有非常多这种谐音或者这个类型的词。然后看到我自己应该就联想到了死，或者是你知道微博它是有那种转发评论的，然后他不是每一个人都会显示他的转发的时候能看到转发的人，它显示的评论，然后有些人他是没有打死妈，他就直接打成了司马？

Is this one *死妈*? If you want to curse someone badly, you will say *死妈*. Yeah, this word sounds very bad. And the first time I saw this word should also be on Weibo, because there are many similar expressions that use homophones. And I should think about the character of *死* as soon as I saw this expression. Or you know on Weibo you can forward the comment, and for everyone who forwarded the Weibo, his or her nickname would be shown along with the comment. Then some people didn't type *死妈*. They typed *司马* instead.

This participant claimed that she had found that many expressions used homophones to create new expressions. She accumulated a lot of experience in understanding TSEs that use homophones to replace the original characters. Although when she understood the TSE in the think-aloud task, she already knew the meaning of the TSE, she shared her understanding experience about the first time she met it. She used the experience of using homophones to identify the meaning to help her figure the

meaning out.

5.3.2.3 When the participants had no knowledge about the TSE

When the participants had no knowledge about the TSE, they would opt to take Path B or Path C. Every time they met the TSE, it would be a brand new experience of understanding a TSE. In the progress of understanding unfamiliar TSEs, participants might use their background knowledge to support their understanding after they found some clues to decode the TSE. Here are three examples where participants met the TSEs for the first time that can illustrate the three different possibilities under the circumstances.

5.3.2.3.1 Giving up understanding

Ex. 39

Background information: Refer to Ex. 10 of this chapter for details.

.

Weibo screenshot of the TSE

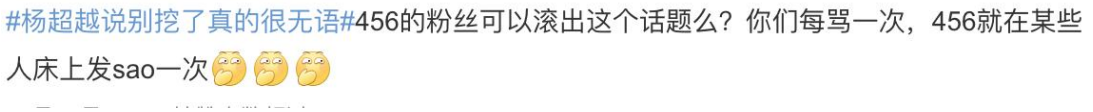


Image 16 Example of *456*

[**Translation**: #Yang Chaoyue said please don't dig, feeling very speechless# *456*'s fans can you please leave this topic? Every time you curse, *456* will be in heat on someone's bed.]

Transcript extract of Participant K's think-aloud protocol:

456，456 什么意思我真不知道呢。

456, 456. What does it mean? I really don't know.

In this example, the participant quickly gave up understanding this TSE after reading this TSE twice. He did not devote much effort to understanding this TSE as he was not familiar with this TSE.

5.3.2.3.2 Referring to external assistance

Ex. 40

Background information: *u1s1* means "frankly speaking" in English. It is the TSE variant of a Chinese expression *有一说一*, which literally means "if you have one thing to talk about, then talk about this thing". This TSE is a relatively complex expression that involves three different

semiotic systems when it is created. The formation mechanism behind this TSE is using acronyms to express a longer expression with many different strategies followed. “u” is the English letter, which is a homophone of the Chinese character *有* as “u” /ju:/ and *有* /you/ are pronounced similarly to each other. In English social media, there are also many cases of using letter “u” to replace “you”. The numeral “1” represents the character 一, which means “one” in English. “1” is the Arabic form of the Chinese character 一. The letter “s” cannot be considered as an English letter. Indeed, it is the pinyin initial of the Chinese character *说* /shuo/. By using the English letter, numeral, and pinyin initial letter, *u1s1* successfully shortens the Chinese expression *有一说一*, which instead needs to type “you yi shuo yi” on typing software.

Image 31 Example of *u1s1*

[**Translation**: *u1s1* net literature on Zhihu tell super sweet love stories.]

Transcript extract of Participant M’s think-aloud protocol:

说实话我一开始看到这种词的时候，我是不知道是什么，于是我就百度了一下，大家就是有一说一的谐音，u 代表有这个是完全谐音的，S 应该代表了说首字母拼音的，然后一就代表一，所以我百度来的。还有包括有很多这种类似的什么阴阳怪气真情实感也都是我百度来的。

To tell the truth, when I first saw this expression, I did not know what it was. So I used Baidu to search for it. People said that it was the homophone of *有一说一*. “u” represented *有*, which was totally a homophone. “s” should represent the pinyin initial of *说*. And then “1” referred to 一. So I used Baidu to search for its meaning. There are also many expressions like this, such as “yygq” and “zqsg”. Both of these were understood by searching for answers on Baidu.

In this participant’s transcript, she mentioned that when she had no clue about TSEs in this form, she used Baidu, a Chinese search engine, to figure out the meaning. Through the search engine, she fully understood the meaning of the TSE and at the same time understood how each element in this TSE was transferred. She also mentioned that for similar TSEs like “yygq” and “zqsg”, she also used path B to help her understand their meanings.

5.3.2.3.3 Using different strategies to understand the TSE

Ex. 41

Background information: The TSE, *桃浦* refers to the English word, “top”, which means being the top. Literally, *桃浦* means peach beach, which is a place name in Shanghai. However, this place name is not very famous in China, especially for people who are not familiar with Shanghai. The pronunciation of *桃浦* is the reason why it can be used to express the English word, “top” because *桃浦* /tao pu/ is pronounced similarly to “top” /ˈtɒp/. Twice is a South Korean girl group, which used to rank first among all K-pop girl groups.

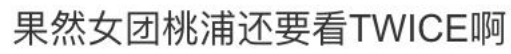


Image 32 Example of *桃浦*

[**Translation**: Sure enough, the top girl group has to be TWICE.]

Transcript extract of Participant T's think-aloud protocol:

果然女团桃浦，桃，浦。Ok，这个我第一次看到，这个词之前确实没有见到过，但是我读下来之后大概可以理解了，桃浦应该是英文 Top 这个音译过来的一个写法。之所以会想到 Top，就是说首先女团什么还要看 twice，这个 twice 应该是一个女团的名称，然后女团一般大家就会去有一些比较或者是怎么样，觉得 Top 在这边好像就挺合适。后来发现 Top 和桃浦的音译是非常相近的，应该就是顶级的意思。

Sure enough, girl group 桃浦, 桃,浦. OK, it's the first time that I saw this. I did not see this word before. But I could barely understand its meaning when I finished reading. 桃浦 should be an expression transferred from the English word "top". The reason why I thought about "top" was that (the text) first mentioned girl group has to look at twice. The TWICE should be a name of a girl group. Then fans would love to compare with each other or something. (So I) thought "top" was very suitable here. After that, I realised that "top" and 桃浦 are pronounced similarly to each other. So 桃浦 should mean being top.

In this example, the participant had not met this TSE before he did the think-aloud task. As he had no direct knowledge about *桃浦*, he started his understanding without any additional assistance. He first read the text and the TSE aloud and noticed the similarity in pronunciation between "top" and *桃浦*. Meanwhile, he activated his background knowledge about girl groups which was not directly about the TSE. He used this information about girl group to help him identify the corresponding English word that is pronounced similarly to *桃浦*. When associating possible words that shared similar pronunciation with *桃浦*, there might be other choices apart from "top" as the corresponding words were not limited in English. Therefore, the participant had to find a suitable word from all languages he knew. The strategies of analysing the TSE's pronunciation and referring to the context together helped the participant understand this TSE, which he met it first time.

Thus, familiarity is a significant factor that might influence the understanding paths. For TSEs that were familiar to the participants, they could either directly refer to their mental lexicon to find the meaning or make use of their former processing experience to simplify the difficulty of understanding new TSEs. For TSEs that they

were not familiar with, they had to devote more effort to the understanding process. Some participants chose to seek external assistance to find a shortcut, while others decided to lean on themselves even though they sometimes failed to understand the TSEs at last. Even for the same TSE, the familiarity with the TSE could exert some influence on the understanding paths of the participants. Taking the example of *u1s1*, I will discuss how familiarity influenced different participants' choice of understanding paths for the same TSE.

Ex. 42

Background information: Refer to Ex. 40 above for details.

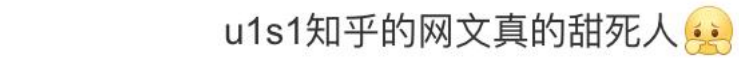


Image 31 Example of *u1s1*

Transcript extract of Participant E's think-aloud protocol:

首先它是 4 个部分。 u1，我也不知道就是因为因为 1 带 1 的 4 字词语。还是就那么几个，然后有一说一我知道了，因为知乎的网文真的舔死人，这是一个很完整的句子结构，然后前面要再加一个什么东西的话，会让我比较能直接联想到有一说一一句可有可无的。

First of all, it was composed of four parts. *U1*. I don't know. Because, because, words in the form of "+1+1" were limited. So I knew it was *有一说一* as the sentence "net literature was so sweet" was a fully structured sentence. If you want to add some words in front of this sentence, I would think of *有一说一* directly, which was not a necessary sentence.

Transcript extract of Participant I's think-aloud protocol:

有一说一，网络热词

有一说一, popular word on social media.

Transcript extract of Participant M's think-aloud protocol:

说实话我一开始看到这种词的时候，我是不知道是什么，于是我就百度了一下，大家就是有一说一的谐音，u 代表有这个是完全谐音的，S 应该代表了说首字母拼音的，然后一就代表一，所以我百度来的。还有包括有很多这种类似的什么阴阳怪气真情实感也都是我百度来的。

To tell the truth, when I first saw this expression, I did not know what it was. So I used Baidu to search for it. People said that it was the homophone of *有一说一*. "u" represented *有*, which was totally a homophone. "s" should represent the pinyin initial of *说*. And then "1" referred to 一. So I used Baidu to search for its meaning. There are also many expressions like this, such as "yygq" and "zqsg". Both of these were understood by searching for answers on Baidu.

Transcript extract of Participant P's think-aloud protocol:

其实就是我一开始当初见到它的时候，第一时间真的没有反应过来，但我是在因为看有一些人有一段时间"有一说一"这个词很火嘛，然后有的在他们讨论段里面可能会存在 u1s1 和有一说一混用的情况，当你同时见到这两个词的时候，你就会很容易的产生联系，把这两次联系在一起。

In fact, when I first met this TSE, I really failed to respond to it. But I finally knew its meaning because this expression was very popular and used by some people. And when they were chatting, they might use both *u1s1* and *有一说一*. When you saw these two words together, you could

build a connection between them.

The four participants (E, I, M, P) described four different understanding paths due to their different levels of familiarity towards *u1s1*. For Participant E, he did not know the exact meaning of this TSE at first. However, he had the information of a similar expression with the same structure, which could be written as "A1B1". He associated Chinese expressions that were in the form of "A1B1" and analysed the sentence structure to narrow down the range of possible answers. For Participant I, she knew the exact meaning of *u1s1*. Therefore, she simply output the meaning and added that it was popular on social media. For Participants M and P, neither of them know the meaning of the TSE at first. They chose two different paths to figure out its meaning. Participant M, as introduced above, reached out to use search engines to find explanations for TSEs that she could not understand. On the contrary, Participant P used his own way to understand the meaning of *u1s1* by observing how other people used this expression and noticed that the TSE was used to replace *有一说一* in real context.

Figure 5.8 describes how these four participants understood the TSE when they had different levels of familiarity with *u1s1*.

When the participants were analysing the TSEs, they were found to apply several associative mechanisms to support their comprehension, which included homophonic association, iconic association, acronymy-based mechanism, and shape-meaning association.

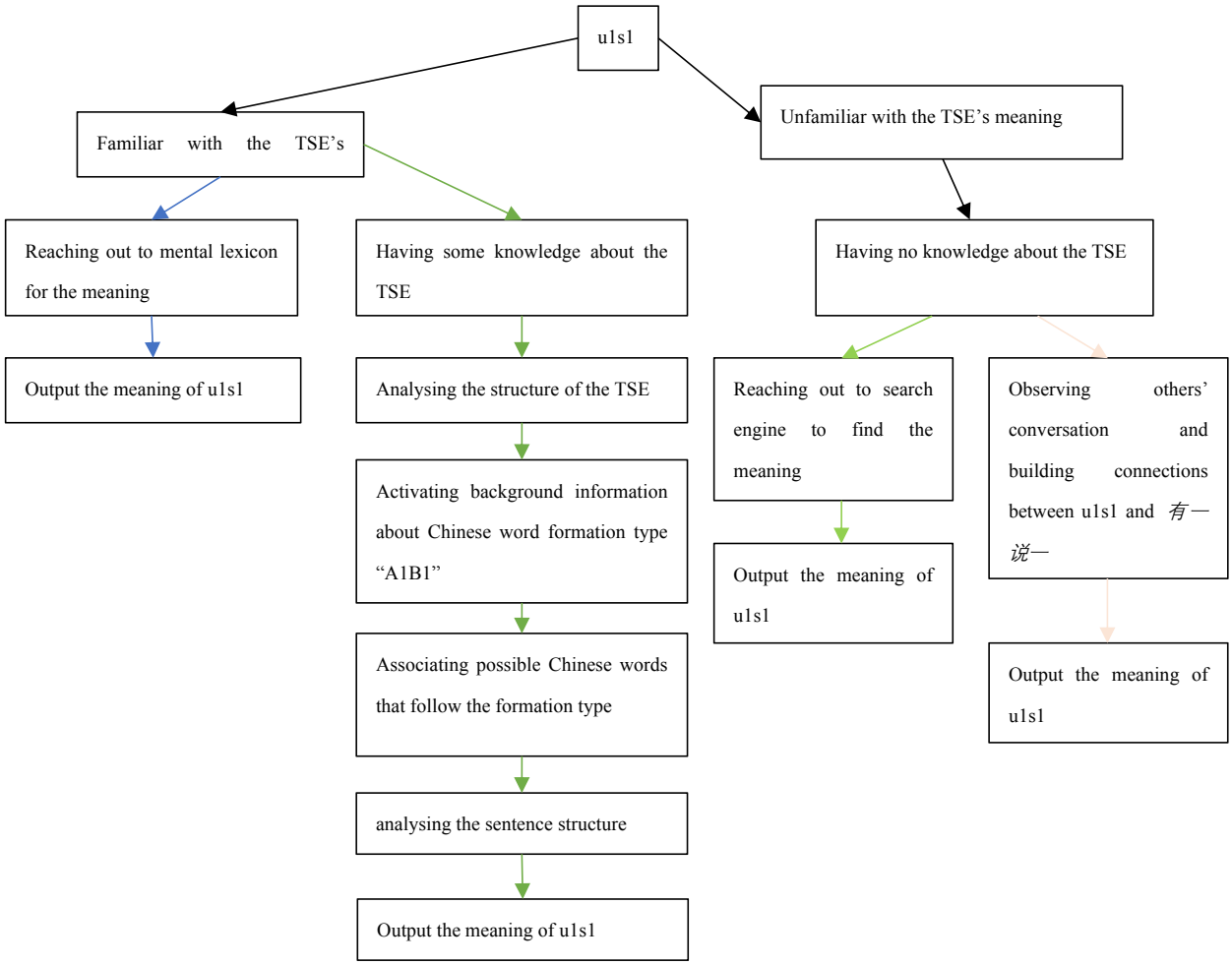


Figure 5.8 Routes for understanding *u1s1* in different familiarity conditions

5.3.3 Associative mechanisms for understanding TSEs

There are four types of associative mechanisms identified in the participants' understanding processes, which are homophonic association, iconic association, acronym-based association, and shape-meaning association. These associations illustrate how participants make association with different information achieved from the decoding processes.

5.3.3.1 Homophonic association

Homophonic association or sound association often occurs in Chinese. For example,

in a study focusing on Chinese two-part allegoric saying, researchers argue that homophonic associations are "realised based on a conceptual connection between the two homophonic expressions: the second part of the sayings and the expression of the idiomatic meaning" (MA *et al*, 2019, p.1). In the TSE understanding process, readers have less contextual support like two-part allegoric sayings. They have to draw association based on the pronunciation of the TSE or some elements of the TSE and then build a conceptual connection with different possibilities. When participants came across a TSE, they might associate it with an expression whose pronunciation was similar or identical to that of the TSE. In the data, the participants analysed the TSE's pronunciation under two different circumstances.

First, analysing the TSE's pronunciation happened when the reader could not find a suitable understanding that could make sense in the context based on the TSE's literal pronunciation. Then the reader would seek to analyse its pronunciation to get more information. Let us see how Participant C understood the TSE of *芜湖 nese*, where the participant used this strategy to analyse *nese*.

Ex. 43

Background information: *芜湖 nese* or *Wuhunese* refers to the people from Wuhu city. Its formation mimics English words like Chinese, Cantonese, and Japanese, which are composed of the country's or region's name and a suffix "-nese". These words as nouns mean people from that country or region. This Weibo post used a dash mark to emphasise that it was a Wuhu's citizen that expressed his or her love for Nanjing. 芜湖, Wuhu, is a city of Anhui Province, China. Nanjing is also a Chinese city, which is the capital city of Jiangsu Province. Geographically, Wuhu is a smaller city near Nanjing. It only takes less than half an hour to travel between Nanjing and Wuhu by speed train. Wuhu, therefore, has much economic and cultural communication with Nanjing. People who live in Nanjing may take a weekend trip to Wuhu, while Wuhu's people have a relatively strong intention to find a job or even migrate to Nanjing. For Wuhu's people, many of them have a positive feeling towards Nanjing. That is why the author of this Weibo post said that he/she loved Nanjing.

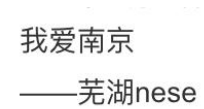


Image 33 Example of *芜湖 nese*

[**Translation**: I love Nanjing.
------Wuhunese]

Transcript extract of Participant C's think-aloud protocol:
芜湖 nese。好吧？可能就是网友前后鼻音不太分，可能就是芜湖垃圾的意思，好。
芜湖 nese. Alright? Maybe it is because the netizens cannot distinguish the front and back nasal. Maybe it means Wuhu is rubbish. Yeah.

As we can see from the transcript, the reader came up with a totally different understanding from the original meaning of this TSE. However, he provided a reasonable explanation based on his analysis of the TSE's pronunciation. The participant mentioned the notion of front and back nasal, which often brought trouble to people who could not speak proper Mandarin. In the understanding process, the participant did not spend effort on the two Chinese characters, *芜湖*. In contrast, he analysed the pronunciation of *nese*. In English, *nese* is usually pronounced as /ni:z/. However, if we pronounce *nese* following the rules of pinyin, it sounds like /nə sə/, which is totally different from English. Then the participant took front and back nasal mistakes into consideration. The consonants /n/ and /l/ are challenging to people from southern parts of China as many of their dialects do not distinguish between these two consonants, with all /n/s pronounced as /l/s. Therefore, in this case, *nese* could be pronounced as /lə sə/. After that, the reader associated the pronunciation of /lə sə/ with a Cantonese word *垃圾*, which means "rubbish" in English. For some people from the southern part of China, they might not be able to pronounce /l/ and /n/ separately. In a study on the classification of Weibo homophones, the researcher mentioned that when using *桑心* (sang xin) to replace *伤心* (shang xin) (meaning "sad"), there may not be a differentiation between these pronunciations as they have the same place or manner of articulation. /n/ and /l/, /f/ and /x/ are not differentiated in Xiang dialect, which is often used by people from Hunan Province (Tang, 2014).

Second, the participants may establish homophonic association across languages and numerals in some cases of TSEs. Let's take Participant T for example.

Ex. 44
Background information: *077* is a nickname created by Chinese NBA fans for Luka Dončić, a Slovenian NBA player. *077* is the Trans-semioticising equivalent of *东契奇* /dong qi qi/, the

Chinese translation of Dončić. *77* (七七) is pronounced as /qi qi/, which is the same as *契奇* /qi qi/. People use the homophones to replace the original Chinese characters. Using *0* to represent *东* is also related to homophonic mechanism. However, *0* is pronounced as /ling/, which is not the same as /dong/. In fact, it is because the numeral *0* looks like a round hole. Hole in Chinese is written as *洞* /dong/. Then since *洞* and *东* are homophones, people use *0* to express the character *东*.

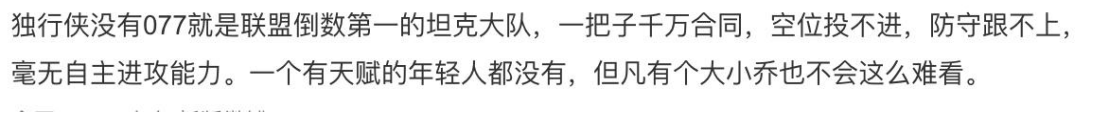


Image 34 Example of *077*

[Translation: Dallas Mavericks, being a team of many tanks, can only rank the last in the league without *077*. They signed so many ten-million-level contracts. However, they cannot score when they shoot without defense and their own defense can never follow the opponents. They barely have talent. No talented young player can be found in this team. It won't be that bad whenever they have senior or junior Jordan.]

Transcript extract of Participant T's think-aloud protocol:

这个我最开始第一次看到这个词的时候，我说实话我没有理解这个词的意思，为什么？因为我一开始想的是东契奇的球衣号码和七有关，所以有没有可能是用 077 来代表他这个人，但是通过球衣号码这种形式，但后来想一下好像也不太对劲，为什么要加个 0？后来想了一下 0 有洞的说法，因为它像一个洞，那么独行独行侠里边球星叫东契奇，东契奇和洞七七谐音非常的相近，所以后来就想这个 077 应该就是东契奇的谐音表示相当于是一个外号。对，中国人给他的一个外号。

I failed to understand the meaning of this TSE when I first saw this word, to be honest. Why? Because I thought Dončić's jersey number is related to seven so maybe it is possible that people use *077* to refer to this person but in the way of jersey number. However, I thought of this afterwards. It did not make sense, because there was a *0*. After that, I thought *0* can be called *洞*, since it looked like a hole. Then the star in Mavericks called *东契奇*. *东契奇* and *洞七七* (077) are pronounced very similarly. They are homophones. So I thought that this *077* should be the homophone of *东契奇*, which was used as a nickname. Yeah. It was a nickname given to him by Chinese people.

In this case, the participant tried different ways to analyse the TSE of *077*. He first referred to his background knowledge of Dončić. He noticed that the jersey's number might be a clue to explain the meaning of *77*. At this point, Participant T split this TSE into two parts, which were *0* and *77*. However, he found it incorrect as *0* could not make sense in this expression. Then he began to analyse the pattern of the TSE, as he mentioned that *0* looked like a hole, which could be pronounced as *洞* /dong/. With *0* transferred to the pronunciation of "dong", Participant T analysed the pronunciation of the TSE and realised that "dong qi qi" was the homophone of the

NBA player *东契奇*, Dončić. Building a connection between numerals and Chinese characters based on their similar pronunciation is also noticed by a previous researcher. In Tang (2014), expressions like “88”, “555”, and “1314” are categorised as numeric letter homophones.

5.3.3.2 Iconic association

In cases that involved the use of emojis, the participants might activate the association between the meaning of the icons or stickers and that of corresponding characters. In this study, when icons or stickers are used in TSEs, the formation process usually involve mapping the pronunciation of the names of the iconic creatures or objects and that of some characters. The original meaning of the icons or stickers is seldom directly used in forming TSEs. Consider Ex. 45:

Ex. 45

Background information: 🐿️🐸 means “Okay” in English. This TSE is composed of two emojis, which represent two different Chinese characters. 🐿️ is an emoji of a monkey, which is called *猴* /hou/ in Chinese. 🐸 refers to frog, which is written as *蛙* /wa/ in Chinese. 猴 and 蛙 together can only mean monkey and frog. Indeed, these two characters are homophones of two other Chinese characters. *猴* is the homophone of *好* in Cantonese, which is pronounced as /hou3/. *蛙* is the homophone of *哇* /wa/, which is an interjection that expresses a feeling of excitement. *好哇* means “Okay” in English. In the Weibo example, the writer was talking about the weather, saying that the weather was very good. So in this case, the writer used 🐿️ to replace *好* in Chinese and 🐸 to replace *哇*.

今天天气真🐿️🐸，准备出去玩啦

Image 35 Example of 🐿️🐸

[**Translation**: Today’s weather is very 🐿️🐸 good, ready to hang out]

Transcript extract of Participant G’s think-aloud protocol:

天气今天天气真🐿️，这个应该也是从广东话那边什么有个好 hou 唉什么,雷猴就是这种，对，然后就把好用猴来表达，然后蛙就是用青蛙的蛙来表达，也是为了表现出来这个音同音不同意，ok。

Weather, today’s weather is very 🐿️. This should also come from Cantonese. They have a saying called *hao* or *hou ei* or something. *雷猴* /lei hou/ is like this. Yeah. Then they use *猴* to express *好*, and 🐸 *to* refer to *蛙* in 青蛙. This is also used to express different meanings with the same pronunciation. OK.

As the shadowed lines show, Participant G clearly mentioned how she associated the

meaning of emojis with that of corresponding Chinese characters. She first read the text and activated her background knowledge of Cantonese, which helped her build a connection between the emoji of monkey and the character of *好*. In this understanding process, the original meaning of emojis was used at the first step of decoding. The participant read the emojis and then further associated the characters of monkey and frog with other homophones.

Ex. 46

Background information: Refer to Ex. 3 of this chapter for detail.

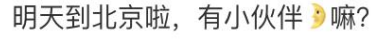


Image 12 Example of *嘛*

[**Translation**: Will arrive at Beijing tomorrow. Any buddy wants to ?]

Transcript extract of Participant T's think-aloud protocol:

<A>明天到北京有小伙伴月亮吗？我是这样理解，这个我是第一次看的，首先<C3>第一反应我看到的这个符号是个月亮的符号，我在想是不是和夜晚和深夜比较有关系，<D>但是好像发现用深夜夜晚这种思维往里边代入有点不太通顺，后来发现应该就是这个表示"约"的意思，<C1>因为"月亮"和"约"是读音上很像的，因为这句话他说明天到北京了，有小伙伴什么吗？<E1>那么大家经常会说有没有人约一下或者是聚一聚这种用约来表示这种聚会的状态，那么月亮应该就是一个谐，也是一种谐音的方式，去代表约这个意思，可以这样。

<A> Will arrive at Beijing tomorrow. Any buddy wants to? /yue/ (moon)? I understand it in this way. It is the first time I met this (expression). <C3>First, at first glance, I saw this emoji, which was a symbol of moon. I was thinking whether it had stronger relationship with evening and night. <D>However, (I realised that) it seemed that the thought of midnight and evening could not make sense in the context. Then I found that the moon emoji should convey the meaning of *约* <C1>as it sounded alike *月* in *月亮* from the perspective of pronunciation. <D>Because he said that he would arrive at Beijing tomorrow, are there any little friends (to do) what(to do) what? <E1>Then people usually say anyone wants to meet or gather around. People use *约* to express the status of holding a party. So the moon's emoji should be a homophone. It's also a way to use homophone to refer to the meaning of *约*. It can work in this way.

In this example, Participant T dealt with the emoji of the moon when he tackled the TSE. Although he finally turned to the pronunciation of the character *月* to understand the TSE's meaning, he tried to analyse the TSE by transferring the emoji's iconic meaning. Based on the presumption of the moon, he associated it with the concept of late night. The meaning of the emoji could not only be transferred to

particular characters like “约” in Ex. 46, but could also to some abstract concepts.

5.3.3.3 Acronymy-based association

Many TSEs identified in this study were in the form of acronyms. Although these acronyms usually only contained Latin letters and sometimes numerals, it was not easy to reverse the processes of fully spelled expressions to initials and numerals. Participants needed to devote extra effort to select an interpretation among many possible ones. Let us take a look at how participants tackled such TSEs and integrated other information to understand them.

Ex. 47

Background information: Refer to Ex. 5 of this chapter for details.

Image 24 Example of *cpdd*

[**Translation**: cpdd@Wangzherongyao]

Transcript extract of Participant T’s think-aloud protocol:

Cpdd 艾特王者，cpdd，cpdd 这个的是第一次看到 cpdd，应该是 4 个汉字的首字母缩写了。Cpdd, Cp，DD 不太能想得到，因为 CP 我一开始想有没有可能和裁判相关，但是好像结合 DD 很难想到 DD 是什么意思。DD 的话 DD 好像不太像的。DD 要么还有一种弟弟的意思在里面，但是弟弟的意思在里边，CP 又是什么意思？Cpcp 要么是 couple 的意思，couple 会不会接弟弟也不太知道，或者就是说他可能跟王者荣耀相关性更强一点，可能有什么战队的名义适合能匹配上也是有可能，但是也不太了解，也只能想要这么多。

“cpdd” @ Wangzhe, cpdd, it was my first time seeing cpdd. It should be the pinyin initials of four Chinese characters. Cpdd, cp, dd. Cannot come up with ideas. Because I first thought about whether “cp” was related to *裁判* /cai pan/. However, when considering “DD” together, it was difficult to think of the meaning of “dd”. For “dd”, “dd” does not look like… Or there is another meaning of *弟弟* /di di/ about “dd”. But with the meaning of “little brother” in it, what does “cp” mean. “cp”, “cp” either means “couple”. Not sure whether couple can be used next to *弟弟*. Or it is more related to Wangzherongyao. Maybe it is the name for an e-gaming club. But I am not familiar with that. That’s all I can think of.

“Cpdd” was an interesting TSE that involved both Chinese and English words’ acronyms. As introduced in the background information, “cp” and “dd” followed different ways of shortening. In the understanding processes, Participant T realised that this TSE was probably formed by pinyin initials at the beginning. Then he began to look for a Chinese word that had two characters and its pinyin initials were “cp”.

He thought about two possible answers like *裁判* and *靠谱* /kao pu/. However, this trial ran into a dead end. Therefore, he tried to find a way out from the latter two letters, "dd". Comparing with "cp", "dd" was relatively easier to find a match as there were relatively few Chinese words whose pinyin initials were "dd". Participant T came up with a thought that "dd" might refer to *弟弟* as this was sometimes used by netizens when communicating on social media. However, *弟弟* was not the right choice, as it did not lead to a meaningful interpretation. In this example, Participant T failed to output an interpretation that made sense in the context. However, he still attempted to use the associative mechanism of acronymy to find an answer.

Ex. 48

Background information: Refer to Ex. 4 of this chapter for details.

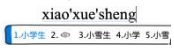


Image 6 Screenshot 1 of the typing software of *xxs*

When typing the full pinyin of the word, the typing system will provide all possible words that share the same pinyin combination in the waiting list. The more times the user types the pinyin, the more precisely the typing system will provide the waiting list.

Image 7 Screenshot 2 of the typing software of *xxs*

As shown in the typing system's screenshot, the waiting list provides four Chinese words whose initials are *xxs*. Apart from *小学生*, *小小思* /**x**iao **x**iao **s**i/, *休息室* /**x**iu **x**i **s**hi/, and *学习上* /**x**ue **x**i **s**hang/ are all words whose characters' pinyin initials are *xxs*. If the user does not want to choose any of these provided words and only keeps the pinyin letters, *xxs*, he or she can press the "Enter" button on the keyboard, while otherwise he or she needs to use the number buttons of keyboard to select corresponding words or use the mouse to click on the target word. Compared with spending additional effort on pressing number button or using the mouse to click, pressing "Enter" button is relatively faster. It has to be mentioned that besides saving efforts, people may also deliberately use initials instead of original words due to diverse motivations such as avoiding mentioning the exact name of the person they want to gossip about, or avoiding typing sensitive words that are rude or abandoned on social media, etc.

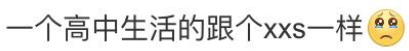


Image 11 Example of *xxs*

[**Translation**: A high school student lives like a *xxs*.]

Transcript extract of Participant B's think-aloud protocol

一个高中生活的跟个 xiao xiao si，小学生一样，对。那就是小学生。是因为一般都是简写，然后他提了高中，所以就肯定是跟高中对比的，要么就是成年人，要么就是小学生。

A high school student lives like a xiao xiao si, 小学生. Yes. This is 小学生. (This is) because usually it is an acronym. And he mentioned high school. So it must be a comparison to high school, (which is) either an adult or primary school student.

When understanding the TSE, Participant B started with reading the text from the first character to the end. The TSE appeared at the end of the sentence with two characters left. When he read the TSE, he did not just read aloud the pinyin letters *xxs* following the pinyin pronunciation like /xi xi si/, or English pronunciation as /eks eks es/. Instead, he tried to transfer the TSE to corresponding Chinese characters for understanding. As shown in the transcript, he read *x* as /xiao/, which meant that he noticed that it was a Chinese pinyin initial of certain character whose consonant is /xi/. He also mentioned that "usually it is an acronym", which proved that he considered *xxs* was following the Chinese pinyin initials' rule to make sense. For *xxs*, he first transferred it to /xiao xiao si/, which exactly matched the 小小思 /xiao xiao si/ mentioned in the background information. Then he read the TSE as /xiao xiao si/ with short pauses between each character and extended the pronunciation time of each character. He did not just cease at this point but continued reading the whole text aloud by still reading 一样 aloud. After he completed reading the whole text, he directly came up with his understanding of the TSE, which means "primary school student". This proved that during the process of reading *xxs* as /xiao xiao si/ slowly, he was trying to figure out the exactly corresponding characters as he noticed that "xiao xiao si" could not make sense in this context. In this case, although the participant referred to his processing history and realised that this was an acronym, he used the strategy of resorting to the text to find the right context, which helped him decide which corresponding characters were correct.

5.3.3.4 Shape-meaning association

Another type of association activated in TSE understanding occurred between the TSEs' characters' shape or font and the corresponding meaning of the TSE. This is a unique situation that usually happened when the TSE contained Chinese characters

which have been modified from the perspective of its shape. Due to the typing system limitation, unless the writer of the TSE uses stickers of drawn pictures, people usually cannot type a character that does not exist in Chinese. For example, Li (2021) mentioned a TSE which was combined by the Chinese character 牛 and the English letter *B* to represent the meaning of *牛逼* (niu bi) in Chinese. However, this type of TSE cannot be directly typed on social media. More commonly, people will use another character which has a similar shape to the original one to convey more meaning and extra pragmatic effects. The case of *𠙵𠙵* can provide a clear explanation of how people deal with this type of TSE through shape-meaning association.

Ex. 49

Background information: Refer to Ex. 13 of this chapter for details.

Image 20 Example of *𠙵𠙵*

[**Translation**: Take a look. Who is the real *𠙵𠙵*?]

Transcript extract of Participant M's think-aloud protocol:

我知道凡凡他有这个，嗯，因为他名字叫吴亦凡，我有见过这个字，但我不知道这个字念什么。我看过这个字，然后因为凡它中间不是有个点，然后这个点就经常大家会可能用中国的汉字各种改造。怎么改造，我现在具体也忘了，但是它放在头上我不知道什么意思，但是我知道他是凡凡的意思，但是肯定就是讽刺他这两个字不一样，讽刺他的某一些特质，但具体什么特质我也看不太出来。

I know that *凡凡* his has this (nick name). Yeah. Because his name is Wu Yifan. I have seen this character before but I don't know how to pronounce it. I have seen this character. And because it has a dot in the middle, does it? So this dot usually is modified by people on the basis of different Chinese characters. I forgot how to modify it in detail now. But I have no idea why it is put on the head. While I do know it refers to fanfan. Anyway, it must be used to satirise him and some of his propensities. But I cannot tell what propensities (this TSE satirises).

In this example, Participant M used the strategy of analysing the TSE's shape to understand the TSE. She could not pronounce the characters of 几几. However, due to the similarity of the shape between 凡凡 and 几几 and the contextual support, she associated these two characters with each other. She also had background knowledge to assist her association as she knew other Chinese characters that experienced transformation before.

It is to be noted that the understanding of some TSEs may involve the operation of more than one type of association. Consider Ex. 50:

Ex. 50

Background information: Refer to Ex. 6 of this chapter for details.

Image 12 Example of 嘛

[**Translation**: Will arrive at Beijing tomorrow. Any buddy wants to ?]

Transcript extract of Participant T's think-aloud protocol:

明天到北京有小伙伴月亮吗？我是这样理解，这个我是第一次看的，首先第一反应我看到的这个符号是个月亮的符号，我在想是不是和夜晚和深夜比较有关系，但是好像发现用深夜夜晚这种思维往里边代入有点不太通顺，后来发现应该就是这个表示“约”的意思，因为月亮和约是读音上很像的，因为这句话他说明天到北京了，有小伙伴什么吗？那么大家经常会说有没有人约一下或者是聚一聚这种，用“约”来表示这种聚会的状态，那么月亮应该就是一个谐音，也是一种谐音的方式，去代表“约”这个意思，可以这样。

Will arrive at Beijing tomorrow. Any buddy moon? I understand it in this way. It is the first time I met this (expression). First, at first glance, I saw this emoji, which was a symbol of moon. I was thinking whether it had stronger relationship with evening and night. However, (I realised that) it seemed that the thought of midnight and evening could not make sense in the context. Then I found that the moon emoji should convey the meaning of 约 as it sounded alike 月 in 月亮 from the perspective of pronunciation. Because he said that he would arrive at Beijing tomorrow, is there any little friends (to do) what? Then people usually say anyone wants to meet or gather around. People use 约 to express the status of holding a party. So the moon's emoji should be a homophone. It's also a way to use homophone to refer to the meaning of 约. It can work in this way.

As we can see from this participant's understanding transcript, he tried two different types of association to derive the meaning of the TSE. He first attempted to find a way out from the actual meaning of the emoji, which was about the notion of evening and midnight, on the basis of iconic association. However, he found that this meaning was not sufficient because it could not make any sense in the context. Then he activated

his background knowledge about what people usually said when they wanted to invite friends to meet. He then built a connection between the two homophones, 月 and 约, which finally helped him understand the TSE. This case illustrates two main destinations readers may choose when transferring the meaning of symbols like emoji. Only when the meaning of the emoji is transferred in the right way can the reader further understand the TSE. Then with other strategies working together, the reader can finally obtain a relatively reasonable understanding of the TSE.

5.4 Summary

In this chapter, Chinese readers' paths, strategies and routes of understanding TSEs on social media were generalised on the basis of the participants' protocols.

According to the data, three main paths were identified, namely direct access, access via external assistance, and access via using comprehension strategies. The first path was relatively simple, which usually happened when the readers were very familiar with the TSEs concerned and therefore could refer to their mental lexicon to get the meaning of the TSEs. The processing of such familiar TSEs is similar to that of regular mono-lingual familiar words which are interpreted literally. It suggests that readers may store the TSEs in their mental lexicon in the same way as they store regular mono-lingual words and that different languages are not stored separately but intertwined in our mind. It might be the case that we store all meaning-form entities including mono-lingual expressions and TSEs in the same warehouse of our mental lexicon. This speculation is consistent with the central argument of Translanguaging theory that there is no boundary between named languages or semiotic systems (Li, 2016a; Li & Zhu, 2019). The second path occurred when the readers were not (very) familiar with the TSEs concerned and yet could search them online. This approach, not much mentioned in related literature, is very much characteristic of online reading, as online search of unknown expressions including TSEs via the computer or the mobile phone is easy and convenient. It shows that when a reader seeks to understand

a TSE on social media, he or she can not only use the information provided from the context, the TSE itself, and the reader's own background knowledge, but also external assistance when all the support above does not work. Apart from looking up the TSE's meaning on search engines and using typing software, there might also be other possible external assistance left to be identified, though, such as consulting people around. Path C happened when the participants were not familiar at all with the TSEs that appeared in the Weibo texts. Yet, despite a lack of familiarity, the participants might benefit from their knowledge about or experience of processing a TSE that looks similar to the current TSE in question.

By analysing the strategies used by the participants who took Path C to decode the TSEs, eight different strategies were found used by the participants, namely resorting to the text, judging from the topic, analysing the TSE, referring to the context, activating background knowledge, associating the meaning of the TSE, referring to the processing history, and Guessing. It is likely that one may employ more than one strategy in tackling the meaning of some TSEs. This not only suggests the potential complexities of many TSEs, but also lends evidence for the great agency and flexibility on the part of the readers in tackling them. It also testifies that understanding TSEs often involve the interaction between the topic and text of the post, the socio-cultural context and our background knowledge.

It is also found that the comprehension of different TSEs may vary depending on post readers' familiarity with the TSEs as well as the formation types of TSEs. For the former, we get to know that For the latter, an interesting thing is that readers may avail themselves of their processing experience with similar TSEs in tackling novel ones. This suggests that language users keep accumulating not only declarative knowledge but also procedure knowledge in the course of learning and processing vocabulary, including TSEs. In addition, it was found that the understanding paths and usage of strategies varied with TSE types and readers' levels of familiarity with the TSEs. In each case, participants showed common and different understanding paths and used different strategies to support their understanding of TSEs.

It must be noted that the keyboard inputting has an important effect on readers' comprehension of the TSEs, as much as in their formation. Readers may just type the pinyin initials of certain TSEs and then arrive at expected interpretations, when the TSEs are already "conventionalised" into a ready mapping between the initials and their pronunciation.

It is worthy of mention that four important associative mechanisms supporting their comprehension of the TSEs were discovered from the participants' report, namely homophonic association, iconic association, acronymy-based association, and shape-meaning association. The finding suggests that association mechanisms play important roles in lexical innovation. In turn, tackling the meaning of innovative lexical items like TSEs also requires the resort to such association mechanisms.

When analysing the understanding paths, strategies and routes of TSE comprehension, I basically depended on the think-aloud protocols provided by the participants. Yet, why the participants took certain paths and routes and applied certain strategies and how the participants chose one meaning out of other possibilities await to be explained. Therefore, to address these questions, in the next chapter, Relevance Theory as a tool will be adopted to provide some plausible explanations on why and how the paths, routes and strategies were selected and relevant meanings of the TSEs were picked by the participants.

Chapter 6 TSE Understanding as a Relevance-Constrained Process

In this chapter, drawing on the data collected from the think-aloud protocol, I will seek to reveal that readers' TSE understanding testifies to the central argument of Relevance Theory that utterance comprehension is constrained by the presumption of optimal relevance. It will serve as a complement to Chapter 5 by exploring their decision making at different stages of comprehension in different situations and their cognitive rewards. Specifically, it will address questions that remain unanswered: 1) What constrains the participants' search for the (right) interpretation of the TSEs although other interpretations are potentially possible? 2) What lexical pragmatic processes are involved and relevance-constrained in analysing unfamiliar TSEs? and 3) What additional gains do the participants obtain in processing the TSEs as a reward for their extra cost entailed by the lexical pragmatic processes?

In Section 6.1, I will display the effect of the constraint of relevance on the lexico-semantic processing of the TSEs. Following up in Section 6.2, I will probe into how lexical pragmatic processes such as semantic narrowing and particularly semantic integration found at work in TSE understanding are relevance-constrained. Section 3 will address the additional gains that emerge from the lexical pragmatic processing of the TSEs in the absence of direct access. Discussions of the three issues will be supported by data from the think-aloud protocols and interview results. Finally, Section 6.4 will generalise the relevance-constrained framework for understanding TSEs on Chinese social media.

6.1 Relevance-constrained lexico-semantic processing in TSE comprehension

In this section, I will demonstrate how the lexico-semantic processing of TSEs via the choice of comprehension strategies lends evidence for the presumption of optimal relevance as the cognitive constraint on TSE comprehension. By lexico-semantic

processing, I refer to the decoding of the components of TSEs to generate their meaning. Many TSEs in my data, such as *酸Q*, *带jio布*, *456*, *桃浦*, *吓skr人*, *司马*, , and *九九*, can be understood simply by means of lexico-semantic processing because as long as decoding is correctly completed, they get correctly understood.

When the participants took Path 3, they would apply various strategies such as resorting to the text, judging from the topic, analysing the TSE, referring to the context, activating background knowledge, associating the meaning of the TSE, referring to processing history, and guessing in order to understand the TSEs, which they were unfamiliar with. As I demonstrate below, the analytical process, including the triggering of the various association mechanisms (i.e., homophonic association, iconic association, acronymy-based association, and shape-meaning association), was also constrained by the presumption of relevance. Consider Ex. 1.

Ex. 1

Background information: *xxs* is the Chinese pinyin initial combination of the word *小学生* /xiao xue sheng/, which means primary school student in English. Using Chinese pinyin initials to replace corresponding characters is often used by people on social media. This path to creating TSE thanks to the mechanism of typing system of simplified Chinese and the pinyin system. When typing a Chinese character on a cell phone or computer, one can only type the pinyin initial of the character and the typing system will automatically provide possible characters which share the same initial being typed in the waiting list for user to choose. It is the same when typing a word that is composed of few characters. In this example, let us use *小学生* to illustrate the mechanism. If someone wants to type *小学生* on his phone or computer, he or she can either type the full pinyin of *小学生* or type the pinyin initials of *小学生*.

Image 6 Screenshot 1 of the typing software of *xxs*

When typing the full pinyin of the word, the typing system will provide all possible words that share the same pinyin combination in the waiting list. The more the user types, the more precisely the typing system will provide the waiting list.

Image 7 Screenshot 2 of the typing software of *xxs*

As shown in the typing system's screenshot, the waiting list provides four Chinese words whose initials are *xxs*. Apart from *小学生*, *小小思* /**x**iao **x**iao **s**i/, *休息室* /**x**iu **x**i **s**hi/, and *学习上* /**x**ue **x**i **s**hang/ are all words whose characters' pinyin initials are *xxs*. If the user does not want to choose any of these provided words and only keeps the pinyin letters, *xxs*, he or she can press the "Enter" button on the keyboard, while otherwise he or she needs to use the number buttons of

keyboard to select corresponding words or use mouse to click on the target word. Comparing with spending additional effort on pressing number button or using mouse to click, pressing "Enter" button is relatively faster. It has to be mentioned that besides saving efforts, people may also deliberately use initials instead of original words due to diverse motivations such as avoiding mentioning the exact name of the person they want to gossip about, or avoiding typing sensitive words that are rude or abandoned on social media, etc.

Image 11 Example of *xxs*

[**Translation**: A high school student lives like a *xxs*.]

Transcript extract of Participant R's think-aloud protocol:

跟个 ss，xxs 一样。一个高中生活的。一个高中，生活的跟个，我操，跟个什么一样。跟个小傻瓜，小。就是我一开始思维肯定是想去拼一个词，跟个潇潇潇洒一样，跟个小傻逼一样。不对。一个高中生活的跟个，肯定感觉是一个贬义的词，一个高中生活肯定是，然后最后配一个 emoji，这个肯定是一个不好的词，大概类似那种跟个监狱。但是这又不是监狱，这是三个字母。三个字母的话就是跟个,一个高中，高中能生活的跟个什么？跟个，监管所，但这也不是监管所的谐音，这是一个 xxs 的谐音。跟个新，谢，受，跟个，跟一个高中生活的，跟个新鲜神，新鲜血，操想不到诶，一个高中生活的跟个，一个高中高中一般生活的就很枯燥，现在刷题对吧？可以往枯燥这种地方想。我感觉就有点像。监管所那种什么监狱一个高中生活的跟个。这个确实不知道，能过吗？

Like a "ss", "xxs". A high school student lives. A high school student, living like. Damn. Like a what. Like a *小傻瓜* /xiao sha gua/, *小* /xiao/. So firstly my mind must be thinking to spell a word. Like a *小小潇洒* /xiao xiao xiao sa/. Like a *小傻逼* /xiao sha bi/. No. A high school student living like a. It must be a derogatory word. A high school life must be, followed by an emoji at the end. This must be a passive word. Maybe like that in prison. But it's not prison (*监狱* /jian yu/). It has three letters. If (it has) three letters, like a high school. How will a high school life gonna to be? Like a supervision department (*监管所* /jian guan suo/). But it is not the homophone of *监管所*. It is a homophone of "xxs". Like a *新*, *谢*, *受*. Like a, like a high school, high school life. Like a *新鲜神* /xin xian shen/, *新鲜血* /xin xian xue/. Damn, I cannot think about its meaning. A high school lives like a. A high school, a life like in high school, which is very boring. So it's about taking quizzes, right? (Maybe I) can think about something boring. I think it's a little bit like supervision department or something. A high school lives like a. I really don't know about this. Can I skip it?

Participant R's think-aloud data provided us with a telling example about how the lexico-semantic processing was constrained by the presumption of relevance. The process of the constraint was actually composed of two steps, namely coming up with an assumption first and then referring to the context to check its relevance. In this example, Participant R repeatedly conducted this process five times. For the first time, the participant analysed the pronunciation of the TSE by reading it as "ss" and then as "xxs". He then referred to the Weibo context, mentioning "high school life", but he

failed to come up with any meaningful interpretation, which annoyed him. Then, he analysed the TSE again by reading the TSE as *小傻瓜* /xiao sha gua/. The second trial was based on the lexical context by considering possible words with the pinyin initials and their order. Obviously, this second trial also could not meet the expectation for optimal relevance. The participant shared his decoding strategy, which was to spell a word that could be written as "xxs". So he used the guessing strategy to interpret the TSE as *潇潇洒洒* /xiao xiao sa sa/, *小傻瓜* /xiao sha gua/. Yet he denied these words by saying "no", because the pinyin initials of the two expressions did not match each other. Then the participant returned to the context. Considering that "high school life" might relate to a passive attitude and there was a crying emoji, he believed that "xxs" should be a word that conveyed a passive meaning. Based on this assumption, he guessed that "xxs" might be *监狱* /jian yu/ ("jail"). However, the pinyin initials of *监狱* were not "xxs" as *监狱* had two characters but "xxs" had three. He believed that the interpretation should be the "homophone" of "xxs", which in effect meant that the participant realised that he needed to find a word whose pinyin initial was "xxs". Based on this assumption, Participant R then tried characters like *新* /xin/, *谢* /xie/, and *受* /shou/ as these characters' pinyin initials could be written as "xxs". However, when bringing these characters to the context, Participant R found that the combination of these characters still could not yield an interpretation that could meet the expectation for relevance. Feeling frustrated, Participant R considered the context for the last time, thinking that the TSE should be related to boring high school life. Finally, he read the sentence again to further understand the TSE. When he still could not get a suitable meaning for this TSE, he gave up understanding. Just as introduced in the background information, Chinese pinyin initials were the combination of the first letters of the characters' pinyin in a word or expression. However, when the process was reversed, pinyin initials could represent more than one word or expression. Therefore, when the reader decoded the pinyin initials, he had to activate his background knowledge about many words or expressions whose

initials were “xxs” and then consider the context to find an interpretation that was consistent with the presumption of relevance. Although Participant R failed to understand this TSE at last, his comprehension processes were constrained by the presumption of relevance. When no suitable meaning was found, his comprehension processes ceased. Now consider Ex. 2.

Ex. 2

Background information: This TSE is the Trans-semioticising equivalent of *看麻* in Chinese. *看麻了* literally means feeling numb when someone sees something. People can use it to express the status of speechless and shocked. *看*🐴 literally means it is time to watch a horse. It is composed of two Chinese characters and one emoji. *看* /kàn/ means to look at, to read, or to see something in Chinese. *了* /le/ is a character that functions as a symbol of the perfect tense in Chinese grammar. It describes the final state when something has been done. The emoji in this TSE is a horse which is running. Horse in Chinese is pronounced as /mǎ/, which is written as *马*. This character is a homophone of *麻* /má/. They share the same consonant and vowel but have different tones. People transfer the character *麻* to *马* and then use the emoji of *马*, horse, to replace the Chinese character. When typing “ma” in pinyin typing system, *马*, 🐴, and *麻* can all be listed in the waiting list (as shown in the screenshot) . It is uncertain why people use the emoji to replace the original Chinese character.

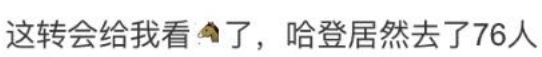


Image 23 Example of *看 了*

[**Translation**: This transfer made me *看*🐴*了*. To my surprise, James Harden went to the 76ers.]

Transcript extract of Participant M’s think-aloud protocol:

看🐴了，这个🐴它有很多种意思，它其实就是 mā má mǎ mà，反正我想的是，应该是，能碰得上，比如说这是看麻了，应该是,看妈了，看马了，看骂了，应该都不是，所以我会猜就是看麻了，但是我应该是知道这个意思，主要是靠意会中文很多都是靠意会。加上这种就那种 Emoji 的表情。我会猜它是看麻了。

看🐴*了*, this 🐴 has many different meanings. It actually should be mā má mǎ mà. Anyway, I think it should be able to match. For example, *看麻*/má/ *了*, *看妈*/mā/, *看马*/mǎ/ *了*,*看骂*/mà/. These should not be. So I would guess it must be *看麻*. But I should know the meaning. Basically Chinese is based on sense. Plus many expressions use the emoji. I would guess it is *看麻*.

In this example, the participant came across a TSE with a Chinese character and an emoji. Participant M first associated the emoji of a horse with a Chinese character’s pronunciation, “ma”. In this decoding process, the constraint of optimal relevance presumption must have worked because when associating the emoji to other semiotic symbols, there might be more than one potential semiotic system to be considered.

For example, the emoji of horse could also be transferred to the English word "horse" or other languages' word that meant "horse". However, in this context, the participant's transferring process triggered by iconic association was constrained by the presumption of optimal relevance as the context was a Chinese context and the other two semiotic elements in this TSE were also Chinese. Therefore, transferring the emoji to a Chinese character might be less effort saving because it can be wholly dependent on the lexical context. Then the participant mentioned that the emoji of horse had many different meanings. These meanings referred to different Chinese characters whose pinyin were "ma". Again, presumably, the presumption of optimal relevance was at work, because the participant's background knowledge supported his analysis. After all, many Chinese characters like *麻*, *妈*, and *马* were potential candidates via iconic association to be chosen for the emoji of horse on social media communication. After the participant determined the corresponding pronunciation and semiotic system type of the emoji, he combined it with the pronunciation and the other characters to locate an appropriate meaning in this context. As the participant showed, "*看麻*/má/. *看妈*/mā/, *看马*/mǎ/, *看骂*/mà/. These should not be". In this process, he took all of the four expressions to the context for consideration in order to find an interpretation consistent with the presumption of relevance. He gave up the latter three meanings as he mentioned that "basically Chinese is based on sense", which actually meant that the meaning of *看麻* was the only one that matched his "sense"; in other words, for Participant M, this meaning met the relevance presumption. Let's take another example.

Ex. 3

Background information: TSE of 🌛 is the Trans-semioticising equivalent of the Chinese word *约* (yue ma). In Chinese, *约* is a short expression asking whether others want to meet and have fun. The emoji of the moon replaces the Chinese character *约* because both *约* and the moon's Chinese words *月* is pronounced as /yue/ in Chinese. In this case, people use the single character to refer to moon. However, in modern Chinese, people seldom use a single character *月* to express the moon. In fact, *月亮* (yue liang) is the often used word to express the notion of moon. This may bring some trouble for readers to understand the meaning of this emoji in this TSE. The character, *约* is pronounced at the first tone of /yue/, which should be /yuē/. *月* should be

pronounced as /yuè/ at the fourth tone. However, in the progress of using emojis to express Chinese characters with similar pronunciation, people can tolerate some mismatching in their pronunciation, such as different tones or front/back nasal, since the amount of emojis is relatively unnecessary for thousands of Chinese characters.]

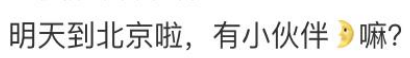


Image 12 Example of *嘛*

[**Translation**: Arriving at Beijing tomorrow, any buddy wants to 🌛?]

Transcript extract of Participant T's think-aloud protocol:

明天到北京有小伙伴月亮吗？我是这样理解，这个我是第一次看的，首先第一反应我看到的这个符号是个月亮的符号，我在想是不是和夜晚和深夜比较有关系，但是好像发现用深夜夜晚这种思维往里边代入有点不太通顺，后来发现应该就是这个表示约的意思，因为月亮和约是读音上很像的，因为这句话他说明天到北京了，有小伙伴什么吗？那么大家经常会说有没有人约一下或者是聚一聚这种，用“约”来表示这种聚会的状态，那么月亮应该就是一个谐音，也是一种谐音的方式，去代表“约”这个意思，可以这样。

Arriving at Beijing tomorrow, any buddy moon? I understand it in this way. It is the first time I have met this (expression). First, at first glance, I saw this emoji, which was a symbol of moon. I was thinking whether it had a stronger relationship with evening and night. However, (I realised that) it seemed that the thought of midnight and evening could not make sense in the context. Then I found that the moon emoji should represent the meaning of *约* as it sounded like *月* in *月亮* from the perspective of pronunciation. Because he said that he would arrive at Beijing tomorrow, is there any little friends (to do) what? Then people usually say anyone wants to meet or gather around. Using *约* to express the status of holding a party. So the moon's emoji should be a homophone. It's also a way to use homophone to refer to the meaning of 约. It can work in this way.

In this example, the presumption of optimal relevance must also have played an important role when Participant T interpreted the TSE. The TSE in this example contained an emoji as well. Participant T had to decode the meaning of the emoji of moon to continue his understanding. At the beginning, he did not associate the TSE with a certain word in a language. Instead, he thought about the literal meaning of the emoji and came up with an interpretation that was related to night time. Then, he brought the meaning of moon and midnight to the context. However, he found that "it seemed that the thought of midnight and evening could not make sense in the context", which proved that for the participant, the meaning of night time was not the most relevant meaning, suggesting that he was driven by the search for optimal relevance. Therefore, he tried another strategy, analysing the TSE's pronunciation. He noticed that the character *约* has a similar pronunciation to that of the Chinese character, *月*,

which meant “the moon”. Then by considering the contextual information about just arriving at Beijing and the younger generation’s new way of expressing holding a party or making an appointment, he believed that the meaning of *约* might be the optimally relevant interpretation supported by iconic association.

6.2 Relevance-constrained lexical pragmatic processing in TSE comprehension

As reviewed in Chapter 2, Relevance Theory has been adopted to account for lexical pragmatic processing in utterance comprehension. Related research mainly focuses on how people undergo narrowing and broadening in processing words in the context of communication. My study shows that the process of narrowing and broadening also applied to the interpretations of some TSEs. For example, in the post “The whole homepages (of the Micro-blog) are worried about the weather, traffic and temperature of Beijing hahahahahaha, *xswl*, which makes me feel so excited too”, *xswl*, which can be an acronym of numerous words, is used to refer specifically to *笑死我了* (“make me laugh to death”); thus, we have a case of narrowing. Here is another example. In the TSE *C 位*, “C” can stand for all the words with “C” as the initial letter. Here, it is used to refer specifically to “*central 位*”, which means “in the central position”. Note that the cognitive process of narrowing in understanding TSEs is also governed by the presumption of relevance. For example, a TSE like *xxs* could be considered as the pinyin acronym of *小学生* (xiao xue sheng), *小小思* (xiao xiao si), and *小小酥* (xiao xiao su) by the participants. However, the latter two possible interpretations were rejected because the participants found that the utterance explicature thus generated could not generate any substantial cognitive effect in the context of the posts in which it occurred. The posts would sound incoherent and even meaningless. Regarding broadening, the processing of *芜湖nese* and that of *笑s* are cases in point. Whereas the former involves metaphorisation (*nese,* which is homophonic to the

equivalent of *垃圾* in Cantonese) needs to metaphorically broadened to refer to useless people), the latter involves hyperbole (*笑s* means "laugh to death").

Apart from semantic narrowing and broadening, another important and even characteristic lexical process constrained by the presumption of relevance, notably semantic integration, is also at work in understanding TSEs. Specifically, when the participants tackled the TSEs, they would integrate the meaning they got from the explicit or implicit semiotic elements of TSEs to generate a possible interpretation. By semantic integration, I mean that in the course of understanding the TSEs consisting of different types of semiotic elements, language users derive their meaning not by way of simply adding up the semantic/literal meaning of their components (in other words, the meaning of these TSEs is not compositional; the meaning of the components cannot be decided alone; rather, the meaning of each component is fixed in relation to that of the other component(s), such that the integrated whole meaning of the TSE would make sense and be relevant in the macro context),. Rather, they do so by way of generating an integral meaning on the basis of component assumptions based on the decoded meaning from the component semiotic elements. Such semantic integration is guided by the search for an interpretation consistent with the presumption of relevance, in that the outcome of the integrating process is expected to make sense or contribute to the discourse of the post in the current context. Guided by the presumption of relevance, the readers of the TSEs need to construct a context in which to process the TSEs.

For the integrating process, there are two different layers of context. One is the macro context, which refers to the readers' assumptions about the content provided in Weibo posts. The other is the micro context, which can be called "lexical context", referring to the readers' assumptions about the TSEs. The readers' assumptions about the semiotic elements and their combination together create a micro context in which they would understand the TSE as appropriate to the macro context`. The readers may weigh over their assumptions about the semantic connection between the two or more elements in comparison with a potential non-Trans-semioticising version of the

expression so as to seek an interpretation consistent with the presumption of relevance.

Semantic integration does not occur in the processing of all TSEs like those I treated in the previous section. Rather, they apply to TSEs such as *打 call*, *niubility, cpdd, 花粉, Up 主, C 位*, *9 敏*, and *book 思议*, which involve components that need to be separately processed and whose component meanings need to be integrated to derive the overall meaning of the entire TSEs. Let us have a look at Ex. 4.

Ex. 4

Background information: *niubility* means extraordinarily good in English. It is the Trans-semioticising equivalent of *牛逼* /niu bi/. Literally, *牛逼* means ox's vagina. People usually use this word as an adjective to describe someone or something that is extremely good. This expression is relatively impolite and often used in oral communication. As we can see, *牛逼*'s pinyin is end up with a "bi", which is also contained in the English suffix "bility". Therefore, people add the suffix to the end of *牛逼*'s pinyin, *niubi*, to form "niubibility". However, two "bi"s does not look right and sounds strange so people eliminate one "bi" in the expression. *niubility* is often used as an adjective in the Chinese context, while it looks similar to "ability" and "capacity", both of which are nouns. Chinese users neglect the English grammar logic when creating this expression.

Image 19 Example of *niubility*

[**Translation**: In the past two years, our diplomacy has received great results. The international environment is getting better and better, along with more foreign friends and increasing respect when travelling abroad. Nowadays, foreigners will not praise *niubility* when they talk about Chinese people. This should be attributed to the successfully established wolf warrior image, which is amiable, adorable, and respectable. Keep fighting, diplomatic star group!]

Transcript of think-aloud task of participant D:

Bility 是通过英语单词的问题是能力的有能力的，然后它加上按中文拼音的话是牛逼，所以说我觉得是表示"非常厉害"的一个词语。应该是中英混合，不然放在帖子里讲不通。

Bility comes from the English word ability, ability. Then it is added with the Chinese pinyin of *牛逼*. So I think it is a word to express "very good at something". A mixture of Chinese and English.

In this example, the participant tried to understand *niubility* from the lexical context.

The reader analysed the expression, and evaluated it in terms of its relevance to the message conveyed by the macro context. The participant mentioned that the suffix part of the TSE came from English, which showed that the reader gave up the interpretation of *niubility* as a pure Chinese expression's pinyin spelling. Then the participant guessed that this TSE was combined by the Chinese pinyin of a Chinese character ("牛") and English suffix ("-bility"), believing that the TSE conveyed a meaning based on that of the Chinese character and that of the English suffix was appropriate to the macro context. Specifically, the Weibo writer was talking about giving a praise comment on the Chinese government. Therefore, the meaning of *niubility* could convey a relevant meaning that fits the macro context. Thus, we could that the participant's choice and integration of the meanings of the two parts were constrained by the presumption of relevance, as it would otherwise be nonsensical if it meant anything else. After all, in this TSE, *niubi* was composed of five Latin letters which could not only be Chinese pinyin but also a word from other languages or might mean nothing to those who knew little about Chinese. But, in the Weibo context, the reader might have assumed that the writer of this Weibo post should be a Chinese speaker, supported by the evidence that the rest of the Weibo post was written in Chinese. Therefore, the cluster of the five letters "niubi" was assumed to be relevant to Chinese, which could be the pinyin of a Chinese word (e.g.,*牛逼, 牛笔, 牛鼻*). Similarly, for the suffix part, the participant mentioned that it was an English word, assuming that "bility" had to be a string of meaningful English letters. In Chinese pinyin system, there is no letter "y", so that the reader could have a clue that "bility" is irrelevant to Chinese or Chinese pinyin. With the reader's English background knowledge, so for the participant, the most relevant understanding of "-bility" was from her English knowledge. Figure 6.1 demonstrates the relevance-constrained integration of *niubility.*

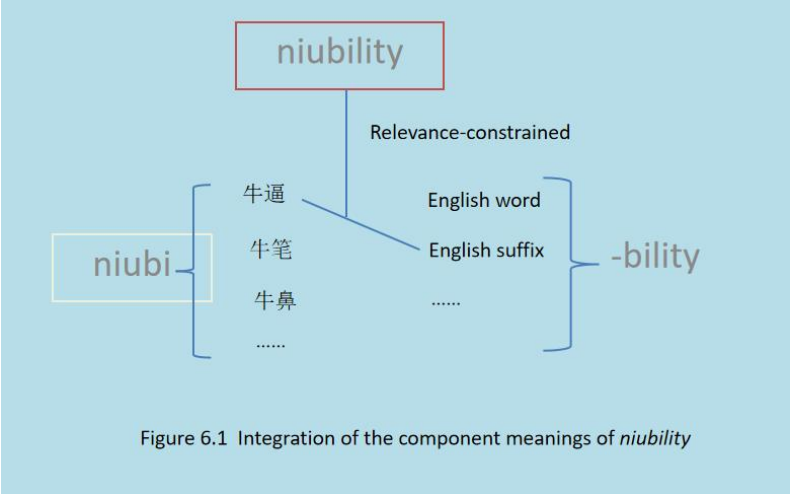


Figure 6.1 Integration of the component meanings of *niubility*

The TSE *cpdd* provides another typical example of how the participant's integrating process was constrained by the optimal relevance presumption. Consider Ex. 5.

Ex. 5

Background information: *cpdd* means "If you want to find an online couple, please send me a message". This Weibo post only contains the TSE without any other text. *@Wangzherongyao* is a tag, which shows that this Weibo is posted under the Super-topic of Wangzherongyao. Super-topic is a function embedded in Weibo. It provides an online forum for people who share the same hobbies to communicate. When adding the tag of *Wangzherongyao*, everyone who follows this Super-topic can see this Weibo post. This TSE needs to be analysed by dividing it into two parts, which are *cp* and *dd*. *Cp* is the shortened form of *couple*, which takes the two initials of the two syllables of ***c**ou**p**le*, /'**k**ʌ**p**l/. Using *cp* to express couple originated from Japanese doujin or fan-fiction culture. Doujin culture was about fans creating their own fictions for the characters that appeared in their beloved cartoons, comics, games, or movies. In Japanese, it was called カップリング, which in English was coupling meaning matching couples, especially between the same gender. After the Japanese culture flowed into Chinese social media, the meaning of *cp* expanded to a wider range. *Cp* now could be used to refer to all types of couple relationships. *dd* is a TSE expression, which means sending messages. *dd* in English is pronounced as /di: di:/. The sound of /di:/ sounds similar to the electronic beep of the message notification sound of some online instant communication media. In China, only QQ uses the sound of /di:/ to notify the users of the coming messages. Therefore, the double *d* can vividly mimic the sound of coming messages of QQ. Using *dd* to represent sending messages is a metonymy trick created by users of this TSE. However, understanding the meanings of these two parts of the TSEs is not enough for understanding the TSE's meaning as a whole. It is because this TSE is actually an extremely shortened sentence with subjects, objects, verbs, and conjunctions all omitted. The full sentence is *if you want to build a **cp** relationship, please **dd** me (send me a message)*. One can hardly understand this TSE without

knowing all of the information above.

Image 24 Example of *cpdd*

[**Translation**: cpdd@Wangzherongyao]

Transcript extract of Participant T's think-aloud protocol:

Cpdd 艾特王者，cpdd，cpdd 这个的是第一次看到 cpdd，应该是 4 个汉字的首字母缩写了。Cpdd Cp，DD 不太能想得到，因为 CP 我一开始想有没有可能和裁判相关，但是好像结合 DD 很难想到 DD 是什么意思。DD 的话 DD 好像不太像的。DD 要么还有一种弟弟的意思在里面，但是弟弟的意思在里边，CP 又是什么意思？Cpcp 要么是 couple 的意思，couple 会不

cpdd王者荣耀

会接弟弟也不太知道，或者就是说他可能跟王者荣耀相关性更强一点，可能有什么战队的名义适合能匹配上也是有可能，但是也不太了解，也只能想要这么多。

"cpdd" @ Wangzhe, cpdd, cpdd was my first time seeing it. It should be the pinyin initials of four Chinese characters. Cpdd, cp, dd. Cannot come up with ideas. Because I first thought about whether "cp" was related to *裁判* /cai pan/. However, when considering "DD" together, it was difficult to think of the meaning of "dd". For "dd", "dd" does not look like… Or there is another meaning of *弟弟* /di di/ about "dd". But with the meaning of "little brother" in it, what does "cp" mean. "cp", "cp" either means "couple". Not sure whether couple can be used next to *弟弟*. Or it is more related to Wangzherongyao. Maybe it is the name an e-gaming club. But I am not familiar with that. That's all I can think of.

In this example, we could see that Participant T considered the TSE's letters when trying to tackle the meaning of the TSE. At the beginning, the participant held an assumption that all the four letters of *cpdd* were the pinyin initials of four Chinese characters. Based on this assumption, he guessed that the first two letters might refer to the pinyin initials of *裁判* /cai pan/ ("referee"). The initial guessing was consistent with the macro context about the popular game of "王者荣耀". However, he then doubted it because he realised that together with "dd", the meaning of "referee" would be hard to be integrated with the meaning of "dd", which was often associated with "弟弟". So he excluded the meaning of *裁判 as* there was limited information provided in this Weibo post about being referees or doing referee work. So he returned to lexical context, hoping that the other part of the TSE could give useful clues. Then he began to figure out the semiotic element "dd". He activated his background knowledge to make the guess that "dd" could refer to *弟弟* /di di/ ("little brother") here. The guessing was constrained by the macro context that this was a

Weibo post, which was posted on Chinese social media and written by a Chinese writer. However, when realizing that the meaning of "referee" (or *差评,* negative comment) or "little brother" (or *滴答*) could not make sense to him, he suspected that not all of the four letters were Chinese pinyin initials. Then, drawing on his background assumptions about the game of @Wangzherongyao, he considered "cp" as the possible acronym for the English word "couple". Integrating the meaning of "couple" and that of "little brother", Participant T referred to the context to check whether the resultant meaning was relevant. He noticed that the macro context showed that the rest of this Weibo was "@Wangzherongyao". So he guessed that this TSE might be related to the MOBA game. Thus, we could see that the constraint of optimal relevance functioned three times during Participant T's decoding process of determining the letters' semiotic types and the process of integrating the component meanings for an overall interpretation. In other words, when Participant T needed to determine which languages these letter come from, he mainly used the lexical context to constrain and exclude less relevant understandings. Although Participant T failed to output a final understanding that could make sense, his process of (decoding the individual meanings of the components and) integrating the component meanings could also illustrate how the presumption of optimal relevance helped him to narrow down the range of possible interpretations. Figure 6.2 captures the integration of *cpdd.*

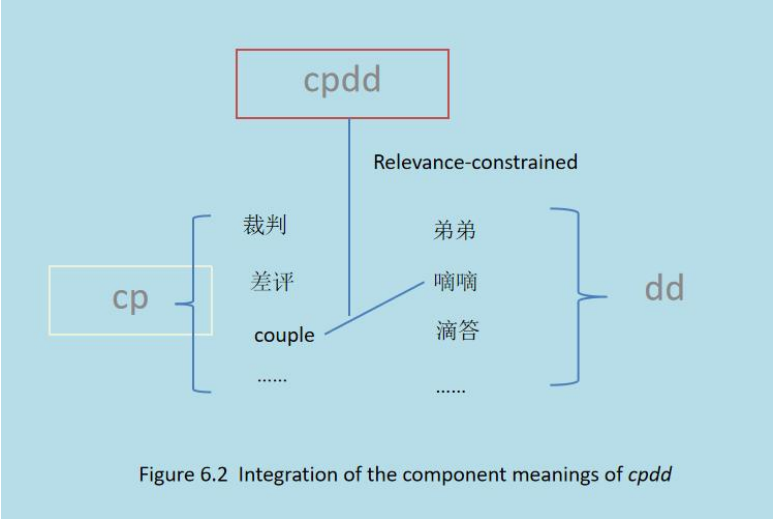


Figure 6.2 Integration of the component meanings of *cpdd*

So far, we have seen how participants' processes of decoding and integrating were constrained by the presumption of relevance. It must be stressed that while the former generally needs the support of the micro context on lexical level of the TSE, the latter process needs to be supported by the readers' background assumptions about the macro context of Weibo post, although both types of context played a crucial role in the process of excluding less relevant meaning. When decoding the TSEs, participants sometimes needed to determine the semiotic systems of semiotic elements involved in the TSEs. In TSEs that were formed by initials, participants had to analyse and decided whether these letters were shortened forms of English words or Chinese pinyin initials or even from other languages. Similarly, after the participants gathered information they believed was enough or usable, they would integrate the information to check whether the information could make sense together. In the integrating process, the search for an optimally relevant meaning might be influenced by not only the context but also other elements in the TSEs. For instance, in example 3, Participant R was thinking about how each character transferred from the initials could form a meaningful word together. In his understanding process, the contextual information also helped him to exclude those less relevant meanings such as *新谢受*.

It must be pointed out that the dependence on the macro context in decoding integrating is not exactly the same as the understanding strategy D -- referring to the context. The former is always at work, whether we take a bottom-up approach or a top-down approach in tackling the TSEs. In both cases, the integration process is guided by the search for an interpretation of the TS as well as the utterance that is consistent with the principle of relevance. The latter is actually a top-down approach. When the participants encounter a TSE, they check what the post is about and then try to figure out the meaning of the TSE. This is different from the case in which we analyse the TSE first and then bring a tentative interpretation to the macro context for checking its relevance.

It is also noteworthy that as my data shows, both decoding and semantic integration were often facilitated by various association mechanisms. For example, to

understand *v587* that occurs in a contextualised utterance, readers associated *v5* (an English letter + a numeral) with "威武" on the basis of their homophony, and associated *87* (two numerals) with "霸气" on the basis of their homophony, such that they got "威武霸气" to mean "domineering". Similarly, to understand *花粉* (which means "fans of Huawei products"), readers associated "花" with "华" on the basis of their homophony, and "粉" with the English word "fan", and then integrated the associations to arrive at the interpretation of "fans of Huawei products").

6.3 Non-propositional effects for expended additional effort in processing TSEs

Relevance Theory argues that verbal communication is guided by the communicative principle of relevance. The presumption of its optimal relevance communicated by utterances stipulates (Sperber & Wilson, 1986, p. 158) that

(a) The set of assumptions I which the communicator intends to make manifest to the addressee is relevant enough to make it worth the addressee's while to process the ostensive stimulus.

(b) The ostensive stimulus is the most relevant one the communicator could have used to communicate I.

For this reason, Relevance Theory warrants us to expect a balance between the cognitive effects we obtain from processing an utterance (or a lexical item in our case) and the processing cost involved. Such balance was at work in TSE processing. For instance, in Ex. 1 of this chapter, since the participant directly retrieved the meaning of the TSE from her mental lexicon by "decoding" and made no additional effort to analyse the TSE for its meaning, she was not rewarded with any additional cognitive effect. Similarly, the presumption of optimal relevance characterised by low cost and small effect prevailed when the participants adopted Path 2 for processing TSEs. As in the case of Path 1, the comprehension of the TSEs did not involve inference. Yet,

they might get also some additional reward in this process for the effort expended for searching. For example, they might obtain some additional information about the TSEs such as their origin and connotation. Several participants mentioned the TSE *酸q* came from the story of Teacher Guo, which gave them a sense of identification and humor. Their knowledge of the back story also helped them derive a sense of achievement.

However, when taking Path 3, the participants had to invest more effort to analyse the TSEs. In this situation, the amount of processing effort required for processing them is higher than that of lexical items that are used literally or those TSEs that could be directly retrieved. Yet different TSEs may vary in the amount of processing effort or cost, depending on how many semiotic elements are to be processed, how explicit the semiotic elements are, how many background assumptions need to be activated, how much support is available from the context, etc. Especially, since integration is not merely adding up the decoded meanings of the component parts of a lexical item like "classroom", it requires more processing effort. According to RT, the extra effort invested in doing the semantic integration is generally rewarded with extra cognitive effects. This is indeed the case, as some participants reported in think-aloud protocols that they found that TSEs had some non-propositional effects (Sperber & Wilson, 1986/1995a; Yus, 2017, 2022) or "poetic effects" (Sperber & Wilson, 1986/1995a, p. 6; 2012b, p. 87) like fun-making, humor, and the like. This fact lends strong evidence to the central argument about optimal relevance in Relevance Theory -- balance between cognitive effects and processing cost.

According to Yus (2017, 2022), the "non-propositional effects" achieved during the understanding processes may cover feelings, emotions, impressions, etc. which are not necessarily intentionally expressed by the writer or speaker but can be generated from the communication acts, and add "positively or negatively" to the cognitive effects derived from the interpretation of the discourse (Yus, 2017, p. 5). By investigating communication in WhatsApp, he summarised 7 types of

non-propositional effects: 1) feeling of connectedness; 2) feeling of group membership and social capital; 3) feeling of increased social presence; 4) feeling of narrowed gap between the physical and the virtual; 5) feeling of increased self-concept and self-esteem; 6) feeling of peer pressure; 7) feeling of emotional involvement.

In this study, I find that when seeking to understand the TSEs as forms of wordplay or, more precisely, Trans-semiotic play, the participants were not only comprehending their semantic meaning, but also noticed and evaluated some non-propositional effects they believed the users of TSEs in the Weibo posts might have conveyed. By analysing the think-aloud protocols and the results of interviews, I categorised two major types of non-propositional effects, notably social associations (such as associations about TSE users' intent of identity construction, solidarity building, sensitivity avoidance, and euphemizing, which go beyond Yus' feeling of connectedness, feeling of group membership) and aesthetic associations (such as humor and irony).

6.3.1 Social associations

By social associations, I refer to additional cognitive effects such as those about identity construction, solidarity building, sensitivity avoidance, and euphemizing. For each type, an example would be supplied to illustrate how the participants responded to the non-propositional effects achieved by the readers in processing the TSEs.

6.3.1.1 Associations about TSE users' intent of identity construction

Such associations relate to the readers' assumptions triggered by the TSEs that the users of the innovative forms might be intentionally or unintentionally constructing their identities by using the expressions. In this way, they would get additional information about the users' identity. In the example below, I will use "cpdd" to illustrate how Participant A got the additional effect when understanding the TSE.

Ex. 6

Background information: *cpdd* means "If you want to find an online couple, please send me a message". This Weibo post only contains the TSE without any other text. *@Wangzherongyao* is a tag, which shows that this Weibo is posted under the Super-topic of Wangzherongyao. Super-topic is a function embedded in Weibo. It provides an online forum for people who share the same hobbies to communicate. When adding the tag of *Wangzherongyao*, everyone who follows this Super-topic can see this Weibo post. This TSE needs to be analysed by dividing it into two parts, which are *cp* and *dd*. *Cp* is the shortened form of *couple*, which takes the two initials of the two syllables of ***c****oup****l****e*, /'**k**ʌ**p****l**/. Using *cp* to express couple originated from Japanese doujin or fan-fiction culture. Doujin culture was about fans creating their own fictions for the characters that appeared in their beloved cartoons, comics, games, or movies. In Japanese, it was called カップリング, which in English was coupling meaning matching couples, especially between the same gender. After the Japanese culture flowed into Chinese social media, the meaning of *cp* expanded to a wider range. *Cp* now could be used to refer to all types of couple relationships. *dd* is a TSE expression, which means sending messages. *dd* in English is pronounced as /di: di:/. The sound of /di:/ sounds similar to the electronic beep of the message notification sound of some online instant communication media. In China, only QQ uses the sound of /di:/ to notify the users of the coming messages. Therefore, the double *d* can vividly mimic the sound of coming messages of QQ. Using *dd* to represent sending messages is a metonymy trick created by users of this TSE. However, understanding the meanings of these two parts of the TSEs is not enough for understanding the TSE's meaning as a whole. It is because this TSE is actually an extremely shortened sentence with subjects, objects, verbs, and conjunctions all omitted. The full sentence is *if you want to build a* ***cp*** *relationship, please* ***dd*** *me (send me a message)*. One can hardly understand this TSE without knowing all of the information above.

Image 24 Example of *cpdd*

[**Translation**: cpdd@Wangzherongyao]

Transcript extract of Participant A's think-aloud protocol:

Participant A：感觉是那种 05 后比较爱用的词，我认识就是处 CP，然后就滴滴。滴滴应该就是找滴滴我，就找我聊天应该是吧？因为之前我也不太懂，我主要是 CP，然后不懂 *dd*，后来听他们说应该是这个意思。

Participant A: I feel that people born after 2005 like to use this word. I knew it. It means to build a *cp* relationship and then *didi*. *didi* should mean come to find, *didi* me, which means come to me and talk to me. Because I did not get it before. I mainly understand what *cp* means but don't understand *dd*. I heard it from others afterwards.

In this example, Participant A was tackling the meaning of the TSE "cpdd". Before she shared her understanding processes, she gave a comment on this TSE, which was "I feel like that people born after 2005 like to use this word". In other words, when she came across a TSE like *cpdd*, she not only cared to get its semantic meaning, but also tried to go beyond in order to know about its user. According to her knowledge,

"cpdd" was a TSE welcomed by the younger generation. Thus, when she saw someone using this TSE, she would have a presumption that the user might be relatively young. She might think that the user wished to construct an identity of being a member of the younger generation in addition to using the TSE to express the meaning of "I want to find a 'couple' or partner in an online game".

6.3.1.2 Associations about solidarity building

Associations of this kind relate to the readers' assumptions triggered by the TSEs that the users might be intentionally or unintentionally attempting to strengthen the feeling of group membership, of belonging to a community of users (Yus, 2017). Unlike the non-propositional associations of identity construction which are about building the identity of the writers or speakers themselves, those of solidarity building focus more on assumptions about the social connections with others who also belong to the circles. Consider Ex. 7.

Ex. 7

Background information: "ikun" is the name of Cai Xukun's fans. Cai Xukun is one of the top idols who are good at singing, dancing, and rapping. "ikun" is composed of one English letter "i" and one Chinese character's pinyin. "i" is the homophone of the Chinese character 爱 /ai/, which means "love" in English. "kun" is the pinyin of the third character of Cai Xukun's name. 爱*kun* is used by the fans to express the meaning of "people who love Cai Xukun". Then they use "i" to replace the Chinese character 爱. The TSE of "ikun" looks more like an English expression such as ipad and iphone.

Image 36 Example of *ikun*

[**Translation**: ikun come on! Here is Cai Xukun's *Lover*'s stage performance mixed cut. Please

come and show your love in order. The dancing on the stage is so handsome. #Weibo Auditorium#Show chaser]

Transcript extract of Participant B's think-aloud protocol:

ikun 快来看蔡徐坤情人舞台混剪。ikun 就是指蔡徐坤的应该是后后援团，就叫 ikun，然后也比较好理解就是爱坤，所以很多明星的后援会其实都有这种的特点，就是带一个偶像的名字里的一个字，然后再加一个那种爱或者什么喜好等等这种词可以。

"ikun" come on! Here is Cai Xukun's *Lover*'s stage performance mixed cut. "ikun" refers to Cai Xukun's support group, which is called "ikun". It is relatively easy to understand it as *爱坤*. So many idols' support groups actually have this characteristic, which takes one character from their idols' name and adds a *爱* or other preferences.

Participant B was familiar with this TSE and knew the meaning of "ikun". He did not simply output the meaning of the TSE and stop at that. Rather, he continued to activate related background assumptions about it. After he output the meaning of this TSE, he provided an extra comment about this TSE. He pointed out that the use of the TSE led him to think of the shared characteristic that many fan groups had. These fans from different support groups would like to give themselves group names. When using group names like "ikun", the writer could narrow down the distance with the audience and attract other fans of Cai Xukun to watch the Weibo and the video. As shown in the Weibo text, the writer was inviting other ikuns to watch and enjoy the great performance of Cai Xukun. Therefore, in this example, the Participant B made additional effort in processing the TSE and in return he derived the thought that when using this TSE, the writer was focusing on building the connections with potential audience instead of constructing the writer's own identity of being an "ikun".

6.3.1.3 Associations about sensitivity avoidance

Associations of this kind relate to the readers' assumptions triggered by the TSEs that the users might be intentionally or unintentionally attempting to avoid sensitive content, such as avoiding platform censorship. They knew that in Weibo, people usually would like to have their posts read by as many people as possible when they posted on Weibo and yet they could not post everything. Such knowledge enabled them to assume that the use of TSE could be a way of avoiding using sensitive expressions, such that the posters would have less chance to get their posts folded by

the platform or experience current limiting. Consider Ex. 8.

Ex. 8

Background information: In this Weibo, the TSE *种花家* refers to China. It is new nick name created by Chinese netizens to kindly call their homeland. *种花家* literally means a family who plants flowers for living, which has nothing to do with the content. Indeed, *种花* is pronounced as /zhòng huā/, which shares a similar pronunciation with *中华* /zhōng huá/. The latter is the traditional name of Chines nation. The character *家* still means family group. Due to the rise of online live streams and online shopping live, many streamers like to call their customers and fans *家人*, which means family members in Chinese. *XX 家* becomes a popular saying referring to possessive nouns such as some brands' or some streamers'. *中华家* together means a family of Chinese nation. The word *种花*, which means planting flowers, also gives others a feeling of being elegant, which can express people's nice expectations toward their country.

Image 21 Example of *种花家*

[**Translation**: Whose family's medal standings rank the first beautifully? Oh it turns out to be our 种花家 .]

Transcript extract of Participant G's think-aloud protocol:

但是我们种花家，之前奥运会的时候，我看到过很多次，然后读出来就明白是我们中华家的意思，可能也是敏感字或者怎么样，反正或者也有可能是小粉红们给这种粉丝群体起的一个名字，我猜是这样。

But our *种花家*. Earlier during this year's Olympic games, I saw this expression for many times. And (if) read it aloud, I could understand that it meant *中华家*. Maybe because it was a sensitive word or something. Anyway, or it was created by little pinks to call the fans, I guess.

In this example, Participant G had already known the meaning of *种花家* before attending the task. She got the meaning of this TSE through analysing the TSE's pronunciation as it was the homophone of *中华家*, which literally meant the home of China. However, when the participant output the meaning of this TSE, she did not just stop and move to next Weibo post. Instead, she evaluated the TSE's communicative

effect (or cognitive effect in Relevance Theory) or tried to figure out why the writer used this expression other than *中国*. She believed that the expression "中华家" might be a sensitive expression on Weibo platform, as it might be a name given by those who were irrationally patriotic. This TSE was actually a positive expression often used by people who were proud of their country. The expression of *中华家* was a popular way to express a close and familiar relationship with *中华*. The TSE followed the form of "xx *家*", which was usually used in topics about online shopping. The "xx" referred to the name of the online shops, such as *耐克家*, which meant "Nike home". People would like to say the product sold in "xx 家". However, the participant got an opposite understanding about its communicative effect. The non-propositional effect might not have been intentionally conveyed by the writer, but the reader still got such an association based on her own understanding.

6.3.1.4 Associations about euphemizing

Associations of this kind relate to the readers' assumptions triggered by the TSEs that the users might be intentionally or unintentionally attempting to replace impolite or aggressive expression for euphemistic purposes. In the data collected from the Weibo super-topics, some TSEs were created in different languages to avoid the use of vulgar or rude expressions. When analysing the participants' understanding paths, it was found that some participants thought of the non-propositional effect of eumphemizing during their understanding processes.

Ex. 9

Background information: *卧* is a TSE formed by a Chinese character and an emoji of clover. It is the Trans-semioticising equivalent of *我操* in Chinese, which is a bad word similar to "fuck" in English. First, *我操* is relatively rude to use when communicate with others online. Therefore, people find two different homophonic Chinese characters to replace *我操*, which are *卧槽*, which literally means "lying in a sink". Based on *卧槽* /wo cao/, this TSE uses the emoji of clover to replace the second character. Clover in Chinese means *三叶草* /san ye cao/, which can be generally considered as an emoji of grass, *草* /cao/ in Chinese. As *草* and *槽* are homophones, using the emoji of clover to replace *槽* can make sense in this TSE.

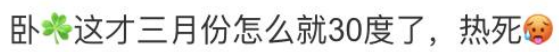


Image 37 Example of *卧*

[**Translation**: 卧 , why the temperature raises to 30°C. It's so hot .

Transcript extract of Participant G's think-aloud protocol:
卧槽，简单。之前在网上也看到过，卧槽就是我操的骂人的脏话的换一种文明的写法，那么也是直接的音译，并且把其中一个字换成小图标，也是一种更文明的写法。
卧槽, easy. I saw this TSE before on social media. *卧槽* is *我操*, a cursing word changing to a civilized way. Then it is a direct transliteration and changing a character to an emoji, which is a more polite way of writing this word.

In this example, Participant G had seen this TSE before on social media. When introducing the understanding process, Participant G mentioned that the changes from 我操 to 卧槽 and to 卧 were motivated by chasing a more civilised way of speaking impolite expression. According to her, to use *我操* was relatively rude. In fact she noticed the euphemizing effect brought about by the use of the TSE. The use of the TSE led her to assume that by employing a literally irrelevant character and an emoji, the user would avoid or reduce the rudeness conveyed by saying *操*, which was similar to the word "fuck".

6.3.2 Aesthetic associations

Aesthetic associations here refer to non-propositional effects readers may derive from processing the TSEs, such as humor and irony.

6.3.2.1 Humor

Humor is a non-propositional effect that the readers assume the writer intentionally or unintentionally conveys in order to entertain when they process the TSEs.

Ex.10
Background information: *9 敏* means "oh my god" in this context. Literally, this TSE is the homophone of *救命* in Chinese, which means "to save my life". The Arabic numeral "9" refers to the character *九*, which is pronounced similarly to *救* /jiu/. The character *敏* /min/ is pronounced similarly to *命* /ming/, while the latter is a back nasal. Using *9 敏* to express the original version of *救命* can be tracked back to a TikTok celebrity called *郭老师* (Teacher Guo).

Her special accent of speaking Mandarin became virus on TikTok. Many people followed her way of speaking both in oral and written occasions. To bring an accent from pronunciation to written form, using homophones became an effective methods. *9 敏* was one of the Guo-language expressions that went popular. Shen Mengyao is a female idol from SNH48, a girl group in China. Public performance is a performance provided by the girl group regularly, which offers a stage to the idols to show their talents and an opportunity to fans to directly watch their idols' live performance.

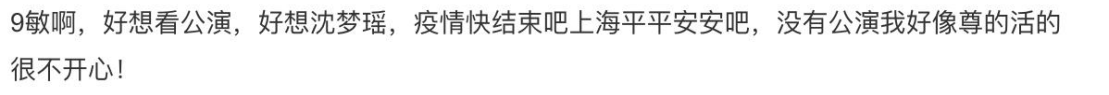


Image 27 Example of *9 敏*

[**Translation**: *9 敏*, I really want to watch public performance. I really miss Shen Mengyao. The pandemic please end as soon as possible! Hope everything good with Shanghai. It seems that I cannot live happily without the public performance.]

Transcript extract of Participant F's think-aloud protocol:

9 敏，看着会很奇怪，但是你把它念出来之后就是救命谐音过去，然后可能想搞笑或者不想好好打字那种。

9 敏 looks very strange but when you read it aloud, (you will find that) it is the homophone of *救命*. And maybe (the writer) wanted to make fun or did not want to type seriously.

In this example, the participant realised that *9 敏* was the homophone of *救命*. Then she commented on why the writer used this Translanguaging expression instead of the original expression. She believed that using this could express a feeling of funny, which might be due to the fact that she knew the background information of Teacher Guo. In her understanding processes of other TSEs, she talked about Teacher Guo and used this background information to support her understanding. The funny tone of speaking *救命* as *9 敏* would produce humor effect. The participant also suggested that there might be another reason of using *9 敏* instead of *救命* was because the writer did not wanted to type the expression in a proper way. The participant realised that the understanding process made her devote extra efforts in reading and decoding the meaning of this TSE. According to RT, the writer should follow the most relevant principle and use the original expression of *救命* to minimise the reader's effort of understanding. However, using this TSE obviously went against the principle. So the reader would consider that the writer might doing this with extra purposes. By devoting extra efforts, the reader should receive additional effects as a reward. In this case, the reader believed that the writer of this Weibo was trying to achieve a humor

effect with readers who could share the same background knowledge. Here is another example. Consider Ex. 11.

Ex. 11

Background information: 🐒🐸, means Okay in English. This TSE is composed of two emojis, which represents two different Chinese characters. 🐒 is an emoji of a monkey, which in Chinese is called *猴* /hou/. 🐸 refers to frog, which writes as *蛙* /wa/ in Chinese. *猴* and *蛙* together can only mean monkey and frog. Indeed, these two characters are homophones of two other Chinese characters. *猴* is the homophone of *好* in Cantonese, which is pronounced as /hou3/. *蛙* is the homophone of *哇* /wa/, which is an interjection that express a feeling of excitement. *好哇* means Okay in English. In the Weibo example, the writer was talking about the weather, saying that the weather was very good. So in this case, the writer used 🐒 to replace *好* in Chinese and 🐸 to replace *哇*.

今天天气真🐒🐸，准备出去玩啦

Image 35 Example of

[**Translation**: Today's weather is very 🐒🐸, good, ready to hang out]

Transcript extract of Participant R's think-aloud protocol:

今天天气真猴准备出去玩，就是一个猴子一个青蛙的 emoji，所以你把他们的 emoji 的原来的词也带进去再念一遍，就发现它是一个谐音，就代表真好，改成就显得很俏皮，增加了这句话的这种风趣的感觉。

Today's weather is very *猴*, going to hang out. These are the emojis of a monkey and a frog. So if you put the original words of the emoji to the expression and read it again, you will find them are homophones, which represent "very good". By making this change, it looks more funny and adds a feeling of humor when saying this sentence.

In this example, Participant R successfully derived the semantic meaning of this TSE's semiotic type and transferred the emojis to Chinese characters. Noticing the homophonic relationship between *猴哇* and *好哇*, he output the meaning of "very good". He could feel, as he commented, that by making the change from Chinese characters to emojis, this TSE conveyed a sense of humor to the readers. As we could see from the context, the writer was planning to hang out to enjoy the nice weather, which might be a delightful activity. Therefore, positive attitudes like happiness and chill were embedded between the lines. Participant R put extra effort in processing this TSE and she was rewarded by the humorous effect she sensed.

6.3.2.2 Irony

Irony is a non-propositional effect that the readers assume the writer intentionally or unintentionally conveys via the use of the TSEs in order to show his or her ironic attitude. Consider Ex. 12.

Ex. 12

Background information: *可拷* literally means be able to be handcuffed. It is the Trans-semioticising equivalent of the word *可靠*, which means "reliable" in English. *拷* and *靠* share the same pronunciation. *拷* is the verb of putting handcuff on criminals while *靠* means "leaning on something". By changing *可靠* to *可拷*, the TSE expresses an ironic attitude towards illegal behaviors.

Image 38 Example of *可拷*

[**Translation**: #There should be someone selling the Epidemic Prevention and Control Pass # Very *可拷*]

Transcript extract of Participant M's think-aloud protocol:

有人卖疫情防控很可拷，可应该是可靠，但是这个应该是拷上那叫什么来着手铐的铐，所以就是说铐上手铐可能就是犯法了，然后就必须进警局了。然后加上 tag，我会觉得说它就是一种讽刺，就把可靠弄个谐音，因为中国像网络上很多这种词汇，这完全是我自己猜的。

Someone sold the pandemic Prevention and Control Pass, which was very *可拷*. *可* should be *可靠*. But this should be the *拷* of *拷手铐*. So when putting a pair of handcuff on, it might mean being against law and must be sent to prison. And considering the tag, I would like to say it is an ironic saying. Making *可靠* changed by homophones. As there were many words used on Chinese social media, I totally guessed.

Participant M resorted to the context when processing the TSE. She transferred *可拷* to *可靠* first. Then she considered the meaning of *拷* and got the assumption that the writer used *可拷* to describe an illegal behavior that was worth a pair of handcuff. She sensed, as she commented, that using *可拷* expressed a note of irony. When reading this Weibo, readers like Participant M could both achieve the meaning of "reliable" and "deserve to be handcuffed". The Weibo was making a comment on a person who sold Epidemic Prevention and Control Pass. The writer used *可拷*, which sounded like *可靠*, to pretend to express an affirmation but actually criticizing this behavior. In the processes of understanding the TSE, Participant M made extra effort to figure out the meaning and attitude that were conveyed by the writer. Apparently, the character *拷* was not the most relevant choice of character, compared with the

conventional one, *靠,* and thus would incur additional processing effort. Nevertheless, the extra effort she paid was rewarded by her derivation of the ironic attitude on the part of the TSE user.

Ex. 13

Background information: The author ironically complained that Jingjiang, a city in Jiangsu Province, got a Covid-19 positive visitor because an Anhui visitor concealed his illness. *可真刑* (ke zhen xing) is a Translanguaging expression which originates from *可真行* (ke zhen xing). *真行* literally means someone is really good at doing something. However, with the character *可* added in front of *真行*, the meaning converts to the opposite side, meaning someone is really good at doing some wrong or bad things. While in this TSE, the character *行* is replaced by its homophone, *刑*. Both these two characters are pronounced as /xíng/ but have different meanings. *行* means being good at doing something in this context, while *刑* means being illegal. So when using *刑* to replace *行*, the former character not only inherits the ironical and passive tone of *可真行* but also enriches original meaning, meaning that someone is good at doing some illegal things (not just wrong). Therefore, the TSE, *可真刑* means that someone does an illegal thing which will bring him or her to prison.

#靖江疫情#恭喜靖江因为一个瞒报的安徽人喜提确诊👏可真刑啊

Image 39 Example of *可真刑*

[**Translation**: #Jingjiang Epidemic# Congratulations to Jingjiang city where people happily got a diagnosed (Covid-19 patient) due to an Anhui visitor's concealment. [clasped fists] That is really *刑*.]

Transcript extract of Participant C's think-aloud protocol:

这应该也是个谐音，然后应该是行不行的行的一个，然后同时可以蕴含瞒报的事情带来的后果，然后相当于这种既能表达它的意思，然后又能表达它的后果。

This should also be a homophone. And it should be the *行* of *行不行*, which can simultaneously convey the consequence of concealment. So it equals to express both the meaning and the consequence.

For Participant C, when dealing with the TSE *可真刑*, he noticed that this TSE was a homophone of *可真行*. Similar to *可拷*, to fully understand this TSE, readers had to not only identify the corresponding original word but also derive the non-propositional effect possibly conveyed by its user. As Participant C mentioned, this TSE could apparently express the meaning and the consequence simultaneously. In other words, the participant could feel that through the use of the TSE, the user

could convey an ironic attitude at the same time. In this process, Participant C devoted additional effort to understand the TSE as *刑* was not the most relevant character as compared to its original expression of *可行,* because had to make a homophonic association and use the context to constrain the range of possible characters. When realizing that *可刑* might express two layers of meaning, he was rewarded with the effect of sensing irony on the part of the user.

To summarise, understanding TSEs in a way that involves the active or passive expenditure of extra effort is often rewarded with some non-propositional effects, social or aesthetic. These effects might have been intended by the users of the TSEs or actively assigned by the readers. Since the additional effort is generally offset by additional effects, we can claim that the comprehension of the TSEs still conforms the communicative principle of optimal relevance (Sperber & Wilson, 1986/1995a).

Note that there are exceptions, though. While the extra effort is generally a rewarding one, it occasionally failed to come through when the participants finally gave up understanding a TSE, as shown by Ex. 7 of Chapter 5. Under such circumstances, the participant did not get any non-propositional effects as a reward. Rather, he would feel so frustrated or disappointed that he chose to give up.

A caveat here is that we have largely depended on the participants' reports to discuss the non-propositional effects conveyed by the TSEs. It does not mean that those who did not report on their derivation of non-propositional effects did not experience them. This might be a limitation of the think-aloud method, as it did not record all the cognitive experiences of the participants. Theoretically, as they decomposed the semiotic elements and integrated the component meanings, they would experience all sorts of associations triggered by the activating related background assumptions as rewards for their expenditure of additional effort.

6.4 A relevance-theoretic framework for understanding TSEs

From the perspective of Relevance Theory, the TSEs readers encounter on Chinese social media are ostensive stimuli that convey the presumption of their own optimal relevance. Therefore, in order to tackle their semantic meaning and sometimes further non-propositional meaning, the readers need to secure an interpretation that is consistent with the presumption of relevance on the one hand, and derive non-propositional effects as a reward if they expend additional processing effort. As our data indicates, there may be three situations of the readers processing the TSEs.

The first situation, as indicated by Path 1, is a most direct and straightforward one: The readers were familiar with the TSE in question and would make direct access to the semantic meaning of the TSEs. They would just reach out to related encyclopedic information stored in their memory system through either logical, encyclopedic, or lexical entry, and map their pre-stored assumption about its semantic meaning onto the form. After they output the semantic meaning, they would move on to the next TSE. Their processing, in the same way they decoded and understood non-Transemioticising words or expressions, did not require additional effort, but was also consistent with the presumption of optimal relevance because it was not rewarded with extra cognitive effect. No inference was involved in this process. Consider Ex. 14:

Ex. 14

Background information: The TSE, *集美* /ji mei/, means sisters in English. It is the Trans-semioticising equivalent of the Chinese word *姐妹* /jie mei/ meaning "sisters" or, more broadly, "female friends". *姐妹*'s pronunciation is similar to the expression of *姐妹*, which can be seen as homophones to each other. However, the creation of using *集美* to replace *姐妹* is not related to any Chinese dialect. This expression became popular due to a web anchor on Douyin. Douyin and TikTok are basically the same created by the same company, while Douyin is provided only for Chinese users. The Douyin anchor was called Teacher Guo, whose video and live streaming became virus due to her humorous jokes, life attitudes, and her special pronunciation. She created her own tone of saying Chinese, which left millions of people a deep impression. Her fans mimicked her accent, which later became a trend among young people. Among all the strange pronunciations of Chinese, *集美* was one of the famous.

[**Translation**: Goodnight *集美 s*. (Which one looks good? Pick one for me! I will see (your comments) tomorrow morning.]

Transcript extract of Participant A's think-aloud protocol:
集美们应该是姐，姐妹们，我知道
集美 s should be sis, sisters. I knew it.

Here, Participant A made no further processing effort as she stopped at the outputting the meaning of the TSE. The interpretation retrieved from her mental lexicon was already consistent with the presumption of optimal relevance, because we can see that it made sense in the macro context of the posts.

When the participants took Path 2 by looking for external assistance, they would resort to various strategies such as looking the TSEs up in the online dictionary or from online resources. Whatever external resources they used, they refrained from analysing the TSEs that they might (not) have met before but were unfamiliar with or unsure about. When they obtained a result, they would directly map the meaning onto the creative forms, as they did in Path 1. They might have found more than interpretation, though. In this case, they would choose the one that fitted the macro context best, that is, the one that was consistent with the presumption of relevance. Since the selection process is relatively easy, the additional effort would be rewarded with very marginal or negligible additional effects. If they spent more time reading about the TSE entry, they would be rewarded with additional effects such as getting to know more about the TSE such as their origin and connotation (Cf. Section 6.3).

When the participants adopted Path 3, they would need to make effort to analyse the TSEs. In this process, they would have to activate related background assumptions triggered by the component parts of the expression. They would have to take into account various word formation mechanisms in Chinese and other languages. They would undergo various lexico-pragmatic processes such as narrowing, broadening and

integration. They would have to check the tentative outcome of integrating the meaning of each component (different TSEs involve different components and modes of integration) against the macro context to determine whether the interpretation was consistent with the presumption of relevance. Since all these steps or processes required some extra effort, the readers would be rewarded with some non-propositional cognitive effects like associations about identity construction, solidarity building, humor, irony, etc. (as discussed in Section 6.3). In cases where the readers were already familiar with certain TSEs and yet expended additional effort on them, as in Ex. 8 and 9, they were still processing the TSEs in a way consistent with the presumption of optimal relevance. But this might have arisen as a result of the think-aloud task. In their regular online reading, they might not stop and speak aloud their associations. However, its occurrence in our data suggests that the use of TSEs may serve to convey one kind or another of non-propositional effects.

Figure 6.3 is intended to provide a relevance-theoretic framework to assist the understanding of TSEs readers encounter on Chinese social media.

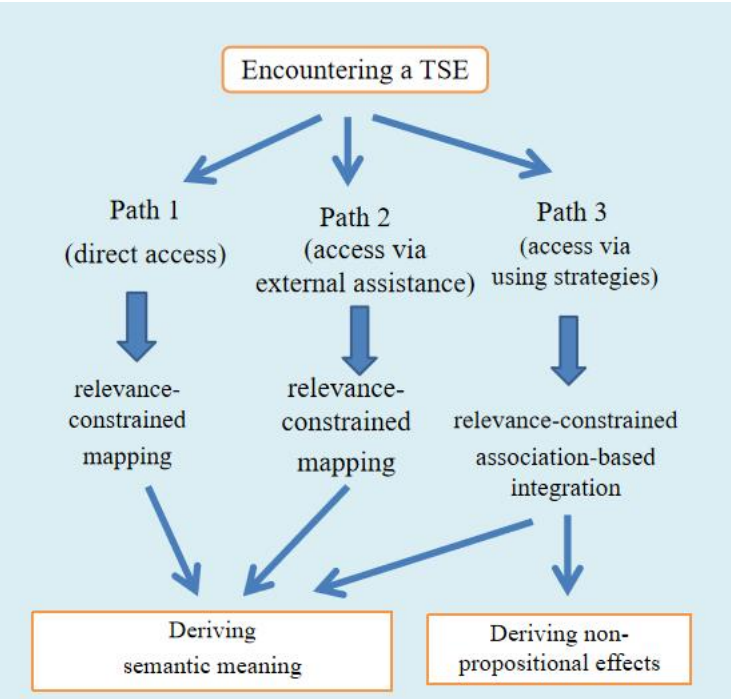


Figure 6.3 A relevance-theoretic account of TSE comprehension

In the understanding processes of TSEs, the constraint of the presumption of

relevance functioned both on the decoding and the integrating processes of TSEs. Participants' efforts on understanding TSEs were devoted to both processes. Compared with reading Weibo with no TSEs, readers might get extra opportunities to achieve more effects when they had to understand the TSEs. As we could see in examples of *cpdd*, *9 敏*, and *可拷*, the participants were not familiar with their TSEs. But during the understanding processes, they got non-proposition effects of one kind or another, as we discussed. When they understood the TSEs, they devoted more efforts to get the information; correspondingly, they got extra effects such as the associations about the writers' identity construction, irony or humor as a reward.

However, in the study of TSEs, it was found that even when the participants knew the meaning of the TSEs before, they could still achieve extra effects, depending on whether they expended additional effort. When the participants were familiar with the provided the TSEs, they could directly reach their mental lexicon the get the exact semantic meaning of the TSEs. In this situation, according to Relevance Theory, the readers could not get further cognitive effects as they did not bother to devote extra understanding efforts. However, as the examples discussed above demonstrated, even when the participants could directly output the meanings of the provided TSEs, they could still achieve non-propositional effects. In other words, the cognitive effects achieved by the participants might not only come from the understanding processes. TSEs could function as keys to activate readers' background knowledge about the TSEs and the background information such as related memes, people who used the TSEs, and how the TSEs were created could also help the readers to obtain cognitive effects. In the examples shared above, participants would laugh when they came across the TSE of *9 敏* as they would think of the sound of Teacher Guo, which was very funny. Also, when a participant saw the TSE of "ikun", he thought of many interesting memes created on social media about Cai Xukun. These memes became the trigger of further non-propositional cognitive effects. This is indeed a situation not well predicted by Relevance Theory, though.

6.5 Summary

In this chapter, I approached the comprehension of TSEs from the perspective of Relevance Theory. In Section 1, I demonstrate how the lexico-semantic processing TSEs via the choice of comprehension strategies lends evidence for the presumption of optimal relevance as the cognitive constraint on TSE comprehension. From the participants' think-aloud protocols, I found evidence that the constraint of the presumption of optimal relevance exerted influence on the understanding processes of Translanguaging expressions. Specifically, when decoding the TSEs, the participants' interpretation of the component semiotic elements was constrained by expectation for relevance. They needed to select the interpretation that helped the utterance containing the TSE make sense in the macro context. It was also found that the analytical process, including the triggering of the various association mechanisms (i.e., homophonic association, iconic association, acronymy-based association, and shape-meaning association), was also constrained by the presumption of relevance.

In Section 2, I explored the lexico-pragmatic processes involved in the comprehension of TSEs. I found that both narrowing and broadening were at play in the comprehension of some TSEs but not others. More importantly, I found a new lexico-pragmatic process, notably semantic integration. Specifically, when the participants tackled the TSEs, they would integrate the meaning they got from the explicit or implicit semiotic elements of TSEs to generate a possible interpretation.

For the integrating process, there are two different layers of contexts. One is the macro context, which refers to the readers' assumptions about the content provided in Weibo posts. The other is the micro context, which can be called "lexical context", referring to the readers' assumptions about the TSEs. The readers' assumptions about the semiotic elements and their combination together create a micro context in which they would understand the TSE as appropriate to the macro context`. The readers may weigh over their assumptions about the semantic connection between the two or more elements in comparison with a potential non-Trans-semioticising version of the expression so as to seek an interpretation consistent with the presumption of

relevance.

Unlike regular expressions composed of a single language such as Chinese, no integration is necessary. However, Trans-semioticising expressions involve elements from two or more different semiotic systems. Therefore, before the participants integrated useful information from the component parts, they had to figure out what types of semiotic elements were actually involved in the TSE. Otherwise, they could not further activate their background information and make use of other information to support their understanding. The constraint of the presumption of optimal relevance also existed in the integrating processes. When the participants were not familiar with the TSEs, they needed to piece together the information collected from the TSEs' formation mechanism, background information and the context provided from time to time and then generated an interpretation that satisfied the expectation for relevance. Very often, participants might come up with more than one understanding about the TSEs. Under such circumstances, the participants relied on the presumption of optimal relevance to choose the most suitable interpretation. In the process, several association mechanisms (i.e., homophonic association, iconic association, acronmy-based association, and shape-meaning association) played an important part, Nevertheless, their operation was still constrained by the presumption of optimal relevance. The participants would bring their interpretations of TSEs to the macro context in order to verify whether they could meet the expectation of optimal relevance.

Furthermore, in Section 3, I discussed the non-propositional effects brought by the efforts devoted to the understanding processes. It was found that when understanding the TSEs as forms of Trans-semiotic play, participants might get non-propositional effects such as social associations, aesthetic associations as rewards for the additional effort the readers invested. Some effects, such as associations about identity construction, solidarity building, and humor are characteristic of TSEs, as applicable to TEs (Li, 2011) or wordplay more generally (Lloyd, 2007). Yet, effects such as assumptions about sensitivity avoidance and euphemizing are new

non-propositional effects found in this study of TSEs. No participant mentioned the function of attracting people's attention (Monsefi & Mahadi, 2016) in the use of TSEs as a form of wordplay.

Finally, in Section 4, A relevance-theoretic framework for understanding TSEs is constructed. In the understanding processes of TSEs, the constraint of the presumption of relevance functioned both on the decoding and the integrating processes of TSEs, where participants' efforts on understanding TSEs were devoted. Comparing with reading Weibo with no TSEs, readers might get extra opportunities to achieve more effects when they had to understand the TSEs. However, in the study of TSEs, it was found that even when the participants knew the meaning of the TSEs before, they could still achieve extra effects, even when they did not spend additional effort. It was because that TSEs could function as the keys to activate readers' background knowledge about the TSEs and the background information such as related memes, people who used the TSEs, and how the TSEs were created could also help the readers to obtain cognitive effects.

Chapter 7 Conclusion

This chapter concludes the whole thesis by summarizing the study and its major findings, discusses its theoretical and methodological implications, and points out the limitations of the study and possible future directions of research.

7.1 Summary of the study

The present study was undertaken to investigate Trans-semioticising expressions (TSEs) on Chinese social media, with the focus placed on their comprehension by online readers. To achieve the goal, three major research questions were proposed:

A. How are the Trans-semioticising expressions (TSEs) Chinese readers encounter in Weibo posts created?

B. How do Chinese readers of the Weibo posts understand the different types of TSEs they encounter in the text of the posts?

C. Why do Chinese readers of the Weibo posts understand the TSEs the way they do?

In order to answer these three main research questions, this thesis designed two parts of data collection and analysis sections.

To categorise different types of TSEs that were used on social media, Weibo, I chose the Weibo super-topics as the source of data collection. Eight Weibo Super-topics were selected and 8000 Weibo posts were collected with TSEs (appeared for 6157 times and 2670 different TSEs identified). By manually marking the TSEs based on the semiotic elements involved in the TSEs, nine types of TSEs were categorised. The combinations of semiotic elements like Chinese character, Chinese pinyin, English letters, emojis, numerals, and other languages resulted in various TSEs.

Then, the study probed into how readers understood TSEs on social media. Three sub-questions were raised under this main question, which were 1) what were the understanding paths, 2) what understanding strategies were used, and 3) what factors influenced the understanding. To address these questions, think-aloud

protocols were adopted to generate answers.

Finally, with evidence from the data collected from the think-aloud protocols, I provided an account of the participants' TSE understanding from the perspective of Relevance Theory. It was discussed and illustrated how the lexico-semantic processing and lexico-pragmatic processing of TSEs were constrained by the presumption of optimal relevance. The various non-propositional effects derived from additional processing effort in tackling the TSEs were also explored.

7.2 Major findings of the study

The analysis of the dataset and the think-aloud protocols and interviews generated the following findings.

First, the TSEs readers encountered on Chinese social media were found to fall into various categories from different perspectives. First, based on the word formation mechanisms, they fell into the following main types, namely morpho-semantic TSEs, semantic TSEs, morphological TSEs, and external TSEs. Second, in terms of the composition of semiotic elements (including five semiotic systems, which were logographic characters, Latin alphabet letters, numerals, emojis, and other symbols), they included the following types: 1) Logographic characters + Latin alphabet letters; 2) Logographic characters + Numerals; 3) Logographic characters + emojis; 4) Latin alphabet letters + Numerals; 5) Logographic characters + other symbols; 6) Logographic characters + other symbols; and 7) Multiple symbol TSE. Thirdly, in terms of the employment of semiotic elements and application of the formation mechanisms, TSEs could be further divided into explicit and implicit ones. If a TSE only had one semiotic element type in its surface form, it was considered as implicit TSE as readers could not determine whether it was a TSE from its form and formation. Nevertheless, it was actually a TSE because its understanding involved the participation of more than one semiotic type.

Second, three paths of TSE understanding were discovered: A. direct access, B.

access via external assistance, C. access through using understanding strategies. Direct access occurred when participants who knew the meaning of the TSE before directly reached out to their mental lexicon to get the meaning of the TSEs. When participants understood TSEs via external assistance, they did not know the meaning of target TSEs. However, instead of understanding the TSEs by themselves, they turned to seek external assistance to help them get the meaning such as using search engines to look up for the exact meaning of the TSEs and using typing systems on their mobile phones or computers to get extra information. The third path was characterised by using different understanding strategies. Compared with understanding paths A and B, the third path was more complicated. Eight different strategies were summarised based on the participants think-aloud protocols. The first strategy was resorting to the text, used unintentionally when the participants needed to find hints from the text. Participants were found to read the text word by word. Judging from the topic was the second strategy used by the participants. Participants might need to consider what the Weibo was talking when they understood the TSEs. For example, a Weibo about online gaming would have less chance to use a TSE related to idol culture for the participants. The third strategy used by the participants was analysing the TSEs with respect to different aspects of the TSEs, including the TSE's pronunciation, the internal structure of the TSE and the shape of the TSE characters. The next strategy found was referring to the context. The participants did not need to read the text character by character. Rather, they relied on information provided or implied in the context. Reading the text can be seen as a prior means for this strategy. The fifth strategy identified was activating background knowledge. Participants as social media users would get information and knowledge about the topics and background stories related to the TSEs and also associated the TSEs with some other words/TSEs they already knew. This strategy provided additional information for the participants so that their understanding would not be limited to the context and the TSEs themselves, which increased the opportunities of understanding TSEs especially when the direct information was limited. The sixth strategy was

referring to processing history. The understanding of TSEs of similar types could offer useful experience for the participants. The last understanding strategy was guessing strategy. The participants directly came up with a guess about the possible meaning or decoding directions when they dealt with the TSEs. These strategies together worked to help the participants tackle the meaning of the TSEs. It was also found that participants sometimes would use more than one strategy to understand one TSE. The order of using the TSEs was also flexible. Participants could use the same strategy for several times. Even for the same TSE, different participants showed different orders of using strategies. The study also discovered some factors that influenced the understanding paths and understanding strategies used. Two main factors identified in this study were discussed, which were the participants' level of familiarity with the TSEs and the types of the TSEs. In the first situation, the participant did not know the meaning of the TSE; however, the participant knew the meaning of a similar TSE (for example, TSEs that shared the same origin expression). In the second situation, the participants knew how to understand this type of TSE based on the experience of understanding similar TSEs. For the third level, they did not know the meaning of the TSE and had no information about how to understand the TSE. For participants that already knew the meaning, they would take direct access to get the meaning of the TSEs by referring to their mental lexicon. For participants who had some information about the TSEs, when the participants came across a TSE that was believed to share the same origin word with another TSE or expression, to understand the meaning of the new TSE would not be very challenging even though they did not know the exact meaning of the TSE. When the participants knew how to understand this type of TSE based on the experience of understanding similar TSEs, the participants would take different paths. The understanding experience could help the participants from different perspectives, including the experience of understanding TSEs with similar word formation type, the experience of associating the semiotic elements with related characters, and the experience of identifying homophones. When the participant had no knowledge about the TSE, they

would opt to take Path B. Access via search. Path C also occurred when the TSEs were unfamiliar to the participants. Every time they met the TSE, it would be a brand new experience of understanding a TSE. In the progress of understanding unfamiliar TSEs, participants might use their background knowledge to support their understanding after they found some clues to decode the TSE. The second factor that was believed to affect the understanding paths and the strategies used was the TSE's type. In this section, the understanding processes were also drawn by the flow charts for types of affixation, phonetic motivation, external TSEs, and semantic TSEs. Furthermore, the study showed several associative mechanisms of relevance search in understanding TSEs. When the participants were analysing the TSEs, they actually followed several associative mechanisms to support their search for optimal relevance, which included homophonic association, splitting mechanism, acronymous mechanism, and pattern-meaning association.

Thirdly, it was argued that the participants' full understanding of TSEs was constrained by the presumption of optimal relevance. It was found that the constraint of relevance took place when participants decoded the TSEs from the perspective of the semiotic elements they contained. When the participants were tentatively decoded the TSEs, they would try to integrate the results of the decoding to see whether the generated interpretation (a possible meaning) was appropriate to the macro context. If not, they would re-decode the component parts of the TSEs to generate a new interpretation. In this process, the constraint of optimal relevance also functioned to exclude those interpretations that were less relevant. Participants needed to devote extra efforts to reach the optimally relevant interpretation among many possible corresponding ones. The study suggested that the extras effort in understanding TSEs may be offset by obtaining some non-propositional effects. By analysing the effects that were mentioned by the participants, two main types of cognitive effects could be categorised, notably social associations and aesthetic associations. The former included associations about identity construction, solidarity building, sensitivity avoidance, and euphemizing. The latter referred to the effects of humor and irony.

Finally, based on the findings from Chapter 5 and Chapter 6 and the discussion of the lexico-semantic and lexico-pragmatic processing of TSEs in 6.2 and 6.3, a relevance-theoretic framework was proposed to capture how the understanding of TSEs in reading posts on social media could be accounted for by Relevance Theory. It was argued that both the interpretation of the meaning and the derivation of non-propositional effects of TSEs are constrained by the presumption of relevance.

7.3 Implications of the study

Theoretically, the current study contributes to the theoretical fields of Translanguaging, dynamic lexicology, and Relevance Theory.

First, with regard to Translanguaging, this study brought the research of Translanguaging from largely the discourse level down to primarily the lexical level. By investigating the formation of Chinese online TSEs, it revealed how elements from different semiotic systems could work together to serve meaning-making purposes. It showed the complexities and diversities of TSEs that go much beyond what we have known about TEs comprising only elements from two languages. Among the few studies that discussed individual cases of Translanguaging expressions online, although there may be more and more new expressions being created by social media users, this study provided a more comprehensive view of the types of TSEs.

Second, with respect to dynamic lexicology, this study expends the scope of applicability of traditional innovative word formation mechanisms. It showed that creative word formation may cross the boundary between different languages and semiotic systems, especially for logographic languages like Chinese, phonetic languages like English, and iconic symbols like emojis. Such breakthrough may take both explicit and implicit forms. It may involve more one innovative word formation mechanisms and the combination of more than two semiotic elements

Third, this study contributes to our understanding and application of Relevance Theory. First, it testified the applicability of Relevance Theory in the study of TSEs. It was found that in the understanding processes of TSEs, the constraint of relevance

functioned both on the decoding and the integrating processes of TSEs, as evidenced by the participants' think-aloud protocols. Second, this study also expanded the application scope of Relevance Theory in the study of lexico-pragmatic processing, as previous studies focuses more on the lexical pragmatic processes in understanding mono-lingual terms. It showed that in addition to relevance-constrained cognitive processes of narrowing and broadening as previously studied, which also operated in understanding TSEs, the comprehension of TSEs involved the cognitive process of integration, which is also relevance-constrained. It provides new evidence for the argument of Relevance Theory that utterance interpretation does not always start from decoding what is said and (if at all) end at inferring the implicature; rather, understanding utterance, including lexical items like TSEs in this study, is a bidirectional process. The processing of what is said provides input for pragmatic inference and pragmatic inference also assists in the decoding of what is said (Sperber & Wilson, 1986/1995a). In addition, this thesis found that the processing of TSEs as creative expressions does not always require additional effort when readers are familiar with them and do not bother to spend more effort on them. As a consequence, no extra cognitive effects (i.e. non-propositional effects) would be obtained as a reward; in contrast, readers would be rewarded with non-propositional effects if they are not familiar with the TSEs, or if they choose to expend more effort on them although they are familiar with them. As my study demonstrated, even when the participants could directly output the meanings of the provided TSEs, they could still achieve non-propositional effects. In other words, the cognitive effects achieved by the participants might not only come from the understanding processes proper. Instead, TSEs could function as triggers to activate readers' background knowledge about the TSEs and the background information such as related memes, people who used the TSEs, and how the TSEs were created, which could also help the readers obtain additional cognitive effects. Basically, these two situations both accord with the presumption of optimal relevance. On top of that, this study enriched Relevance Theory and Yus' research on non-propositional effects by displaying an array of

social and aesthetic associations. Whereas social associations such as those about identity construction and solidarity building have been addressed by Yus (2017, 2022) and aesthetic associations such as humor and irony have been discussed by Sperber and Wilson (1986/1995a, 2012a, 2012b), the associations about sensitivity avoidance and euphemizing are brand new types of non-propositional effects revealed in the present study about TSEs as forms of wordplay or Trans-semiotic play, as I choose to rename it. The new discovery is a natural outcome as online communication is inevitably subject to various restrictions. It turns out that using TSEs is a suitable way of coping with the restrictions.

From the perspective of methodology, this study has implications in two directions. For one thing, it introduced the method of think-aloud tasks into the study of TSE understanding. Previous research on Translanguaging phenomena was primarily theory-driven and related discussions were largely speculative. By adopting the think-aloud method, this study managed to make observable at least some proportion of the participants' information processing that took place during their tackling of the TSEs. By qualitatively coding the participants' think-aloud protocols, the study revealed in a fairly objective way the participants' cognitive processes in TSE comprehension. Unlike previous studies that usually used think- aloud protocols as a tool to analyse the participants' thoughts during writing tasks, using a product, etc., this study focused on the context of using think-aloud protocols and made it a useful tool to showcase the readers' understanding processes that could not have easily been displayed directly. By applying think-aloud protocols, the present study was able to look into the understanding strategies and paths used by the readers. Compared to methods like interview or eye-tracking, think- aloud protocols used in this study provided more detailed, real-time, and direct information about the participants' thought without imposing the researchers' subjective perceptions upon them. For another, this study worked on the basis of a self-built sizable database in the categorisation of TSEs instead of depending on a very limited number of cases or instances. The big size of naturally occurring data made it possible to generalise as

fully as possible the types of TSEs with regard to the application of formation mechanisms, combination types of varied semiotic elements, and the explicitness of the formation mechanism used and semiotic elements involved. The generalisation on the basis of a big-size database enabled the display of the richness, diversity, and complexity of TSEs Chinese new media users create for meaning-making purposes. It lent strong evidence for the evolution of human communication from mono-lingual and mono-modal interaction to multilingual and multi-modal sense-making, which is characteristic of online communication supported by the Internet, computer, and smartphone technological affordances.

7.4 Limitations and future directions

Despite the provision of a whole picture of the TSEs used on Weibo and revelation of the paths and strategies of TSE understanding, this study suffers several important limitations.

First, the collection of TSEs was limited in several ways. First, the TSEs identified were all found on Weibo posts from only 8 super-topics only. As a consequence, there are still chances that the categorisation of TSEs was not complete. What's more, these TSEs were all used in the monologue contexts. Although some Weibo posts were written in the tone of talking to the audience, most of the Weibo posts were not happening in conversations between netizens. Thus, the context of the TSEs provided to the participants was relatively limited and the understanding processes were more related to reading than to communicating. In addition, due to the fast update speed of social media language, this study failed to include the most up-to-date expressions since 2021 into the database. Moreover, the TSEs included were mainly restricted to the topics of online gaming, idol-fan relationship, and a few daily expressions. Future studies could diversify the context of TSE research and expand the scope of data collection to track more.

Secondly, the number of participants attending the think-aloud task was relatively homogeneous, as they were all university students and familiar to the researcher. Yet,

social media users come from all walks of life. Thus, the think-aloud reports from the participants might not be representative of online readers from other social backgrounds, because they might be more talkative and imaginative than less well educated people, particularly younger generation like primary and middle school students. Moreover, the familiarity of the participants with the researcher might have some impact on their performance, particularly that of encouraging them to be more cooperative or even more supportive than strangers.

Thirdly, although think-aloud protocols used could reflect participants' thoughts by encouraging them to speak aloud while thinking, not every thought could be shown through this method. Participants sometimes would fall into silence when doing the think-aloud tasks. Also, for some participants who thought actively, they might finish the understanding within very short time, even less than one second. To address this problem, this study had to add a few follow-up questions after the participants finished understanding each TSE when necessary. There was another drawback of this design, which laid in the think-aloud method. It was not easy to rebuild the authentic reading context according to the participants as different participants might surf Weibo on different social media. Using the designed Weibo post display tool developed through Python, this study tried to mimic the real Weibo reading situations, while there was still difference comparing with authentic reading. Also, when selecting materials that were used for the think-aloud tasks, the level of contextual support was very difficult to control. For different readers, the contextual support varied a lot. Therefore, I could only invite more participants to reduce the influence of this factor. Future studies could diversify the methodology and gather other types of evidence to triangulate the research.

References

Aduda, K. 2013. *A Relevance Theoretic Analysis of Dholuo Neologisms*. PhD thesis. University of Nairobi, Kenya.

Algeo, J. 1977. 'Blends, a structural and systemic view'. *American Speech,* 52(1/2): 47-64.

Antrushina, G. B., Afanaseva, O. V., & Morozova, N. N. 1999. *Lexicology of the English language*. M.: Дрофа.

Baker, C. 2011. *Foundations of bilingual education and bilingualism* 79. Bristol: Multilingual matters.

Balota, D. A., Pilotti, M., & Cortese, M. J. 2001. 'Subjective frequency estimates for 2,938 monosyllabic words'. *Memory & cognition,* 29: 639-647.

Baron, N. S. 2003. 'Language of the Internet'. *The Stanford handbook for language engineers*, 59-127.

Baumann, Z. 2000. *Liquid modernity*. Cambridge: Polity

Becker, A. L. (1991). Language and languaging. *Language and communication*, 11 (1–2), 33–35.

Blommaert, J. 2010. *The sociolinguistics of globali*sa*tion*. Cambridge: Cambridge University Press.

Canagarajah, S. 2011. 'Translanguaging in the classroom: Emerging issues for research and pedagogy'. *Applied Linguistics Review,* 2: 1-28.

Carmel, E., Crawford, S., & Chen, H. 1992. 'Browsing in hypertext: A cognitive study'. *Systems, Man, and Cybernetics,* 22: 865-884.

Carston, R. 2002. *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Wiley-Blackwell.

Carston, R., & Powell, G. 2006. 'Relevance theory: New directions and developments'. *Mind & Language* 341-360.

Carston, R. 2010. 'Lexical pragmatics, ad hoc concepts and metaphor: From a relevance theory perspective'. *Italian Journal of Linguistics,* 22(1): 153-180.

Cenoz, J., & Santos, A. 2020. 'Implementing pedagogical translanguaging in trilingual schools'. *System*, doi: https://doi.org/10.1016/j.system.2020.102273.

Chaffin, R., Morris, R. K., & Seely, R. E. 2001. 'Learning new word meanings from context: A

study of eye movements'. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1): 225.

Chaffin, R., Morris, R. K., & Seely, R. E. 2001. 'Learning new word meanings from context: A study of eye movements'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1): 225–235.

Charmaz, K. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*. New York: Sage.

Chen, S. Y. 2014. 'From OMG to TMD–Internet and Pinyin acronyms in Mandarin Chinese'. *Language@ Internet* 11(3).

Christison, M., & Murray, D. E. 2020. 'An overview of multilingual learners' literacy needs for the 21st century'. *Handbook Of Research On Cultivating Literacy In Diverse And Multilingual Classrooms*, 1-21.

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. 1990. 'Word familiarity and frequency in visual and auditory word recognition'. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 16(6): 1084.

Corbin, J. M., & Strauss, A. 1990. 'Grounded theory research: Procedures, canons, and evaluative criteria'. *Qualitative Sociology,* 13(1): 3-21.

Corbin, J., & Strauss, A. 2008. 'Qualitative research'. *Techniques and procedures for developing grounded theory*. 3.

Crampton, J. 1992. 'A cognitive analysis of wayfinding expertise'. *Cartographica,* 29(3/4): 46–65.

Creese, A., & Blackledge, A. 2010. 'Translanguaging in the bilingual classroom: A pedagogy for learning and teaching?'. *The Modern Language Journal,* 94(1): 103-115.

Crystal, D. 2006. *Language and the internet (2nd ed.)*. Cambridge: Cambridge University Press.

Darken, R. P., & Sibert, J. L. 1996. 'Navigating large virtual spaces'. *International Journal of Human-Computer Interaction,* 8(1): 49–71.

Doran, M. 2004. 'Negotiating between Bourge and Racaille: Verlan as youth identity practice in suburban Paris'. *Negotiation of Identities in Multilingual Contexts,* 45: 93-124.

Dovchin, S. 2016. 'Multilingual wordplays amongst Facebook users in Mongolia: Crossing languages to play with words'. *Multidisciplinary Perspectives*, 97-113.

Dolgunsöz, E. 2018. 'Word familiarity effects in EFL reading: An eye tracking study'. *International Online Journal of Education and Teaching,* 5(2): 252-265.

Dürscheid, C., & Siever, C. M. 2017. 'Jenseits des Alphabets–Kommunikation mit Emojis'. *Zeitschrift für germanistische Linguistik,* 45(2): 256-285.

Ericsson, K. A., & Simon, H. A. 1993. *Protocol analysis: Verbal reports as data*. Boston: The MIT Press.

Ferraro, F. R., & Sturgill, D. 1998. 'Lexical effects and lexical properties associated with National Adult Reading Test (NART) stimuli in healthy young adults and healthy elderly adults'. *Journal of Clinical Psychology,* 54(5): 577-584.

García, O. 2009. *Bilingual education in the 21st Century: A Global perspective.* Malden, MA: John Wiley & Sons.

Garcia, O. & Li, W. 2014. *Translanguaging: Language, bilingualism and education*. Basingstoke: Palgrave Macmillan.

García, O., Flores, N., Seltzer, K., Wei, L., Otheguy, R., & Rosa, J. 2021. 'Rejecting abyssal thinking in the language and education of racialised bilinguals: A manifesto'. *Critical Inquiry in Language Studies,* 18(3): 203-228.

García, O., & Kano, N. 2014. 'Translanguaging as process and pedagogy: Developing the English writing of Japanese students in the US'. *The Multilingual Turn in Languages Education: Opportunities and Challenges*, 258-277.

Glaser, B. G. 1978. *Theoretical sensitivity: Advances in methodology of grounded theory*. Mill Valley, CA: Sociological Press.

Glaser, B.G. 1998. *Doing grounded theory: Issues and discussions*. Mill Valley, CA: Sociology Press.

Glaser, B. G., & Strauss, A. L. 1967. *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine Publishing Company.

Godin, M. N. 2006. 'Urban youth language in multicultural Sweden'. *Scandinavian-Canadian Studies,* 16: 126-141.

Gontšarova, J. 2013. *Neologisms in Modern English: study of word-formation processes*. PhD Thesis.Narva College of the University of Tartu, Estonia.

Gordon, P. 1985. 'Level-ordering in lexical development'. *Cognition*, 21(2): 73-93.

Grice, G. R. 1967. *The grounds of moral judgement*. Cambridge: Cambridge University Press.

Grice, H. P. 1989. *Studies in the way of words*. Boston: Harvard University Press.

Haspelmath, M. 2009. 'Lexical borrowing: Concepts and issues'. *Loanwords in the world's languages: A comparative handbook,* 35-54.

Hill, J. R., & Hannafin, M. J. 1997. 'Cognitive strategies and learning from the World Wide Web'. *Educational Technology Research & Development,* 45(4): 37–64.

Jourdenais, R., Ota, M., Stauffer, S., Boyson, B. & Doughty, C. 1995. Does textual enhancement promote noticing? A think aloud protocol analysis. In R. Schmidt (ed.), *Attention and awareness in foreign language learning*. Honolulu, HI: University of Hawaii.

Juhasz, B. J., & Rayner, K. 2003. 'Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading'. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 29(6): 1312.

Kabatek, J. 2015. 'Wordplay and discourse traditions'. *Wordplay and Metalinguistic/ Metadiscursive Reflection. Authors, Contexts, Techniques, and Meta-Reflection*. 213-228.

Knospe, S., Onysko, A., & Goth, M. 2016. *Crossing languages to play with words:* M*ultidisciplinary perspectives*. Walter de Gruyter GmbH & Co KG.

Lehrer, A. 2007. 'Blendalicious'. In J. Munat (ed.), *Lexical Creativity, Texts and Contexts*. Amsterdam: John Benjamins Publishing. 115–133.

Leppänen, S., Pitkänen-Huhta, A., Piirainen-Marsh, A., Nikula, T., & Peuronen, S. 2009. 'Young people's translocal new media uses: A multiperspective analysis of language choice and heteroglossia'. J*ournal of Computer-mediated communication,* 14(4): 1080-1107.

Lewis, G., Jones, B., & Baker, C. 2012. 'Translanguaging: Origins and development from school to street and beyond'. *Educational Research and Evaluation,* 18(7): 641-654.

Li. W. 2011. 'Moment analysis and translanguaging space: Discursive construction of identities by multilingual Chinese youth in Britain'. *Journal of Pragmatics,* 43(5): 1222-1235.

Li, W. 2016a. 'New Chinglish and the post-multilingualism challenge: Translanguaging ELF in China'. *Journal of English as a Lingua Franca,* 5(1): 1–25.

Li, W. 2016b. 'Multi-competence and the translanguaging instinct'. In V. Cook & W. Li (Eds), *The Cambridge handbook of multi-competence*. Cambridge: Cambridge University Press. 533–43.

Li, W. 2018. ‘Translanguaging as a practical theory of language’. *Applied linguistics,* 39(1): 9-30.

Li, W., 2021. Translanguaging as a political stance: Implications for English language education. *ELT J.* , 1-11.

Li, W. 2022. ‘Translanguaging as method’. *Research Methods in Applied Linguistics,* 1(3): 100026.

Li, W., Tsang, A., Wong, N., & Lok, P. 2020. ‘Kongish Daily: Researching translanguaging creativity and subversiveness’. *International Journal of Multilingualism,* 17(3): 309-335.

Li, W. & Zhu, H. 2013. ‘Translanguaging identities and ideologies: Creating transnational space through flexible multilingual practices amongst Chinese university students in the UK.’ *Applied Linguistics,* 34: 516–535.

Li, W., & Zhu, H. 2019. ‘Tranßcripting: playful subversion with Chinese characters’. *International Journal of Multilingualism,* 16(2): 145-161.

Lipka, L. 2002. *English lexicology: Lexical structure, word semantics & word-formation*. Gunter Narr Verlag.

Lloyd, M. 2007. ‘Rear gunners and troubled privates: Wordplay in a dick joke competition’. *Journal of Sociolinguistics*, 11(1): 5-23.

Ma, L., Ma, Y., He, X., Liu, H., & Zhang, J. 2019. ‘Processing of Chinese homophonic two-part allegoric sayings: Effects of familiarity and homophone’. *Acta Psychologica Sinica,* 51(12): 1306.

Monsefi, R., & Mahadi, T. S. T. 2016. ‘Wordplay in English online news headlines’. *Advances in Language and Literary Studies*, 7(2), 68-75.

Moreno, R. 2007. *Creativity and convention: The pragmatics of everyday figurative speech.* John Benjamin Publishing Company.

Nunberg, G. 2004. ‘The Pragmatics of Deferred Interpretation’. *The handbook of pragmatics,* 15: 344.

O’Dell, F. 2016. ‘Creating new words: Affixation in neologisms’. *ELT Journal,* 70(1): 94-99.

O'Neill, M., & Roberts, B. 2019. *Walking methods: Research on the move*. London: Routledge.

Ortega Y Gasset, J.1957. ‘What People Say: Language. Toward a New Linguistics’. In *Man and People*, *Ch. II*, edited by W. R. Trask, 222–257. W.W. Norton.

Ortner, L. 2013. 'Visuell markierte Wortbildungen'. Plädoyer für eine visionslinguistische Nominationsforschung. *Wenn die Ränder ins Zentrum drängen. Außenseiter in der Wortbildung (sforschung)*. Berlin: Frank & Timme 43-83.

Partington, A. 2009. 'A linguistic account of wordplay: The lexical grammar of punning'. J*ournal of Pragmatics*, 41(9): 1794-1809.

Patton, M. Q. 2002. 'Two decades of developments in qualitative inquiry: A personal, experiential perspective'. *Qualitative Social Work,* 1(3): 261-283.

Peprník, J. 2006. *English Lexicology*. Olomouc: Univerzita Palackého v Olomouci.

Qi, F., & Zhang, K. 2020. 'Translanguaging hybrids on Chinese gateway websites'. *Asian Englishes,* 1-14.

Ren, W. & Guo, Y. 2022. 'Translanguaging in self-praise on Chinese social media'. *Applied Linguistics Review*, 1-22. /doi.org/10.1515/applirev-2021-0169.

Ren, Guo, & Li, 2022. 'Chinese young people's attitudes towards translanguaging in self-praise on social media'. *Journal of Pragmatics, 198:*93-103.

Renner, V. 2015. 'Lexical blending as wordplay'. *Wordplay and metalinguistic/metadiscursive reflection. Authors, contexts, techniques, and meta-reflection.* 119-133.

Rúa, P. L. 2007. 'Keeping up with the times: lexical creativity in electronic communication'. In J. Munat (Ed.), *Lexical creativity, texts and contexts*. John Benjamins. 137-159. DOI:10.1075/sfsl.58.18lop.

Schoonen, R., & Appel, R. 2005. 'Street language: A multilingual youth register in the Netherlands'. *Journal of Multilingual and Multicultural Development,* 26(2): 85-117.

Setyaningrum, R., Setiawan, S. & Anam, S. 2022. 'Translanguaging as a scaffolded practice in a primary school content and language integrated learning context during the COVID-19 pandemic'. *European Journal of Educational Research*, 11(4): 2043 - 2055.

Smith, J. A., & Osborn, M. 2015. 'Interpretative phenomenological analysis'. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods*. London: Sage. 25–52.

Sperber, D., & Wilson, D. 1986. *Relevance: Communication and cognition (Vol. 142)*. Cambridge, MA: Harvard University Press.

Sperber, D., & Wilson, D. 1995. *Postface to the second edition of Relevance: Communication and Cognition*. London: Blackwell.

Sperber, D., & Wilson, D. 1997. 'Remarks on relevance theory and the social sciences'. *Multilingua Journal of Cross-Cultural and Interlanguage Communication,* 16(2-3): 145-152.

Sperber, D., & Wilson, D. 2002. 'Pragmatics, modularity and mind-reading'. *Mind & Language,* 17: 3-23.

Sperber, D., & Wilson, D. 2005. 'Reply to rajagopalan'. *Intercultural Pragmatics,* 2(1).

Sperber, D., & Wilson, D. 2008. 'A deflationary account of metaphors'. In R. Gibbs (ed.) *The Cambridge Handbook of Metaphor in Language and Thought*. Reprinted in Wilson, D. & Sperber, D. 2012, *Meaning and relevance*. Cambridge. Cambridge University Press. 84-105.

Sperber, D., & Wilson, D. 2012a. Rhetoric and relevance. In Deirdre Wilson and Dan Sperber (Eds.), *Meaning and relevance*. Cambridge: Cambridge University Press. 84-96.

Sperber, D., & Wilson, D. 2012b. A defationary account of metaphors. In Deirdre Wilson and Dan Sperber (Eds.), *Meaning and Relevance*. Cambridge: Cambridge University Press. 97-122.

Sultana, S. 2014. 'Heteroglossia and identities of young adults in Bangladesh'. *Linguistics and Education* 26: 40-56.

Tai, K. & Li, W. 2020. Constructing playful talk through Translanguaging in English medium instruction mathematics classrooms. *Applied Linguistics*, 1–34. doi:10.1093/applin/amaa043.

Tanaka, K. 1992. 'The pun in advertising: A pragmatic approach'. *Lingua*, 87(1–2): 91-102.

Thaler, V. 2012. *Sprachliche Höflichkeit in computervermittelter Kommunikation (Vol. 70)*. Stauffenburg-Verl..

Thornberg, R., Pozzoli, T., Gini, G., & Jungert, T. 2015. 'Unique and interactive effects of moral emotions and moral disengagement on bullying and defending among school children'. *The Elementary School Journal,* 116(2): 322-337.

Tebaldi, C. 2019. 'Mock translanguaging and the production of whiteness'. *Translinguistics: Negotiating Innovation and Ordinariness,* 4.

Unicode Standard Version 12.1. 2019. http://www.unicode.org

Vogel, S., & García, O. 2017. 'Translanguaging'. In G. Noblit & L. Moll (eds.), *Oxford research encyclopedia of education*. Oxford: Oxford University Press.

Walaszewska, E. 2010. 'Chapter three narrowing and taboo contamination: relevance theory and euphemisms'. In *Mind and across Minds: A Relevance-Theoretic Perspective on Communication and Translation,* 61.

Whalen, D. H., & Zziga, E. 1994. 'Subjective familiarity of English word/name Homophones'. *Behavior Research Methods, Instruments, and Computers*, 26: 402–408.

Wilson, D. 2004. 'Relevance and lexical pragmatics'. *UCL Working Papers in Linguistics,* 16: 343-360.

Wilson, D., & Carston, R. 2006. 'Metaphor, relevance and the 'emergent property' issue'. *Mind & Language,* 21: 404-433.

Wilson, D. , & Carston, R. 2007. 'A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts'. In N. Burton-Roberts (Ed.), *Pragmatics*. *(Palgrave Advances in Linguistics)*. Basingstoke: Palgrave Macmillan. 230-259

Wilson, D., & Sperber, D. 2012a. 'Explaining irony'. In D. Wilson,& D. Sperber (Eds.), *Meaning and relevance.* Cambridge: Cambridge University Press. 123-145.

Wilson, D., & Sperber, D. 2012b. 'Linguistic form and relevance'. In D. Wilson & D. Sperber (Eds.), *Meaning and relevance.* Cambridge: Cambridge University Press. 149-168.

Williams, R., & Morris, R. 2004. 'Eye movements, word familiarity, and vocabulary acquisition'. *European Journal of Cognitive Psychology,* 16(1-2): 312-339.

Yang, X. U. 2007. 'On numerals in words and expressions of *Jinpingmei*'. *Journal of Henan University (Social Science),* 1-22.

Yuan, E. J. 2011. 'Language as social practice on the Chinese internet'. In K. Amant (Ed.), *Computer-mediated communication across cultures: International interactions in online environments*. *IGI Global*, 266– 281.

Yayli, D. 2010. A Think-Aloud study: Cognitive and metacognitive reading strategies of ELT department students. *Eurasian Journal of Educational Research*, 10: 234-251.

Yus, F. 2017. 'Contextual constraints and non-propositional effects in WhatsApp communication'. *Journal of Pragmatics,* 114: 66-86.

Yus, F. 2022. *The pragmatics of smartphone communication: Interactions in the App ecosystem*. Abingdon: Routledge.

Zhang, Y. 2017. 'The semiotic multi functionality of Arabic numerals in Chinese online

discourse'. *Language@Internet* 14(2).

Zhang, Y. 2020. 'Adopting Japanese in a popular Chinese video-sharing website: Heteroglossic and multilingual communication by online users of Bilibili.com'. *International Multilingual Research Journal,* 14(1): 20-40.

Zhang, Y., & Ren, W. 2022. ''This is so skrrrrr'–creative translanguaging by Chinese micro-blogging users'. *International Journal of Multilingualism*, 1-16.

Zheng, Y. 2015. 'A metaphorical study on Chinese neologisms'. *Journal of Language Teaching and Research,* 6(6): 1379-1383.

Xinlang Weibo Data Center. 2021. *Weibo 2020 Annual report on user development*. https://data.weibo.com/report/reportDetail?id=456

Appendix A TSE Random Selection Progamme

```
import cv2
import os
import random

base_gallery = ['Affixation', 'Complex', 'Composition', 'External_TEs',
'Morphological_TEs', 'Phonetic_Motivation', 'Semantic_TEs']
gallery2img = {}
for name in base_gallery:
    gallery2img[name] = os.listdir(name)
real_gallery = base_gallery * 3
random.shuffle(real_gallery) ## 目录打乱
for item in gallery2img:
    random.shuffle(gallery2img[item]) ## 图片打乱

cot = 1
while len(real_gallery):
    print('当前次数: ', cot)
    select_menu = real_gallery.pop(0)
    print('选择了目录: ', select_menu)
    select_pic = gallery2img[select_menu].pop(0)
    print('选择了图片: ', select_pic)
    name = select_pic
    cv2.namedWindow(name)
    cv2.moveWindow(name, 600, 300)
    # cv2.setWindowProperty(name, cv2.WND_PROP_FULLSCREEN,
cv2.WINDOW_FULLSCREEN)
    cv_show = cv2.imread(os.path.join(select_menu, select_pic))
    cv2.imshow(name, cv_show)
    cv2.waitKey(0)
    cv2.destroyAllWindows()
    print('按下空格进入下一张图片...')
    cot += 1
```

Appendix B Ethical Approval From UCL Institute of Education

Doctoral Student Ethics Application Form

Anyone conducting research under the auspices of the Institute of Education (staff, students or visitors) where the research involves human participants or the use of data collected from human participants, is required to gain ethical approval before starting. This includes preliminary and pilot studies. Please answer all relevant questions in simple terms that can be understood by a lay person and note that your form may be returned if incomplete.

Registering your study with the UCL Data Protection Officer as part of the UCL Research Ethics Review Process

If you are proposing to collect personal data i.e. data from which a living individual can be identified **you must be registered with the UCL Data Protection Office before you submit your ethics application for review**. To do this, email the complete ethics form to the UCL Data Protection Office. Once your registration number is received, add it to the form* and submit it to your supervisor for approval. If the Data Protection Office advises you to make changes to the way in which you propose to collect and store the data this should be reflected in your ethics application form.

Please note that the completion of the UCL GDPR online training is mandatory for all PhD students.

Section 1 – Project details

a. Project title: Understanding Translanguaging expressions on Chinese social media (Weibo): A Relevance-Theoretic Account

b. Student name and ID number (e.g. ABC12345678): Chen Tiancheng 17050183

c. *UCL Data Protection Registration Number: Z6364106/2020/11/47

   a. Date Issued: 16/11/2020

d. Supervisor/Personal Tutor: Li Wei

e. Department: Institute of Education

f. Course category (Tick one):

PhD ☒

EdD ☐

DEdPsy ☐

g. **If applicable**, state who the funder is and if funding has been confirmed.

Bloomsburry Colleges Scholarship

h. Intended research start date: 21/05/2022

i. Intended research end date: 28/05/2022

j. Country fieldwork will be conducted in: the UK

k. If research to be conducted abroad please check the Foreign and Commonwealth Office (FCO) and submit a completed travel risk assessment form (see guidelines). If the FCO advice is against travel this will be required before ethical approval can be granted: *UCL travel advice webpage*

l. Has this project been considered by another (external) Research Ethics Committee?

Yes ☐

External Committee Name: Enter text

Date of Approval: Enter text

No ☒ ***go to Section 2***

If yes:

- Submit a copy of the approval letter with this application.
- Proceed to Section 10 Attachments.

Note: Ensure that you check the guidelines carefully as research with some participants will require ethical approval from a different ethics committee such as the National Research Ethics Service (NRES) or Social Care Research Ethics Committee (SCREC). In addition, if your research is based in another institution then you may be required to apply to their research ethics committee.

Section 2 – Research methods summary (tick all that apply)

☒ Interviews

☐ Focus Groups

☐ Questionnaires

☐ Action Research

☐ Observation

☐ Literature Review

☐ Controlled trial/other intervention study

☐ Use of personal records

☐ Systematic review – **if only method used go to Section 5**

☐ Secondary data analysis – ***if secondary analysis used go to Section 6***

☐ Advisory/consultation/collaborative groups

☒ Other, give details: Think aloud protocols

Please provide an overview of the project, focusing on your methodology. This should include some or all of the following: purpose of the research, aims, main research questions, research design, participants, sampling, data collection (including justifications for methods chosen and description of topics/questions to be asked), reporting and dissemination. Please focus on your methodology; the theory, policy, or literary background of your work can be provided in an attached document (i.e. a full research proposal or case for support document). *Minimum 150 words required.*

This study includes a pilot study and a main study. The aim of the pilot study is to test whether the research design is feasible. The aim of the main study is to reveal how readers comprehend Translanguaging expressions in Chinese Weibo posts. Before conducting the think aloud study, I have collected and categorised 6,500 TEs from Weibo Super-topic. Then I plan to run an experimental study to investigate readers' understanding processes when they comprehend the meaning of TEs that are formed based on different types of word formation processes. The hypothesis of this task was that people may employ different ways to understand different types of TEs. This method may illustrate the nuanced differences in the processes of understanding different TEs. To ensure the research design of the main study will minimise the influence of the pariticipants' understanding processes, a pilot study will be conducted before the main study.

Main research questions:

1. What types of TEs are used in digital communication contexts, and when? This involves a detailed account of the ways TEs are constructed, and the contexts in which they appear.

2. How do people interpret novel TEs in different communicative contexts? And do they follow the steps of interpretation proposed in Relevance Theory? Does the level of familiarity with different types of TEs matter, and how?

1. Pilot Study

1.1 Participants

To ensure the validity of the study, I will run a pilot think-aloud study on 5 participants (with 2 males and 3 females), who are not included in the participants that would take part in the final study (10 male and 10 female students). These participants are all Chinese Weibo users, who ages between 20 and 30. According to the *Weibo Users' Development Annual Report 2020*, the main users of Weibo are born after 1990 (48%) and 2000 (30%). Therefore, I will recruit participants who can represent the majority of Weibo users. The recruitment does not take into consideration the participants' major, job, or gender. What mattered is that all of them use social media on regular or even daily basis. When I recruit the pilot study's participants, I will invite my friends and former classmates. This choice can bring great convenience to the selection of participants. Because firstly, I can easily contact them and arrange time slots to do the task; secondly, they will not be nervous when talking with me during the think aloud tasks as we are very familiar with each other; thirdly, I can choose those who are talkative in daily life in advance, so that they may provide better quality think aloud data when sharing their mind.

All the 5 participants volunteered to anonymously take part in the pilot study and all their personal information will be carefully protected. They will be rewarded with one-year membership of Weibo for participating in the project, which amounts to the value of 118 CNY for one year. Considering that all the participants' first language is Chinese, to ensure they can best express themselves during the think-aloud task without any language difficulties, the pilot study will be done in Chinese. The pilot study will not record any of the participants' audio and video of thinking aloud. The purpose of this pilot study is to test the designed procedures so that the researcher can avoid unexpected problems during the main study.

1.2 Pilot Procedure

1.2.1 Introduction and trial task

Before the pilot study, I will give the participants instructions about the think-aloud method in Chinese one by one before the task, followed up with a training on how to participate in a think-aloud task, including a training video clip and a trial task. The training material is made into a video, which contains two sections. In the first section, I orally introduce the concept and definition of think-aloud in Chinese with explanations shown on the screen (in the format of PowerPoint). The second section is a video record example of a participant who is doing a reading aloud task. By watching the instructional video, the participants can have a vivid idea of what is think-aloud and what they need to do during the think-aloud task. Then a Q&A section is provided for each participant, if they still feel confused on any details. Then I go on with some specific requirements about the upcoming task.

To help the 5 participants to get familiar with the think-aloud task, I then invite them to conduct a trial task after they watch the training video. To avoid practice effect, I used 2 Chinese expressions that are once very popular among Chinese young people, namely "喜大普奔" (*xi da pu ben*) and "奥利给" (*ao li gei*). These two Chinese expressions are not Translanguaging expressions. "喜大普奔" (*xi da pu ben*) is a contracted form of four Chinese idioms, namely "喜闻乐见", "大快人心", "普天同庆" and "奔走相告", meaning that people should share the great news and celebrate it together. In its formation, the first four character of each idiom is picked out and then combined as a new expression. People endow a stronger and richer meaning to this expression as it represents four different but similar idioms. "喜闻乐见" literally means happy to hear and see (something happens). "大快人心" literally means something makes people feel extremely happy sincerely. "普天同庆" means people who are under the same sky celebrate for something together. "奔走相告" means people rush about telling others a good news. All these idioms illustrate people's reactions when something awesome happens vividly. So wtih these four

idioms added together as an acronym, “喜大普奔” can express a “four-times” stronger joyful feeling. “奥利给”, however, is not a Chinese expression that can be found in Chinese dictionary.

Unlike “喜大普奔”, which is formed based on other existing Chinese idioms, “奥利给” is a recently created utterance-final particle used to cheer people up. It was first used by a Chinese streamer who loved to spread positive energy to his audience. Every time when he finished uttering an inspiring sentence, he would add an “奥利给” to end his speech. For example, “one shall never give up when life beats you! Fight back and overcome all the barriers! Come on, 奥利给”. The streamer always used a very powerful and loud voice to articulate “奥利给”, which left people a deep impression.

“喜大普奔” is an expression that follows a certain word-formation rule (clipping), while “奥利给” can only be understood through the context. These two Chinese expressions can be understood based either on their formation or context; however, neither of them belongs to the category of TEs as they do not contain other semiotic elements in their forms apart from Chinese and they do not need other semiotic elements to help understand their meaning. Thus, “喜大普奔” and “奥利给” can be used to help the participants to get familiar with thinking aloud their understanding processes of special expressions without causing a practice effect on their processes of understanding the target TEs.

1.2.2 Think aloud task materials

During the think aloud task, each participant was requested to understand 18 different TEs with 3 different TEs from 6 sub-types. The TEs used in the pilot study of think aloud task are not restricted to those in the data collected from Weibo super-topic. They are randomly selected from Weibo. The TEs used for the pilot study

fall into 4 main groups and 6 sub-types on the basis of the outcomes of my typological analysis that have been completed before the pilot study, as shown in Table 3.1. For each sub-type of TEs, I choose 6 different TEs with different contexts. In all, 36 TEs are selected.

Table 3.1 Examples of TEs used in the pilot study

TE types		**Test items**
morpho-semantic TEs	affixation	*ungeilivable, niubility*
	composition	*打 call，气 skr 人*
	phonetic motivation	 *, 666*
semantic TEs		*华粉，* 
morphological TEs		*xswl, cpdd*
external TEs		*买米当卡，欧克*

The TEs used for the pilot study are not presented in isolation to the 5 participants but rather they occurr in the context of the original Weibo posts. In each Weibo post, there is only one TE, so that the participants will not be confused about which TE they should understand. Some TEs used in this task might belong to more than one types but they were still used. Many TEs involved different types of word formation methods, for example, *Book 思议* belongs to both the category of phonetic motivation and composition. It is very difficult to find enough TEs for each type that only involved one word formation sub-type. Therefore, if a TE involves more than one formation methods, it can only be put in one sub-type group once. For instance, I put *book 思议* into the group of composition so that it will not be put in phonetic motivation any more. Readers may apply one or both formation strategies when understanding the TEs meaning. So with each participant working with 18 TEs from six different types, in the main study, all participants will understanding enough

amout of TEs. Therefore, there will still be enough data of each sub-type of TEs to support the analysis of the understanding processes and patterns.

Another consideration is that not all types of word formation could include enough examples. Types like affixation and composition can only find five to six examples of TEs, while the number of phonetic motivated TEs is huge. So for each group, I selected 6 TEs. However, not all of the 36 TEs will be used in the think aloud task. Each participant will complete the think aloud task with 6 TEs from different sub-types. TEs under the same sub-type are comparable in their word formation process. For example, all TEs under the type of "affixation" are formed through using affixations. even though they are different TEs, they all contain affixations in them. I will introduce how I selected target TEs from the 36 TEs for the participants to complete the task later. Here are 6 examples of the posts in which test items occurred (The personal information of the users has been pixelated for ethical consideration).

Example 1 ungeilivable

薛教授这一季的镜头，也太太太少了吧……求多切几个嘉宾的镜头补给教授吧……反正这一季的嘉宾个个ungeilivable的…

[**Translation**: *Professor Xue has so little footage in this season… Please provide more footage for Professor Xue. Since other guests in the season are all ungeilivable.*]

[**Annotation**: The expression of ungeilivable refers to not being helpful/awesome enough. Compared with Professor Xue, other guests' performance in the TV show was mediocre.]

Example 2 *666*

[**Translation**: *It's not going well this afternoon. Haven't won a single game. In this game I used the champion of Zed. I got the highest kill. When my teammates were all terminated, I managed to kill Jax (full health) with only one quarter of my health. Two of my teammates said "**666**" and the other said "nb". However, it's still useless. We still lost the game. All my teammates' KDAs are negative and they contributed little damage to the game. What was left to us was 持て余す鼓動に (for the heartbeat that I cannot hold).*]

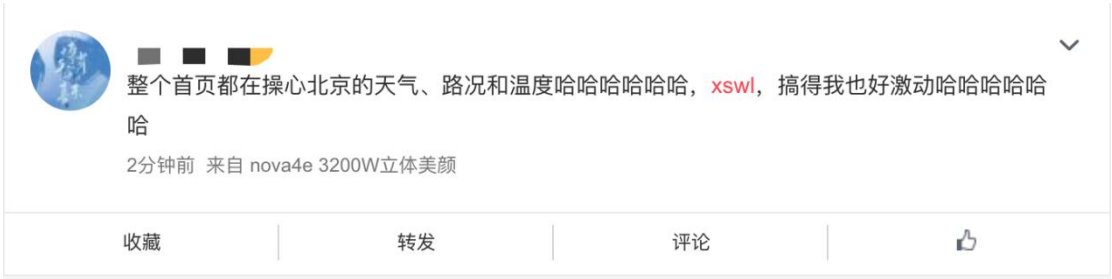
[**Annotation:** The *666* here conveys praise. The author performed well in a League of Legend game.]

Example 3 *气skr人*

[**Translation**: *# Eliminated# What a pity for the team to be eliminated! They did well in the two games. However, G's team lost due to the misjudgment in last episode. If they are eliminated in this episode, it can really* ***气 skr 人****(make me piss off). #This is Slam Dunk#*]

[**Annotation:** The TE of *气 skr 人* in the example can be rephrased as "气死个人", meaning "making someone piss off". "skr" inside the TE is interpreted as "死个" instead of "是个" or other words because the interpretation is consistent with the speaker's intended meaning for the whole utterance (i.e., the author believed that it was utterly unfair for the team to be eliminated). Thus, the interpretation of the TE in this way will convey not only an emphasis (which "气死个人" can also achieve), but also some additional effect, because the processing of "skr" will call forth the fun story of Wu Yifan, a pop star, who was involved in awkward scandal.]

Example 4 *xswl*



[**Translation**: *The whole homepages (of the Micro-blog) are worried about the weather, traffic and temperature of Beijing hahahahahaha,* ***xswl****, which makes me feel so excited too.*]

[**Annotation:** In this case, *xswl* is interpreted as the acronym of the Chinese expression of "笑死我了", which means something is so funny and is killing the hearer. This interpretation is in line with the context, unlike other potential interpretations such as "学生未来" or "雪山无路". The use of *xswl* is more effective

than its direct equivalent “笑死我了” in that it saves typing effort and looks exotic as well.]

Example 5 *花粉*

[**Translation**: *As a* ***花粉****, all I want to say is @华为终端官方微博 (The Official Weibo Account of Huawei Terminals) that I love every product (you presented).*]

[**Annotation**: **花粉** (hua fen) refers to the fans of Huawei. **花粉** (hua fen) is the homophone of 华粉 (hua fen), which is the Chinese acronym of 华为的粉丝 (the fans of Huawei).]

Example 6 *欧克*

[**Translation**: *I drew my loved boy in spare time. I've been a fan of him for seven years. Fangfangzi is the most handsome guy. Do you think the painting is* ***欧克*** *(OK)?*]

[**Annotation**: In this example, the TE expression of *欧克* is written in the form of Chinese characters but actually its interpretation depends on one's knowledge of the English word "OK" as well as its pronunciation. The innovativeness of the TE conveys additional special effects because it builds up a homophonic effect.]

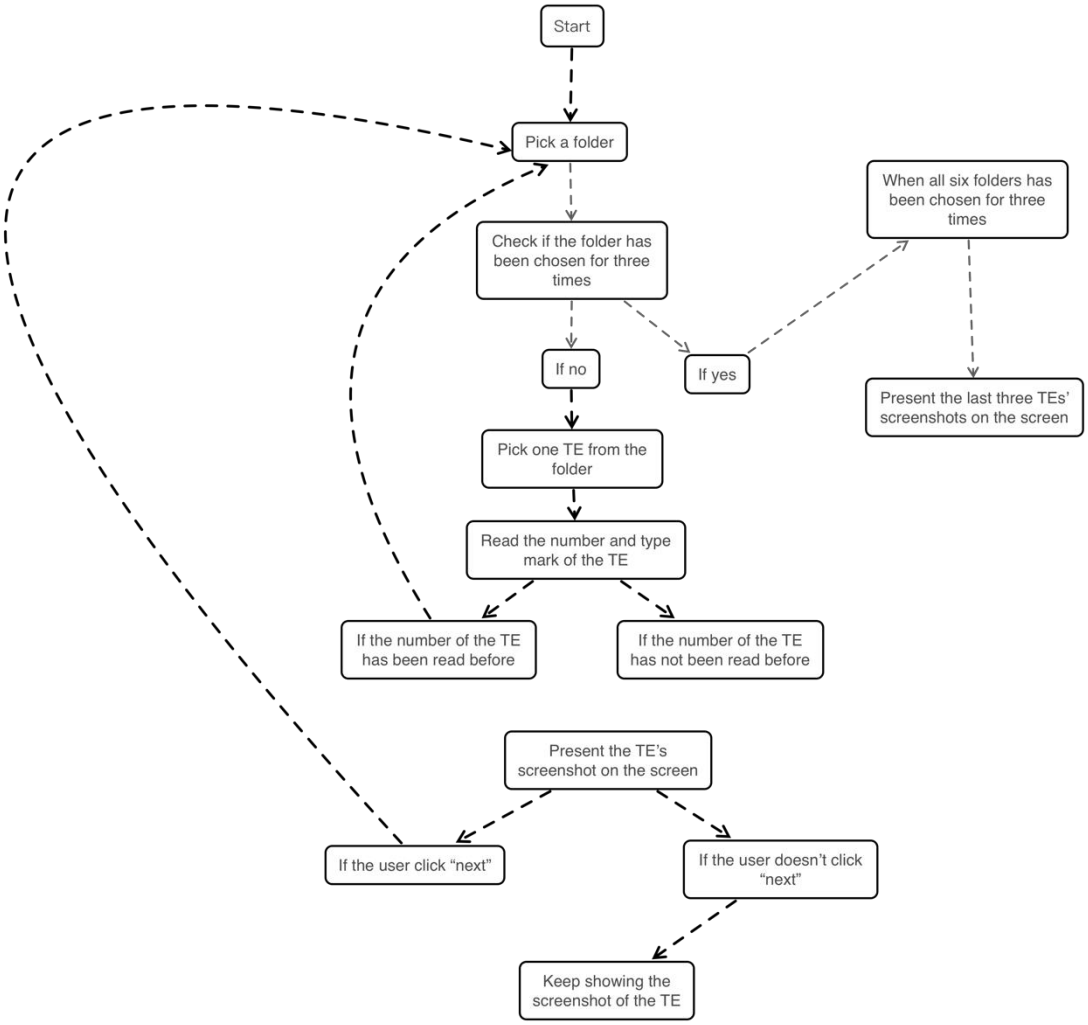
Here are some interesting situations that need to be considered. First of all, during the task, the participants might come across TEs that they had already known before the task. This situation might lead to a very short process of understanding. In this study, all kinds of situations were welcomed. Even if the participant knew the meaning of the selected TE, it was fine because this situation could reflect a more natural and real outcome that might appear in the participant's daily life. So before asking the participants to think aloud the TE's meaning, I will first ask the participant whether they know the TE before. There was still a little chance that the participant knew all 18 TEs presented to him or her. But by recruiting 5 participants, the chance would be reduced a lot, getting alone there would be 20 participants in the formal task.

Secondly, since all TEs were randomly selected from Weibo, the Weibo writers might provide different levels of background information or context that might support the understanding of TEs. In some extreme cases, the participants might come across TEs with no context at all. For example, in the case of *cpdd*, there were many Weibo posts containing nothing else but a string of the four letters, so that the participants had no contextual support. In example 4 listed above, after using *xswl*, the writer wrote many *哈* s, which is an onomatopoeia of laughing. In this case, the participant could get some related information from the context. In some other examples, the writer may directly provide both the TEs and their meaning at the same time. For instance, a writer mentioned both *cxk* and *蔡徐坤* (cxk) in her Weibo. *Cxk*

is the Pinyin initialism of *蔡徐坤*, which became an obvious hint for the participants to understand the TE. As we can see, in different cases, the participants may have different levels of contextual information for support while understanding the TEs' meaning, which may influence their understanding processes to certain degrees. According to relevance theory (Sperber & Wilson 1986/1995a), when interpreting the meaning of an utterance, the reader will turn to the context to find the most relevant explanation. For this reason, if the context information is limited, the readers may get into trouble when they need information support from the context. However, it is very difficult to evaluate the exact level of support that each context provided to the target TEs. Firstly, different readers may not rely on the context similarly to understand TEs. That is to say, even though a context may be enough for Reader A, it may still be insufficient for Reader B. Secondly, it is very difficult to find enough original TEs with relatively the same level of contextual information. Thirdly, some contextual information is time-sensitive and social-sensitive. For example, if a Weibo containing a TE is posted when a social event happens, it may become very difficult for a reader to relate the social background to the Weibo many months later. Therefore, in this study, I planned to minimise the influence of the level of contextual information support rather than take it as a variable to measure in the task at the very outset. In effect, with 20 participants involved in the main study, the outcome of the think aloud task will cover TEs with different levels of contextual support. This will be an interesting topic that can be discussed in the discussion chapter.

To minimise the influence of the two factors mentioned above (whether the participants know the TE before and how much information support they can get from the context), I use Python to design a small programme to automatically and randomly select target TEs for the participants to work with. As mentioned earlier, I collect 6 different TEs for the six sub-types with their own Weibo context. There are in all 36 TEs being selected, screenshot, and stored in the form of .PNG. The file name of each TE is created in the form of "serial number_TE's name". For instance, "00_xswl". I create six folders for the six sub-types with each of them containing the screenshots of

six different TEs. When I start the programme, the programme will randomly pick one folder and select one screenshot of TE from the folder. Each folder can only be picked for three times. If one folder has been picked for three times, the programme will pick from the rest folders that have not been picked for three times. In each folder, one serial number can only be picked for once. Then, on the screen, the screenshot of the Weibo will be presented to the participant. When the participant finishes the think-aloud task of understanding this TE, I will click the "next" button, and the programme will continue to randomly pick a new TE. The programme will stop automatically when all six folders have been picked for three times. Here is the logic flow diagram of the programme:



Here is the python codes that are written to achieve the above purpose:

```
import cv2
import os
import random
```

```
base_gallery = ['Affixation', 'Complex', 'Composition', 'External_TEs',
'Morphological_TEs', 'Phonetic_Motivation', 'Semantic_TEs']
gallery2img = {}
for name in base_gallery:
    gallery2img[name] = os.listdir(name)
real_gallery = base_gallery * 3
random.shuffle(real_gallery) ## 目录打乱
for item in gallery2img:
    random.shuffle(gallery2img[item]) ## 图片打乱

# print(real_gallery)
# print(gallery2img)

cot = 1
while len(real_gallery):
    print('当前次数: ', cot)
    select_menu = real_gallery.pop(0)
    print('选择了目录: ', select_menu)
    select_pic = gallery2img[select_menu].pop(0)
    print('选择了图片: ', select_pic)
    cv_show = cv2.imread(os.path.join(select_menu, select_pic))
    cv2.imshow(f"Word Formation [{cot}]", cv_show)
    cv2.waitKey(0)
    cv2.destroyAllWindows()
    print('按下空格进入下一张图片...')
    cot += 1
```

1.3 Data Collection

The pilot study will not record the participants' video or audio. Generally, I asked the participants to report on how they understand the various TEs presented to them one after another. In the pilot study, the researcher will encourage the participants to keep talking and say as much as they can about how they understand the expressions. Specifically, the researcher will tell the 5 participants to talk continuously about how they understand the sample TEs during the task. The researcher will inform them that they will be reminded to keep talking whenever they fall into silence for up to 15s. The researcher will only say "请继续讲" (Keep talking,

please) to break silence. The researcher will not give the participants any hint that might help them understand the target TE. Any type of report from the participants is a meaningful outcome, which will be considered as a possible situation when people might come across TEs. The think-aloud task for each TE will stop when the participants signale they have nothing more to say (by raising their hands up and saying they are done). Then the researcher will move to the next TE expression.

Considering this Think-aloud task of this study is conducted online, the participants' attention is focused on the computer while reading the Weibo posts and doing Think-aloud. Therefore, the researcher can not use light or other visual signals to remind the participants to continue speaking their minds aloud when they become silent. The researcher can use language to remind and encourage the participants as this study focuses on the comprehension process instead of their language. The researcher's words will not influence the way participants talked. As the study focuses on the thinking procedures the participants followed in comprehending the TEs, direct reminder in Chinese will not influence them or provide additional information. The participants may not help asking questions to confirm whether they properly understand the TEs. The researcher will not provide any information related to the comprehension. No hints should be given to help the participants understand the TEs.

This study will use Tencent Meeting as the online communicating tool during the study. Due to network issues, Zoom is unstable in China, so that use Zoom cannot guarantee the success and quality of the recording. In addition, Chinese participants are unfamiliar with Microsoft Teams, and they need to buy the software before using its full functions. During the pandemic, the most popular online meeting and teaching software in Chinese is Tencent Meeting, provided for free by Tencent supporters. Chinese students are generally familiar with this software and will not feel anxious while using it.

Stage 1: Preparation

The researcher confirms a meeting time with each of the 5 participants. For personal reasons, maybe not all of the participants can participate in the task on the same day. It may take 2 days to have all the participants complete the task.

The researcher sends a consent form and information sheet to the participant through email. The researcher explains the general purpose of the research and what the researcher expects them to do. Also, the researcher introduces the procedure of Think-aloud task they are invited to attend, and answers the participant's questions in understanding the task.

The researcher double checks that the participants really understand what they are expected to do. After they have agreed to participate in the think-aloud task, the researcher asks each of them to sign on the consent form (The participants can use e-signature to sign the consent forms) and send it back to the researcher through email.

Then, the reseacher will instruct them to download and install Tencent Meeting on their PC if they have not installed such software, and make sure the participant can use the screen sharing function.

Meanwhile, the researcher prepares the TE Selection Programme to show the Weibo posts to the participants. Before the task starts on the scheduled day, the researcher will check and make sure that the Internet connection on both sides is stable. In addition, both the participant and the researcher are in a quiet and undisturbed environment to ensure the success of recording and the recording quality. The participants can attend the think-aloud task either at home or in any other quiet place where they feel relaxed. They need to open up the Tencent meeting software on their PC and get ready for taking the task.

The researcher needs to contact the participants through platforms like WeChat, QQ, email, etc., and use Tencent Meeting for Think-aloud tasks. In addition, the researcher needs to help the participants deal with all technique problems such as Internet connection, PC software download, etc. These jobs should be done before the day of Think-aloud tasks.

Stage 2: Task performance

The participant is asked to turn on the cameras so that the researcher can have better control over the experimental procedure.

When the researcher receives the signed consent form, the researcher will play the introduction video to the participant and then invite the participant to ask questions if they feel confuse. When the participant reports that he or she is ready, the researcher will inform the participant: "We will begin the Think-aloud task. "

Then the researcher turns on the screen sharing function and starts the TE Selection Programme.

When the Weibo post with the TE is shown on the screen, the researcher asks the participant to start the think aloud task. The researcher informs the participant that the Think-aloud task of this TE begins: "Now you can read the Weibo. When you start reading, please speak all thoughts that appear in your mind aloud. You should follow your reading habits when you surf on Weibo. If you believe that you have completed the understanding, please raise your hand and say *I have completed*. Please try your best to figure out its meaning, but also feel free to quit. When I notice that you are not speaking your thoughts aloud, I will remind you. You do not need to feel stressed when I remind you by say *Keep on talking*. Because it is just an encouragement that reminds you to speak aloud."

When the participant becomes quiet and does not speak, the researcher will directly remind the participant to speak his or her mind aloud: "Don't forget speaking your thoughts aloud. Just speak whatever that appears in your mind."

When the participant informs the researcher that the reading procedure has been done or he or she can no longer figure out the TE's meaning, the researcher will stop the Think-aloud task of this TE and move to the next TE: "Okay, thank you for your cooperation. Let us move to the next TE."

The researcher continues the think aloud task by clicking the “next” button until the programme stops. The researcher tells the participant: "Thank you for your cooperation. We have finished today's Think-aloud task."

Stage 3: After the task

The researcher stops closes Tencent Meeting.

2. Main study

In the main study, I will make some changes based on the experience get from the pilot study. The main procedures of the main study are similar to the pilot study.

Here are some changes that differ from the pilot study:

1) participants

A total of 20 participants will be recruited to join the think aloud task. People who have took part in the pilot study will not be invited to join the main study in case they meet the same TEs they have already done in the pilot. These participants are all Chinese Weibo users, who ages between 20 and 30. According to the *Weibo Users' Development Annual Report 2020*, the main users of Weibo are born after 1990 (48%) and 2000 (30%). Therefore, I will recruit participants who can represent the majority of Weibo users. The recruitment does not take into consideration the participants' major, job, or gender. What mattered is that all of them use social media on regular or even daily basis. When I recruit the pilot study's participants, I will invite my friends and former classmates. This choice can bring great convinence to the selection of pariticipants. Because firstly, I can easily contact them and arrange time slots to do the task; secondly, they will not be nervus when talking with me during the think aloud tasks as we are very familiar with each other; thirdly, I can choose those who are talktive in daily life in advance, so that they may provide better quality think aloud data when sharing their mind.

All the 5 participants volunteered to anonymously take part in the pilot study and all their personal information will be carefully protected. They will be rewarded with one-year membership of Weibo for participating in the project, which amounts to the value of 118 CNY for one year. Considering that all the participants' first language is

Chinese, to ensure they can best express themselves during the think-aloud tasks without any language difficulties, the pilot study will be done in Chinese.

The whole think aloud tasks of the participants will be recorded in the form of screen clips recording with both the researcher and the participant's voice being recorded. The participants will be given a consent form and an information sheet before the think aloud tasks. The main think aloud study of each participant will not start until receiving the participant's signed consent form and information sheet.

2) Data collection

The data collection section in the main study will add an extra follow up interview.

Stage 1: Preparation

The researcher confirms a meeting time with each of the 20 participants. For personal reasons, maybe not all of the participants can participate in the task on the same day. It may take 10 days to have all the participants complete the task.

The researcher sends a consent form and information sheet to the participant through email. The researcher explains the general purpose of the research and what the researcher expects them to do. Also, the researcher introduces the procedure of Think-aloud task they are invited to attend, and answers the participant's questions in understanding the task. The researcher informs the participants that after the think aloud task, there will be short follow-up interview regarding to places where the participants do not talk clearly.

The researcher double checks that the participants really understand what they are expected to do. After they have agreed to participate in the think-aloud task, the researcher asks each of them to sign on the consent form (The participants can use e-signature to sign the consent forms) and send it back to the researcher through email.

Then, the researcher will instruct them to download and install Tencent Meeting on their PC if they have not installed such software, and make sure the participant can use the screen sharing function.

Due to network issues, since Zoom is unstable in China, using Zoom cannot guarantee the success and quality of the recording. In addition, Chinese participants are unfamiliar with Microsoft Teams, and they need to buy the software before using its full functions. During the pandemic, the most popular online meeting and teaching software in Chinese is Tencent Meeting, provided for free by Tencent supporters. Chinese students are generally familiar with this software and will not feel anxious while using it.

Meanwhile, the researcher prepares the TE Selection Programme to show the Weibo posts to the participants. Before the task starts on the scheduled day, the researcher will check and make sure that the Internet connection on both sides is stable. In addition, both the participant and the researcher are in a quiet and undisturbed environment to ensure the success of recording and the recording quality. The participants can attend the think-aloud task either at home or in any other quiet place where they feel relaxed. They need to open up the Tencent meeting software on their PC and get ready for taking the task.

The researcher needs to contact the participants through platforms like Wechat, QQ, email, etc., and use Tencent Meeting for Think-aloud tasks. In addition, the researcher needs to help the participants deal with all technique problems such as Internet connection, PC software download, etc. These jobs should be done before the day of Think-aloud tasks.

Stage 2: Task performance

The participant is asked to turn on the cameras so that the researcher can have better control over the experimental procedure.

When the researcher receives the signed consent form, the researcher will play the introduction video to the participant and then invite the participant to ask questions if they feel confuse. When the participant reports that he or she is ready, the

researcher will start screen recording and informed the participant: "I have opened the screen recording. Then, we will begin the Think-aloud task. "

Then the researcher turns on the screen sharing function and starts the TE Selection Programme.

When the Weibo post with the TE is shown on the screen, the researcher asks the participant to start the think aloud task. The researcher informs the participant that the Think-aloud task of this TE begins: "Now you can read the Weibo. When you start reading, please speak all thoughts that appear in your mind aloud. You should follow your reading habits when you surf on Weibo. If you believe that you have completed the understanding, please raise your hand and say I have completed. Please try your best to figure out its meaning, but also feel free to quit. When I notice that you are not speaking your thoughts aloud, I will remind you. You do not need to feel stressed when I remind you by say Keep on talking. Because it is just an encouragement that reminds you to speak aloud."

When the participant becomes quiet and does not speak, the researcher will directly remind the participant to speak his or her mind aloud: “Don’t forget speaking your thoughts aloud. Just speak whatever that appears in your mind.”

When the participant informs the researcher that the reading procedure has been done or he or she can no longer figure out the TE's meaning, the researcher will stop the Think-aloud task of this TE and move to the next TE: “Okay, thank you for your cooperation. Let us move to the next TE.”

The researcher continues the think aloud task by clicking the “next” button until the programme stops. The researcher tells the participant: “Thank you for your cooperation. We have finished today’s Think-aloud task.”

When the researcher does the Think-aloud task with the participants, the researcher needs to record or mark every silence point and place where the participants fails to express clearly on a notebook, so that the researcher can ask the participants later about these unclear places.

Then the researcher will start the follow-up interview. The researcher will say, “Now let us start the follow-up interview.”

Here are the questions that may be asked in the follow-up interview.

1. When we did the think aloud task, I could not hear you very clearly when you say…”
2. what were you saying here?
3. what did you mean by…? Etc.

When the follow-up interview finished, the researcher stops the screen recording, closes Tencent Meeting, and saves the record to an encrypted PC. All records will be deleted when the researcher finishes the thesis.

Stage 3: After the task

After the study, all the files of the recordings are stored in an encrypted computer and UCL research log e-disk. The records of each participants will be transcribed into Chinese version and be sent to the participants to check if the content is correct. Only the sections that will be used in the discussion chapter will be translated into English.

Section 3 - research Participants (tick all that apply)

☐ Early years/pre-school

☐ Ages 5-11

☐ Ages 12-16

☐ Young people aged 17-18

☒ Adults please specify below

☐ Unknown - specify below

☐ No participants

Adults that are Weibo users who age between 22 – 32

Note: Ensure that you check the guidelines carefully as research with some participants will require ethical approval from a different ethics committee such as the National Research Ethics Service (NRES) or Social Care Research Ethics Committee (SCREC).

Section 4 – Security–sensitive material (only complete if applicable)

Security sensitive research includes: commissioned by the military; commissioned under an EU security call; involves the acquisition of security clearances; concerns terrorist or extreme groups.

a. Will your project consider or encounter security–sensitive material?

Yes* ☐ No ☐

b. Will you be visiting websites associated with extreme or terrorist organisations?

Yes* ☐ No ☐

c. Will you be storing or transmitting any materials that could be interpreted as promoting or endorsing terrorist acts?

Yes* ☐ No ☐

* *Give further details in* ***Section 8 Ethical Issues***

Section 5 – Systematic reviews of research (only complete if applicable)

a. Will you be collecting any new data from participants?

Yes* ☐ No ☐

b. Will you be analysing any secondary data?

Yes* ☐ No ☐

* *Give further details in* ***Section 8 Ethical Issues***

If your methods do not involve engagement with participants (e.g. systematic review, literature review) ***and*** *if you have answered* ***No*** *to both questions, please go to* ***Section 8 Attachments.***

Section 6 – Secondary data analysis (only complete if applicable)

a. Name of dataset/s: Enter text

b. Owner of dataset/s: Enter text

c. Are the data in the public domain?

Yes ☐ No ☐

If no, *do you have the owner's permission/license?*

Yes ☐ No* ☐

d. Are the data special category personal data (i.e. personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation)?

Yes* ☐ No ☐

e. Will you be conducting analysis within the remit it was originally collected for?

Yes ☐ No* ☐

f. **If no**, was consent gained from participants for subsequent/future analysis?

Yes ☐ No* ☐

g. **If no,** was data collected prior to ethics approval process?

Yes ☐ No* ☐

** Give further details in **Section 8 Ethical Issues***

*If secondary analysis is only method used **and** no answers with asterisks are ticked, go to **Section 9 Attachments**.*

Section 7 – Data Storage and Security

Please ensure that you include all hard and electronic data when completing this section.

a. Data subjects – Who will the data be collected from?

The participants involved in the think aloud protocols

b. What data will be collected? Please provide details of the type of personal data to be collected

audio records

Is the data anonymised? Yes ☒ No* ☐

Do you plan to anonymise the data? Yes* ☒ No ☐

Do you plan to use individual level data? Yes* ☐ No ☒

Do you plan to pseudonymise the data? Yes* ☒ No ☐

** Give further details in **Section 8 Ethical Issues***

c. **Disclosure** – Who will the results of your project be disclosed to?

The researcher, supervisors, and UCL

Disclosure – Will personal data be disclosed as part of your project?

No

d. Data storage – Please provide details on how and where the data will be stored i.e. UCL network, encrypted USB stick**, encrypted laptop** etc. Encrypted desktop PC

*** Advanced Encryption Standard 256 bit encryption which has been made a security standard within the NHS*

e. **Data Safe Haven (Identifiable Data Handling Solution)** – Will the personal identifiable data collected and processed as part of this research be stored in the UCL Data Safe Haven (mainly used by SLMS divisions, institutes and departments)?

Yes ☐ No ☒

f. How long will the data and records be kept for and in what format?

The data will be kept for a period of five years for further study.

Will personal data be processed or be sent outside the European Economic Area? (If yes, please confirm that there are adequate levels of protections in compliance with GDPR and state what these arrangements are)

NO

Will data be archived for use by other researchers? (If yes, please provide details.)

NO

g. If personal data is used as part of your project, describe what measures you have in place to ensure that the data is only used for the research purpose e.g. pseudonymisation and short retention period of data'.

This study will not use participants' personal data. Their names will be replaced by "A,B,C, …"

* *Give further details in* ***Section 8 Ethical Issues***

Section 8 – Ethical Issues

Please state clearly the ethical issues which may arise in the course of this research and how will they be addressed.

All issues that may apply should be addressed. Some examples are given below, further information can be found in the guidelines. *Minimum 150 words required.*

- Methods
- Sampling
- Recruitment
- Gatekeepers
- Informed consent
- Potentially vulnerable participants
- Safeguarding/child protection
- Sensitive topics
- International research
- Risks to participants and/or researchers
- Confidentiality/Anonymity
- Disclosures/limits to confidentiality
- Data storage and security both during and after the research (including transfer, sharing, encryption, protection)
- Reporting
- Dissemination and use of findings

The method of think aloud protocols will recruit 5 participants for pilot study and 20 participants for main study. All participants will be recruited by the researcher. They will be the researcher's friends so that they are familiar with the researcher and they will not be nervous or anxious during the study. The study requires the participants' voice to be recorded. All participants will be given a information sheet and consent form to sign before they participate the project. The audio recordings will be stored

in encrypted desktop PC and will be kept for five years after the study for further study. The participants will be informed and asked for permission when the researcher uses the data in the future. Only the data used in discussion chapter will be transcribed and translated to English. All the personal information will be pseudonymised by letters like "A, B, C, …". The data collected from the think aloud tasks will only be used by the researcher. All materials used for think aloud tasks will not contain any personal information of the Weibo users. Sensitive or impolite Weibo content will not be selected as the material used for the tasks.

Please confirm that the processing of the data is not likely to cause substantial damage or distress to an individual

Yes ☒

Section 9 – Attachments.

Please attach your information sheets and consent forms to your ethics application before requesting a Data Protection number from the UCL Data Protection office. Note that they will be unable to issue you the Data Protection number until all such documentation is received

a. Information sheets, consent forms and other materials to be used to inform potential participants about the research (List attachments below)

Yes ☒ No ☐

Information sheet and consent form

b. Approval letter from external Research Ethics Committee Yes ☐

c. The proposal ('case for support') for the project Yes ☐

d. Full risk assessment Yes ☐

Section 10 - Declaration

I confirm that to the best of my knowledge the information in this form is correct and that this is a full description of the ethical issues that may arise in the course of this project.

I have discussed the ethical issues relating to my research with my supervisor.

Yes ☒ No ☐

I have attended the appropriate ethics training provided by my course.

Yes ☒ No ☐

I confirm that to the best of my knowledge:

The above information is correct and that this is a full description of the ethics issues that may arise in the course of this project.

Name Tiancheng Chen

Date 17/ 05/2022

Please submit your completed ethics forms to your supervisor for review.

Notes and references

Professional code of ethics

You should read and understand relevant ethics guidelines, for example:

British Psychological Society (2018) *Code of Ethics and Conduct*

Or

British Educational Research Association (2018) *Ethical Guidelines*

Or

British Sociological Association (2017) *Statement of Ethical Practice*

Please see the respective websites for these or later versions; direct links to the latest versions are available on the Institute of Education Research Ethics website.

Disclosure and Barring Service checks

If you are planning to carry out research in regulated Education environments such as Schools, or if your research will bring you into contact with children and young people (under the age of 18), you will need to have a Disclosure and Barring Service (DBS) CHECK, before you start. The DBS was previously known as the Criminal Records Bureau (CRB). If you do not already hold a current DBS check, and have not registered with the DBS update service, you will need to obtain one through at IOE.

Ensure that you apply for the DBS check in plenty of time as will take around 4 weeks, though can take longer depending on the circumstances.

Further references

Robson, Colin (2011). *Real world research: a resource for social scientists and practitioner researchers* (3rd edition). Oxford: Blackwell.

This text has a helpful section on ethical considerations.

Alderson, P. and Morrow, V. (2011) *The Ethics of Research with Children and Young*

People: A Practical Handbook. London: Sage.

This text has useful suggestions if you are conducting research with children and young people.

Wiles, R. (2013) What are Qualitative Research Ethics? Bloomsbury.

A useful and short text covering areas including informed consent, approaches to research ethics including examples of ethical dilemmas.

Departmental Use

If a project raises particularly challenging ethics issues, or a more detailed review would be appropriate, the supervisor must refer the application to the Research Development Administrator via email so that it can be submitted to the IOE Research Ethics Committee for consideration. A departmental research ethics coordinator or representative can advise you, either to support your review process, or help decide whether an application should be referred to the REC. If unsure please refer to the guidelines explaining when to refer the ethics application to the IOE Research Ethics Committee, posted on the committee's website.

Student name: Tiancheng Chen
Student department: CCM
Course: PhD
Project Title: see above

Reviewer 1

Supervisor/first reviewer name: Li Wei
Do you foresee any ethical difficulties with this research?
No

Supervisor/first reviewer signature [redacted]
Date: 2/6/2022

Reviewer 2

Second reviewer name: John Gray
Do you foresee any ethical difficulties with this research?

No

Second reviewer signature: 

Date: 01/06/2022

Decision on behalf of reviewers

Approved ☐

Approved subject to the following additional measures ☐

Not approved for the reasons given below ☐

Referred to the REC for review ☐

Points to be noted by other reviewers and in report to REC:

Comments from reviewers for the applicant:

Once it is approved by both reviewers, students should submit their ethics application form to the Centre for Doctoral Education team: IOE.CDE@ucl.ac.uk.

Dear Tiancheng

Thank you for sending in your ethics application.

I am writing to confirm that ethical approval has been granted by the UCL Institute of Education for your doctoral research project titled:

Understanding Translanguaging expressions on Chinese social media (Weibo): A Relevance-Theoretic Account

This ethical approval has been granted from 4th June 2022 and the document you provided has been saved to your student file.

Please can you also upload the approved ethics form to your UCL Research Student Log https://researchlog.grad.ucl.ac.uk/.

I wish you all the best for your forthcoming research.

Regards,

Ms Michelle Brown

Programme Administrator | Centre for Doctoral Education |

Appendix C Information Sheet

Participant Information Sheet For Think Aloud Pariticipants

UCL Research Ethics Committee Approval ID Number: _____

YOU WILL BE GIVEN A COPY OF THIS INFORMATION SHEET

Title of Study:

Understanding Translanguaging expressions on Chinese social media (Weibo): A Relevance-Theoretic Account

Department: Insititue of Education

Name and Contact Details of the Researcher(s):

Tiancheng Chen

dtnvttc@ucl.ac.uk

+44 07930905902

+86 17612566331

Name and Contact Details of the Principal Researcher: Tiancheng Chen

dtnvttc@ucl.ac.uk

+44 07930905902

+86 17612566331

1. Invitation Paragraph

You are being invited to take part in a Ph.D research project. Before you decided, it is important for you to understand why the research is being done and what participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask me if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. Thank you for reading this.

2. What is the project's purpose?

The study focuses on the interesting expressions (Translanguaging expressions) that appear in daily Weibo reading. Readers may come across different types of Translanguaging expressions, while they cannot always figure out the meanings of all of them, which sometimes bothers the readers. By investigating in the understanding processes of different types of Translanguaging expressions, this study plans to find out different understanding patterns for different types of Translanguaging expressions, which will help more people to understand new expressions during their reading. This project will take each participants around 30 minutes to complete.

3. Why have I been chosen?

This project looks for people who are Weibo users and are between 20 to 30 years old as this age period accounts for the largest Weibo users according to Weibo annual report. Participants are expected to be familiar with the researcher and would like to share their thoughts with the researcher without being nervous or shy.

You are chosen to attend this study because you are friend of the researcher and also meet up with the criteria of Weibo using and age. You are one of the 20 participants that are invited to join this study.

4. Do I have to take part?

You should explain that taking part in the study is entirely voluntary and that refusal to agree to participate will involve no penalty or loss of benefits to which the participant is otherwise entitled. The participant may discontinue participation at any time without penalty or loss of benefits to which the participant is otherwise entitled.

Example paragraph:

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. You can withdraw at any time without giving a reason and without it affecting any benefits that you are entitled to. If you decide to withdraw you will be asked what you wish to happen to the data you have provided up that point.

5. What will happen to me if I take part?

The project will last around 30 minutes for once. The whole process of the project will take place online. No fees will be charged to participate the project.

This project is a think aloud task. You will be given 21 Weibo screenshots which include different Translanguaging expressions. You are expected to speak out your thoughts of understanding these target expressions. You are encouraged to share every detail that appears in your mind while you are understanding the expression that is shown on the screen. The audio of the think aloud tasks will be recorded. You will not be contacted after you have finished the tasks. All the data will be kept for five years for further study. If the researcher will use the data in the future, you will be informed.

6. Will I be recorded and how will the recorded media be used?

The audio recordings of your activities made during this research will be used only for analysis and for illustration in conference presentations and lectures. No other use will be made of them without your written permission, and no one outside the project will be allowed access to the original recordings.'

7. What are the possible disadvantages and risks of taking part?

There will not be any risks of taking part in this project.

8. What are the possible benefits of taking part?

Whilst there are no immediate benefits for those people participating in the project, it is hoped that this work will inspire you how to understand or figure out new expressions that you meet online. This work may also practice your ability of expressing yourself.

9. What if something goes wrong?

If you have any complaint, it will be handled. You can contact my supervisor Professor Li Wei, if you consider the researcher cannot address your complaints. You can email to li.wei@ucl.ac.uk

Should you feel your complaint has not been handled to your satisfaction (e.g. by the PR or the supervisor) that you can contact the Chair of the UCL Research Ethics Committee –

ethics@ucl.ac.uk

10. Will my taking part in this project be kept confidential?

All the information that we collect about you during the course of the research will be kept strictly confidential in an encrypted desktop. Only the researcher has access to your data. You will not be able to be identified in any ensuing reports or publications.

11. Limits to confidentiality

- lease note that assurances on confidentiality will be strictly adhered to unless evidence of wrongdoing or potential harm is uncovered. In such cases the University may be obliged to contact relevant statutory bodies/agencies.
- Please note that confidentiality will be maintained as far as it is possible, unless during our conversation I hear anything which makes me worried that someone might be in danger of harm, I might have to inform relevant agencies of this.
- Please note that confidentiality may not be guaranteed; due to the limited size of the participant sample.
- Confidentiality will be respected subject to legal constraints and professional guidelines.
- Confidentiality will be respected unless there are compelling and legitimate reasons for this to be breached. If this was the case we would inform you of any decisions that might limit your confidentiality.
- Confidentiality may be limited and conditional and the researcher has a duty of care to report to the relevant authorities possible harm/danger to the participant or others.

12. Use of Deception

Research designs often require that the full intent of the study not be explained prior to participation. Although we have described the general nature of the tasks that you will be asked to perform, the full intent of the study will not be explained to you until after the completion of the study [at which point you may withdraw your data from the study]".

13. What will happen to the results of the research project?

You should be able tell the participants what will happen to the results of the research (i.e. when the results are likely to be published, where they can obtain a copy of the published results, whether they be told which arm of the project they were involved in) and add that they will not be identified in any report or publication.

The result of the research project will be presented in the researcher's PhD thesis and stored till the thesis is passed.

14. Local Data Protection Privacy Notice

Notice:

The controller for this project will be University College London (UCL). The UCL Data Protection Officer provides oversight of UCL activities involving the processing of personal data, and can be contacted at data-protection@ucl.ac.uk

This 'local' privacy notice sets out the information that applies to this particular study. Further information on how UCL uses participant information can be found in our 'general' privacy notice:

For participants in research studies, click here

The information that is required to be provided to participants under data protection legislation (GDPR and DPA 2018) is provided across both the 'local' and 'general' privacy notices.

Your personal data will be processed so long as it is required for the research project. If we are able to anonymise or pseudonymise the personal data you provide we will undertake this, and will endeavour to minimise the processing of personal data wherever possible.

If you are concerned about how your personal data is being processed, or if you would like to contact us about your rights, please contact UCL in the first instance at data-protection@ucl.ac.uk.

15. Who is organising and funding the research?

This study is sponsored by Bloomsburry Colleges Scholarship

16. Contact for further information

Chen Tiancheng

dtnvttc@ucl.ac.uk

+44 07930905902

+86 17612566331

Flat 108, 79 Pershore Street, The Forum, Birmingham, B5 4RW

You should give the participant a contact point for further information. This can be your name, address and telephone number or that of another researcher in the project (if this is a supervised-student project, the address and telephone number of the student's supervisor).

Finally the information sheet should state that the participant will be given a copy of the information sheet and, if appropriate, a signed consent form to keep and remember to thank the participants taking part in the project.

Thank you for reading this information sheet and for considering to take part in this research study. You will have a copy of this information sheet and a signed consent form to keep.

Appendix D Consent Form

CONSENT FORM FOR Think Aloud Pariticipants IN RESEARCH STUDIES

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Title of Study: Understanding Translanguaging expressions on Chinese social media (Weibo): A Relevance-Theoretic Account

Department: Insititue of Education

Name and Contact Details of the Researcher(s):

Tiancheng Chen

dtnvttc@ucl.ac.uk
+44 07930905902
+86 17612566331
Name and Contact Details of the Principal Researcher:
Tiancheng Chen
dtnvttc@ucl.ac.uk
+44 07930905902
+86 17612566331
Name and Contact Details of the UCL Data Protection Officer: Data Protection Team
data-protection@ucl.ac.uk
This study has been approved by the UCL Research Ethics Committee: Project ID number: _____

Thank you for considering taking part in this research. The person organising the research must explain the project to you before you agree to take part. If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

I confirm that I understand that by ticking/initialling each box below I am consenting to this element of the study. I understand that it will be assumed that unticked/initialled boxes means that I DO NOT consent to that part of the study. I understand that by not giving consent for any one element that I may be deemed ineligible for the study.

		Tick Box
1.	*I confirm that I have read and understood the Information Sheet for the above study. I have had an opportunity to consider the information and what will be expected of me. I have also had the opportunity to ask questions which have been answered to my satisfaction and would like to take part in the think aloud task.	
2.	*I understand that I will be able to withdraw my data up to the end day of the study	
3.	*I consent to participate in the study. I understand that my personal information (name and age) will be used for the purposes explained to me. I understand that according to data protection legislation, 'public task' will be the lawful basis for processing.	
4.	*I understand that all personal information will remain confidential and that all efforts will be made to ensure I cannot be identified unless: 1. lease note that assurances on confidentiality will be strictly adhered to unless evidence of wrongdoing or potential harm is uncovered. In such cases the University may be obliged to contact relevant statutory bodies/agencies. 2. Please note that confidentiality will be maintained as far as it is possible, unless during our conversation I hear anything which makes me worried that someone might be in danger of harm, I might have to inform relevant agencies of this. 3. Please note that confidentiality may not be guaranteed; due to the limited size of the participant sample.	

	3. Confidentiality will be respected subject to legal constraints and professional guidelines. 4. Confidentiality will be respected unless there are compelling and legitimate reasons for this to be breached. If this was the case we would inform you of any decisions that might limit your confidentiality. 5. Confidentiality may be limited and conditional and the researcher has a duty of care to report to the relevant authorities possible harm/danger to the participant or others. I understand that my data gathered in this study will be stored anonymously and securely. It will not be possible to identify me in any publications.	
5.	*I understand that my information may be subject to review by responsible individuals from the University and the Bloomsburry Colleges Scholarship for monitoring and audit purposes.	
6.	*I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason. I understand that if I decide to withdraw, any personal data I have provided up to that point will be deleted unless I agree otherwise.	
7.	I understand the potential risks of participating and the support that will be available to me should I become distressed during the course of the research.	
8.	I understand the direct/indirect benefits of participating.	
9.	I understand that the data will not be made available to any commercial organisations but is solely the responsibility of the researcher(s) undertaking this study.	
10.	I understand that I will not benefit financially from this study or from any possible outcome it may result in in the future.	
11.	I understand that I will be compensated for the portion of time spent in the study (if applicable) or fully compensated if I choose to withdraw.	
12.	I agree that my pseudonymised research data may be used by others for future research. [No one will be able to identify you when this data is shared.]	
13.	I understand that the information I have submitted will be published as a report and I wish to receive a copy of it. Yes/No	
14.	I consent to my interview being audio/video recorded and understand that the recordings will be: stored anonymously, using password-protected desktop and destroyed when the study ends. To note: If you do not want your participation recorded you can still take part in the study.	
15.	I hereby confirm that I understand the inclusion criteria as detailed in the Information Sheet and explained to me by the researcher.	
16.	I hereby confirm that: (a) I understand the exclusion criteria as detailed in the Information Sheet and explained to me by the researcher; and (b) I do not fall under the exclusion criteria.	
17.	I agree that my GP may be contacted if any unexpected results are found in relation to my	

	health.	
18.	I have informed the researcher of any other research in which I am currently involved or have been involved in during the past 12 months.	
19.	I am aware of who I should contact if I wish to lodge a complaint.	
20.	I voluntarily agree to take part in this study.	
21.	I would be happy for the data I provide to be archived at the researcher's password–protected desktop. I understand that other authenticated researchers will have access to my pseudonymised data.	

If you would like your contact details to be retained so that you can be contacted in the future by UCL researchers who would like to invite you to participate in follow up studies to this project, or in future studies of a similar nature, please tick the appropriate box below.

	Yes, I would be happy to be contacted in this way	
	No, I would not like to be contacted	

_____	_____	_____
Name of participant	Date	Signature
_____	_____	_____
Researcher	Date	Signature