科技与中文教学

Journal of Technology and Chinese Language Teaching

Volume 9 Number 1, June 2018
二〇一八年六月　第九卷第一期

Volume 9 Number 1, June 2018

Articles

Columns

Shijuan Liu, Indiana University of Pennsylvania
Song Jiang, University of Hawaii at Manoa
Copy editor: Steven Paul Day, Benedictine University

Contacts

Exploration of Affordances of Open Data for Language Learning and Teaching

Jozef Colpaert
Universiteit Antwerpen
(安特卫普大学)
jozef.colpaert@uantwerpen.be

Abstract: The new wave of huge data sources, also referred to as big data, entails new forms of information and content for Computer Assisted Language Learning (CALL). Open data is the latest appeal in the open data philosophy and trends. In this article, we explore the affordances of (linked) open data for education in general and for CALL in particular. Learning analytics, adaptive learning environments, and open educational resources are the obvious beneficiaries, but contextualization of the learning process appears to be a relatively unexplored affordance. And this while available literature already touts its merits.

Keywords: Open data, contextualization, computer assisted language learning (CALL)

1. Introduction

The most vulnerable aspect of Computer Assisted Language Learning (CALL) as a discipline is its multifaceted nature (Levy, 1997), which I have coined the "pluriness" of CALL (Colpaert, 2018). CALL is dealing with a multitude of different disciplines, actors, contexts, theories, cultures, languages, skills, levels, subjects, technologies, learning goals, content types, and standards.

The underlying foundations—technology and pedagogy—are subject to rapid changes. The technological evolution is obvious: from mainframe, personal computer, multimedia CD-ROM, networking, Internet, mobile devices to current smart devices. Recent pedagogical paradigms are mushrooming, rooted in theories and approaches such as Constructivism (sociocultural theory and activity theory, among others), Project-Based Learning, Connectivism, Problem-based Learning, Connectionism, Task-Based Language Teaching (TBLT), Deep Learning, Complex Dynamic Systems, Self-Determination Theory, Universal Design for Learning, Dörnyei's L2 SELF model, 4CD/ID (van Merriënboer & Kirschner, 2007) Complex Dynamic Systems (Ellis & Larsen-Freeman, 2009), skills (Bloom digital taxonomy, 21st century skills, and higher order thinking skills), or technology (TAM, TPACK, and SAMR).

A third stream, data, strongly linked to available technology, is currently changing the CALL landscape imperceptibly. Such phenomena include big data, smart cities, the Internet of Things and augmented reality. In this context, a recent movement has emerged which has not been linked frequently yet with language learning and teaching: open data. In this article, we explore the affordances of open data for CALL.

As this article is a continuation of an earlier article in this journal (Colpaert, 2016), and a small step in our long-term research project, we apologize for the inevitable extensive self-referencing.

2. CALL and Data

We will first provide an overview of the various instantiations of data in the world of CALL, and the problems associated with some of them.

Traditional textbooks by publishers (Decoo, 2010), or self-authored materials produced by teachers for use with their own students only, have been the predominant language-learning content in the 20th century.

Since the advent of affordable digital technology in the 1980s, and of the Internet around the turn of the century, more artefacts—developed by design for education or not—have emerged. Such artefacts include: authentic documents found on the Web, adapted (easy) readers, corpora, interactive (tutorial) courseware, Massive Open Online Courses (MOOCs), Open Educational Resources (OERs), learning analytics, virtual worlds (such as Second Life and OpenSim), serious games, augmented reality (the view of a real-world environment whose elements are augmented by sensory input such as sound, video, graphics, or GPS data), and the Internet of Things (real-world objects and artefacts which carry readable data that can be used as content in tasks).

This wide array of educational artefacts opens up many possibilities for language teachers worldwide. But at the same time, we already know that affordances also entail limitations and challenges, certainly for interactive courseware, MOOCs, OERs and learning analytics. Interactive language courseware, also called tutorial CALL, has become an endangered species due to the labor-intensiveness of content and software development, the complexity of the required linguistic-didactic functionality, and the lack of generic, reusable, and exchangeable content (Colpaert & Decoo, 1999; Colpaert, 2004; Colpaert, 2013). MOOC platforms, such as Moodle and OpenLearning, allow teachers to create and use language courses that are accessible worldwide. But to what extent are these materials massive, open, online, and interactive enough to be called a genuine course (Colpaert, 2014)? OERs are materials that are shared, reused, improved and shared again. They are supposed to reduce workload for teachers and to increase learning efficacy considerably. But their success seems to be hampered (Colpaert, 2012) by various challenges, namely psychological ("what will others say about my content?"), technological ("what should I use to share my content?"), epistemological ("what does 'open' mean exactly?") and

juridical ("is Creative Commons enough to protect my content?'"). Finally, learning analytics: in the early years of CALL, referred to as "Tracking and Logging" and later known as (e-)portfolios (Little, 2002), the latest trend is in gathering and collecting data about the learner. But how should we analyze and use all these data? Is it just about testing and certification? Or should we not focus more on how learners build knowledge together in their cultural and social settings (Ferguson & Buckingham Shum, 2012)?

3. Data, Information, and Content

This brings us to the definition of data. As already explained in Colpaert (2016), data are series of raw tokens which can turn into information when they reduce uncertainty (Shannon & Weaver, 1949), or into content when they contribute to learning.

We distinguish four types of information:

- Information about the learner (learn**er** analytics): information provided by the learner, the teacher, the school, parents, e-portfolios, social networks (e.g. through data scraping) and smart phones (e.g. by providing the learner's geotemporal location). This information can be useful and even necessary for the adaptation and personalization of the learning process.
- Information about the learning process (learn**ing** analytics): data gathered by a system or electronic learning environment with a view to: a) assessing and supporting the learner; b) analyzing the learning process; c) improving the learning environment, and; d) predicting learner behavior (in the case of big data or educational data-mining on a large scale).
- Pedagogical metadata: sets of data that may facilitate the reusability and discoverability of digital learning resources, with a view to supporting the learner (hyperlink glosses, captions, just-in-time information, and procedural information) or the teacher (tags, readability indexes, CEFR levels, and pedagogical instructions). These metadata are mostly integrated in content for learning or teaching.
- Research data: with recent open access policies also comes the tendency to facilitate access to the datasets with any publication.

We distinguish four types of content:

- Published materials: textbooks, courseware, and MOOCs.
- Self-created (or co-authored with colleagues or students) materials: OERS, LMS-embedded exercises, sound files, subtitles, captions, corpora, knowledge clips, fan fiction, and textual or audiovisual content produced in online communities of practice.
- Authentic documents found on the Web, especially the semantic web, or level-adapted materials (e.g., easy readers).

- Content found in virtual worlds, serious gaming, ambient intelligence, augmented reality and the Internet of Things.

This availability of huge amounts of data for language learning and teaching under the form of information or content is often called big data. But how discoverable, accessible, usable, useful, interoperable, reusable, and sustainable are these data? A recent phenomenon sheds new light on this question: open data.

4. The Open Philosophy

Before exploring open data in more detail, it is interesting to mention other frequent collocations with "open."

The term "open" probably appeared for the first time in a discussion about open versus closed courseware in the 1980s (Van Elsen et al., 1991). In closed courseware, learners and teachers were able to make sophisticated selections, but they were not allowed to make any changes to the content itself. With open courseware (such as Adam and E*ve* and Hot Potatoe*s* later), teachers could enter their own content.

The term "open source" originated in the context of software development to designate a specific approach to creating and distributing source code for computer programs. It is known as the Open Source Initiative[1]. It mainly applied to licensing and redistribution aspects. Currently, it also refers to any artefact users can modify and share because its design is publicly accessible. It "designates a broader set of values—what we call 'the open source way.' Open source projects, products, or initiatives embrace and celebrate principles of open exchange, collaborative participation, rapid prototyping, transparency, meritocracy, and community-oriented development" (c.f., Open Source Way[2]).

In this respect, the "open" attitude was applied very quickly to content development in general and in educational content development more specifically. Open educational resources "are any type of educational materials that are in the public domain or introduced with an open license. The nature of these open materials means that anyone can legally and freely copy, use, adapt and re-share them. OERs range from textbooks to curricula, syllabi, lecture notes, assignments, tests, projects, audio, video and animation" (UNESCO, 2018).

Open education is a philosophy and a movement about the way people should produce, share, and build on knowledge. Everyone in the world should have access to high-quality educational experiences and resources, and proponents work to eliminate barriers

---

[1] c.f. https://www.opensource.org
[2] c.f. https://www.opensource.com

to this goal. Open universities, open courses, such as OpenLearn, and many publicly accessible MOOCs are concrete instantiations of this movement.

Open access (OA) refers to online research outputs that are free of restrictions for access and free of restrictions for use. Open access can be applied to all forms of published research output, including articles, papers, theses, book chapters, and monographs. The FAIR principle summarizes the criteria: research findings should be Findable, Accessible, Interoperable and Reusable.

Many other collocations arise in this respect, such as open standards, open government, open science, open gaming, open hardware, and open knowledge.

5. Open Data Initiatives

The Open Knowledge Initiative[3] is a global non-profit organization focused on realizing the value of open data to society by helping civil society groups access and use data to take action on social problems. Open Knowledge International does this by showing the value of open data, by providing organizations with the tools and skills to effectively use open data, and by making government information systems responsive to civil society. "Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)" (Open Definition, 2018).

The Global Open Data Index[4] is an annual effort to measure the state of open government data around the world. The crowdsourced survey is designed to assess the openness of specific government datasets according to the open definition.

The Belgian Open Data Initiative[5] offers more than 7,000 datasets, including environment, public sector, science, technology, economy, finance, population, transport, regional matters, culture, sports, energy, health, education, international matters, agriculture, fisheries, and justice.

China is opening up as a potential huge data repository for information and content. "With the release of the Big Data Development Action Plan by the State Council in September 2015, open data was officially recognized and listed as one of ten key national projects in the plan. The Plan not only gave a concrete timeline for opening government data, but also explained the motivation behind the national initiative: it is expected that big data and open data may drive economic transition, improve China's competitiveness, and create a modern governance model" (Ma, 2017).

Open Data China[6] is an advocate for open data by providing information on open data initiatives, supporting training and dissemination activities, and by stimulating

---

3 c.f. https://okfn.org/about

4 c.f. https://index.okfn.org

5 c.f. https://data.gov.be/en

6 c.f. http://opendatachina.com/en/

innovation for citizens. Many challenges, however, lie ahead, such as the fact that open data initiatives are limited to major cities (Liu et al., 2015), and "the absence of clear guidance and strong enforcement, the lack of transparency culture in the government, weak data literacy among both officials and the public, and limited tracking of data usage and impact" (Ma, 2017).

Opening up existing data sets is just the first step and does not automatically lead to a democratic government: "Publishing open data is of course not sufficient for open governments or open societies. It is just one ingredient in the mix, and no replacement for other vital elements of democratic societies, like robust access to information laws, whistleblower protections and rules to protect freedom of expression, freedom of the press and freedom of assembly" (Jonathan Gray, as cited in Ross, 2015).

6. (Linked) Open Data

Before looking into the affordances of open data for education in general and for Computer Assisted Language Learning in particular, let us first have a closer look at the more technical side of open data, without losing ourselves in irrelevant details.

Strangely enough, open data are not open to the extent that we can access a data source and change its content. Open data is based on the principle that a system or app contacts a data source using its URI (Uniform Resource Identifier) and RDF (Resource Description Framework).

The client system requests a dataset from the data server by calling an API (similar to web services) or by "scraping" the data source. The discussion between advocates of both approaches is still quite lively. Data sources can be in various formats, such as XLS, JSON, CSV, XLSX, HTML, PDF, XML, ODS, or TXT. The client system can now use these data and link them to other data sources. These linked open data can become complex and powerful sources of information for new services. The Linked Open Data Cloud[7] shows a diagram of all linked open data, also per subcloud such as geography, government, life sciences, linguistics, media, publications, social networking, or a user-generated subcloud. A good example is DBpedia[8], a public data infrastructure for a large, multilingual, semantic knowledge graph.

Besides the mere legal requirement of the open definition mentioned above, the goal of open data has also been described as maximizing the reuse of data. The more machines can access public datasets, the more value it creates for the public.

When creating applications or answering queries with public datasets today, a developer will make a pragmatic decision on reusing a limited set of datasets. Each dataset

---

[7] c.f. https://lod-cloud.net
[8] c.f. http://wiki.dbpedia.org

comes with a certain cost to adopt, even when publicly and freely available, as the developer needs to invest time to understanding how to reuse the data. The field of open data research has been characterized by looking for how technical principles can foster data adoption, cost-efficiently scaling up the average number of adopted datasets per application. This research happens on the four data interoperability levels: legal, technical, syntactical, and semantic. Principles such as the FAIR principles, the data on the Web best practices, or the linked data principles, advocate raising interoperability when publishing data on the Web (Colpaert, 2017).

The idea of open data, now well adopted by governments globally as evidenced by the many open data portals that can be found worldwide, ultimately puts forward a vision where the data published on the Web by many organizations can create power for the many, not the few.

7. Affordances of (Linked) Open Data for Education

The (linked) open data phenomenon entails several affordances for education in general. Large learning management systems, interactive courseware applications, and mobile apps will benefit from open data as new functionalities will be added and new systems will be conceptualized.

On the level of data as information we can distinguish:

- *Personalization* of the learning process: adaptation of parameters such as level of difficulty, activity type, speed, topic, feedback, and help functionalities on the basis of open learner data (Walkington, 2013; Brusilovsky & Millán, 2007). This adaptation is formulated in the form of explicit rules.

- *Intelligence* of the learning process: traditional neural networks become far better performing through the availability of huge data sources, which give a boost to Artificial Intelligence (Luckin et al., 2016). The intelligence of a system in terms of complex tutoring depends on the data of many learners. Rules remain implicit.

- *System upgrades* on the basis of extensive data mining about learner behavior and performance (Koedinger et al., 2008).

- Better *integration of research findings* in education and in the development of educational systems.

- Evidence-based *educational policy*: all data sources combined (learner, school, and community) should allow policy makers to draw up a more justifiable education policy plan.

On the level of content we can distinguish:

- By linking existing content to other data sets, learning content can be *enriched.* A typical example would be to link sentences of existing exercises to grammar topics in another data set.
- Existing content can also be *reused* in order to generate new systems, services, and products. Just as textbook content has been transformed into CD-ROM databases for interactive courseware and later into SQL repositories for online distributed systems, this content can now also be regenerated into other products as open data. The generation of mobile (and even personalized) exercises on the fly based on old textbook content is technologically speaking quite straightforward.
- Open data may rejuvenate the old concept of *open educational resources.* We mentioned epistemological, psychological, technological, and juridical barriers for OERs to break through. New systems based on an open data approach may offer a safer, comprehensible, and trusted environment for finding, accessing, using, editing, and sharing learning content.
- *Contextualization* of the learning process. Depending on the geotemporal location of the learner, a system can look up content, interactive or not, from open data sources, which is relevant for learners regarding their learning goals, interests, and preferences. This approach fits perfectly within current phenomena such as Augmented Reality, Smart Cities, and the Internet of Things.

In recent literature on the topic, the main focus is on two of these functionalities: learning analytics and open educational resources (Mouromtsev & d'Aquin, 2016). Let us now have a closer look at affordances for CALL.

8. Open Data for CALL: Focus on Contextualization

The considerations on affordances for education seem to be valid for CALL as well. Literature on the topic also mainly focuses on learning analytics and OERs (e.g. De Meester et al., 2018). As far as Chinese is concerned, two examples are worth mentioning. Zhishi.me is an effort to build Chinese linking open data, covering the three largest Chinese encyclopedias: Baidu Baike, Hudong Baike and Chinese Wikipedia (Niu et al,. 2011), and an online database for Chinese verb compounds (Zhan et al., 2015).

There is, however, one function that further deserves our special attention: the contextualization of the learning process.

Let us first look at an example of a new, imaginary app for Contextualized Language Learning:

> *I am in the neighborhood of a fisherman's village on the coast. It is about noon. The app gives me information on the type of fish sold here. It*

explains that I first have to buy the fish and then have to take it to one of the local restaurants across the street. It provides reviews of the restaurants so I can choose. It explains that the waiter, who does not speak English, will ask me if I want knife and fork or chopsticks, and if I want my fish steamed, grilled, or cooked. The waiter will also ask if the fish should be cut in small pieces or not, and served with cheese or not. It will explain how to do the dishes with tea. Finally, it will explain how to ask for the bill and thank the staff.

This app will not only turn my meal into the best possible lunch, but also into an experience with maximum exposure and learning effect.

The context of the learner can be defined as the collection of data available for, about, and at a specific location at a given moment in time. These data can become information or even content. Information enables the learner to perform a task and to focus on task execution (e.g., when the learner knows the departure time, bus number, and platform number, she or he will be able to take that bus), while content refers to data the learner has to process in some way.

Contextualization refers to the adaptation of the learning process to the specific context of the learner. Mobile devices such as smart phones and tablets can establish our location, date, and time exactly, while related services can inform us about weather and traffic conditions, cultural events, where to eat, and what to visit at that specific location (de Jong, Specht, & Koper, 2008). When we approach a railway station, these devices can automatically list arriving and departing trains, depending on our plans. When we are in the neighborhood of an Italian restaurant around noon, they can show the lunch menu, even when the learner does not ask for it. Given the potential of current paradigms such as Web 3.0 (semantization and enrichment), open knowledge and open data, Augmented Reality, Ambient Intelligence, and Sustainability of Learning Content, we need to investigate under which conditions online data can become useful as information or content.

The notion of context, as described above, becomes particularly important in light of the current pedagogical evolutions. Constructivist and task-based approaches strongly focus on learner autonomy with some degree of teacher and system support. The Situated Learning approach (Lave & Wenger, 1991) states that learning should be situated in a specific context and embedded within a particular social and physical environment. Complexity Theory and the Dynamic Systems Approach (Larsen-Freeman, 2002) see learning as a non-linear process, where rich input is considered crucial for learning to occur. Based on these arguments, we need to investigate to what extent the context of the learner, as defined above, can be implemented in a meaningful, useful, and enjoyable way within a state-of-the-art pedagogical approach in a mobile language learning environment.

The main objective of CALL research in this respect should be to create new opportunities for language learners to practice their language skills, wherever and

whenever they are, "temporarily halted… they can divert their attention to language learning" (Darren, Searle, Chiu, Zhao, & Landay, 2011). Literature seems to suggest (there is no substantiated evidence in terms of generalizable findings yet) that learners are tempted to learn and practice more frequently when the learning content is adapted to their immediate surroundings.

Contextual language learning via mobile devices is a relatively new research topic. The didactic origins for contextual learning can be retraced to Situated Learning and Communities of Practice (Lave & Wenger, 1991). Contextual learning is also based on findings from cognitive psychology: Tulving and Thompson refer to contextualization as "encoding specificity" (Tulving & Thompson, 1973), which means that we link a word to the location where we learned it, thus making it easier to recall when being asked for it at the same or in a similar location. Godwin-Jones (2010) also advocates for learning language in real-world locations, as the learner will naturally space repetitions over both time and the places where we need it most (Darren et al., 2011). Godwin-Jones refers to the algorithms for review as specified in the Leitner system, developed in the 1970s.

Contextualization is also found in the concept of situated cognition (Brown, Collins, & Duguid,1989), which argues that knowledge is situated in activities, contexts, and culture in which it is developed and used. However, Larsen-Freeman (2012) warns that the dynamism as expressed in her Complexity Theory and the Dynamic Systems Approach is not a natural characteristic of many online language learning courses. In a similar vein, de Jong, Specht, and Koper (2008) point out the shortcomings of current learning content. They refer to Ogata and Yano, who identified the characteristics of contextualized learning that make it suitable for learning: permanency of the recorded learning processes, immediate access and accessible content, interactivity, and the situating of instructional activities (Ogata & Yano, 2004). Next, they translate these characteristics into a technical framework, which specifies context as information about objects in the real world (including the learner), time, location, activity (goals and tasks), and relations.

Much like exploratory projects, such as MicroMandarin (Darren et al., 2011) and LOCH (Ogata, Hui, & Yin, 2008), we should define context in terms of time and location. But we should also distinguish ourselves from previous attempts to contextualize content, as it also will take into account problems encountered when trying to use various data sources, such as authentic documents, enriched data, open data, Open Educational Resources, and existing interactive materials, as well as psychological issues, such as self-efficacy, locus of control, resistance, and motivation. Motivation will be a key topic, confronting Self-Determination Theory (Deci & Ryan, 2000; Deci & Vansteenkiste, 2004) with the L2 SELF model (Dörnyei & Ushioda, 2009) and our Personal Goal Theory (Colpaert, 2010).

Information-based tasks, such as "This museum is now open. Go inside and ask for a museum hand-out" or "The next bus will arrive in three minutes. Ask the driver when the next bus will arrive, " are exciting to conceptualize, but they are difficult to evaluate and

often entail the involvement of native speakers in a real-world setting. They cannot be defined in advance and should be generated "on-the-fly" based on an analysis of the context. Content-based tasks, on the contrary, involve some action to be undertaken by the learner based on the content itself. The advantage of content-based tasks is that they can be prepared in advance, and that any learner action can be monitored and logged in detail.

9. Consequences and Caveats

The open data phenomenon affords several unique opportunities for education worldwide. It will allow teachers and publishers to find, retrieve, edit, and share content worldwide without having to change the structure of the content.

However, based on our experience with previous technological "revolutions," we need to point out the following caveats. Teachers will still need to fundamentally rethink their attitude towards learning content and their innate notion of intellectual property. Researchers might need to focus more on designing interfaces for exchanging data in the most effective way. Policy makers may need to stop providing for repositories of learning content, and instead promote the emergence of start-ups around the development of learning apps. Publishers should finally understand that they have to fundamentally rethink their obsolete production chain, adopt a new business-to-business model, and radically select for sustainable learning content. This choice will not only entail consequences for authoring content, but it will also open up a new product range based on the same data sets: traditional textbooks, tailor-made ("on-the-fly") textbooks, learning content on demand, learning content for specific purposes, Open Educational Resources, interactive app(lication)s, mobile intelligent apps (with personalization and contextualization of the learning process), learning support, business-to-business products, and research data.

In this respect, there is a need for more transdisciplinary research focusing on the ontological specification of common concepts for guiding the design of interfaces for accessing learning content.

Only a transdisciplinary approach can succeed. If we want to learn from the past, then we should know that interdisciplinarity is not the solution to solve the "pluriness" of CALL. Open data is a focal point where it all comes together: technology, pedagogy, policy, business, linguistics, and psychology. Transdisciplinarity defined by Colpaert (2018) implies the construction of mental and physical artefacts on a higher boundary-transcendent level of abstraction. We can start with conceptualizing the learner, the teacher, and the learning environment in a new way.

References

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1)*, 32-42.

Brusilovsky, P., & Millán, E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The adaptive web* (pp. 3-53). Berlin: Springer-Verlag.

Colpaert, J. (2004). *Design of online interactive language courseware: Conceptualization, specification and prototyping. Research into the impact of linguistic-didactic functionality on software architecture* (Unpublished doctoral dissertation).   University of Antwerp, Antwerpen, Belgium.

Colpaert, J. (2010). Elicitation of language learners' personal goals as design concepts. *Innovation in Language Learning and Teaching*, *4*(3), 259-274.

Colpaert, J. (2012, March 29-30). Open Educational Resources for language teachers: A goal-oriented approach. Paper presented at the Joint CMC and Teacher Education EuroCall SIGS Workshop Bologna. Retrieved from http://eurocallsigsbologna.weebly.com/papers.html

Colpaert, J. (2013). Ontological specification of an authoring interface for creating sustainable language learning content. In L. H. Wong, C. C. Liu, T. Hirashima, P. Sumedi & M. Lukman (Eds.), *Proceedings of the 21st International Conference on Computers in Education* (pp. 644-653). Indonesia: Asia-Pacific Society for Computers in Education.

Colpaert, J. (2014). Reflections on present and future: towards an ontological approach to LMOOCs. In E. Martín-Monje & E. Bárcena (Eds.), *Language MOOCs: Providing learning, transcending boundaries* (pp. 162-173). Berlin: Mouton de Gruyter.

Colpaert, J. (2016). Big content in an educational engineering approach. *Journal of Technology and Chinese Language Teaching, 7*(1*),* 1-14.

Colpaert, P. (2017). *Publishing transport data for maximum reuse* (Unpublished doctoral dissertation). Ghent University, Flanders, Belgium.

Colpaert, J. (2018). Transdisciplinarity revisited. *Computer Assisted Language Learning 31*(1-2). doi:10.1080/09588221.2018.1437111

Colpaert, J., & Decoo, W. (1999). The role of didactic functions in CALL design. In K. Cameron (Ed.), *CALL & the learning community* (pp. 64-74). Exeter: Elm Bank Publications.

De Jong, T., Specht, M., & Koper, R. (2008). Contextualized media for learning. *Educational Technology & Society*, *11*(2), 41-53.

De Meester, B., Lieber, S., Dimou, A., & Verborgh R. (2018). Interoperable user tracking logs using linked data for improved learning analytics. In *CALL your Data, XIXth International CALL Research Conference-Proceedings*. Antwerp: Universiteit Antwerpen.

Deci, E. L., & Vansteenkiste, M. (2004). Self-determination theory and basic need satisfaction: Understanding human development in positive psychology. *Ricerche di Psichologia*, *27*, 17-34.

Deci, E. L., & Ryan, R.M. (2000). The "what" and "why" of goal pursuits: Human needs and self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.

Decoo, W. (2010). *Systemization in foreign language teaching: Monitoring content progression*. London: Routledge Research.

Dörnyei, Z., & Ushioda, E. (2009). *Motivation, language identity and the L2 self.* Bristol: Multilingual Matters.

Edge, D., Searle, E., Chiu, K., Zhao, J., & Landay, J. A. (2011). MicroMandarin: Mobile language learning in context. *Conference on Human Factors in Computing Systems – Proceedings* (pp. 3169-3178).

Ellis, N.C., & Larsen-Freeman, D. (2009). *Language as a complex adaptive system.* Chichester; West Sussex, U.K: Wiley-Blackwell.

Ferguson, R., & Buckingham Shum, S. (2012). Social learning analytics: Five approaches. In *2nd International Conference on Learning Analytics & Knowledge-Proceedings*, (29 Apr-2 May, Vancouver, Canada). ACM Press: New York.

Godwin-Jones, R. (2010). Emerging technologies from memory palaces to spacing algorithms: approaches to second-language vocabulary learning. *Language Learning & Technology*, *14*(2), 4-11.

Koedinger, K., Cunningham, K., Skogsholm, A., & Leber, B. (2008). An open repository and analysis tools for fine-grained, longitudinal learner data. In *1st International Conference on Educational Data Mining — Proceedings* (pp. 157-166).

Larsen-Freeman, D. (2002). Language acquisition and language use from a chaos/complexity theory perspective. In C. Kramsch (Ed.), *Language acquisition and language socialization: Ecological perspectives* (pp. 33-46). London: Continuum.

Larsen-Freeman, D. (2012). Complex, dynamic systems: A new transdisciplinary theme for applied linguistics*? Language Teaching*, *45*(2), 202-214.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: University of Cambridge Press.

Levy, M. (1997*). Computer-assisted language learning: Context and conceptualization*. Oxford: Oxford University Press.

Little, D. (2002). The European language portfolio: structure, origins, implementation and challenges. *Language Teaching*, *35*(3), 182-189.

Liu, X., Song, Y., Wu, K., Wang, J., Li, D., & Long, Y. (2015). Understanding urban China with open data. *Cities*, *47*, 53–61.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. London: Pearson.

Ma, Y. J. (2017). Why Open Data is good for China. *Global Investigative Journalism Network*. Retrieved from https://gijn.org/2017/03/28/why-open-data-is-good-for-china/.

Mouromtsev, D., & d'Aquin, M. (Eds.) (2016). *Open Data for education. Linked, shared, and reusable data for teaching and learning*. Springer, Switzerland. Retrieved from https://link.springer.com/book/10.1007%2F978-3-319-30493-9.

Niu, X., Sun, X., Wang, H., Rong, S., Qi, G., & Yu, Y. (2011) Zhishi.me–Weaving Chinese linking Open Data. In L. Aroyo, C. Welty, H. Alani, J. Taylor, A. Bernstein, L. Kagal,...E. Blomqvist (Eds.), *The Semantic Web–ISWC 2011. Lecture Notes in Computer Science*, *7032*. Berlin; Heidelberg: Springer.

Ogata, H., & Yano, Y. (2004). Context-aware support for computer-supported ubiquitous learning. In *Proceedings of the 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education* (pp. 27–34). IEEE Computer Society, Washington, DC.

Ogata, H., Hui, G. L., & Yin, C. (2008). LOCH: supporting mobile language learning outside classrooms. *International Journal of Mobile Learning and Organisation*, *2*(3), 271-282.

Open Definition (2018). The Open Definition. Retrieved from https://opendefinition.org/.

Ross, E. (2015, December 2). Why open data doesn't mean open government. *The Guardian*. Retrieved from https://www.theguardian.com/media-network/2015/dec/02/china-russia-open-data-open-government.

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68-78

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: The University of Illinois Press.

Tulving, E., & Thompson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352-373.

UNESCO. (2018). What are Open Educational Resources (OERs)? Retrieved from http://www.unesco.org/new/en/communication-and-information/access-to-knowledge/open-educational-resources/what-are-open-educational-resources-oers/.

Van Elsen, E., Van Deun, K., & Decoo, W. (1991). Wordchip: The application of external versatility to an English lexical CALL program. *System*, *19*(4), 401-417.

Van Merriënboer, J. G., & Kirschner, P. (2007). *Ten steps to complex learning*. New York: Taylor & Francis.

Walkington, C. A. (2013). Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology*, *105*(4), 932-945.

Zhan, W., Ma, T., Tian, J., & Sunaoka, K. (2015). Development of an online database for Chinese resultative verb compounds and its application based on data visualization technology. *Journal of Technology and Chinese Language Teaching*, *6*(1), 1-15.

Performing *WeChat* Recording Tasks in Mixed-Ability Study Abroad Content Courses
(案例研究：海外混合程度班以微信录音功能完成语言任务)

Zhan, Hong
(战红)
Embry-Riddle University
(安博瑞德航空大学)
zhan121@erau.edu

Chen, Leeann
(陈丽安)
Embry-Riddle University
(安博瑞德航空大学)
chenl@erau.edu

Abstract: This case study explores the use of *WeChat*'s recording tool as a solution to the challenges of teaching mixed-ability students in content courses offered in study abroad programs. The tool successfully reduced anxiety and boredom, created opportunities for students to engage in personalized learning tasks in real time, enabled instructors to provide individualized feedback, and helped course curriculums stay on track. Data was collected from an online survey, email interview, and students' recordings of topics and instructor's feedback. The study determined that using *WeChat*'s recording function to complete linguistic tasks is a useful instructional tool for a mixed-ability classes in study abroad, despite limited technical difficulties.

摘要：本案例研究旨在探索以微信录音功能完成语言任务，作为解决学生水平差异较大之教学问题的一个有效办法。微信录音可减少焦虑、厌烦之情绪，提供个性化的口语练习，收到个性化反馈，以助学生更好地学习专业内容课程。数据来自网上及笔头调查，学生及教师的微信录音。虽然微信录音技术有些缺陷，但它却可以成为海外混合程度班的成功教学工具之一。

Keywords: Differentiated instruction, *WeChat* for content courses, study abroad

关键词：差异教学，微信应用，专业内容课程，留学中国

1. Introduction

Students populating Chinese classes taught in American universities have become increasingly diverse. One of the biggest challenges in teaching a foreign language is to meet a variety of student needs to maximize individual learning potential. This challenge becomes even more apparent in a mixed-ability class, where students have different language proficiencies, cultural backgrounds, and learning styles. To overcome this

challenge, differentiated instruction is generally recommended as a "philosophy of teaching and learning" (Theisen, 2002, p. 2) to accommodate diverse learners of different linguistic and cultural backgrounds, personal interests, and readiness levels of learning (Tomlinson, 2014). Through differentiating curricular elements, i.e. content, process, or product (output of language learning), teaching can be modified in response to each student's particular learning needs. Differentiated instruction further promotes equality and engagement of learning (Theisen, 2002).

However, differentiated instruction is not easily implemented in foreign language content courses because "many subject-area teachers want to maintain strong control over their particular courses and subject matter" (Grabe & Stoller, 1997). Ideally, learning linguistic knowledge while simultaneously learning content knowledge in content courses mutually enhances the learning process. As mentioned in Stryker and Leaver (1997, p. 5), "language proficiency is achieved by shifting the focus of instruction from learning language *per se* to learning language through learning content." When teaching content courses in the target language, the target language is primarily the medium through which students learn an academic subject matter (Crandall, 1994). Meanwhile, learning the content knowledge helps students learn how the target language is used in authentic and specific content areas. The instructor lectures and the students listen. This type of content course poses additional challenges for desired differentiated instruction because in traditional content courses, teacher-centered teaching deprives students of real communication opportunities (Lü, 2014), which are imperative for developing language proficiencies.

Additionally, limited class time is another challenge most teachers face when implementing differentiated instruction. The dual task of learning both language and content increases the demand for more time in foreign language classes. However, foreign language classes in American universities normally last only 50 to 60 minutes. This limited class time constrains language teachers to employing a "one-size-fits-all" type of instruction and feedback to students of different interest and readiness levels (Reese, 2011; Theisen, 2002; Tomlinson, 2014).

Time becomes even more constrained during intensive summer language courses taught abroad, where instructors are pressured with other tasks in addition to teaching and providing timely feedback on student work. Students often demand prompt responses from their teacher because future learning tasks generally depend on the instructor's feedback from previous assignments. Moreover, when studying foreign languages in the target language environment abroad, students are exposed to more learning materials and learning contexts where they can employ their content knowledge. Thus, providing timely feedback on students' practice of the content knowledge guides them in the learning process. Teachers must turn to technological means to enhance differentiated instruction and resolve the dilemma of providing individualized, differentiated, and personally meaningful instruction and feedback to language learners in content courses.

WeChat, a communication tool for mobile phones, has become the Chinese app for nearly everything—from text and voice messages to "friend circles" social media (similar to Facebook or Twitter) and online mobile payments for individual vendors or

shops. *WeChat* has become an important aspect of modern Chinese culture. Because of several useful functions embedded within the *WeChat* app, many language educators have discovered the benefits of employing *WeChat* in foreign language education.

This case study explores the benefits of using the *WeChat* voice message function in teaching content courses in Mandarin Chinese. Specifically, this case study observed students' involvement during the class, student-teacher contact time in the target language, and students' oral proficiency development via timely practice and individualized feedback from the instructor. Furthermore, this study explores the effective use of class time for content instructors. Data were collected from an online survey at the end of a summer program in Beijing, transcriptions of students' recorded homework, and email interviews with two content instructors. This case study finds that, despite occasional technical difficulties, using *WeChat*'s recording function to complete oral linguistic tasks increased students' overall involvement in class by reducing anxiety and boredom, improving oral proficiency development by creating opportunities for students to carry out personalized tasks simultaneously, increasing contact time for instructors to provide individualized feedback, and helping to ensure that content courses stayed on track.

2. Literature Review

2.1 Pedagogical Considerations for this Case Study

Adult second language acquisition is a long and complex process. All learners have unique differences that are important for instructor consideration. When implementing technology tools in the classroom, it is crucial to ensure that the technology does not surpass the pedagogy. In order to help students learn and reach their full potential through technology-mediated Chinese language learning, sound pedagogical principles and learning theories must be considered meticulously. In this case study, the researchers designed *WeChat* learning activities based on pedagogical principles derived from differentiated instruction, best teaching practices in providing feedback, and mobile-assisted language learning theory.

2.1.1 Differentiated instruction

Tomlinson, the leading educator advocating differentiated instruction, defines differentiated instruction as an educational approach to responding to "students' differences in readiness, interest, or learning profiles" (Tomlinson, 2014, p.103). While paying attention to a broad range of differences in students' background, differentiated instruction engages students in the learning process through a supportive learning environment, quality curriculum, assessment that guides teaching and learning, differentiated instruction that is responsive to students' needs, and by leading students and managing routines (Tomlinson & Moon, 2013). By differentiating curriculum content, learning processes, and products demonstrating the mastery of such content, the learning environment becomes flexible and adaptive to students' learning needs. The following chart best illustrates the underpinnings of differentiated instruction.
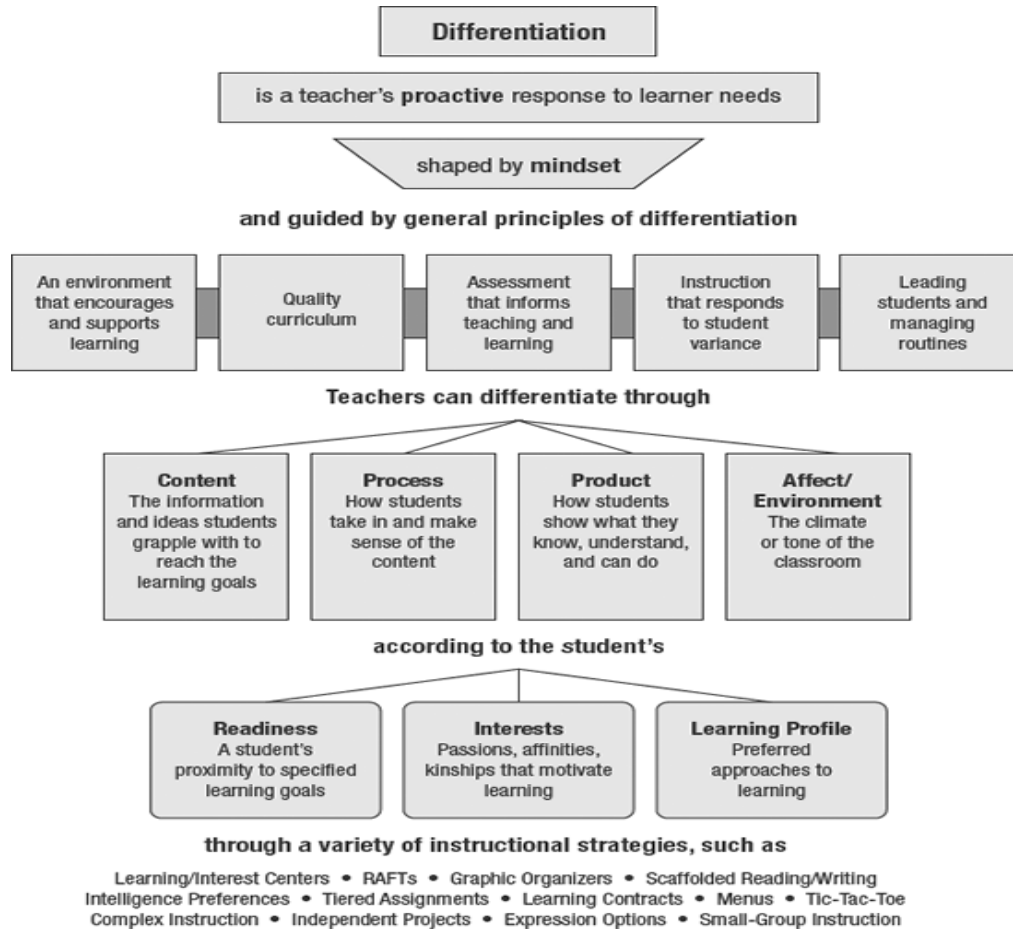
(Tomlinson & Moon, 2014)

Tomlinson's differentiated instruction principles provide a powerful and practical guideline for teaching. In the foreign language classroom, the researchers recognized that learners' characteristics may differ in proficiency levels; therefore, instructional strategies must be adjusted accordingly (Theisen, 2002; Roiha, 2014). The higher a student's level, more explicit and individualized instruction is needed to develop lexical precision, syntactical complexity, and organized speech (Leaver and Shekhtman, 2002; Ingold, 2002). Thus, providing differentiated instruction is necessary to advance students' language proficiency.

2.1.2 Feedback through assessment

As shown in the illustrative chart above, assessment is the central part of differentiated instruction. Normally, assessment is divided into two categories: formative and summative. These two types of assessment for learning and instruction serve different purposes in differentiated instruction. As Tomlinson and Moon explain (2014), whereas formative assessments serve to adjust course design in content, procedure, and product, summative assessments measure and evaluate student learning outcomes. From these two types of assessment, "differentiation places particular emphasis on pre-ass

differentiation places particular emphasis on formative assessment pre-assessment and formative assessment" (Tomlinson & Moon, 2013, p.20).

Feedback as an effective strategy for formative assessment has a powerful impact on learning. A meta-analysis study of the effect of feedback in educational research ranked feedback highest among hundreds of educational practices (Goodwin & Miller, 2012). Providing specific and timely feedback provides students with opportunities to identify their strengths and weaknesses, further revise and improve their work, and ultimately deepen their understanding. Particularly, research has found that feedback is most effective when provided immediately. For example, Opitz, Ferdinand, and Mecklinge's 2011 study found that in learning artificial language grammar, participants who were provided immediate feedback showed a significantly larger gain in performance compared to the participants who had received delayed feedback. All of these studies provide empirical findings of the importance of timely feedback on effective learning.

2.1.3 Mobile-assisted language learning (MALL)

Today's advanced technology provides many options for responsive teaching and differentiated instruction (Reese, 2011). Mobile devices, such as smart phones, have become an integral part of student life. The Pew Research Center (2018) reports that 91% of American college students own a smart phone. College student smartphone owners have almost tripled the total number of smartphone owners in 2011 (35%), when smartphones first became widely available and affordable on the market. With a surge in mobile devices (particularly smart phones), advanced mobile technologies, and wireless network accessibility, mobile-assisted language learning (MALL) is considered an ideal solution to language learning constrains in terms of place and time (Burston, 2013).

Research across disciplines and subjects has found that mobile devices have the potential to enhance language learning since they can easily connect users with a variety of online multi-level learning resources through a variety of apps. An annotated bibliography reviewing the historical background of MALL from 1994 to 2012, shows that MALL studies covered a variety of topics, including "technical specifications, mobile device ownership, pedagogical design, learning theory, user attitudes, motivational effects, institutional infrastructure, and teacher training" (Burston, 2013, p. 157). In addition, Burston (2015) studied the result of learning outcomes related to MALL project implementation in the past twenty years through a meta-analysis report and found that, even though MALL studies focusing on vocabulary did not have significant difference, those that investigated reading, listening, and speaking found that MALL had contributed to the development of target language skills in these aspects. The research findings encouraged more language educators to explore best practices and to investigate persisting or emerging issues related to MALL. With the advancement of mobile and other emerging technologies, such as augmented or virtual reality, MALL will certainly remain in demand and continue to grow as a field in its own right.

2.2 *WeChat* and Related Studies of *WeChat* Applications in CFL Instruction

2.2.1 *WeChat* development

MALL would not have been possible without the development of hand-held computing and mobile technology devices. From pocket dictionaries and PDAs to MP3 players, tablets, and smart phones, mobile technologies have launched the trend in MALL studies. A remarkable number of language apps have experienced surges in popularity over the years. Among many widely used smart phone apps, *WeChat* has become more popular worldwide in the past four years.

WeChat, a free instant message app lunched by Chinese company Tencent in 2001, has reached nearly 800 million users as of July 2017 (TechNews Report, 2017). *WeChat* is available on multiple platforms for mobile phones, tablets, or desktop computers. It is a worldwide social networking platform where users can not only post images and text as well as share photos and files, but can also chat via audio or live video. Additional functions, such as "Moments" (a function similar to a combination of Facebook and Blogger, where users post photos and circulate information) and "Subscription Accounts" (a large group chat of up to 500 people), enable *WeChat* users to interact simultaneously with large groups of people. Most importantly, *WeChat*'s mobile payment function via QR code scanning or direct linking to a user's bank card, has made *WeChat* an inseparable part of modern life for Chinese nationals. It is not an exaggeration to say that *WeChat* has become a lifestyle in China. The following table lists major developmental stages and features of *WeChat* and its functions.

Table 1 *WeChat*'s Development Features, Tools, and Functions

Year	Development Features	Tools and Functions
2011	Start of development	
	First launch for iPhone	Private messages in text and voice to communicate Video editing to create videos Search other *WeChat* users nearby Group chat to interact with many users *WeChat* Moments to visually share information

2012	Reaches 100 million users; *WeChat* 微信 in Chinese	Added more foreign languages (Thai, Vietnamese, Indonesian, and Portuguese） Voice and video chat to talk live QR code scanning to quickly add people in *WeChat* Photo sharing *WeChat* payments Subscriptions or public platform (公众号) Desktop *WeChat*
2013	Reaches 300 million users *WeChat* for Android and Windows phones *WeChat* International	Voice and video chat with multiple users Mobile payment (bank card) Game center Scan function *WeChat* pay in Jingdong Store
2014		Didi taxi Red Envelope (to transfer money)
2015	Reaches 500 million users Huge scale in advertising	
2016	Reaches 700 million users Optimization of existing features	
2017	Controls Chinese mobile device market	Plugins (small programs)

2.2.2 Related studies of *WeChat* applications in CFL instruction

WeChat was not designed for learning foreign languages. However, many *WeChat* tools contain powerful functions and applications in support of language learning. For example, functions like text and voice messages, videos, text translation into the user's interface language, as well as the ability to switch between traditional and simplified Characters, have given *WeChat* the title of the most favorable mobile app among CFL learners. In Liu's 2014 report, 93% of beginner CFL learners already have *WeChat* accounts.

Chinese language educators have shown increasing interest in integrating *WeChat* in their Chinese classroom. Empirical studies related to CFL, though scarce, have found that *WeChat* is a very effective tool in helping CFL learners learn the Chinese language. For example, Hu (2014) looked into the use of *WeChat* Moments (akin to a mini blog to share photos and information publicly with *WeChat* friends) in Chinese reading and writing instruction. After one-month training in speed reading Tencent News (腾讯新闻)

posted in *WeChat* Moments, students' reading speed and motivation in writing had greatly increased. Similar results have also been reported in Wang (2015), whose semi-experimental study found that students who intensively applied *WeChat* in their daily learning outperformed in reading, creating sentences, pronunciation, and accuracy.

Studies have also found *WeChat* tools especially helpful in developing oral proficiency. Yang (2014) designed the Chinese Suisuinian ("汉语碎碎念") public platform in *WeChat* based on current CFL pedagogy and as well as other *WeChat* features. The platform provides intermediate level oral instructional materials, including voice, video, graphic, and hyperlinks. After field testing the platform for about 4 months, CFL teachers and learners were asked to take an online survey to share their experience. 83% of users reported that the platform helped teaching and learning spoken Chinese.

In a 2016 mix-method design study, Luo and Yang explored the benefits of using *WeChat* to teach lower level CFL students. Through five types of *WeChat* learning activities, including asking/answering questions, mini-writing tasks, mini-oral projects, socializing and information sharing, and non-graded extracurricular input, participants reported five major benefits. It was found that participants had expanded their time in learning, increased their linguistic gains, experienced more cultural learning, developed higher learning motivation, and had established a supportive Chinese language learning community. Particularly, mini-oral tasks through *WeChat* were considered the most useful. Students commented that mini-oral tasks were a fun way to practice communicating in Chinese and to develop their oral skills.

Nowadays, many study abroad programs in China choose *WeChat* as a primary learning and communication tool because almost all Chinese are using it for daily life. Li (2017) reported on his case study, observing what *WeChat* could provide to two CFL learners who were studying in Shanghai in an intensive summer program. Based on affordance theory, or "the opportunities for action offered by specific object or environment," Li's case study found that although the two participants were different in linguistic meaning-focused communication, linguistics resources and multiliteracies, and space for new identity creation, both had reported that *WeChat* afforded them a fun and casual space to have instant and direct communication with native speakers of Chinese. These affordances helped the participants develop further communicative competences needed in real-life conversations, and improved their confidence as users of Mandarin Chinese. Obviously, when studying abroad in China, *WeChat* can be a very convenient tool to learn the authentic use of Chinese and to build and maintain a connection with native speakers of Chinese.

The aforementioned empirical studies reported that *WeChat* is a powerful tool for teaching and learning the Chinese language, and is particularly helpful in oral proficiency development and providing access to authentic social interactions. However, no studies have thoroughly investigated how *WeChat* can be used in the classroom within a study abroad context. Furthermore, no studies have researched how this convenient app can be used to provide immediate and individualized instruction. These two applications of *WeChat* deserve more research attention. Thus, the focus of the present study is how to

use *WeChat* to enhance differentiated practice and feedback in a mixed-ability content class.

3. Present Research

3.1 Research Questions

The purpose of this study is to investigate *WeChat*'s function in students' involvement in learning, the contact time between teacher and student in the target language, oral speaking ability development, instruction time for content, and the limitations of *WeChat* for performing oral tasks. In particular, the case study investigates the following questions:

1. How does the use of *WeChat* recording help provide an anxiety-free environment or reduce boredom for mixed-ability students?
2. How does *WeChat* recording provide a more efficient way of oral practice?
3. How does *WeChat* help provide timely and individualized feedback?
4. How can *WeChat*'s recording tool assist a content instructor in using class time more efficiently?
5. What are the drawbacks of using *WeChat* recording for performing oral tasks?

3.2 Research Context, Participants, and *WeChat* Recording Tasks

This research was conducted from June 15 to August 15, 2015 in a study abroad program in Beijing. The program lasted 9 weeks during which three content courses were taught: Speech, Cross-Cultural Communication, and Introduction to Geography. Students had 3 hours of class daily, Monday to Friday, and 1-2 hours of homework. There was also a required 30-minute tutoring session Monday through Thursday. Students all stayed with a Chinese host family and spent at least 2 hours commuting to and from the classroom via public transportation.

The Speech and Cross-Cultural Communication courses were taught from the beginning to the end of the program. The Introduction to Geography course was taught from July 13 to August 9. During this period, there were 2 hours dedicated to geography daily, and 1 hour dedicated to the 2 remaining courses.

The program included one CFL instructor, one geography instructor, two tutors, and seven students. The CFL instructor was a Chinese native speaker who works as a professor in the United States. This instructor taught Speech and Cross-Cultural Communication courses. The Geography Instructor was a native Chinese high school geography teacher in Beijing who did not speak English. The seven students (two females and five males) from four U.S. institutions differed in their cultural and linguistic background, as well as their proficiency levels as determined by the ACTFL Oral Proficiency Interview via Computer (OPIc). After data screening, 5 students were selected for case studies. The participants' proficiency information is detailed below.

Table 2 Participants

Participants	Gender	Pre-Program ACTFL OPIc Level	Years of Chinese learning
S1	F	Had not taken OPIc	2
S2	M	Intermediate-Mid (IM)	2
S3	M	Intermediate-Mid (IM)	2
S4	M	Advanced-Low (AL)	Heritage; left China at age 6
S5	F	Advanced-High (AH)	Heritage; left China at age 10

All three courses required students to create speech. As a powerful and ubiquitous communication tool in Chinese speaking communities, *WeChat* can be integrated in Chinese language education in both domestic and study abroad contexts to connect language learners at various levels of proficiency with native speakers of Chinese (Jin, 2017). In addition, since other popular tools, such as SnapChat, Google Chat, and Line are not accessible in mainland China, from the second week of the program, the participants began using *WeChat* to perform recording tasks inside and outside of class.

During the program, students created multiple oral recordings for their three classes. The topics varied from self-introductions to descriptions of a geographic location and natural resources. Speech tasks were differentiated to meet students' proficiency levels. For example, in a descriptive speech, lower level students were asked to describe their homestay house and the neighborhood, including the information about their address, the environment, how to get there by public transportation, neighborhood features, and the type of home (apartment or house). For the same topic, higher level students were asked to describe a landmark building in Beijing, including the design of the building, materials used, and its functions. Table 3 lists speech topics and corresponding differentiations.

Table 3 Differentiation in Speech Tasks

Speech type	Speech topic	Low Level	High Level
Description	Describe a place	Homestay house and neighborhood	A landmark building in Beijing
Introduction	Introduce a person	Homestay families or friends	A famous person in China
Information	Inform others about a Chinese province	Basic geographic information	Geographic information of a specific area (e.g. economic development of a province)

3.3 Data Collection and Analysis

Data were collected from online student surveys, instructor email inquiries, and students' recorded tasks. Through these three data sources, the research aimed to gather various information to answer the research questions.

The online survey was sent out at the end of the program to collect information about students' opinions of and experiences using *WeChat* for performing speaking tasks. The survey included 12 questions with 5 Likert scales: Strongly Agree, Mostly Agree, Somewhat Agree, Disagree, Strongly Disagree. Survey questions were classified into five categories: involvement in learning indicated by reducing learning anxiety and boredom (questions #1 and #2); oral proficiency development in terms of creating spontaneous speech at the student's own level (questions #3 and #4), individualized feedback and instruction in terms of receiving timely and individualized feedback (questions #5, #6, and #7), time for content learning in class (question #8), and *WeChat* limitations for learning Chinese (questions #9, #10, #11, and #12). In addition, there was one open-ended question to allow students to address any other thoughts they had on *WeChat* use during study abroad. The specific research questions can be found in Appendix 1.

A total of 111 recordings (78 from students and 33 from the language instructor) were transcribed. For each recording, length, speed, and characters were analyzed. Each speech was a unit of analysis. When analyzing each speech, the total length of the speech was counted in minutes and the speed was counted in number of characters per second.

The instructors' input about the use of *WeChat* in the class for content instruction was gathered through emails. The inquiry question for the two teachers was: *WeChat was used in your summer classes. What did you think of the use WeChat in your class for content instruction?* The emails were sent out after the completion of the summer program. Email responses from the two instructors were analyzed to find common themes about their opinions and experiences using *WeChat* in their content instruction.

4. Discussion of Results

The preliminary analysis shows that *WeChat* recordings could be used effectively to assist individualized instruction for mixed-ability content courses, but should not replace or reduce in-class student interactions. Further discussion of the findings related to research questions is presented in the following section.

4.1 About *WeChat* Use Regarding Student Involvement in the Class

Research questions #1 and #2 investigated students' perspectives on learning involvement in the class. Table 4 presents the survey questions and responses. The survey results found that 100 % of students agreed that when they were in a class with different proficiency levels and cultural backgrounds, *WeChat* use in class recording tasks helped reduce intimidation or anxiety because students only talk to themselves, not to the whole class. Furthermore, 80% of students agreed that in the same situation, the use of *WeChat* also helped them reduce boredom because they were only talking to themselves, instead of just sitting there waiting for other students to finish the task in class. Students can receive individual feedback as well. These results demonstrate that students were more engaged in class due to using *WeChat* functions.

Table 4 *WeChat* for Student Involvement in the Class

	Q1: *WeChat* helps reduce intimidation or anxiety		Q2: *WeChat* helps reduce boredom	
	Number of responses	Percentage	Number of responses	Percentage
-Strongly Agree	3	60		
-Mostly Agree	1	20	2	40
-Somewhat Agree	1	20	1	20
-Disagree	0	0	1	20
-Strongly disagree	0	0	1	20

(Q=Survey Question.Q1-Q11 means Survey Questions 1-11).

Responses from the open-ended question on the survey also support this finding. For Q1 and Q2, students also wrote comments. For Q1, 100% of students chose "Agree," indicating that anxiety exists for all students, despite proficiency differences. One advanced student said, "Although using *WeChat* recording may not be natural in a class, it is easier than speaking to the whole class." At the time of the program, some students even asked to make the recording outside of the classroom door. One other advanced student commented, "Discussion in groups according to students' proficiency levels also helped to reduce anxiety, and was more interesting than *WeChat* recording." This comment serves as a reminder to the researchers that students were aware of mixed ability challenges. There are multiple ways of reducing anxiety in such a class. The use of *WeChat* should not be too frequent, and should be used as only one methods among others.

For Q2, it seemed that advanced students were more likely to be bored and felt that *WeChat* was a tool to help them solve that problem. One intermediate student said, "Situational discussions in class better involved me, and thus reduced my boredom." As with some comments for Q1, one should not ignore other types of in-class student interactions. If using *WeChat* can save class time, perhaps the saved time should be used to create more in-class activities when feasible.

4.2 About *WeChat* Use Regarding Oral Proficiency Development

Survey result showed that 100% of students agreed that *WeChat* was a useful tool that provided more opportunities for them to speak at their own language level, as opposed to a situation where everyone took turns speaking in class. Moreover, by doing spontaneous recordings in class, *WeChat* helped students improve their ability to make impromptu speech in the target language. Corresponding statistical analysis is presented in Table 5.

Table 5 *WeChat* Use for Development of Oral Proficiency

	Q3: *WeChat* provides more opportunity to speak in class at one's own level		Q4: *WeChat* helps improve spontaneous speech	
	Number of responses	Percentage	Number of responses	Percentage
-Strongly Agree	4	80	1	20
-Mostly Agree	1	20	1	20
-Somewhat Agree	0	0	3	60
-Disagree	0	0	1	0
-Strongly disagree	0	0	1	0

Transcribed recordings were analyzed to understand how students at varying levels differed in their use of *WeChat* inside and outside the class to develop oral proficiency. Table 6 presents the number of total recording tasks each student completed, the length of recording tasks, number of total feedback the instructor provided, and the length of the feedback.

Table 6 Number and Length of Recorded Tasks or Feedback

Participants	Pre-Program OPIc	# of tasks	Average Minutes	# of teacher feedback	Average Feedback Minutes
S1	N/A	19	2.26	6	5.6
S2	IM	11	1.44	6	3.2
S3	IM	11	1.64	3	2
S4	AL	16	3.39	8	4.1
S5	AH	21	2.97	10	6
Total		78	2.46	33	20.9

Table 6 shows variations in speech length and task completion rates throughout the program. IM students spoke fewer minutes and completed fewer out-of-class tasks than the advanced students. The instructor *WeChat* feedback showed a parallel tendency. The data excluded those given in person or in class. It seems that the use of *WeChat* created a space where each student could express themselves at their own proficiency levels and speed, but the students themselves needed to take the initiative to use the tool for learning.

4.3 About *WeChat* Use Regarding Individualized Feedback and Instruction

Based on survey results, 100% of students agreed that *WeChat* helped them receive more individualized and timely feedback. Students became more aware of their strengths and weaknesses in their speech through listening to instructor feedback. Table 7 shows the statistics from the survey responses.

Table 7 *WeChat* Use for Individualized Feedback

	Q5: *WeChat* helps receive more individualized feedback		Q6: *WeChat* helps receive faster feedback from the instructor	
	Number of responses	Percentage	Number of responses	Percentage
-Strongly Agree	4	80	1	20
-Mostly Agree	1	20	1	20
-Somewhat Agree	0	0	3	60
-Disagree	0	0	1	0
-Strongly disagree	0	0	1	0

Table 8 *WeChat* Use for Individualized Instruction

	Q7: *WeChat* provides more opportunities to speak at one's own language level	
	Number of responses	Percentage
-Strongly Agree	4	80
-Mostly Agree	1	20
-Somewhat Agree	0	0
-Disagree	0	0
-Strongly disagree	0	0

In addition, in order to see if and to what extent *WeChat* recordings allowed for individualized speech and instruction, the researchers analyzed one in-class recording task and feedback.

On July 2, students in the Speech course were asked to speak on the use of one natural resource as part of the preparation for their speech on a Chinese province. Table 9 shows the variation in the topic, speech length, speed, and feedback length and speed.

Table 9 July 2 Students' In-Class Recording Task and Teacher Feedback

Participants	Pre-Program OPIc	Speech Content	Speech Length	# of Words per Second/ characters	Feedback Length	# of Words per Second Feedback/ characters
S1	N/A	Solar energy	2 mins	1.8/ 210	0.29 min	3.5/101
S2	IM	Forest	1.32 mins	1.0/135	0.31 min	3.7/116
S3	IM	Land	1.54 mins	1.5/152	0.29 min	4.0/115
S4	AL	Resource distribution	2.47mins	1.6/263	0.14 min	4.2/59
S5	AH	Resource distribution	2.25mins	1.9/279	0.30 min	6.0/178

As illustrated in Table 9, students 1-3 spoke about the use of one resource as the assigned topic requested, whereas students 4 and 5 chose to vary the topic and spoke about the distributions of several resources. In addition to variations in topics, students

also varied in speed. The AH student orally conveyed more content than the rest, at 1.9 characters/second; whereas student 2's speed was at 1 character/second. Students 1-3 also struggled with expressing the content. Errors occurred in pronunciation, word choice, and syntax. One IM student, for instance, could not distinguish between "ask for help" and "ask a question," a difference that students had been taught in their first semester of Chinese in college. Another student repeated the same words and sentences several times, as if to allow for more time to compose other sentences internally. Students 4 and 5, by contrast, clearly expressed the topic and even used formal language such as " 均衡," a word that some of the other students did not know.

The instructor's feedback speed similarly varied from 6 characters/second for the AH student to 3.5 characters/second for student 1. The feedback focused on one or two main characteristics in the recording. For the AH student, the instructor explained the difference between "assigning tasks" and "resource distribution;" both phrases begin with the same Chinese character. For the AL student, the instructor explained the idiomatic yet formal expression of "the western part" of a country. For the non-advanced students, the instructor focused on explaining some basic lexical and syntax issues.

4.4 About *WeChat* Use Regarding Saving Class Time for Content Instruction in Class

The survey results showed that 100% of students felt that opportunities for them to speak in Chinese about geography content would have been significantly less without the use of *WeChat* in the Geography class. Table 10 presents the statistical results.

Table 10 *WeChat* Use for Saving Class Time for Content Instruction

	Q8: *WeChat* provides more opportunity to speak about content knowledge	
	Number of responses	Percentage
-Strongly Agree	4	80
-Mostly Agree	1	20
-Somewhat Agree	0	0
-Disagree	0	0
-Strongly disagree	0	0

The survey results are echoed in the emails between the geographer and the CFL instructors. Their responses show a strong preference for the use of *WeChat* to record content tasks because it helped both instructors carry out content tasks in class more efficiently.

The geography instructor only needed to call on one or two students to speak about a geography topic, make a comment, and continue on to the lecture since every student had already interacted with the CFL instructor regarding the topic via a *WeChat* recording task. The reduced speed of the geography instructor's lecture facilitated student comprehension and provided more time for questions.

The CFL instructor commented that if every student took turns to perform an oral task in class, and the teachers took turns to provide feedback, it could take 20 minutes of class time. Using *WeChat* to do the same tasks simultaneously took only 5-6 minutes, and the teacher could provide feedback after class. As a result, the instructor could use the class time for more discussion and listening comprehension and reading practice.

4.5 About *WeChat* Use Regarding the Technical Limitations of *WeChat* for Learning Chinese

Although *WeChat* has been a popular communication tool for Chinese people in daily life, *WeChat* is not designed for the purpose of learning foreign languages. A few limitations are identified in *WeChat* if using it as a language learning tool. To name a few: *WeChat*'s "Hold and Talk" function can create only a one-minute recording per time; one can accidently let go of the audio message button before finishing the recording, and; sometimes, mainly due to web connection problems, *WeChat* messages are not actually sent to the receiving party. Because of these limitations, the researchers of this study are interested in finding out how students perceived the use of *WeChat* for their course work.

The survey results show very positive opinions regarding the use of *WeChat* in the classroom. 60% of the students, despite minor technical problems, strongly recommended the use of *WeChat* in CFL education. Among those 60%, 40% were advanced heritage students. These results indicate that in the study abroad context, students became more tolerant of the flaws of *WeChat*. They considered *WeChat* as the best technological tool for maximizing student-teacher contact. Table 11 presents the specific responses to survey questions #9 to #10.

Table 11 Students' Opinions about Technical Limitations of *WeChat* Recording Function

	Q9		Q10		Q11		Q12	
	#	%	#	%	#	%	#	%
Strongly agree	1	20	2	40	1	20	2	40
Mostly agree	3	60	3	60	3	60	2	40
Somewhat agree	1	20	2	0	1	20	1	20
Disagree	0	0	0	0	0	0	0	0
Strongly disagree	0	0	0	0	0	0	0	0

5. Conclusion

This case study has much room for expansion and further analysis. The scope should be expanded to include more students. The majority of recording content awaits further analysis. Yet, the initial findings are encouraging, as demonstrated in the survey results, the instructor feedback, and the recording transcriptions. The use of *WeChat* recording to complete linguistic tasks helped provide: 1) an anxiety-free environment or the reduction of boredom for mixed-ability students; 2) a more efficient way of oral

practice, and; 3) timely and individualized feedback. In addition, the use of *WeChat* maximized teacher-student target language contact, which enabled content instructors to use class time more efficiently.

At the end of the program, the same four students who took OPIc at the beginning of the program were tested again. Of the five students, three had made progress. One Intermediate-Mid student reached Intermediate-High, and two Advanced-Low students reached Advanced-High. Perhaps *WeChat* recordings played a role in helping these students raise their proficiency levels.

Since summer 2015, the researchers have started demonstrating to students how to use *WeChat* to make sure they know that each recording can only be one minute long. Using *WeChat* to record tasks will continue to be used in future study abroad programs in China. Even if Google and other web recording devices are available, *WeChat* still appears to be most convenient, and thus will continue to be used in other CFL courses.

Regardless of the many limitations of this case study, the findings of this research can contribute to the CFL field through an in-depth investigation of the use of *WeChat* for teaching students in mixed-level content courses focused on oral proficiency development that are taught in a study abroad context. However, more future research is needed to address many unanswered questions in this study. For example, this study finds that students appreciated the instructors' timely and individualized feedback. But how exactly such feedback helps students remains unknown. Did the feedback help students most in pronunciation, ways of expression, or oral text structure? Future research is needed to answer these questions.

References

Burston, J. (2013). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL 27*(1), 4-20.

Crandall, J. (1994). Content-centered language learning. *ERIC Digest.* Retrieved from http://files.eric.ed.gov/fulltext/ED367142.pdf.

Pew Research Center. (2018). Mobile fact sheets. Retrieved from http://www.pewinternet.org/fact-sheet/mobile/.

Grabe, W., & Stoller, F. L. (1997). Content-based instruction: Research foundations. In M. A. Snow & D. M. Brinton (Eds.), *The content-based classroom: Perspectives on integrating language and content* (pp. 5-21). White Plains, NY: Longman.

Goodwin, B., & Miller, K. (2012). Research says good feedback is targeted, specific, timely. *Educational Leadership*, *70*(1), 82-83.

Hu, Y. (2014). Teaching reading and writing Chinese through *WeChat*: Experiments, evaluation, and discussion. In J. Da, S. Jiang, P. Zhang & S. Liu (Eds.), *Proceedings of the 8th International Conference & Workshops on Technology & Chinese Language Teaching* (pp. 57-66). Boston, MA: Tufts University and Hamilton College. [胡越. (2014). 基于 *WeChat* 网络平台的中文阅读与写作学

习实验及评析. In J. Da, S. Jiang, P. Zhang, & S. Liu (Eds.), *Proceedings of the 8th International Conference & Workshops on Technology & Chinese Language Teaching* (pp. 57-66). Boston, MA: Tufts University and Hamilton College.]

Ingold, C. W. (2002). The LangNet "Reading to the four" project: Applied page technology at higher levels of language learning. In A. L. Leaver & B. Shekhtman (Eds.), *Developing professional-level language proficiency* (pp. 141-155). New York: Cambridge University Press.

Leaver, B., & Shekhtman, B. (2002). Principles and practices in teaching superior-level language skills: Not just more of the same. In A. L. Leaver & B. Shekhtman (Eds.), *Developing professional-level language proficiency* (pp. 3-33). New York: Cambridge University Press.

Li. J. (2017). Digital affordances on *WeChat*: Learning Chinese as a second language. *Computer Assisted Language Learning*, *31*(1-2), 27-52.

Lü , Y. Q. (2014). Application of the *WeChat* public platform in teaching Chinese as a second language. *The Journal of Modernization of Chinese Language Education*, *3*(2), 51-57. [吕亚其. (2014). 微信公众平台在辅助对外汉语课堂教学中的应用研究. *中国教育现代化学报*, *3*(2), 51-57.]

Luo, H., & Yang, C. (2016). Using *WeChat* in teaching L2 Chinese: An exploratory study. *Journal of Technology and Chinese Language Learning, 7*(2), 82-96.

Opitz, B., Ferdinand, N. K., & Mecklinger, A. (2011). Timing matters: The impact of immediate and delayed feedback on artificial language learning. *Frontiers in Human Neuroscience*, *5*(8), 1-9. doi.org/10.3389/fnhum.2011.00008.

TechNews Report. (2017, September 6). Active *WeChat* users reached 100 million account, but its' projected increase in the next three to five years is only 4%. *TechNews*. Retrieved from http://www.sohu.com/a/190052789_115161.

Reese, S. (2011). Differentiation in the language classroom, *The Language Educator*, *6*. Retrieved from https://www.actfl.org/sites/default/files/pdfs/TLEsamples/TLE_Aug11_Article.pdf.

Riaz, S. (2016, March 17). *WeChat* nears 700M users. *Mobile World Live*. Retrieved from http://www.mobileworldlive.com/apps/news-apps/wechat-hits-697m-users/

Roiha, A. S. (2014) Teachers' views on differentiation in content and language integrated learning (CLIL): Perceptions, practices and challenges. *Language and Education*, *28*(1), 1-18.

Smith, A. (2015). U.S. Smartphone use in 2015. *Pew Research Center*. Retrieved from http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/.

Stryker, S. B., & Leaver, B. L. (1997). *Content-based instruction in foreign language education: Models and methods*. Washington, D.C: Georgetown University Press.

Theisen, T. (2002). Differentiated instruction in the foreign language classroom: meeting the diverse needs of all learners. *LOTE CED Communiqué, 6*. Retrieved from http://www.sedl.org/loteced/communique/n06.html.

Tomlinson, C. A. (2000). Differentiated instruction: Can it work? The Education Digest, 65(5), 25-31.

Tomlinson, C. A. (2014). *Differentiated classroom: Responding to the needs of all learners* (2nd Ed.). Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

Tomlinson, C.A., & Moon, T. (2013). *Assessment and student success in a differentiated classroom*. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

VanPatten, B. (2002) *From input to output: A teacher's guide to second language acquisition.* New York: McGraw-Hill.

Wang, Z. Y. (2015). Study of *WeChat* applications for language learning. *Open Education Research*, *21*(2), 113-119. [王卓玉. (2015). 微信平台在语言学习中的应用研究, *开放教育研究*, *21*(2), 113-119.]

Yang, X. (2014). *Study of teaching oral Chinese in the WeChat public platform* (master's thesis). Sichuan University, Chengdu, China. Retrieved from http://cdmd.cnki.com.cn/Article/CDMD-10636-1014336786.htm. [杨夏. (2014). 基于微信公众平台的对外汉语口语教学研究(硕士论文). 四川大学, 成都, 中国. Retrieved from http://cdmd.cnki.com.cn/Article/CDMD-10636-1014336786.htm.]

Zhou, X. Q. (2015). Study of *WeChat*-based interaction models for teaching Chinese as a foreign Language. *Education and Vocation 25*, 97-100. [邹小青. (2015). 基于微信的对外汉语互动教学模式探究. *教育与职业*, *25*, 97-100.]

Appendix

WeChat Recording Task Survey Questions

1. In a class where your classmates were of different Chinese proficiency levels and backgrounds, do you agree that the use of *WeChat* recording helped reduce intimidation or anxiety because you are only talking to yourself, not to the whole class?

2. In such a class, do you agree that the use of *WeChat* recording helped reduce boredom any, if at all, because you were only talking to yourself and receiving individualized feedback?

3. Do you agree that the use of *WeChat* provided more opportunities for you to speak at your own language level than a situation in which everyone takes turns to speak in class?

4. Do you agree that by doing spontaneous recordings in class *WeChat* helped you to improve your ability to make spontaneous speech?

5. Do you agree that by using *WeChat* you received more individualized feedback?

6. Do you agree that by using *WeChat* you got faster feedback?

7. Do you agree that by using *WeChat* to hear your teacher's feedback you became more aware of your strengths and weaknesses in your speaking?
8. Do you agree that in the geography class without the use of *WeChat* the opportunities for you to speak in Chinese about geography content would have been significantly less?
9. Do you agree that *WeChat*'s limitation of one-minute recording per time was not really a problem since one can make multiple recordings one after another?
10. Do you agree that accidently letting go of the audio message button before one finishes recording is a problem to be aware of but does not hinder oral communication?
11. Do you agree that sometimes *WeChat* messages not being actually sent to the receiving party (mainly due to web connection problems) is a problem to be aware of but does not hinder oral communication?
12. Do you agree that in China, even without the availability of Google and a much slower Internet speed, the use of *WeChat* is the best technological tool for maximizing student-teacher contact?

Learning Chinese Colloquialisms through Mobile Technology
(基于移动通信技术的口语习语教学)

Yang, Jia
(杨佳)
University of Dayton
(代顿大学)
jyang2@udayton.edu

Yin, Chengxu
(尹承旭)
University of Notre Dame
(圣母大学)
cyin@nd.edu

Abstract: This is a study of the effectiveness of learning colloquialisms through the use of mobile technology. A within-subjects research design was adopted to expose all participants to learning colloquialisms through mobile phones and printed materials. The results indicate that distributing vocabulary-learning materials via mobile messages could suit learners' preference for using mobile devices and increase the possibility of incidental vocabulary learning. However, such a practice only promotes learners' short-term retention of the vocabulary. The study also reveals that the length of previous language learning experience did not have a significant effect on the outcome of mobile-assisted language learning. Furthermore, there were no statistical correlations between learners' attitudes toward mobile learning and their learning outcomes either. The results suggest that mobile-assisted language learning can be an effective teaching approach for L2 learners, regardless of their language proficiency levels. Learners can be engaged in the learning of new words even when they do not have positive attitudes toward the perceived usefulness of the mobile-assisted approach, as long as the mobile-assisted learning activities are designed to connect with in-class instruction.

摘要：本文旨在研究学习口语习语的过程中使用手机是否行之有效。本研究采用了组内设计的方法，以确保每个被试者既接受了手机辅助的教学，又接受了基于纸面教材的传统教学。研究结果显示，由于学生本来就有使用移动通信设备的习惯，通过短信来发送词汇学习材料可以因势利导，增加学生偶然性学习的机会。但是，这种手段只能促进学生对口语习语的短期记忆。本研究也发现，移动辅助语言学习的效果如何，在统计学意义上并不受学生之前的语言学习时间长短的影响；另外，从统计学意义上来说，学生对移动辅助语言学习的态度和教学效果之间也没有显著的相关性。由此可以得出结论：移动辅助语言学习的效果并不受学生外语水平的影响。只要把移动通信辅助教学和课堂教学有机地结合起来，即使学生对移动辅助语言学习的态度不甚积极，其教学效果也不会受到太大影响。

Keywords: Mobile assisted language learning, colloquialisms, *WeChat*

关键词: 移动辅助语言学习，口语习语，微信

1. Introduction

According to Kubler (2007), some colloquialisms that are used by native speakers even at the age of 2 or 3 are still challenging for non-native speakers to master. This is mainly because such colloquialisms rarely appear in the classroom or in textbooks and because L2 learners usually have very limited exposure to them in their interactions with native speakers, who tend to avoid using them when talking to non-native speakers. However, advancements in mobile technology can create opportunities to expand informal learning outside the classroom. These opportunities provide new ways to integrate the instruction of those colloquialisms into regular classroom teaching. This study examines the effectiveness of learning colloquialisms through the use of mobile phones. The data were collected from classes that were designed for the utilization of mobile technology.

2. Literature Review

A number of studies have appeared in recent years on the application of mobile technology in L2 vocabulary learning. One of the most investigated areas has been vocabulary lessons sent via mobiles phones or emails (Cavus & Ibrahim, 2009; Kierran & Aizawa, 2004; Lu, 2008; Miangah & Nezarat, 2012). Preliminary evidence suggests that text-based vocabulary learning messages sent via mobile devices generally lead to better learning outcomes than email-based or paper-based learning (Ma, 2017). However, as pointed out by Burston (2014), a number of mobile-assisted language studies fail to discuss some essential factors, such as the participant's level of language proficiency or the duration of the project. Few of those studies have language acquisition as their primary focus. More empirical studies are needed to examine whether the effectiveness of mobile-based learning varies across learners at different language proficiency levels and whether mobile-based instruction has a long-term effect on students' learning.

Studies have also found that certain factors might affect the learners' use of mobile technology for learning. For example, Stockwell (2007) reported that learners using mobile devices spent more time to complete vocabulary activities (such as choosing/writing a word to complete a sentence and choosing/writing a definition for a word) and achieved slightly lower scores than those completing the same activities on desktop computers. The difference in the two learning platforms might be caused by the limitations of the mobile phone interface, such as small screens and inconvenient keypads, and the environmental issues that students encounter when using mobile devices in a non-ideal learning environment (Stockwell, 2008). Similarly, Liu, Li, and Carlsson (2010)

reported that perceived near-term/long-term usefulness and personal innovativeness significantly influence learners' inclination to adopt mobile learning. Learning styles and needs, encouragement and support from peers and teachers, and learners' attitudes toward technology all factor in to the learners' disposition to integrate mobile technology in learning. From these findings arises the question of whether learners' attitudes toward mobile-based learning might influence their perceived usefulness of mobile-based instruction and consequently affect their learning outcomes. The present study aims to examine three questions:

- In comparison with the use of printed materials, does mobile-learning more effectively promote students' short-term and long-term retention of Chinese colloquialisms?
- Do students' previous instructional contact hours have an effect on the effectiveness of learning colloquialisms via mobile technology?
- Is there a correlation between learners' attitudes toward mobile learning and their learning outcomes?

3. Instructional Design

3.1 Theoretical Framework

As vocabulary knowledge is considered the core component in better comprehension in L2 learning (Nation, 1990; 2001), a number of studies have discussed effective methods in vocabulary instruction (Hulstijn, 1992; Rott, 1999; Schmitt & McCarthy, 1997). Research has shown that there is no one-size-fit-all approach. Effective instruction should include both explicit and implicit word learning through a variety of methods. The effective teaching methodology should encourage and enable learners to review and study the words taught in class on a more regular basis. Explicit and implicit learning in vocabulary acquisition can be effectively combined by taking advantage of students' interest in using mobile devices.

Ma (2014) proposed a memory-based strategic framework for vocabulary learning through a large-scale questionnaire study. In this framework, vocabulary learning involves four stages of processing: 1) perceiving the new word form in visual or/and auditory input; 2) accessing the meaning of the new word from the mental lexicon; 3) building the new word entry in the mental lexicon through the connection to the existing meaning; and 4) retrieving the new word from the mental lexicon for receptive or productive use so that the new word entry will be consolidated. According to Ma's framework, in-class instruction can be used for the first three stages to establish the new word entry in learners' mental lexicon while mobile technology can be used in the last stage to create opportunities for learners to retrieve the new word. Moving further in the direction suggested by Ma, we designed the following instruction method for learning colloquialisms.

3.2 Instructional Cycle

First, two colloquialisms were introduced during the first five minutes of each class period to help students comprehend their meaning and connect the meaning with existing knowledge. After class, instructional materials of these colloquialisms were sent to participants either via *WeChat* or via printed materials to create opportunities for retrieval. We chose to use *WeChat* as our app for its capability in multi-communication in text messaging, hold-to-talk voice messaging, broadcast (one-to-many) messaging, photo/video sharing, location sharing, and contact information exchange as well as its featured function of group discussion. (For more details about this app, see http://www.wechat.com/en/)

Students were encouraged to use these instructional materials to review the colloquialism taught in class. A within-subjects research design was adopted to expose all participants to learning colloquialisms by *WeChat* and by printed materials. Specifically, participants came from two classes. Students in class A received instruction via *WeChat* in week 1 and via printed materials in week 2. Students in class B received instruction in reverse order.

3.3 Instructional Materials

The researchers first selected 8 target colloquialisms that are commonly used in spoken language: 土 (tǔ) , Adjective + 成狗 (chénggǒu), 牛 (niú), 也是醉了(yěshì zuìle), 宅 (zhái)，给力 (gěilì), 二 (èr), and 打酱油 (dǎ jiàngyóu). A survey of eight native speakers (including three instructors of Chinese) was conducted to elicit usage examples. Each participant was asked to provide context-based examples for each of the target colloquialisms. In the end, only examples related to students' lives were selected for inclusion in the instructional materials. For each target colloquialism, the in-class instructional materials included a brief explanation (in English) of its usage, a few context-based examples, followed by two or three practice questions. The instructional materials sent via *WeChat* or handed out as printed materials included a brief summary of the usage in English and two example sentences that were not used in classroom instruction.

4. Research Design

4.1 Participants

Forty-six English-speaking learners of Chinese at two Midwest universities in the United States participated in this study. Participants in class A were from university A while participants in class B were from university B. Their ages ranged from 17 to 23. According to the length of formal Chinese instruction they received before the time of experiment, the participants were categorized into three groups: less than one year, 1-2

years, and more than 3 years. Participants' significant background information and the instructional sequence are summarized in table 1.

Table 1 Summary of Participants' Background Information and Instructional Sequence

Length of Formal Chinese Instruction Before treatment	Class	N	Instructional Sequence
Less than one year	A	7	Instruction via *WeChat* -> Instruction via printed materials
	B	7	Instruction via printed materials -> Instruction via *WeChat*
1-2 years	A	8	Instruction via *WeChat* -> Instruction via printed materials
	B	8	Instruction via printed materials -> Instruction via *WeChat*
More than 3 years	A	8	Instruction via *WeChat* -> Instruction via printed materials
	B	8	Instruction via printed materials -> Instruction via *WeChat*

4.2 Instruments

The vocabulary knowledge scale (VKS) developed by Paribakht and Wesche (1993) was used to design the pre-, post-, and delayed post-tests to examine students' retention of the knowledge of the instructed colloquialisms. The VKS included five statements to assess how well students know a target word:

1. *I've never seen/heard this word/phrase before.*
2. *I've seen/heard this word/phrase before, but I don't know what it means.*
3. *I've seen/heard this word before and I think it means _____.*
4. *I know what this word means. It means _____.*
5. *I know what this word/phrase means (fill the meaning in the blank in number 4) and I can use it in a sentence. Write your sentence here _____.*

The questionnaire developed by Suwantarathip and Orawiwatnakul (2015) was adapted to create 11 Likert items (see Table 2) to investigate students' attitudes toward

learning vocabulary through *WeChat*. Three open-ended questions were also included in the questionnaire to investigate students' general perceptions of the vocabulary learning experience via *WeChat*: 1) What in your experience are the benefits of using *WeChat* for learning new words/expressions? 2) What are the drawbacks of using *WeChat* for learning new words/expressions? 3) Do you have any suggestions for the use of *WeChat* in learning new words/expressions?

Table 2 Likert Items to Investigate Students' Attitudes Toward *WeChat* Assisted Vocabulary Learning

Learning vocabulary through *WeChat* was convenient	1	2	3	4	5
Learning vocabulary through *WeChat* made learning materials easier to access	1	2	3	4	5
Learning vocabulary through *WeChat* enabled me to review the vocabulary more frequently	1	2	3	4	5
Learning vocabulary through *WeChat* made learning more engaging	1	2	3	4	5
Learning vocabulary through *WeChat* was a more effective way to learn	1	2	3	4	5
Learning vocabulary through *WeChat* provided flexibility of learning	1	2	3	4	5
Learning vocabulary through *WeChat* made me feel more motivated to learn new words/expressions	1	2	3	4	5
Learning vocabulary through *WeChat* made me learn more and better	1	2	3	4	5
Learning vocabulary through *WeChat* made the learning fun	1	2	3	4	5
I prefer the learning through *WeChat* to a printed handout.	1	2	3	4	5
Learning vocabulary through *WeChat* should be integrated into our course	1	2	3	4	5

4.3 Data Collection

The four-week study was conducted in spring of 2018. One day before the instruction started, all participants received a pre-test to examine their prior knowledge of the target colloquialisms. A post-test was conducted for participants after the two-week-long instruction was completed. Two weeks after the post-test, participants received the delayed post-test and the attitudinal questionnaire. All the pre-, post-, and delayed post-tests and the attitudinal questionnaire were conducted in class. Participants were asked to complete the tests without using dictionaries or class notes.

5. Results

This section presents the descriptive statistics (see Table 3) comparing the *WeChat*-based approach and the paper-based approach on the pre-test, post-test, and delayed post-test. The results show that the mean scores of the *WeChat*-based approach were higher than those of the paper-based approach on the post-test and delayed post-test. The paired-samples *t*-test was used to further examine whether the differences of mean scores between the *WeChat*-based approach and the paper-based approach were statistically significant. One-way MANOVA analysis was employed to examine if students' previous classroom instruction length has any effect on the mean test scores of the *WeChat*-based approach. As shown in Table 4, the paired-samples *t*-test results indicate that the difference between the post-test scores of the *WeChat*-based instruction and the paper-based instruction was at a significant level ($p < .05$). Since the mean test scores of *WeChat*-based instruction are higher than that of the paper-based instruction, the results indicate that while both methods of instruction enhanced the participants' knowledge of the target colloquialisms, *WeChat*-based instruction contributed to significantly better learning outcomes. In the delayed post-test, however, the test scores of *WeChat*-based instruction were not significantly better than those of the paper-based instruction. The conclusion, therefore, is that *WeChat*-based instruction promoted students' short-term retention of Chinese colloquialisms but did not contribute to better long-term retention.

Table 3 Descriptive Statistics of Pre-, Post- and Delayed Post-Tests

Test	Instruction	N	Minimum	Maximum	Mean	Std. Deviation
Pre-test	*WeChat*	46	5.00	12.00	7.17	1.64
	Paper	46	4.00	12.00	6.74	1.54
Post-test	*WeChat*	46	8.00	20.00	16.20	3.40
	Paper	46	8.00	20.00	14.50	3.76
Delayed post-test	*WeChat*	46	6.00	20.00	13.96	3.49
	Paper	46	6.00	20.00	12.98	3.95

Table 4 The Paired-Samples *T*-test Results

		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pre-test	*WeChat Paper*	.43	1.67	.25	-.06	.93	1.77	45	.084
Post-test	*WeChat Paper*	1.70	3.85	.57	.55	2.84	2.99	45	.005
Delayed post-test	*WeChat Paper*	.98	4.67	.69	-.41	2.37	1.42	45	.163

The one-way MANOVA analysis results (see Table 5) indicate that there were no statistically significant differences in the mean pre-, post-, and delayed post-test scores of the *WeChat*-based approach among students with different lengths of previous classroom instruction. This suggests that the length of students' previous exposure to the target language did not significantly affect the learning effectiveness via *WeChat*.

Table 5 The One-way MANOVA Analysis Result

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Length of Instruction	**Pre-test: *WeChat***	12.14	2	6.07	2.40	.102
	Post-test: *WeChat*	.84	2	.42	.04	.966
	Delayed post-test: *WeChat*	24.87	2	12.43	1.02	.368

The results of the attitudinal questionnaire show that students generally held a positive attitude toward *WeChat* learning (see Table 6). Most students agreed that learning vocabulary through *WeChat* provided greater convenience and easier access to learning materials. The majority of them also supported the integration of learning vocabulary via *WeChat* into the regular course. Two items ("learning vocabulary through *WeChat* enabled me to review the vocabulary more frequently" and "I prefer the [sic] learning through *WeChat* to a printed handout") received a mean lower than three. Students' responses to the open-ended questions provided more information regarding why they gave relatively lower ratings on those two items. For example, some students reported that mobile learning might not always provide the expected mobility of learning at any time and in any place. Some students preferred to read hard copies of the instructional materials while it was not convenient to print materials sent by *WeChat*. Students also reported that the *WeChat* assisted learning did not increase opportunities for vocabulary review because they sometimes forgot to check the sent messages. Some first-time users of *WeChat* also reported that unfamiliarity with the interface of *WeChat* prevented them from fully taking advantage of the functions provided by *WeChat*, such

as checking the translation of the sentences, and copying and pasting words. These findings echo Stockwell's (2008) finding that the limitations of the mobile phone interface might affect learners' perception of the effectiveness of mobile learning.

In order to further investigate whether students' perceived usefulness of the *WeChat* assisted instruction might affect their learning outcomes, our study examined the correlations between the means of students' attitudes toward the *WeChat*-based approach and their mean scores on the post-test and delayed post-test. The results found that none of those correlations was significant (see Tables 7 and 8). This indicates that even though students might have negative attitudes toward mobile learning, these negative attitudes did not significantly affect the effectiveness of mobile learning.

Table 6 Descriptive Statistics of Students' Attitudes Toward Learning via *WeChat*

	N	Min.	Max.	Mean	Std. Deviation
Learning vocabulary through *WeChat* was convenient	46	2.00	5.00	3.85	.76
Learning vocabulary through *WeChat* made learning materials easier to access	46	1.00	5.00	3.83	1.08
Learning vocabulary through *WeChat* enabled me to review the vocabulary more frequently	46	1.00	5.00	2.98	.93
Learning vocabulary through *WeChat* made learning more engaging	46	2.00	5.00	3.22	1.03
Learning vocabulary through *WeChat* was a more effective way to learn	46	2.00	5.00	3.07	.74
Learning vocabulary through *WeChat* provided flexibility of learning	46	1.00	5.00	3.65	.87
Learning vocabulary through *WeChat* made me feel more motivated to learn new words/expressions	46	1.00	5.00	3.09	.94
Learning vocabulary through *WeChat* made me learn more and better	46	1.00	5.00	3.15	.89
Learning vocabulary through *WeChat* made the learning fun	46	1.00	5.00	3.28	.96
I prefer the learning through *WeChat* to a printed handout	46	1.00	5.00	2.74	1.25
Learning vocabulary through *WeChat* should be integrated into our course	46	1.00	5.00	3.83	.90

Table 7 Pearson Correlations Between the Means of Attitudes and the Means of *WeChat* Post-test Scores

	Correlations	Sig- (2-tailed)
Learning vocabulary through *WeChat* was convenient	.23	.12
Learning vocabulary through *WeChat* made learning materials easier to get access to	.11	.46
Learning vocabulary through *WeChat* enabled me to review the vocabulary more frequently	.06	.70
Learning vocabulary through *WeChat* made learning more engaging	.22	.15
Learning vocabulary through *WeChat* was a more effective way to learn	.19	.21
Learning vocabulary through *WeChat* provided freedom of learning	.06	.699
Learning vocabulary through *WeChat* made me feel more motivated to learn new words/expressions	.09	.57
Learning vocabulary through *WeChat* made me learn more and better	.09	.54
Learning vocabulary through *WeChat* made the learning fun	.28	.06
I prefer the learning through *WeChat* to a printed handout.	.20	.18
Learning vocabulary through *WeChat* should be integrated into our course	.29	.05

Table 8 Pearson Correlations Between the Means of Attitudes and the Means of *WeChat* Delayed Post-test Scores

	Correlations	Sig- (2-tailed)
Learning vocabulary through *WeChat* was convenient	.03	.84
Learning vocabulary through *WeChat* made learning materials easier to get access to	.08	.60
Learning vocabulary through *WeChat* enabled me to review the vocabulary more frequently	.14	.34
Learning vocabulary through *WeChat* made learning more engaging	.26	.09
Learning vocabulary through *WeChat* was a more effective way to learn	.22	.13
Learning vocabulary through *WeChat* provided freedom of learning	.14	.35
Learning vocabulary through *WeChat* made me feel more motivated to learn new words/expressions	.12	.44
Learning vocabulary through *WeChat* made me learn more and better	.12	.41
Learning vocabulary through *WeChat* made the learning fun	.28	.06
I prefer the learning through *WeChat* to a printed handout.	.-01	.93
Learning vocabulary through *WeChat* should be integrated into our course	.65	.07

6. Discussion

Our research finds that sending vocabulary-learning materials via mobile messages can take advantage of learners' preference for using mobile devices, and thereby increase the opportunity of incidental vocabulary learning. However, such a mechanism only promotes learners' short-term retention of the vocabulary. In order to promote learners' long-term retention, a greater variety of instructional activities should be designed to engage learners in the retrieval of the newly learned words on a regular basis. For example, learning materials and practice questions can be sent to students via messages on an regular interval basis. This study also finds that the length of previous language learning experience does not have a statistically significant role in the effectiveness of mobile-assisted vocabulary learning. This suggests that mobile-assisted learning can be effective for teaching L2 learners at beginning, intermediate, and advanced levels. Moreover, the study reveals that there are no statistical correlations between learners' attitudes toward mobile learning and their learning outcomes. Learners can still be engaged in the learning of new words even when they do not have very positive attitudes toward the perceived usefulness of the mobile-assisted approach, as long as the mobile-assisted learning activities are designed to connect with in-class instruction.

Comments by students on their learning experiences reveal that information was delivered fast and direct via *WeChat* messages so that they could look at the learning materials sooner and more easily than the printed materials. Mobile-assisted learning provided learners with easy access and mobility for reviewing the learning materials outside of class. Several students mentioned that the mobile device enabled them to get feedback more quickly since they could ask questions about the new words and get responses from teachers or peer students instantaneously through *WeChat* messages. A few students also reported that they could use the translation function of *WeChat* to see what a word means, which helped their understanding of the learning materials. This suggests that mobile devices not only provided mobility and accessibility for vocabulary learning, but also offered other functions (such as instant chatting, dictionary, or translation) to help learners resolve questions encountered in self-learning.

7. Conclusion

The learning of new vocabulary requires students to invest time and effort in order to have the new information processed in the brain. Mobile learning provides great potential to engage learners in learning activities to process new information via both audio and visual sensory channels in or outside of class. Moreover, the prevalent use of mobile devices among the young generation of learners facilitates the creation of an environment in which their daily routines and learning activities coalesce to help them notice or retrieve the new words. This study confirms that mobile-assisted learning can promote students' short-term retention of vocabulary regardless of their previous language learning experiences or their attitudes towards mobile learning. Therefore,

mobile learning can be utilized effectively to assist vocabulary learning in foreign language instruction at all levels, even when students may not all have a positive attitude toward mobile learning. Our study only employed mobile devices as a means to present and deliver the learning materials and found that, compared to the traditional method of delivering paper-based instructional materials, mobile devices are a more effective medium to promote short-term retention. More investigation is needed to ascertain if adding more variety to the types of learning activities (such as using the new words to interact with other peer students) would more actively engage students and promote long-term retention of the new words.

References

Burston, J. (2014). The reality of MALL: Still on the fringes. *CALICO Journal*, *31*(1), 103–125.

Cavus, N., & Ibrahim, D. (2009). M-Learning: An experiment in using SMS to support learning new English language words. *British Journal of Educational Technology*, 40, 78-91.

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London: MacMillan.

Kiernan, P.J., & Aizawa, K. (2004). Cell phones in task based learning--Are cell phones useful language learning tools? *ReCALL*, *16*(1), 71-84.

Kubler, C. (2007). Teaching advanced-level spoken Chinese grammar and vocabulary to non-native learners. In X. Cui (Ed.), *Chinese language pedagogy: Interaction and mutual complementation in China and overseas* (pp. 169-184). Beijing, China: The Commercial Press. [Kubler, C. (2007). 高级汉语口语语法与词汇的训练：对汉语非其母语学生的教学法研讨. 崔希亮（主编），*汉语教学：海内外教学的互动与互补* (pp. 169-184), 北京：商务印书馆.]

Liu, Y., Li, H., & Carlsson, C. (2010). Factors driving the adoption of M-learning. An empirical study. *Computers and Education*, *55*(3), 1211–1219.

Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of Computer Assisted Learning*, *24*(6), 515-525.

Ma, Q. (2014). A contextualized study of EFL learners' vocabulary learning approaches: Framework, learner type and degree of success. *Asia TEFL*, *11*(3): 33–71.

Ma, Q. (2017). Technologies for teaching and learning vocabulary. In C. A. Chapelle & S. Sauro (Eds.), *The handbook of technology and second language teaching and learning* (pp. 45-61). Hoboken, NJ: Wiley-Blackwell.

Miangah, T. M., & Nezarat, A. (2012). Mobile-assisted language learning. *International Journal of Distributed and Parallel Systems*, *3*(1), 309-319.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.

Paribakht, T.S., & Wesche, M. (1993). The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal*, *11*(1), 9-29.

Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies of Second Language Learning*, *21*(4), 589-619.

Schmitt, N., & McCarthy, M. (1997). *Vocabulary: Description, acquisition and pedagogy*. New York: Cambridge University Press.

Stockwell, G. (2007). Vocabulary on the move: Investigating an intelligent mobile phone-based vocabulary tutor. *Computer Assisted Language Learning*, *20*(4), 365–383.

Stockwell, G. (2008). Investigating learner preparedness for and usage patterns of mobile learning. *ReCALL*, *20*(3), 253–270.

Suwantarathip, O., & Orawiwatnakul, W. (2015). Using mobile-assisted exercises to support students' vocabulary skill development. *The Turkish Online Journal of Educational Technology, 14*(1), 163-171.

Modular Approaches in eLearning Design: Computer-Assisted Pronunciation Training Design and Evaluation
(模塊化電子教學設計：電腦輔助發音訓練設計與評測)

Lee, Siu-lun
(李兆麟)
The Chinese University of Hong Kong
(香港中文大學)
slee@cuhk.edu.hk

Abstract: Pronunciation is one of the more difficult aspects in teaching and learning Chinese and Cantonese as a second language. This paper suggests a modular approach in designing pronunciation training through eLearning. This paper discusses two models in computer-assisted pronunciation training (CAPT) modules, namely the accuracy model and the fluency mode. This paper also presents Pre-test-Intervention-Post-test experimental research, which is intended to evaluate these modules. This research uses experimental research to examine how the configurations of linguistic knowledge teaching and language skills practice affect the effectiveness of learning Cantonese as a second language.

摘要: 掌握正確發音一直是華語二語和粵語二語教學的教學重點。本文建議以電子教學模塊概念設計電腦輔助發音訓練模型。本次研究用了"前測-介入-後測實驗"去評估兩套電腦輔助發音訓練模型：發音準確度模型和流暢度模型的可用性。並且查考語言知識教學和語言技能練習的排列和佈局對粵語二語學習的影響。

Keywords: Modular approach, teaching Cantonese as a second language, computer-assisted pronunciation training (CAPT), eLearning design and evaluation

關鍵詞: 模塊概念設計，粵語二語學習，電腦輔助發音訓練，電子教學設計和評估

1. Introduction

Training for pronunciation accuracy is an important area in foreign language teaching (Chen, 1983; Macdonald, Yule, & Powers, 1994; He, 2011; Gilakjani & Sabouri, 2016; Huensch & Thompson, 2017). Pronunciation accuracy is also one of the targets set by Chinese as a second language (CSL) teachers and learners (Hsiao & Shicock, 2006;

Wu & Miller, 2007) and for learners learning Cantonese as a second language (Ball, 1883). It has been pointed out that Cantonese pronunciation (Ball, 1883; Lee 2005) is one of the most difficult areas when learning the language. Research on studying methods to train pronunciation in the classroom (Tominaga, 2009; Foote, Trofimovich, Collins, & Urzúa, 2013) and by using educational technologies (Mushangwe, 2014; Hanna & Gao, 2016) has been conducted already. In literature on second language teaching and learning, many researchers believe that there is a multi-componential aspect to the construct of L2 proficiency and L2 performance. The principal dimensions can be captured by the notions of "complexity," "accuracy," and "fluency" (Skehan, 1998; Ellis, 2003, 2008; Housen & Kuiken, 2009; Larsen-Freeman 2009). These three notions, "complexity," "accuracy," and "fluency" became major research variables in applied linguistic research starting from the 1980s, and were also used as performance descriptors in L2 assessment as well as indicators for progress in language learning.

This project proposes two models in computer-assisted pronunciation training (CAPT) and uses experimental research to examine how the configuration of linguistic knowledge teaching and language skills practice in CAPT module designs affect the effectiveness of learning Cantonese as a second language. The eLearning design under investigation consists of two models, namely the "accuracy model" and the "fluency model." In my research, 12 non-Cantonese speaking international university undergraduate students were recruited to participate in a "Pre-test-Intervention-Post-test" experiment to investigate the effectiveness of the eLearning model designs. The "Pre-test-Intervention-Post-test" experiment focused on pronunciation accuracy and fluency. Preliminary findings provide evidence to support web-based computer-assisted pronunciation training (CAPT) and the hypothesis that CAPT should include both audio-visual feedback and annotated PowerPoint videos in order to enhance pronunciation accuracy.

2. Literature Review: Research on eLearning Design

Studies confirming that eLearning has an important role in higher education and has positive effects on students' motivation have been published (Kaewkiriya, 2013; Harandi, 2015). Researchers further investigated the content presented in eLearning (Boyle, 2003; Kanuka, 2006) or how content was presented (Alsadhan, Alhomod, & Shafi, 2014). Kanuka (2006) points out that the design of eLearning needs to connect content and pedagogy. Steen (2008) suggests that there is no "one-size-fits-all" eLearning software and that eLearning designers should take into account learning and/or training theories and should understand the knowledge or skills to be taught in order to achieve effective eLearning results. Some researchers suggest that eLearning design should be grounded on a core principle of learning theories (Sims, 2006; Pange & Pange, 2011), and should take into account learners' learning styles (Cooze and Barbour, 2007; Cercone, 2008; Sangsawang, 2015). eLearning practitioners have tried to build different eLearning models based on different theories and different learning needs (Koohang et al., 2009; Alonso et al., 2005). The similarities among these models were learner-centered and theoretically based.

3. eLearning Mode: Design and Presentation

The computer-assisted pronunciation training (CAPT) designed in this study consists of two models, namely the "accuracy model" and the "fluency model." The concept of this dual-model design is based on the dichotomy of "accuracy and fluency" used by Brumfit (1984), who distinguished between "accuracy-oriented activities," which focus on linguistic forms, and "fluency-oriented activities," which foster spontaneous oral L2 production. "Accuracy-oriented activities," such as pronunciation drills and vocabulary drills, were used in the teaching of a new target item. "Fluency-oriented activities," such as extensive reading and question-and-answer (Q/A) exercises, are aimed at developing students' spontaneous application of what they have already learned. Table 1 below shows the purposes, content, and activities based on accuracy and fluency.

Table 1. The purposes and content of accuracy activities and fluency activities (adopted from Brumfit (1984)

Accuracy Activities	Fluency Activities
Purpose: the primary purpose is to help students achieve accurate perception and production of language forms.	**Purpose**: the primary purpose is to help students develop language fluency in using the language in spontaneous communication.
Content: the texts are usually composed of discrete items, such as sentences or words; the target items are usually practiced out of context.	**Content**: the texts are usually discursive length, such as conversations and stories; dialogues are spoken and articles and/or written stories are read; an effort is made to use authentic material from real life.
Activities: students' attention is focused on a particular target item; their output is usually predictable; student performance is assessed on how few language mistakes are made; students' errors are corrected; tasks do not usually simulate real-life situations.	**Activities**: students' attention is focused on communicating information and expressing ideas; their output may not always be predictable; their performance is assessed on how well ideas are expressed or understood; students' errors are not corrected when they interfere with communication; tasks often simulate real-life situations.

Other research has been undertaken which examines effective user interface design for eLearning software (Faghih, Azadehfar, & Katebi, 2013). Apart from addressing some of the issues discussed in previous studies, the eLearning interface design of the two models in this study has taken into account three major aspects discussed in the literature on eLearning design in foreign language teaching. The 3 aspects are: 1) Presentation of the speaker's face; 2) presentation of waveforms; 3) teaching of spectrographic analysis.

1) Presentation of the speaker's face

Hardison's (2003) study presents experimental research on the effect of the presentation of the speaker's (teacher's) face in eLearning design for pronunciation training. He studied two experimental groups. The first group used an eLearning model with audio-visual presentation in which the speaker's (teacher's) face appeared visually, and the second group used an eLearning model with audio-only presentation. Hardison concluded that the audio-visual group showing the speaker's face outperformed the audio-only group when training Japanese and Korean learners to distinguish /r/ and /l/. Hardison (2006) further confirmed that students using eLearning models with faces familiar to the subjects enhanced their comprehension.

2) Presentation of waveforms

Okuno (2013) suggests that waveform display helps learners acquire Japanese geminates. Okuno and Hardison (2016) further elaborate that waveform display in eLearning pronunciation training models is useful to perceive and produce vowels with durational difference.

3) Teaching of spectrographic analysis

Researchers (Olson, 2014; Quntana-Lara, 2014) conclude that teaching students to analyze the spectrograms and to compare their own pronunciation with native speaker's pronunciation led to an improvement on consonant and vowel production. Olson's (2014) study focuses on L1 English speakers acquiring Spanish intervocalic stops. The results show that teaching students to analyze spectrographs in eLearning design improved students' consonant production. Quintana-Lara (2014) shows positive effects of spectrographic analysis toward the acquisition of Spanish pre-service English teachers' vowel production.

Based on studies of the presentation of a speaker's face, the presentation of waveforms, and the teaching of spectrographic analysis, the user interface for the accuracy model and fluency model were developed. Figure 1 shows the design and presentation of the "accuracy model" and Figure 2 shows the presentation of the "fluency model" used in this study.
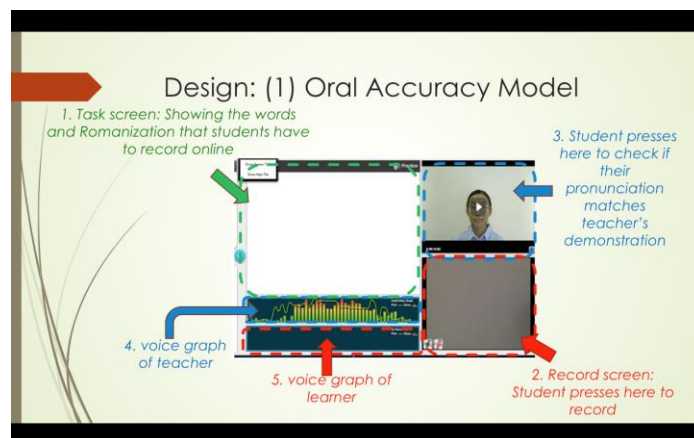

Figure 1 Design and presentation of the "accuracy model"

The design of the "accuracy model" contains five major parts, as shown in Figure 1 above. The first component (upper-left corner on the computer screen) contains a task screen, which shows the words (with Chinese characters and Cantonese romanization) that students are required to read out loud and record online. The second component (lower-right corner) shows a recording screen. The third component (upper-right corner) shows the teachers' face while pronouncing the word on the task screen. Students can watch the teachers' demonstration before recording themselves. The fourth and fifth components (lower-left corner) contain the voice graph of the teacher and the voice graph of the student. Students can compare their pronunciation with the teachers.' Students were trained to read voice graphs (spectrographs) when doing the tasks in this "accuracy model."

Figure 2 below shows the design and presentation of the "fluency model." The "fluency model" took the form of a simulation exercise. The first component (upper-left corner on the computer screen) shows visual prompts while the second component (upper-right corner) shows the face of the teacher who was asking questions. The third component (lower-left corner) shows the teacher's voice graph and the student's voice graph. The voice graphs indicate turn-taking timing and the flow of the conversations required by the task. The student's face is shown in the fourth component (lower-right corner) and the student's responses were recorded.

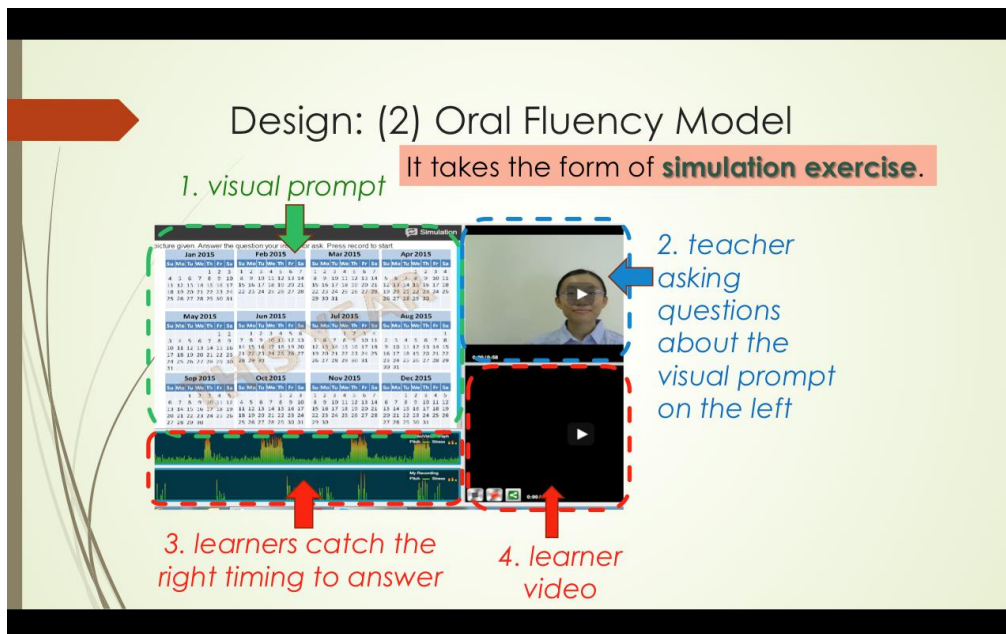**Figure 2 Design and presentation of the "fluency model"**

4. Research Questions and Methodology

This research uses a single group design with repeated measurements. The research design uses a one-group pre-test—post-test design. The pre-test—post-test design was adopted for evaluating the impact of computer-assisted pronunciation training (CAPT) used when students were learning Cantonese as a second language. In the

pre-test, students were given 25 Cantonese two-syllable words and 10 sentences to read. The 25 Cantonese words were terms appearing in beginners' textbooks. For example, "附近 fuhgahn (nearby)," or "瞓覺 fangaau (to sleep)." The 10 sentences including, "中文書，英文書，我都想買 jūngmàhnsyū, yīngmàhn syū, ngóh dōu séung máaih ('I want to buy both Chinese and English books')", were also selected from beginners' text books. Both Chinese characters and Cantonese romanization were given to students in the pre-test.

After finishing the pre-test, students were asked to undertake training sessions (intervention) using two eLearning pronunciation training models, namely the "accuracy model" and the "fluency model." There were 15 sessions using the "accuracy model" and 20 sessions using the "fluency model." The 15 "accuracy model" sessions included: 1 overview session; 6 sessions on Cantonese tones; 7 sessions on Cantonese finals, and; 1 review session. The 20 "fluency model" sessions included oral Q/A practice, where prompts (presented in English, Chinese characters, and Cantonese Romanization on screen) were given to students. 4 questions used in everyday life, such as, "What is your name?" and "What time is it?" were posed to students in each session. Students watched and observed the teacher's demonstration on screen after their trials and students could record themselves again and make corrections when they found that their output had deviated from the teacher's demonstration.

Table 2 below shows the interventions used in the training sessions in this research. Each training session lasts around 30 minutes. Subsections specified in the content last around 5-10 minutes. Pronunciation videos were used to complement the audio-visual feedback training in the two models.

Table 2 Interventions used in the training sessions

Training session	Content	Format
1.	1.1 Introduction to Cantonese pronunciation	PowerPoint video with narration
	1.2 General Cantonese pronunciation exercises	Web-based matching exercises
	1.3 High level tone	PowerPoint video with narration
	1.4 High level tone pronunciation exercises	Audio-visual feedback exercises
2.	2.1 Review of training 1	Oral response on explicit knowledge
	2.2 High rising tone	PowerPoint video with narration
	2.3 High rising tone pronunciation exercises	Audio-visual feedback exercises
	2.4 Long "e" and short "e"	PowerPoint video with narration
	2.5 Long "e" and short "e" pronunciation exercises	Audio-visual feedback exercises

3.	3.1 Review of training 2	Oral response on explicit knowledge
	3.2 Mid level tone	PowerPoint video with narration
	3.3 Mid level tone pronunciation exercises	Audio-visual feedback exercises
	3.3 Long "u" and short "u"	PowerPoint video with narration
	3.4 Long "u" and short "u" pronunciation exercises	Audio-visual feedback exercises
4.	4.1 Review of training 3	Oral response on explicit knowledge
	4.2 Long "yu"	PowerPoint video with narration
	4.3 Long "yu" pronunciation exercises	Audio-visual feedback exercises
	4.4 Long "i" and short "i"	PowerPoint video with narration
	4.5 Long "i" and short "I" pronunciation exercises	Audio-visual feedback exercises
5.	5.1 Review of training 4	Oral response on explicit knowledge
	5.2 Low falling tone	PowerPoint video with narration
	5.3 Low falling tone pronunciation exercises	Audio-visual feedback exercises
	5.4 Long "o" and short "o"	PowerPoint video with narration
	5.5 Long "o" and short "o" pronunciation exercises	Audio-visual feedback exercises
6.	6.1 Review of training 5	Oral response on explicit knowledge
	6.2 Low rising tone	PowerPoint video with narration
	6.3 Low rising tone pronunciation exercises	Audio-visual feedback exercises
	6.4 Low level tone	PowerPoint video with narration
	6.5 Low level tone pronunciation exercises	Audio-visual feedback exercises
7.	7.1 Review of training 6	Oral response on explicit knowledge
	7.2 Long "eu" and short "eu"	PowerPoint video with narration
	7.3 Long "eu" and short "eu" pronunciation exercises	Audio-visual feedback exercises
	7.4 Long "a" and short "a"	PowerPoint video with narration
	7.5 Long "a" and short "a"	Audio-visual feedback exercises

8.	pronunciation exercises	
	8.1 Review of training 7	Oral response on explicit knowledge
	8.2 Overall review - Cantonese tones	PowerPoint video with narration
	8.3 Six tones pronunciation exercises	Audio-visual feedback exercises
	8.4 Overall review – final lengthening	PowerPoint video with narration
	8.5 Final lengthening pronunciation exercises	Audio-visual feedback exercises

After students completed the training sessions, they were given a post-test, which required them to read 10 bi-syllabic Cantonese words and 5 Cantonese sentences.

12 full-time university undergraduates participated in this research. The subjects were recruited through social media and all the subjects were asked to fill out a language background form. None of them had any Cantonese background at all at the time the research was conducted. All the subjects filled out a research consent form (required by the university's Research Ethics Committee) and the students' biographic data was not disclosed in the research. All 12 participants finished the pre-test. 2 students dropped out during the training sessions. Eventually 10 students completed the entire research process and completed the post-test.

5. Data Analysis and Discussion

This research examines whether the combination of the accuracy model and the fluency model improved second language learners' Cantonese pronunciation. In what ways may the models be useful? The research also looks at whether the models are useful in improving vowel accuracy, consonant accuracy, or rhyme duration and what the representational disfluency types may be among elementary Cantonese learners.

This study employed a one-group Pre-test—Intervention—Post-test design to study the effect of audio-visual training on second language learning, with a focus on pronunciation training in teaching Cantonese as a second language.

For the "accuracy model," 200 test tokens in the pre-test and post-test were extracted through PRAAT, with rhyme duration changing between pre-test and post-test. Instead of formant values and trajectories, these changes are the focus. The results of this study are in line with the plausibility of previous computer-assisted pronunciation training (CAPT) studies (Neri et al, 2008; Luo, 2014).

Based on the analysis, there are several findings on the subjects' accuracy.

1. The duration of the second syllable of the learners was lengthened significantly with a medium effect size.
2. The ratio of Rhyme2/Rhyme1 was closer to the native speaker norm after

training.
3. There is less variance in bi-syllabic word ratio after intervention.
4. The results of two-way ANOVA show that both token-type (Tone1, Tone3, Tone6) and previous Mandarin level of participants (intermediate level or above versus beginning level) do not moderate statistically the effects of training.

Rhyme duration gains as shown in Figure 3 were different (ranked from Tone6, Tone1 and Tone3). However, the interaction effect was not significant, indicating that the gains were not statistically different. Partial eta squared is 0.55, indicating a medium effect size.
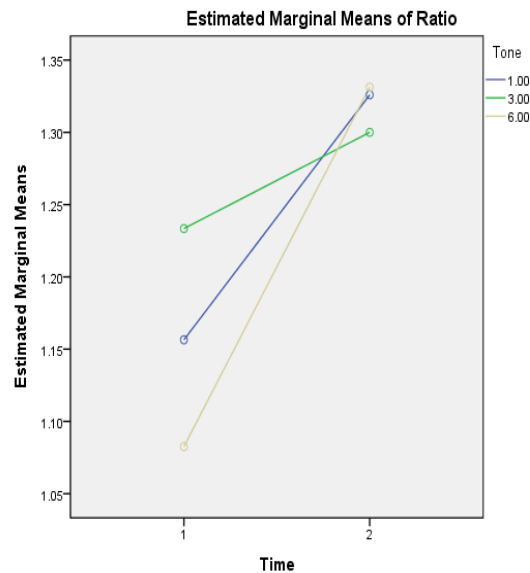

Figure 3 Rhyme duration gains of Cantonese tone

Rhyme duration gains, as shown in Figure 4, were different among subjects. Subjects with lower Mandarin level (M2) benefited more than subjects with higher Mandarin level (M1). However, the interaction effect was not statistically insignificant. Partial eta squared is 0.6, indicating a medium effect size.
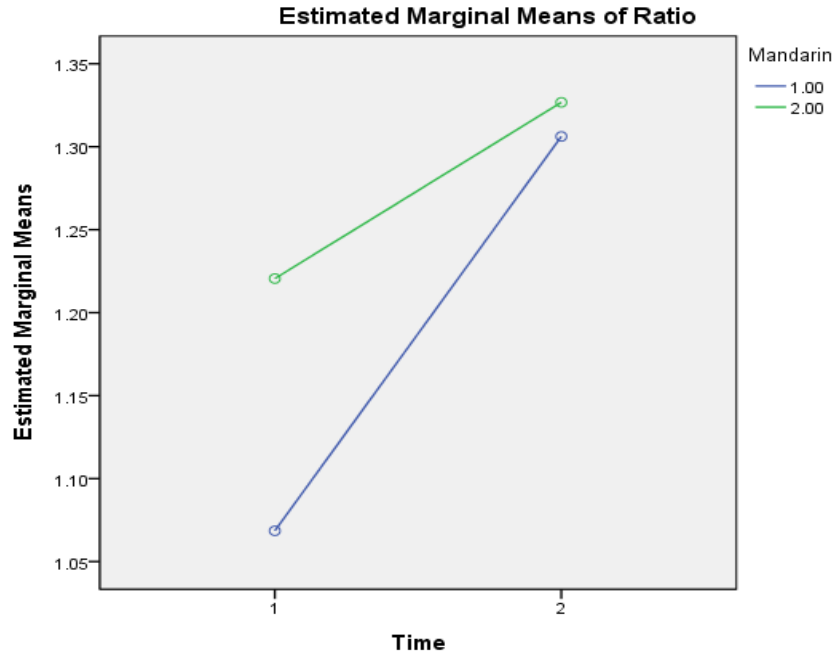
Figure 4 Rhyme duration gains with respect to Mandarin background

For the "fluency model," trials from the 10 students in this research cohort were extracted for analysis. The overall question completion rate is around 60%. This may suggest that the "fluency model" may need to accommodate slower learners in order to raise the completion rates. In order to identify representational disfluency types among elementary learners, trials of the students have been transcribed and disfluency types (filler, breakdown, and pauses) were analyzed. Based on the preliminary analysis, the study concludes that: 1) the number of words uttered correlates with a breakdown having a medium effect size ($r=0.6$, $p<0.05$), and; 2) the number of words uttered does not correlate with conversational fillers or pauses.

Further investigations need to be carried out in order to confirm whether or not the conversational fillers used were a transfer of first language speaking style. Moreover, professional raters' and native speakers' perception and judgement on accuracy/intelligibility may be used in further research to test whether or not a relationship exists between disfluency types and listener perception and whether there are correlations between raters'/native speakers' judgement and rhyme duration gains.

6. Conclusion

In this research, the subject's rhyme duration and first syllable to second syllable rhyme ratio have improved. The current results encourage a more extensive use of web-based computer-assisted pronunciation training (CAPT) to enhance Cantonese language teaching and learning.

As to the improvement of learners' pronunciation accuracy, the subject's rhyme

duration and first syllable to second syllable rhyme ratio have improved in this study. The results could be extended to other second language teaching and learning contexts. In some languages, such as English and Japanese, duration also played an important role. The current results encourage a more extensive use of web-based computer-assisted pronunciation training (CAPT) to enhance language teaching and learning, especially in pronunciation training courses at higher educational levels. The results show that teacher pronunciation demonstration videos were beneficial. Web-based audio-visual feedback is also useful to support pronunciation training along with classroom instruction. Moreover, pronunciation knowledge could also be delivered in video format so that students could learn according to their own needs.

As to learners' fluency, this study, through investigating fluency tokens produced by elementary Cantonese learners, concludes that the frequency of disfluency types (conversational fillers, breakdown, and pauses) differed among the participating subjects. The current study suggests that breakdown type disfluency occurred more frequently if learners tried to speak more, while filler type disfluency did not have such a correlation. Further investigation could be carried out to confirm whether or not filler type disfluency tends to be a transfer of first language speaking style. The results of this research have implications for designing overall oral assessment rubrics by considering learners' first language interference. The results could be extended to disciplines such as education studies, second language teaching, and second language assessment.

References

Alonso, F., Lopez, G., Manrique, D., & Viñes, J. M. (2005). An instructional model for web-based e-Learning education with a blended learning process approach. *British Journal of Educational Technology*, *36*(2), 217-235.

Alsadhan, A. O., Alhomod, S., & Shafi, M. M. (2014). Multimedia based e-Learning: design and integration of multimedia content in e-Learning. *International Journal of Emerging Technologies in Learning*, *9*(3), 26-30. Retrieved from http://dx.doi.org/10.3991/ijet.v9i3.3308.

Ball, J. D. (1883). *Cantonese made easy.* Hong Kong: China Mail Office.

Boyle, T. (2003). Design principles for authoring dynamic, reusable learning objects. *Australian Journal of Educational Technology*, *19*(1), 46-58.

Brumfit, C. J. (1984). *Communicative methodology in language teaching*. Oxford: Cambridge University Press.

Cercone, K. (2008). Characteristics of adult learners with implications for online learning design. *AACE Journal*, *16*(2), 137-159.

Chen, Y. (1983). Principles and methods of the teaching of pronunciation. *Language Learning and Communication*, *2*(1), 63-67.

Cooze, M., & Barbour, M. K. (2007). Learning styles: A focus upon e-Learning practices and their implications for successful instructional design. *Journal of Applied Educational Technology*, *4*(1), 7-20.

Czwenar, I. (2014). Analysing spoken language for complexity, accuracy and fluency: Some methodological considerations. In W. Szubko-Sitarek, L. Salski, P. Stalmaszczyk (Eds.), *Language learning, discourse and communication* (pp. 81-92). New York, NY: Springer.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.

Faghih, B., Azadehfar, M. R., & Katebi, S. D. (2013). User interface design for e-Learning software. *The International Journal of Soft Computing and Software Engineering*, *3*(3), 786-794.

Finger, G. (2014). eLearning: Exploring digital futures in the 21st century. *Education Sciences*, *4*, 209-212.

Foote, J. A., Trofimovich, P., Collins, L., & Urzúa, F. S. (2013). Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal*, *44*(2), 181-196.

Gilakjani, A. P., & Sabouri, N. B. (2016). How can EFL teachers help EFL learners improve their English pronunciation? *Journal of Language Teaching and Research*, *7*(5), 967-972.

Hanna, B. E., & Gao, Y. (2016). Exploring optimal pronunciation teaching integrating instructional software into intermediate-level EFL classes in China. *CALICO Journal*, *33*(2), 201-230.

Harandi, S. R. (2015). Effects of e-Learning on students' motivation. *Social and Behavioral Sciences, 181*, 423-430.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, *24*(4), 495-522.

Hardison, D. M. (2006). Effects of familiarity with faces and voices on second-language speech processing: Components of memory traces. In Stern R. M. (Ed.). (2006) *Proceedings of Ninth International Conference on Spoken Language Processing*, Pittsburgh, PA, USA, September 17-21, 2006.

Hardison, D. M. (2007). The visual element in phonological perception and learning. In M.C. Pennington (Ed.), *Phonology in context* (pp. 135–158). Basingstoke: Palgrave Macmillan.

Hardison, D. M. (2010). Visual and auditory input in second-language speech processing. *Language Teaching*, *43*(1), 84-95.

He, L. (2011). Metacognition in English as a foreign language pronunciation learning among Chinese tertiary learners. *Applied Language Learning*, *21*(1-2), 1-30.

Hirumi, A. (2002). The Design and sequencing of eLearning interactions: A grounded approach. *International Journal on E-Learning, 1*(1), 19-27.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30*(4), 461–473.

Hsiao, J. H., & Shilcock, R. (2006). Analysis of a Chinese phonetic compound database: Implications for orthographic processing. *Journal of Psycholinguistic Research*, *35*(5), 405-426.

Huensch, A., & Thompson, A. S. (2017). Contextualizing attitudes towards pronunciation: Foreign language learners in the United States. *Foreign Language Annals*, *50*(2), 410-432.

Kaewkiriya, T. (2013). A design and development of e-Learning content for multimedia technology using multimedia games. *International Journal of Software Engineering & Applications*, *4*(6), 61-69.

Kanuka, H. (2006). Instructional design and eLearning: A discussion of pedagogical content knowledge as a missing construct. *E-Journal of Instructional Science and Technology*, *9*(2), 1-17.

Koonhang, A., Riley, L., Smith, T., & Schreurs, J. (2009). E-learning and constructivism: From theory to application. *Interdisciplinary Journal of E-learning and Learning Objects*, *5*, 91-108.

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30*(4), 579-589. doi.org/10.1093/applin/amp043.

Lee, J., & Lee, Y. (2012). Development and application of e-Learning content for advertising education. *International Journal of Advanced Science and Technology*, *47*, 1-12.

Lee, S. (2005). *History and current trends of teaching Cantonese as a foreign language: Investigating teaching approaches* (Ed.D. dissertation). University of Leicester, Leicester, UK.

Luo, B. (2014). Evaluating a computer-assisted pronunciation training (CAPT) technique for efficient classroom instruction. *Computer Assisted Language Learning*, *29*(3), 451-476.

Macdonald, D., Yule, G., & Powers, M. (1994). Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning*, *44*(1), 75-100.

Mushangwe, H. (2014). De-foreignizing a sound: Computer-assisted pronunciation practice in learning a foreign language. *Theory and Practice in Language Studies*, *4*(2), 303-312.

Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning*, 21(5), 393-408.

Okuno, T. (2013). *Acquisition of L2 vowel duration in Japanese by native English speakers* (Unpublished Ph.D. dissertation). Michigan State University, East Lansing, MI.

Okuno, T., & Hardison, D. M. (2016). Perception-production link in L2 Japanese vowel duration: Training with technology. *Language, Learning and Technology*, *20*(2), 61-80.

Olson, D. J. (2014). Benefits of visual feedback on segmental production in the L2 classroom. *Language Learning & Technology*, *18*(3), 173-192.

Pange, A., & Pange, J. (2011). Is e-Learning based on learning theories? A literature review. *International Journal of Educational and Pedagogical Sciences*, *5*(8), 932-936

Quintana-Lara, M. (2014). Effect of acoustic spectrographic instruction on production of English /i/ and /l/ by Spanish pre-service English teachers. *Computer Assisted Language Learning*, *27*(3), 207-227.

Sangsawang, T. (2014). Instructional design framework for educational media. *Social and Behavioral Sciences*, *176*, 65-80.

Sims, R. (2006). Beyond instructional design: Making learning design a reality. *Journal of Learning Design*, *1*(2), 1-7.

Skehan, P. (1998). *A Cognitive Approach to Language Learni*ng. Oxford: Oxford University Press.

Steen, H. L. (2008). Effective eLearning design. *MERLOT Journal of Online Learning and Teaching*, *4*(4), 526-680.

Tominaga, Y. (2009). An analysis of successful pronunciation learners: In search of effective factors on pronunciation teaching. *Journal of Pan-Pacific Association of Applied Linguistics*, *13*(1), 127-140.

Wang, Z. (2014). Developing accuracy and fluency in spoken English of Chinese EFL learners. *English Language Teaching*, *7*(2), 110-118.

Wu, H., & Miller, L. K. (2007). A tutoring package to teach pronunciation of Mandarin Chinese characters. *Journal of Applied Behavior Analysis*, *40*(3), 583-586.

Using Educational Technology to Enhance the Three Modes of Communication
(應用科技提高二語學習者三种沟通模式的能力)

Wu, Ching-Hsuan	Childs, Lily	Hsu, I-Ling
(吳青璇)	(林秀桂)	(徐翊玲)
Western Kentucky University	Mary Institute and St Louis Country Day School	University of Rhode Island
(西肯塔基大學)	(瑪麗聖路易斯學校)	(羅德島大學)
ching.hsuan.wu@gmail.com	lchilds@micds.org	ilinghsu@uri.edu

Abstract: This study explores recently innovated language learning tools to emphasize their use in engaging learners to develop second language skills in the three modes of communication: the interpersonal, interpretive, and presentational. A literature review and our own use of these tools in the classroom suggest that Chinese learners across age groups and proficiency levels find freely accessible technology, including Go Formative, Edpuzzle, and Google Sites, motivating. This study presents evaluative data, focuses on each tool's pedagogical effects on the three communicative modes, and provides instructors with evaluation rubrics and task examples as a foundation for incorporating these tools into classroom use. In addition, the present study recognizes limitations in using pedagogical technology, and discusses them from both instructor and learner perspectives.

摘要：本研究探討如何將近年來創新的科技教學應用於培養二語學習者的三種溝通技能：(1) 雙向溝通、(2) 理解詮釋、以及 (3) 表達演示。根據相關文獻以及本篇文章作者們收集的研究資料顯示，總體而言不同年齡群和不同中文程度的學習者均認為像是 Go Formative、Edpuzzle 跟 Google Sites 等之類的開放式科技應用有益于提升他們的學習動機。為了能讓教師們容易地將這些科技工具運用在他們的課堂之中，本研究提出評量性的數據、說明每個科技工具在三種溝通模式方面的教學效果，并提供測評的評分表和溝通任務的實例。此外，本研究從教師及學習者的角度探討科技使用的限制與挑戰。

Keywords: Second language acquisition, pedagogical technology, three modes of communication

關鍵詞：二語習得，教學科技，三項溝通技能

1. Introduction

The use of pedagogical technology to enhance second language acquisition (SLA) has received attention from both researchers and classroom teachers. For the latter, it is often the practical needs in teaching and learning that drive the selection of technology tools. A teacher's decision can be made based on ways that tools can potentially help learners accomplish tasks, achieve proficiency goals, and support student needs in becoming world-ready. To leverage pedagogical technology in an attempt to support second language (L2) learning is an overall goal for its users. Specifically, analyses and discussions on how different tools yield varying learning outcomes in diversified learning contexts can inform L2 instructors on the applications of pedagogical technology in their classrooms and the aspects of one's language skills that these tools can improve.

This study aims to report the extent to which three technology tools, Go Formative, Edpuzzle, and Google Sites, address the development of language functions in the three modes of students' communication skills: the interpersonal, interpretive, and presentational. These three applications are selected to address learners' language development in the three modes of communication based on their common and independent features. Go Formative, Edpuzzle, and Google Sites share two important features in common: cloud computing and content creation. Cloud technology enables all the participating members, including instructors and learners, to store and access data and programs over the internet, instead of on their computer's local drive. The anytime-anywhere open accessibility on any internet-connected device makes these three tools welcoming and user-friendly. In addition, teachers are able to conveniently use these three applications to create learning content that aligns with the desired pedagogical outcomes in order to hone students' skills.

Moreover, the individual features of these tools further their utility. Go Formative provides opportunities for learners to improve both writing and speaking interpersonal communication skills. Individual students can interact with teachers and peers through question-answer exchanges based on prompts designed by the teachers or they can form discussion forums to engage more peers. Edpuzzle is a useful tool to improve learners' interpretive language skills because it allows teachers to conveniently edit authentic videos and embed comprehension questions that meet the objectives of their courses. On the other hand, students can replay the edited videos as many times as they need to achieve the desired comprehensibility. They can then review their progress along with their wrong answers any time of the day and wherever they are. Finally, Google Sites offers opportunities for students to collaboratively practice their presentational skills and receive teachers' feedback. It is also an archived platform, which shows student work (with privacy settings) and uses a variety of mashup materials, such as texts, images, audios, and videos.

According to ACTFL (2012), the three modes of communication provide the organizing principle for classroom L2 instructors to describe learners' language performance. The distinctive language functions, one of the parameters for the language learner's performance in each of the three communicative purposes, can guide the design

of classroom activities. This paper, therefore, analyzes each of the three technology tools with a focus on task examples, student comments, assessment rubrics, and their potential to develop language functions in the target mode of communication.

2. Literature Review

The three modes of communication are a cluster assessment in Integrated Performance Assessment (IPA). Research conducted on the three modes of communication sheds light on the influence of IPA on instruction and student performance. The initial project conducted by ACTFL that involved approximately 40 language teachers and 1,000 students of Chinese, French, German, Italian, Latin, and Spanish across grade levels 3-12, indicated the usefulness, feasibility, and challenges of IPA. Teachers indicated that IPA makes them more aware of the need to integrate the three modes of communication into their lesson planning on a regular basis. For example, the IPA prototype encourages them to use authentic documents in designing standards-based interpretive tasks, integrate more open-ended speaking tasks for interpersonal speaking tasks, and use more standards-based rubrics to assess student language performance (Adair-Hauck, Glisan, Koda, Swender, & Sandrock, 2006). On the other hand, teachers also reported the challenges of implementing IPA to enhance the three modes of communication. This type of assessment lacks age-appropriate authentic texts, and as a result, it is difficult to help students prepare for oral presentation tasks in which they are expected to react spontaneously rather than read from a prepared text (Adair-Hauck, Glisan, & Troyan, 2013).

More research was conducted subsequently on IPA at the post-secondary level to examine its feasibility and to analyze student performance across the three modes of communication. The study was conducted in an advanced Spanish course to examine how many students exceeded, met, or did not meet expectations (Glisan, Uribe, & Adair-Hauck, 2007). The researchers revealed that, in the interpersonal mode, the students did not perform as well as in the presentational mode, which is related to the previous finding regarding the challenge of teaching students to communicate spontaneously. Secondly, the interpretive mode was the only mode in which some students did not meet expectations. They reported that they were not exposed to the listening strategies enough. The students performed the best on the presentational tasks, which may have something to do with the reality that presentational tasks tend to be the predominant mode in most language instruction (Adair-Hauck et al., 2013).

The previous research has shown the benefits and challenges in integrating the IPA prototype in instruction. Although the three modes are widely recognized in language education, very little research has investigated how they can be enhanced through the use of technology for Chinese learners. In order to fill this gap in pedagogy, this paper aims to provide Chinese teachers an overview of how ACTFL's three modes of communication are aligned with three technology tools.

2.1 Second Language Pedagogical Technology and Student Learning: Benefits and Limitations

The use of technology in assessing projects and activities and its importance are not new in the study of language teaching and learning. Researchers such as Klopfer, Osterweil, Groff, and Haas (2009) claim that using technology can create a more deeply engaging learning experience. Hmard (2006) also discusses how computer-assisted interaction in a web-based environment helps develop assessment strategies in the learning process. In addition, many researchers favor technology use because it also promotes autonomous learning. These studies inform us that the use of technology plays a crucial role in assessment. However, there is little written about how to use computer-aided assessment tools for assessing Chinese proficiency in terms of the three modes of communication. Lai (2017) suggests that technology tools enable teachers to differentiate instruction and adapt classroom activities and thus enhance language learning experience. For example, the implementation of ThingLink, Padlet, and HomeStyler helps students create, analyze and synthesize materials by developing multimedia projects that draw on multiple literacies. With technological tools applied to facilitate curricular activities, students can actively learn the target language in multiple authentic contexts (Dema & Moeller, 2012). In addition, in investigating the benefits of the planned and purposeful use of technology for L2 learning, Shrum and Glisan (2010) find three advantages. Technology can: 1) enable learning to happen anywhere or anytime, which thus results in a better and more effective use of class time; 2) individualize learning at the learners' own pace, and; 3) empower learners through more accessible assessment tools. Yang (2001) suggests that online experiences invite learners to participate in the culture of the target language, which in turn enables them to compare the target culture with their own. Furthermore, a technology-supported environment enlarges the scope of the learners' language learning and opens a broader range of connections and meaning-making among learners (Lai, 2017).

While previous research has provided a foundation for understanding the significance of using technology in SLA, it is equally critical that language educators are aware of the limitations in the use of technology. The effectiveness of any technological tool often depends on the knowledge and expertise of a qualified language teacher who manages and facilitates the language learning environment (ACTFL, 2017). It can be challenging for many instructors to find the best way to teach with technology, to figure out what students like or dislike, and to determine what works or does not work (Wu, 2013). A teacher's attitude can also lead to students' frustration in integrating technology tools in their language learning. Bourgerie (2013) finds that students' negative attitudes toward technology are the result of unenthusiastic teacher support and failure to integrate the materials into a larger learning environment. On some occasions, the majority of classroom students simply surfed the Internet instead of participating in language learning (Dema & Moeller, 2012). Taking the three modes of communication into account, this paper presents some benefits and limitations of technology tools used in the Chinese language classrooms to measure students' progress and to maximize their learning experience.

2.2 Second Language Pedagogical Technology and Three Modes of Communication

In 1952, the Civil Service Commission developed a register of language skills for government employees. But the Commission had neither a system of proficiency tests nor outlined criteria for test construction. It also had no form of standardized speaking tests across academic institutions. As a result, ACTFL began developing speaking assessments in the late 1980s and 1990s with college-level students and even older learners in mind. It now also includes assessments of listening, reading, and writing (Cox, Malone, & Winke, 2018). The ACTFL Proficiency Guidelines were first published in 1986 for the academic community of the U.S. Government's Interagency Language Roundtable (ILR) Skills Level Descriptions. Updated and revised during the past several decades, the current ACTFL Proficiency Guidelines (2012) provide descriptions of what learners can do in terms of various functional aspects of language ability (listening, reading, speaking and writing) in real-world spontaneous and non-rehearsed contexts. This current iteration underlies the development of the ACTFL Performance Guidelines for K-12 learners to describe how well learners meet content standards.

While the Proficiency Guidelines and Performance Descriptors accompany each other, their approach to describing the function of language learning is somewhat different. The Proficiency Guidelines are organized in terms of individual language skills (listening, reading, speaking and writing) with limited elaboration on how each skill is used in the three modes. For example, under the speaking section, the Proficiency Guidelines elaborate the characteristics of speakers' functional language, with a brief explanation that these speaking guidelines can be used to assess speech that is either interpersonal (two-way communication) or presentational (one-way communication). In contrast, the Performance Descriptors outline how language learners demonstrate performance of the three modes of communication in explicit instructional settings. The Performance Descriptors help educators implement the standards and offer descriptions of how language learners can perform with their language skills in terms of the three modes of communication.

Features of the three modes described in the Performance Descriptors (ACTFL, 2012) and by Kissau and Adams (2016) to assess learner performance are identified as: 1) the interpretive mode is one-way communication without negotiation of meaning with the producers as evidenced in the reading (websites, articles, or stories), listening (messages, speeches, or songs), or viewing (videos, movies, or TV shows) of authentic materials; 2) the interpersonal mode is a two-way active negotiation of meaning, including adjustments and clarifications, as seen in speaking and listening (voice messages and conversation) as well as reading and writing (text messages, emails, or social media), and; 3) the presentational mode is one-way communication with limited opportunities for feedback (ACTFL & P21, 2011) as seen in writing (articles or reports), speaking (telling a story, giving a speech, or performing a rehearsed skit), or visually presenting something (videos or PowerPoint). Adair-Hauck et al. (2006) suggest classroom achievement tests and standardized instruments still rely on easily quantifiable, non-contextualized testing procedures that isolate only single skills. In contrast, integrating the three communicative modes requires multiple skills that are typical of real-world communication.

The ACTFL Statement on the Role of Technology in Language Learning (2017) strongly recommends that language educators leverage technology to support learning and it emphasizes that technology should be used to enhance human interactions. Rather than being an isolated tool, technology should be the key curricular component woven into language learning in multi-modalities such as text, audio, video, image, and a variety of mixes. These tools can be identified as either synchronous or asynchronous. Communication using synchronous tools tends to involve face-to-face communication, such as instant video or audio chats, and written communication, such as texting or group chats. In contrast, asynchronous communication involves time lapses of turn-taking, such as emails, online discussions, and blogging (Kessler, 2018).

In summary, there have been abundant studies that propose theoretical frameworks about what learners can do in terms of single skills in listening, speaking, reading, and writing. In contrast, less attention has been paid to integrating the three modes of communication, especially with regard to teaching Chinese as Foreign Language (CFL). To better understand the role instructional technology can play to enhance the three modes of communication while aligning with national standards, this study illustrates the use of technology tools in instructional settings, and provides feedback from learners.

3. Theory to Practice: Description, Analysis, and Discussions

3.1 Interpersonal Mode: Go Formative

Go Formative is a digital formative assessment tool that teachers can use to assess students by adding content and questions, checking answers instantly, and giving timely feedback. A traditional assessment is simply a check for understanding, such as a vocabulary quiz or unit test. Go Formative is a platform where the teachers can add both content and questions. Content includes image, text, video, and whiteboard. Adding content allows teachers to add learning materials from YouTube (such as a movie clip about traveling) or images (such as a picture of a train ticket). Students in turn gain more exposure to learning in authentic contexts. Teachers can also add a variety of question types, such as open-ended, multiple-choice, short answer, true or false, and audio responses. Although language teachers traditionally have given students vocabulary lists and pictures for memorization, Ousselin (2013) argues that a simple picture with a basic word or sentence does not engage students to their fullest capacity. For instance, the "Show Your Work" feature in Go Formative goes beyond a list or a picture and can annotate an image with video, audio, and textual tags to strengthen vocabulary learning. The "Audio Responses" and "Feedback" features help teachers interact with students' responses verbally and non-verbally. They give and receive feedback, creating interpersonal communication. Hattie and Timperley (2007) found that the most effective type of feedback is to offer assistance (e.g., cues or reinforcement), and utilize technology (video, audio, or computer) in feedback delivery. "View Responses" in Go Formative can be used to promote interpersonal communication by showing student responses in front of class. Based on responses, students can form discussion groups to complete tasks such as think-pair-share. Go Formative therefore serves as a platform for

students to easily interact and negotiate meaning in written form and spoken conversations. Furthermore, learners share information, reactions, feelings, and opinions in multiple directions, i.e., teacher-student and student-student. Thus, one aim of this study is to examine the impact of feedback with the implementation of Go Formative on enhancing students' interpersonal communication skills.

The two learning outcomes for the interpersonal mode are: 1) students observe and monitor one another's responses in authentic contexts through the annotation of key words and highly contextualized formulaic phrases, and; 2) students utilize supporting details and contextualized clues to provide feedback to each other by initiating, negotiating, and sustaining the conversation spontaneously. Examples of the tasks used in this study were video clips, images, and websites across a variety of topics.

To understand the strengths and limitations of Go Formative, twelve college students in a Chinese language course at the Novice level were recruited for the study. At the end of the semester, student feedback on their use of Go Formative as part of the teaching and learning materials was collected through a survey and interviews with the instructor. The Go Formative assignments comprised 15% of their final grade. The focus of this tool is on interpersonal communication (two-way, verbal, and non-verbal). Therefore, students were required to provide answers, investigate a question, and exchange communication through body language and text messages. Go Formative is used to break down information into short segments, which makes it easier for the students to memorize. Sections designed to assess students' performance in Go Formative include: the instructor's summary video, grammar points, handwriting practice, basic questions, dialogue practice, and interpersonal activities (think-pair-share and interview). After the students learn the basic vocabulary and grammar, they seem more comfortable in using them to create with language. Student performance was evaluated by the interpersonal rubric (Adair-Hauck et al., 2013).

Sample Question: Read the Chinese website (Figure 1) for booking flights below and answer the following question based on the information given:

Figure 1 A screenshot from an authentic Chinese website for booking flights

Question: 从上海到昆明差不多几个小时？

(Hint: Don't change the word order; just replace the question word with your answers.)

Interpersonal Task: You are studying in a language school in Shanghai. Summer vacation is coming up. You want to explore another city in China. Pick a city, book a round-trip ticket, and share your travel plans with your classmates.

Choose round-trip ticket "往返"
出发城市：上海
到达城市：？？(Pick a city of your choice!)
出发日期：2018/8/1

Go online: http://www.ctrip.com (Try to play around!)

Tell us: topics include but are not limited to...
1) A city you want to visit in China. 你要去哪个城市？
2) How many hours from Shanghai? 要坐几个小时的飞机？
3) How much is the airfare? 飞机票多少钱(RMB)？
4) Your own question

With a partner, interview each other about your travel plans!

  In the example task given above, the students were learning about traveling. First, the students went through a series of questions related to the topic and were given authentic materials (a site in Chinese to book flights) to complete the questions before they were able to complete the interpersonal task. The interpersonal task is given to assess students' speaking ability. Students were required to initiate a conversation with classmates and to talk about their travel plans. The student performances were evaluated using the ACTFL interpersonal rubric (Adair-Hauck et al., 2013), and the aspects that were included in the evaluation were as follows: language function, text type, communication strategies, comprehensibility, and language control. Student performance was qualitatively ranked on a continuum: Exceeds Expectations; Meets Expectations-Strong; Meets Expectations-Minimal; Does Not Meet Expectations. A sample answer to prompt #1 that would be given a rating of "Meets Expectations-Strong" is: "我去香港，日本，也去台灣" because the student is able to understand and produce highly practiced words and phrases and make a list. An example of a "Meets Expectations-Minimal" answer is "票，多少五十錢," which shows that the learner is unable to fluently use highly predictable and formulaic phrase, but is able to minimally complete the exchange based on prompt #3 in the above example.

  The students' feedback on the use of Go Formative is positive. Findings reveal that 92% of the students indicated that it is a very helpful tool. More than 90% of the

students found that most of the sections are very helpful, except for the handwriting practice (53%). In follow-up interviews, 50% of the students indicated that Go Formative was the most effective aspect of this language course. They stated that Go Formative helped them build up the vocabulary needed to ask, understand, and answer questions. In addition, because Go Formative can integrate listening, reading, typing, and handwriting practice in a single platform, it helps them to create with the language. Go Formative enables learners to learn oral, written, and non-verbal communication skills and encourages them to interact with each other. Learners do not just learn basic vocabulary and grammar; they also strive to create with language and move beyond memorized phrases in an interactive manner.

3.2 Interpretive Mode: Edpuzzle

Edpuzzle is a free online free tool that is designed for teachers to easily choose and edit video clips from the Internet with only a few limitations. Edpuzzle engages students with authentic videos and enables instructors to assess students' receptive language skills and to collect learning data. With Edpuzzle's user-friendly editing functions, instructors can shorten or crop videos to be an appropriate length or size that focus student attention on the key points of the videos and meet the needs of the lesson. The quiz feature of Edpuzzle allows instructors to individualize popup questions as a means to assist different students and subsequently review how they fared on the questions and the amount of time they spent on completing each of the embedded questions in a video. In addition, the audio voiceover feature makes it possible for instructors to record their own voiceover in an online video. Its benefits include allowing instructors to have a voice in the material while students independently watch the video and align content with the learning objectives of a lesson. Moreover, Edpuzzle videos can be easily embedded in learning management systems, such as Blackboard and Moodle, and they allow instructors to place restrictions on videos so learners have to watch the video and answer all the associated questions without being able to fast forward.

The primary objective of the implementation of Edpuzzle in this study is focused on enhancing students' interpretative listening skills and developing specific L2 learners' language functions in comprehending: 1) meaning through the recognition of key words and highly contextualized formulaic phrases at the Novice level; 2) the main ideas and some supporting details at the Intermediate level, and; 3) the main idea along with supporting details of narrative, descriptive, and straightforward persuasive texts as well as inferences derived from context and linguistic features (ACTFL, 2012). Examples of the video clips used in the present study were TV commercials, news, talk shows, movies, and documentaries across a variety of topics with four to eight multiple-choice and/or open-ended questions each added to the use of authentic materials in learning Chinese.

The implementation of Edpuzzle in a CFL classroom can take different formats depending on the conditions and resources of each teaching and learning context. In this study, the fourteen student users of Edpuzzle were enrolled in an advanced content-based language course in an intensive Chinese program at the college level. The fourteen students were either in their junior or senior year of college, and their oral proficiency was Intermediate-High or Advanced-Low on the ACTFL scale at the time data were

collected. The students were required to watch the assigned Edpuzzle clips and answer the embedded popup questions on their own, then discuss any linguistic and extended content questions from the videos in class. The popup questions on Edpuzzle were designed by the instructor specifically to evaluate two sets of the learners' comprehension skills: literal and interpretive. The former assessed the students' ability to literally understand: 1) keywords; 2) main ideas, and; 3) supporting details of the given video clips. The latter explored the learners' interpretative comprehension of: 4) organization; 5) contextualized meaning; 6) inferences; 7) narrator's perspective, and; 8) cultural perspectives. These were the eight criteria that the instructor intended to evaluate.

Ms. Yuan Yuan Liu's 2014 speech[1] was one of the Edpuzzle clips that was used in the present study. Part of her speech in the video and the subsequent embedded question are as follows.

片段：
現實生活是一種很神奇的生活，在現實生活中那些尊重規則的老實人往往一輩子都默默無聞，反倒是那些弄虛作假的人到最後會名利雙收，於是乎像我這樣的年輕人就經常有那些看著很有經驗的前輩過來拍拍你的肩膀跟你說"年輕人你還不懂。

題目(1)：為什麼 "現實生活是一種很神奇的生活"?
a.　尊重規則的老實人卻不會名利雙收。
b.　弄虛作假的人卻會默默無聞。
c.　有經驗的前輩常拍著年輕人的肩膀

題目(2)：演講者提及"年輕人你還不懂". 從演講者的觀點看來，年輕人不懂的是什麼？

In the preceding example, the popup questions appeared in order on the screen at the end of this paragraph of Ms. Liu's speech. Students were required to answer both the multiple-choice and open-ended questions before they could continue watching the video. Question 1 was intended to evaluate students' literal comprehension skill with regard to the main ideas and supporting details in the paragraph. Question 2 set out to assess students' interpretive comprehension skills pertaining to making inferences about the speaker's perspective.

On the teacher account of Edpuzzle, the instructor was able to access an overview of a quantitative measure of the students' comprehension based on the answers to the multiple-choice questions. In addition, the students' comprehension was qualitatively evaluated on a continuum using the IPA interpretive mode rubric (Adair-Hauck et al., 2013, p. 125). The IPA rubric articulated the instructor's expectations for her students' interpretative comprehension by listing the eight criteria and describing levels of quality from outstanding to poor. This helped the instructor facilitate the implementation of

---

[1] c..f. https://www.youtube.com/watch?v=-X-rCTkmMbI

Edpuzzle tasks—from designing the popup questions to giving specific feedback on the students' skills in the interpretive mode of communication.

More specifically, an example answer to the preceding multiple-choice question was "C" (有經驗的前輩常拍著年輕人的肩膀). In addition to being informed that the answer was wrong, the student was also given a qualitative rating of "Did Not Meet Expectations." The student's answer seemed to suggest that he was unable to identify the key words or main idea appropriately within the context of the text and that he possessed minimal interpretive literal comprehension in the categories of word recognition and main idea detection. An example answer to the preceding open-ended question "演講者提及 '年輕人你還不懂'. 從演講者的觀點看來，年輕人不懂的是什麼？" was "年輕人不懂現實生活和我們想的不太一樣。我們覺得好人會有好的結果，可是常常壞人卻過得更好." This answer on Edpuzzle was rated as "Exceeds Expectations" for both categories of guessing meaning from context and inferences between the lines. The rating of "Exceeds Expectations" was given because the student successfully interpreted contextualized meaning and accurately inferred the meaning of unfamiliar words and phrases in a highly plausible manner.

In order to investigate how students and instructors perceived the usefulness of Edpuzzle in terms that improve learner interpretative language skills, the researcher interviewed the participants individually. The interviews began with the participants' narrative regarding their overall feedback on their use of Edpuzzle and continued with follow-up questions from the researcher. The use of Edpuzzle at the advanced level was generally well received. The three comments with regard to language functions in the interpretative mode of communication that were most commonly shared among the fourteen students are as follows. First, the popup questions directed the students to reevaluate what they had thought they knew about the content of the clips and to learn to make inferences beyond what was directly presented to them. Second, Edpuzzle practice focused student attention on the details of the supporting arguments in addition to the main ideas and directed them to identify and justify the standpoints in the videos. Third, cultural authenticity in the assigned Edpuzzle videos imparted knowledge of cultural norms to the students in context and connected cultural products, practices, and perspectives. Other student reflections include remarks that the instructor's audio voiceover and the back-tracking feature gave them a sense of being supported as they independently rose to the challenge of tackling authentic materials. Moreover, inserted questions along the way engaged the student participants and guided them to discover new concepts that they might have otherwise overlooked. Finally, the use of Edpuzzle changed the landscape for the student participants and added novelty in the selection of learning materials.

Edpuzzle is not without limitations as a classroom pedagogical technology tool and, like any other tool, it is only as effective as the teacher makes it. One of the students commented that Edpuzzle video clips can be most engaging when their length ranges from one to two minutes. It thus requires teachers to selectively edit down most available online videos to be concise enough to sustain student interest but also long enough to cover the learning objectives. In addition, the pool of videos for Edpuzzle is wide,

including most major public websites, such as Youtube, TedTalk, National Geographic, and the Khan Academy. The wealth of available online videos can pose challenges for instructors because they will need to search patiently for suitable clips that have appropriate content, allow purposeful popup questions intermittently, and provide linguistic features that push them upward on their proficiency scale.

Despite certain drawbacks, Edpuzzle engages learners to hone their language functions in the interpretative mode of communication by using authentic materials. It also offers an opportunity for instructors to collaboratively edit videos and share teaching resources. In addition, because of the authenticity in the materials, students' comprehension of Edpuzzle videos can inform instructors about students' proficiency independent of any specific teaching materials, and it is indicative of the gap between what is covered in class and what students need to know to be world-ready.

3.3 Presentational Mode: Google Sites

According to Performance Descriptors (ACTFL, 2012), an essential function of presentational communication is the ability to present information by creating with language. Learners present information by writing (journals, articles, or reflections), speaking (telling a story or giving a speech), or visually representing (student-made videos or PowerPoint). Can-Do Statements (NCSSFL & ACTFL, 2017) reflect the continuum of growth in communication skills which learners: 1) identify and set learning goals to chart their progress towards learning proficiency; 2) reflect an interactive process that captures interaction between learners and teachers/facilitators, and; 3) promote self-feedback and assessment. Technology paves the way to implement integration with the pedagogical curriculum, and Google Sites is a powerful tool to support the presentational mode in L2 learning. It is useful for educators who wish to show students' learning experience, and also for learners to demonstrate their creation of messages to inform, narrate, and reflect on their learning that can be shared with various audiences of listeners, readers, or viewers.

Following the aforementioned guidelines and statements, Google Sites is a structured web-page creation tool where multiple people can collaborate and share files. Google Sites is therefore an effective tool to be used as an ePortfolio that shows learners' learning experience in context. An ePortfolio allows students to display evidence that demonstrates their communicative skills and allows teachers to construct measures of assessment and feedback. Students choose what to integrate from assignments and products into their ePortfolios, acting as autonomous thinkers. Google Sites has the following features: 1) as an archive for students' work in the form of their learning ePortfolio; 2) as anecdotal evidence of the presentational mode to show learners' creation in writing, speaking, or videos, either in public or private group settings for collaboration, and; 3) as easily embedded external content, including texts, images, audios, and videos (audios and videos use third-party online tools due to storage limitations). Additionally, Google Sites can be organized based on themes, classes (groups of students), or language tasks.

Three classes of a total of 38 high school students were informed that Google Sites would be used in language learning. Their ages ranged from fifteen to seventeen years old. Everyone took the ACTFL Assessment of Performance toward Proficiency in Languages (AAPPL) test one month before the data were collected, and their scores in the presentational skills (speaking and writing) ranged from Novice-High to Intermediate-High. Students knew that they would use Google Sites in the following three ways:

- To write a paragraph summary and/or reflection: students could choose to listen to a podcast or watch a video clip or TV show in Chinese, then write a summary paragraph of their reflection.
- To write an essay based on oral or written prompts: for example, students could write about their favorite travel destination and include travel photos.
- To perform a rehearsed skit: students could plan and work in groups to make videos which they would upload to YouTube and then embed on Google Sites.

Before they began, students were given presentational rubrics to understand expectations. Grading criteria include contexts/content, text type, language control (i.e. vocabulary use and grammatical structures), and cultural awareness. Expectations also included that students would be able to communicate information, make presentations, express thoughts about familiar topics, and use sentences as well as a series of connected sentences at a functional level through spoken and written language while making comparisons with the target culture (Can-Do Statements, 2017). Depending on the assignment, students sometimes had the opportunity to engage in peer review by posting comments in response to their peers' work. The teacher also gave feedback based on the rubric provided.

A questionnaire survey was used to collect feedback at the end of the year-long course after the students had used Google Sites for one academic year. The survey consisted of questions in relation to the participants' experiences with and their attitudes toward their general use of Google Sites and their perceptions on how effective the tool was in their learning. The feedback was positive. For example:

Student 1: "I like Google Sites because I can see my own personal development."
Student 2: "Google Sites is easy to use, I can work independently or with friends."
Student 3: "I can see my previous work from previous years, including the feedback all on one page."

Although Google Sites is a user-friendly tool, a considerable time investment is required to figure out how to set it up for student access. Further drawbacks for teachers who have not used it before include the need for initial guidance to create the sites and pages for students to access, and the fact that some students are uncomfortable allowing peers to view and comment on their work. However, for teachers who are willing to learn

the system and be sensitive to student reservations about sharing, Google Sites is an engaging and useful tool that enhances classroom learning.

4. Conclusion and Pedagogical Implications

According to Lai (2017), the use of technology can broaden the horizon of L2 learning. To further understand how technology tools can specifically add to teaching and learning in a classroom, this study reviews the use of Go Formative, Edpuzzle, and Google Sites in relation to the three modes of communication. The objective is to demonstrate the facilitative role of technology in classrooms and to offer ideas that encourage teachers to incorporate free online tools to engage their students in learning Chinese. This article includes classroom task examples and rubrics to evaluate student performance in an attempt to show teaching professionals how technology and its associated language proficiency assessments can be incorporated into day-to-day teaching practices. The preparation for the effective use of technology requires teachers to identify technology tools that meet their teaching objectives and to explore student interests in terms that motivate them to use technology to assist their learning. In addition, teachers will need to familiarize themselves with the technology tools of their choice, make plans to appropriately embrace technology to maximize their pedagogical efficacy in the classroom, collect feedback from their students on their use of pedagogical technology, and continue to improve the ways that technology is incorporated into the traditional classroom to enhance the student experience in learning foreign languages.

While the student comments show that technology can engage them in learning Chinese, it is not yet clear to what extent the use of these tools affects students' acquisition of Chinese in the three modes of communication. Student feedback collected in this study shows that Go Formative, Edpuzzle, and Google sites are useful tools in motivating students to develop their language skills in the three modes of communication. The tools provide a platform for a variety of learning materials to be stored and processed. The three applications also engaged the participating learners in practicing their language skills through various activities either independently or with different groups of audiences, such as viewers, listeners, and readers. While the benefits of using these three tools were acknowledged in the students' feedback, the current study was not designed to empirically measure the effect of these three tools on the learners' acquisition of Chinese or their language improvement in the three modes of communication. Because these technologies are user friendly for both instructors and students, are well received by students, and provide instructors with robust evaluative opportunities, their effectiveness in fostering SLA is a topic that merits further research.

References

Adair-Hauck, B., Glisan, E. W., Koda, K., Swender, E. B., & Sandrock, P. (2006). The integrated performance assessment (IPA): Connecting assessment to instruction and learning. *Foreign Language Annals*, *39*(3), 359-382.

Adair-Hauck, B., Glisan, E. W., & Troyan, F. J. (2013). *Implementing integrated performance assessment.* Alexandria, VA: American Council on the Teaching of Foreign Languages.

American Council on the Teaching of Foreign Languages. (2012). *Performance descriptors for language learners.* Retrieved from https://www.actfl.org/sites/default/files/pdfs/PerformanceDescriptorsLanguageLearners.pdf.

American Council on the Teaching of Foreign Languages. (2012). *Proficiencyguidelines for language learn*ers. Retrieved from https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012.

American Council on the Teaching of Foreign Languages. (2017). *Statement on the role of technology in language learning*. Retrieved from https://www.actfl.org/news/position-statements/statement-the-role-technology-language-learning.

American Council on the Teaching of Foreign Languages, & National Council of State supervisors for Languages. (2017). *Can-do statements*. Retrieved from https://www.actfl.org/publications/guidelines-and-manuals/ncssfl-actfl-can-do-statements.

American Council on the Teaching of Foreign Languages. (2011). *21st Century skills Map*. Retrieved from https://www.actfl.org/sites/default/files/pdfs/21stCenturySkillsMap/p21_world languagesmap.pdf.

Bourgerie, D. S. (2013). Computer aided language learning for Chinese: A survey and annotated bibliography. *Journal of the Chinese Language Teacher Association*, *38*(2), 17-47.

Cox, T., Malone, M., & Winke P. (2018). Future directions in assessment: Influences of standards and implications for language learning. *Foreign Language Annals*, *51*(1),104-115.

Dema, O., & Moeller, A. K. (2012). *Teaching culture in the 21st century language classroom*. In T. Sildus (Ed.), *Touch the world: Selected papers from the 2012 central states conference on the teaching of foreign languages* (pp. 75-91). Eau Claire, WI: Crown Prints.

Glisan, E. W., Adair-Hauck, B., Koda, K., Sandrock, S. P., & Swender, E. (2003). *ACTFL integrated performance assessment*. Yonkers, NY: ACTFL.

Glisan, E., Uribe, D., & Adair-Hauck, B. (2007). Research on integrated performance assessment at the post-secondary level: Student performance across the modes of communication. *Canadian Modern Language Review*, *64*(1*)*, 39-67.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.

Hmard, D. (2006). Design issues related to the evaluation of learner-computer interaction in a web-based environment: Activities vs. tasks. *Computer Assisted Language Learning, 19*(2–3), 261–276.

Kessler G. (2018), Technology and the future of language learning. *Foreign Language Annals, 51* (1), 219-224.

Kissau S. & Adams M. (2016). Instructional decision making and IPAs: Assessing the modes of communication. *Foreign Language Annals*, *49* (1), 105-123.

Klopfer, E, Osterweil, S., Groff, J., & Hass, J. (2009) *Using the technology of today in the classroom today*. Retrieved from https://education.mit.edu/wp-content/uploads/2015/01/GamesSimsSocNets_EdArcade.pdf.

Lai, A. (2017). Implementing online platforms to promote collaborative learning in Chinese language classrooms. *Journal of Technology and Chinese Language Teaching, 8*(2), 39-52.

Ousselin, C. (2013). *ThingLink for world language educators and learners.* Retrieved from http://aatftech.blogspot.com/2013/06/thinglink-for-world-language-educators.html.

Shrum, J. L., & Glisan, E. W. (2010). *Teacher's handbook: Contextualized language instruction*. Boston, Mass: Heinle & Leinle.

Wu, J. (2013). Students in the new millennium: How much do we know about them? In B. Zou, M. Xing, Y. Wang, M Sun, & C. H. Xiang (Eds.), *Computer-assisted foreign language teaching and learning: Technological advances* (pp. 118-137). Hershey, PA: Information Science Reference.

Yang, S. C. (2001). Integrating computer-mediated tools into the language curriculum. *Journal of Computer Assisted Learning*,*17*(4), 85-93.

机器翻译对中文传统翻译作业的挑战
(The Challenge of Machine Translation to Traditional Translation Homework in Chinese Language Learning)

田野
(Tian, Ye)
宾夕法尼亚大学
(University of Pennsylvania)
tianye1@sas.upenn.edu

摘要：本研究在一所常春藤盟校的一门中级汉语课中进行，分析了 13 名学生 10 篇作业中约 500 道翻译练习题。通过内容分析、访谈以及参与观察，本研究发现机器翻译在很大程度上挑战了传统的中文作业（尤其是翻译作业）的意义。本研究发现，学生会在不同程度上依赖机器翻译来完成翻译作业，而且教师很难发现这种行为。在完成大量作业时，要取得好成绩的压力是迫使使学生使用机器翻译的主要原因之一。机器翻译，尤其是搜狗翻译的准确性极高，这一结果破除了机器翻译不够准确的迷思。学生完全可以将机器翻译作为获得高分的捷径。

Abstract: This study was conducted at an Ivy League University and analyzes approximately 500 translation exercises that were collected from 10 homework assignments of 13 students in an intermediate Chinese language course. Through content analysis, interviews, and participant observation, this research finds that machine translation poses challenges to the traditional format of Chinese language assignments, especially translation homework. This research shows that students rely on machine translation to varying degrees to do these translation exercises. The stress of receiving a good score after finishing a huge amount of homework is one of the main reasons that motivates students to use machine translation. The accuracy of machine translation, especially Sogou Translate (搜狗翻译), is extremely high. This finding is likely to shatter the myth that machine translation is not accurate enough hence students cannot use it as a short cut to receive a good score.

关键词：中文教学，机器翻译，翻译作业，教学法

Keywords: Chinese language teaching, Machine Translation, Translation homework, Pedagogy

1. 引言

　　尽管近年来人工智能和语言处理技术发展迅速，但许多中文教师仍然沿用传统的作业形式，如抄写汉字、翻译练习等等。从传统教学法的意义上来讲，这些作业，特别是翻译练习，是教师检测学生学习进度和分析教学效果的有效工具（李泉，2006；赵金铭，2010）。学习中文的学生常常要花大量的时间来完成作业里的这些翻译练习，教师们也常常必须投入同样的甚至更大的精力来对学生的答案进行批改和打分。然而，新的科学技术，例如机器翻译，对传统翻译作业的意义、价值和可靠性都提出了挑战。一个很明显的问题就是，中文教师怎么知道自己的学生是否在课下借助了机器翻译工具，例如谷歌翻译[1]，来完成他们布置的作业。从严格意义上来讲，机器翻译并不是一个新鲜事物。早在上世纪七、八十年代机器翻译的概念和应用就已经出现了，但是受限于昂贵的设备以及令人担忧的准确程度，机器翻译并未对外语教学产生特别大的影响（Garcia & Pena, 2011; Somers, 2003）。

　　现在外语教学界对以谷歌翻译为代表的机器翻译的研究主要集中在以下三个不同的方面。第一个方面是如何培训未来的翻译人员熟练使用专业翻译软件，比方说 Somers （2003）的研究就探讨了如何在初、中、高级的翻译课程中传授使用机器翻译。第二个方面是谷歌翻译的准确程度。比方说 Ducar 和 Schocket （2018）和 Correa（2014）介绍了谷歌翻译在将英语翻译成西班牙语时的一些优点及缺点。Case （2015）和 Groves & Mundt （2015）的研究都指出谷歌翻译的准确率已经接近了中级水平的英语学习者，而且谷歌的准确率会随着时间的推移而越来越高。第三个方面则是外语学习者使用机器翻译的情况，以及外语教师和外语学习者对使用机器翻译的不同态度。在这方面，很多研究都证明即使在教师明令禁止的情况下还是有很多学生会使用机器翻译（Clifford, Merschel, & Munné, 2013; Correa, 2011, 2014; García & Pena, 2011; Niño, 2009）。比方说，Clifford 等（2013）用问卷的形式调查了杜克大学罗曼语学系的老师和学生。他们发现老师和学生对机器翻译的态度是不同的。尽管学生能意识到机器翻译是不完美的，但是超过 88% 的被调查学生都承认他们在学习中使用过机器翻译，而且大多数学生都认为机器翻译有助于他们学习新的语言。可是，他们调查的老师则大多对机器翻译的有效性持怀疑的态度，认为机器翻译的出现干扰了传统的课堂教学，因此会在课程大纲中明令禁止学生使用机器翻译来回答要给分的作业。值得注意的是并不是所有的外语教师都反对使用机器翻译，比方说，Niño （2009）和 Baker （2013）的研究就都揭示了还是有很多外语教师和学生，尤其是有志于电脑辅助教学的教师和高年级的外语学习者，对机器翻译是持较为正面的态度的。他们认为机器翻译的缺点甚至可以被用来提高学生对翻译和语言学习复杂性的认识。

　　与其他语种相比，北美中文教学领域对机器翻译的关注度还不是很高。但是随着智能手机与免费机器翻译工具（例如谷歌翻译、百度翻译 [2]、搜狗翻译 [3]）的普

---

[1] https://translate.google.com/

[2] http://fanyi.baidu.com/

[3] http://fanyi.sogou.com/

及，笔者认为中文教学的研究者和从事一线教学的教师都将面对并有必要正视这个问题。结合外语教学界其他语种对机器翻译研究的情况，本研究旨在回答以下三个问题。1）学中文的美国学生在多大程度上使用机器翻译（如谷歌翻译）来完成作业中的翻译任务？2）这些学生对使用机器翻译做作业有什么看法？3）机器翻译在中英文互译方面的准确度如何？

　　本文以下将分为三个部分来回答这三个问题。第二节会介绍本研究的具体研究方法，包括本研究所涉及到的中文课的课程安排、学生情况、数据资料的收集和分析方法。第三节将提供本研究关于中文课翻译作业和机器翻译的三大发现。第四节将是本文的讨论总结部分。

2. 研究方法与研究背景

　　本文的研究方法是三种在教育学领域最为常见的研究方法：内容分析（content analysis）、深度访谈（interview）以及参与观察（Alridge, 2006; Becker, 1996; Ochs, 1979）。本研究在笔者自己所教的一门中级（三年级下学期）中文课[4]中进行。这门中文课共有 13 名学生。他们各自主修的专业不同，年级不同，背景也不尽相同，既有华裔学生，也有非华裔学生，还有来美国留学的国际学生。有的学生在高中已经学过了两三年中文，有的学生是在大学才开始学习中文。大部分学生都有在中国留学和旅游的经历。这些学生的中文水平比较接近，在参与本研究时基本均属于ACTFL 中低至中高水平（intermediate-low to intermediate-high）。

　　笔者所在大学为常青藤盟校之一，其中文项目的教学传统悠久，教学质量较高，学生素质极高。本研究所关注的中级中文课每周共四课时，所用教材为《新的中国》（Chou, Chiang, & Eagar, 2011）从第 21 课开始的后半部分。学完每篇课文之后，学生要自己打印老师提前放在课程网站上的作业单（参见附录一），完成作业以后再交给老师。每一课的作业单约有 3 到 5 页，均由该课程以前的任课老师们精心设计。在本研究进行之时，这些作业单已经由不同的任课老师沿用了至少 5 个学年。这些作业的形式比较传统，前半部多是多项选择、填空、完成句子、回答问题等，侧重于检测该课所学到的重点词语、语法和句型。后半部一般则是三个英译中的翻译题以及一个中译英的翻译题。英译中的翻译题每道都是以两三个句子组成的语段为主，每道题后面都会列出学生应该使用的句型和词汇，学生被要求在作业单上用铅笔写汉字作答。而中译英的翻译题则多为以两三个语段组成的长达三四百字的语篇。因为中译英当中有较多学生没有学习过的生词，所以每个词汇上边都会有相应的拼音标识。学生被要求在电脑上用英语作答，再将答案打印出来，附在作业单上，然后交给老师[5]。从传统意义上来看，这些作业题，尤其是翻译题的设计质量都相当高。它们话题紧紧围绕课文，检测了学生新学习的词汇、语法结构。鉴于我们使用的教材比较陈旧，同时为了遵循克拉申 "i + 1" 的输入假说理论（Krasehn, 2003），任课教师对这些翻译题的内容也进行了特意设计，既给学生补充了有关中

---

[4] 如果从零起点算起，参与本研究的学生要在大学连续学习 5 个学期的中文课才能上这门中文课。
[5] 附录一提供了第 35 课《老年人的生活》的作业单，以便读者有一个直观的感受。

国的新信息与新知识，又能对学生带来适度的挑战。

　　笔者在批改学生作业时发现，这些作业中前半部的题型比较容易批改，但是后半部的大段翻译题往往会占用笔者大量的时间，而且学生的答案时常会出现一些"匪夷所思"、"驴唇不对马嘴"的翻译。这种翻译很类似于一些很蹩脚的机器翻译，而且笔者曾经也无意中在图书馆撞见过自己的学生利用手机上翻译软件的照相功能来做作业。这些现象都促使笔者思考这样一个问题，到底在多大程度上学生其实是借助了机器翻译来做完成这些作业？为了回答这个问题，笔者研究分析了这两个班 10 次作业中的 40 道翻译练习题。具体研究步骤如下：1）笔者对学生提交的作业先进行照相留底，然后照常批改；2）将这 40 道翻译练习题分别放入谷歌翻译、搜狗翻译和百度翻译中生成机器翻译的答案；3）将这些答案同学生提交的答案进行分析对比研究；4）发现问题之后，笔者会同个别学生进行谈话，以期了解学生的心态。

3. 研究发现

　　本研究主要有三个发现。1）学生在不同程度上是会依赖机器翻译来完成这些翻译作业的。2）包括中文课在内，学生的课业压力其实极大。尽快完成大量作业并获得令人满意的成绩是迫使学生使用机器翻译的主要原因之一。3）机器翻译的准确度相当高，尤其是英译中的准确度在本研究进行之时已经超过了参与本研究的学生的翻译水平。而且机器翻译具有极强的自我学习能力，其准确度是在日益提高的。中国大陆公司开发出的搜狗翻译、百度翻译在中英文互译方面的准确率更高，比谷歌翻译的表现要更好。学习中文的学生其实可以通过课下偷偷使用机器翻译来获得更高的成绩。

3.1 学生在不同程度上使用了机器翻译

　　在完成课后作业（尤其是大段翻译题）的过程中，学生使用参考资料（例如老师的教案、教材、字典等等）是不可避免的事情。通过对比学生的作业和机器翻译的结果，以及通过对学生的访谈，笔者发现自己的很多学生其实也不列外。而且随着科技的发展，几乎没有学生会使用传统纸质的字典。最近几年，以谷歌翻译为代表的机器翻译，因为其更为强大的整句翻译功能更是取代了传统的在线字典。这种新的趋势就给中文教师带来了一个直接的挑战，那就是我们是否应该允许学生使用谷歌翻译？使用谷歌翻译在哪种程度上可以被认定为是作弊？这个问题其实在整个外语教学界也还没有达成共识（Correa, 2011, 2014; Ducar & Schocket, 2018; Garcia & Pena, 2011; Somers, 2003）。

　　本研究发现学生对机器翻译的使用，按照程度来分，大致可以分为以下三种情况。第一种比较传统，学生把谷歌翻译当做字典，用其来查找某些生词的意思。常识和经验告诉我们说，在做作业的时候，我们应该允许甚至鼓励学生的这种行为，因为这种情况非常符合学生在非课堂环境里所遇到的情况。第二种情况是有的学生

会把翻译段落中某个比较难的句子放在谷歌翻译里面，然后对谷歌翻译的答案进行修改、翻抄（因为要使用老师要求的语法和词汇）。这种行为和谷歌翻译的功能有关，同传统的在线字典不同，谷歌翻译的一大优势就是它可以提供整句、整段，乃至整篇文章的翻译。而教学经验告诉我们，这个行为似乎就已经踩到了"作弊"的边缘地带了。因为这毕竟是在完成课后作业，老师想看到的是学生自己的学习水平，而不是谷歌翻译的水平。第三种情况则最为极端，本研究发现有的学生根本不经过自己的思考，不管谷歌翻译的质量怎么样，直接大段照抄谷歌翻译的结果，甚至将一些明显错误的内容都抄写上来。笔者认为这种情况等同于作弊。

先举一个例子来解释一下第三种情况，因为这种直接抄袭的情况最为直观，很容易理解。这个例子来自笔者的学生学完《新的中国》第 33 课《从"发福"到"减肥"》后所完成的翻译作业。例 1.1 是该课作业里面中译英的一部分中文题干。例 1.2 是笔者进行本研究的前期探索时（2017 年 2 月）谷歌翻译对这段话的翻译。

例 1.1

D. Translation: Chinese to English (Please type and attach a separate piece of paper)
随着社会的发展和生活水平的提高，对于美的标准也在逐渐变化。**在现代社会里人们大多以瘦为美，越来越多人觉得自己太胖。**打开报纸和杂志，到处都可以看到减肥的广告，电视上也有越来越多教人怎么减肥的报道。

减肥的方法很多，最常见的就是加强锻炼和节制饮食。许多人认为跑步是消耗热量最有效的方法，也有人选择用快走或者骑自行车的方式来减肥，除了多运动以外，少吃油腻的食品，少喝含糖量高的饮料也是减肥的重要事项。近来有研究指出，睡眠不足也可能造成体重增加，因此想减肥的人也应该尽量避免熬夜。总的来说只要能养成良好的生活与饮食习惯，体重自然就会下降……

例 1.2

谷歌翻译（2017 年 2 月）
With the development of society and the improvement of living standards, the standard for beauty is also gradually changing. **In modern society, <u>people are mostly thin for the United States</u>, more and more people feel too fat.** Open newspapers and magazines, everywhere you can see the <u>ads lose weight</u>, there are more and more people on television to teach people how to lose weight<u>, and reported</u>.
There are many ways to lose weight, the most common is to strengthen the exercise and diet, many people think that running is the <u>consumption of calories, the most effective way</u>, some people choose to brisk walking or cycling way to lose weight, in addition to more exercise, <u>eat greasy Food</u>, drink less high-sugar drinks is also important to lose weight. Recent studies have pointed out that lack of sleep, may also cause <u>weight gain</u>, so people who want to lose weight should try to avoid staying up

---

从例 1.2 可以看出，其实谷歌翻译的质量并不尽如人意。例 1.2 中下划线部分是谷歌翻译里一些翻译得不是很准确的地方[6]。其中谷歌对"人们以瘦为美"这句话翻译得最为离谱——"people are mostly thin for the United States"。在实际学生提交的作业中有母语为英文的美国学生（下称学生 A）原封不动地照抄了这句错误的英文。例 1.3 是学生 A 交上来的这部分作业的照片，从中我们可以发现该学生第一段的翻译其实是一字不差地粘贴复制谷歌翻译的结果，完全可以被判定为抄袭行为。

但是学生 A 对第二段翻译的处理就属于上文提到的第二种情况。该学生虽然在基本上也是照抄谷歌翻译，但是还是修改了个别词汇、语法的表达。比方说，谷歌对"许多人认为跑步是消耗热量最有效的方法"的翻译是"many people think that running is the consumption of calories, the most effective way"，而学生 A 则将这个明显错误的翻译改写为"many people believe that running is the most effective way to burn consumption"。如上文所述，老师很难判定这种行为是否为抄袭行为。

例 1.3
学生 A 第 33 课作业里面中译英的答案(第一段完全照抄谷歌翻译；第二段还是以谷歌翻译为主，但是进行了修改。)

With the development of society and the improvement of living standards, the standard for beauty is also gradually changing. In modern society, people are mostly as thin as the United States, more and more people feel too fat. Open newspapers and magazines, and everywhere you can see ads to lose weight, there are more and more people on TV reporting on how to lose weight.

There are many ways to lose weight, the most common is to intensify the exercise and diet. Many people believe that running is the most effective way to burn calories, some people choose to brisk walk or cycle as a way to lose weight. In addition to more exercise, eating less greasy food, drinking less high-sugar drinks is also important to lose weight. Recent studies have pointed out that lack of sleep may also cause weight gain, so people who want to lose weight should try to avoid staying up late. In general, as long as you can develop good living and eating habits, weight will naturally decline.

---

[6] 下文均用下划线标出各种机器翻译有问题的地方。

　　这里笔者可以再举一个例子。例 2.1 是《新的中国》第 34 课《从 "温饱" 到 "小康"》英译中的翻译作业。例 2.2 是笔者进行本研究的前期探索时（2017 年 3 月）谷歌翻译对这段话的翻译。

例 2.1

C. Translation: English to Chinese

1. In the past ten years, it can be said that Chinese people's lives have gradually changed from a state of "comparatively well-off" to a state of "richness." You can tell from the change in people's leisure activities that people's expectations (期望, qīwàng) toward life are getting higher and higher. For example, in the past people seldom travel far, but now traveling abroad is no longer a rare thing. （可以说是；从 A 变到 B；看得出来；远处；出国）

例 2.2
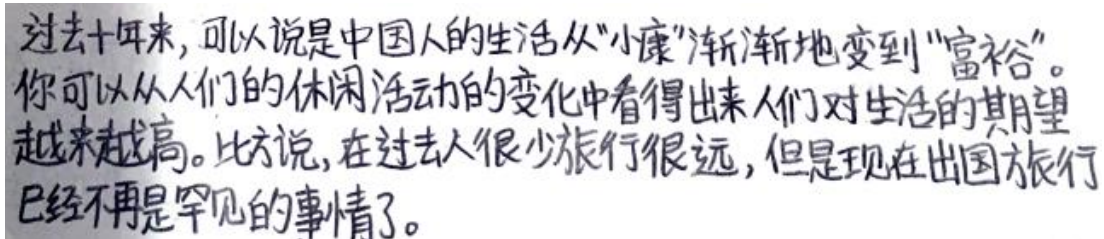
谷歌翻译（2017 年 3 月）

在过去十年中，可以说中国人的生活已经从 "<u>相对富裕</u>" 状态逐渐转变为 "<u>圭富</u>" 状态。**你可以从人们的休闲活动的变化中看出人们对生活的期望 越来越高。 例如，在过去人很少<u>旅行很远</u>，但现在出国旅行已不再是罕见的事情。**

　　在实际学生的作业中笔者同样又发现了有学生直接照抄谷歌翻译的情况。例 2.3 是学生 B 这部分作业的照片。除了个别词汇以外（例如 "小康"、"富裕" 是当课的生词），学生的答案和谷歌翻译完全一样。学生不但使用了从未学过的生词 "罕见"，而且也犯了和谷歌翻译同样的语法错误（"旅行很远"）。这都进一步证明学生在完成这部分作业时其实是使用了谷歌翻译。

例 2.3

学生 B 第 34 课作业英译中答案

　　类似的例子其实还有很多，笔者这两个班的十三名学生中，至少有五六名学生有明显使用谷歌翻译的迹象。而且，这种现象也并不局限在笔者任教的这门三年级中文课。在进行本研究时，笔者同时也在教一门二年级的中文课，笔者用同样的研

究方法抽查了二年级个别学生的翻译作业，同样也发现了学生照抄或者改写谷歌翻译的现象。这些发现是否意味着，任课教师在过去很长的时间里一直花了大量时间批改的其实是谷歌翻译，而非学生自己的语言产出。这种现象直接挑战了传统翻译作业的意义和作用。

3.2 学生使用机器翻译的原因

笔者发现上述学生的情况后，分别跟"涉嫌"使用了谷歌翻译的几名学生进行了访谈。谈话的目的在于从学生的角度来理解这个情况，而不是批评学生有作弊的行为等。笔者发现，因为在课上老师并没有特意强调过不可以使用谷歌翻译，所以这些学生都声称他们不知道不可以用谷歌翻译。其实笔者特意没有强调的一个原因也是担心会适得其反，警告学生不可以用谷歌翻译，也许反而会提醒一些原本不知道谷歌翻译的学生去偷偷使用谷歌翻译。

在和这些学生进行访谈的过程中，也有学生承认在完成大量的家庭作业时获得高分的压力是他们使用谷歌翻译的主要原因之一。一个学生坦白说，他也知道用谷歌翻译不好，但是有的时候实在是没有时间完成作业，而他们每一课的作业都有大段大段的翻译，谷歌翻译可以帮助他们节省很多时间。也有一个学生坦承，他的基础不好，中文课对他来说越来越难，本门中文课会是他最后一门中文课，他非常担心这节中文课会不及格，没有谷歌翻译，他完全不知道这些翻译应该怎么办。

从使用工具辅助语言学习这个角度来讲，这些学生利用谷歌翻译帮助其完成作业也不能全部被指责。谷歌翻译非常方便，对很多学生来说是非常大的诱惑。试想一下，如果在电脑上简单的复制粘贴一下就可以得出答案，对这个答案稍加修改之后还会比他们自己花半个小时得出的答案更好，得分更高，那么利用机器翻译来做作业其实就是很自然的事情了。这里需要反思的其实应该是任课教师，是教学安排、作业设计没有跟上时代的发展。

这里也要强调一下笔者作为研究者/任课教师这个双重身份带来的矛盾。从研究准则上来说，笔者不应该在研究资料收集完成之前主动告知学生笔者已经发现他们使用了谷歌翻译来写作业。可是，从教师责任的角度来说，笔者在发现学生照抄谷歌翻译后不能装作什么都没有发生。事实上，在笔者跟学生谈完话之后，那种非常明显的照抄谷歌翻译的情况就大大减少了，这在一定程度上也影响了本研究的代表性。

3.3 机器翻译的准确度

本研究的初衷本来是看看学生在多大程度上使用了谷歌翻译，但是在研究过程中，笔者同时也对机器翻译的准确度进行了初步的比较和探讨，并得到了——在笔者看来——更有价值的几个发现。

3.3.1 谷歌翻译的英译中比中译英要更加准确

本研究发现谷歌翻译的英译中和中译英的准确度是不同的，至少在本研究考察的语言层次和范围内，前者的准确度要高于后者。比方说，上文例 1 中谷歌翻译对"在现代社会里人们大多以瘦为美"这句话的翻译在 2017 年 2 月就错得离谱（"people are mostly thin for the United States"）。事实上笔者在对比检查学生作业的时候，大部分"错得离谱"乃至于"驴唇不对马嘴"的翻译都来自谷歌的中译英翻译。而相对而言，谷歌的英译中翻译要准确得多。虽然谷歌的英译中翻译时常也有一些错误，但是如果不是刻意去对比，老师很难将这些机器翻译的错误同学生自己犯的错误区分开来。

笔者并不了解谷歌翻译背后的算法，但是常识告诉我们如果要将中文翻译成英文，那么这些机器翻译首先要了解中文。考虑到谷歌公司的美国背景似乎不难解释为什么谷歌翻译的英译中功能要强于中译英的功能。这也是值得未来研究者进一步探索的方向。

3.3.2 谷歌翻译的自我进化能力

很多外语教师，也包括中文老师，在提到机器翻译时总会有一种错误的认知，认为机器翻译的准确度不高，不应该给学生推荐机器翻译，或者学生并没有办法通过使用机器翻译来获得高分（Groves & Mundt, 2015）。诚然，上文也提到了谷歌翻译有时会出现"错得离谱"的情况，但是笔者在进行本研究时发现谷歌翻译其实具有极强的自我学习以及进化能力，谷歌翻译的准确度提高得极快。还是以例 1.1 的那段关于减肥的中译英为例，例 3.1、3.2 和 3.3 列出了谷歌翻译在不同时间对同一段中文的不同英文译文。笔者还是用下划线标出了译文里翻译不准确的地方。

例 1.1

在现代社会里人们大多以瘦为美，越来越多人觉得自己太胖。打开报纸和杂志，到处都可以看到减肥的广告，电视上也有越来越多教人怎么减肥的报道。

例 3.1

谷歌翻译（2017 年 2 月）

In modern society, people are mostly thin <u>for the United States</u>, more and more people <u>feel too fat</u>. Open newspapers and magazines, everywhere you can see the <u>ads lose weight</u>, there are more and more <u>people on television</u> to teach people how to lose weight, and reported.

例 **3.2**

谷歌翻译（2017 年 3 月）

In modern society, most people are <u>thin and beautiful</u>, more and more people feel that they are too fat. Open newspapers and magazines, everywhere you can see the weight loss <u>of advertising</u>, <u>television</u>, there are more and more teach people how to lose weight <u>reported</u>.

例 **3.3**

谷歌翻译（2018 年 3 月）

In modern society, people are mostly <u>lean and beautiful</u>, and more and more people feel that they are too fat. With newspapers and magazines open, advertisements for weight loss can be seen everywhere, and there are more and more reports on how to teach people how to lose weight.

       这三个取自不同日期的例子清楚地显示，谷歌翻译的准确度是一直在提高的，而且提高的速度非常快。在 2017 年 2 月笔者刚刚开始这个研究的时候，谷歌翻译的错误很多。而且还有一些错得离谱的翻译，比方说，把"人们大多以瘦为美"翻译成"people are mostly thin for the United States"。但是仅仅在一个月之后，这个句子就被纠正成为"most people are thin and beautiful"，虽然这个翻译仍然是错误的，但是已经非常接近一个外语学习者也会犯的错误了。同时，谷歌翻译在翻译这段话时，总的错误数量也在一直减少。当笔者在一年以后再次使用谷歌翻译来翻译同一段中文的时候，发现谷歌的错误量已经减少到了一个。除了"以瘦为美"还是无法正确翻译出来以外，其他的句子、词汇几乎已经是完全没有错误了。

       而且笔者发现不断进化的谷歌翻译常常已经做的比笔者班上最优秀的学生要好。比方说下面的例子。例子 4.1 是第 33 课课后作业里面英译中的一道题。例 4.2、4.3 和 4.4 是谷歌翻译在不同时期对同一段英文的翻译，而例 4.5 则是笔者班上中文程度最好的一个学生（下称学生 C）对该段英文的翻译。

例 **4.1**

C. Translation: English to Chinese

In order to lose weight, he wakes up before 5 o'clock every day (不到…) to exercise. In addition, he has changed his dietary habits, and does not eat sweet and greasy foods. As a result, he has lost 10 kilograms in less than one month. He is about 175cm tall and weighs 65 kilograms. Obesity is no longer something he is worried about.（饮食，又…又…，结果，不再…了）

例 4.2

谷歌翻译（2017 年 2 月）

为了减肥，他每天 5 点钟前醒来<u>锻炼</u>。此外，他改变了他的饮食习惯，不吃<u>甜的</u>和油腻的食物。结果，他在不到一个月内<u>失去</u>了 10 公斤。他身高约 175 厘米，体重 65 公斤。肥胖不再是他<u>所担心的</u>。

例 4.3

谷歌翻译（2017 年 3 月）

为了减肥，他每天五点钟前醒来锻炼身体。另外，他改变了他的饮食习惯，不吃<u>甜</u>和油腻的食物。 结果，在不到一个月的时间里，他已经<u>输</u>了 10 公斤。 他身高 <u>175 厘米</u>，重 65 公斤。 肥胖不再是他担心的事情。

例 4.4

谷歌翻译（2018 年 4 月）

为了减肥，他每天早上 5 点起床锻炼身体。 此外，他改变了饮食习惯，不吃甜食和油腻食物。 结果，他在不到一个月的时间里就已经<u>损失</u>了 10 公斤。 他身高约 175 厘米，重 65 公斤。 肥胖不再是他所担心的事情。

例 4.5

学生 C 第 33 课作业英译中答案

从例 4.2、4.3 和 4.4 能看出来谷歌翻译英译中的准确度还是很高的，虽然笔者在例 4.2 划出了四个翻译得不是很恰当的地方，但是严格来说，只有"失去了 10 公斤"这个翻译是完全错误的；另外三处（"锻炼"、"甜的"、"所担心的"）在表达上其实都是可以接受的。而且从这三个例子中同样能看出来谷歌翻译的自我纠错、自我进化能力极高。在例 4.4 中，虽然谷歌翻译还是未能将"lost"翻译得很准确，但是它将另外三个表达修改的得非常自然、得体（"锻炼身体"、"甜食"、"所担心的事情"）。

而且谷歌翻译，即使在自我进化前，就已经比学生 C 的翻译要好了。学生 C 是笔者当时三年级班上中文水平最好的一个学生。他去中国大陆旅游过两次，中学学过三年中文，在大学上了二年级中文课，学习态度极好，非常认真努力。从例 4.5 中，我们可以看到学生 C 的翻译里面有很多错误，例如"关联"、"营养习惯"、"又甜又腻的饮食"、"到一个月"、"不再他的烦恼了"。

其他学者也有类似的发现。Groves 和 Mundt（2015）也认为谷歌翻译中译英的准确度达到了中级英语学习者的水平。笔者认为这个发现给语言学习的目的和教学方法、手段都提出了新的问题。如果说谷歌翻译的的质量和准确度已经超过了最好的学生，那么学习和教授外语的意义与目的在哪里，传统的教学法又应当如何面对科技发展带来的挑战？

3.3.3 搜狗翻译与百度翻译更准确

谷歌翻译是世界上知名度最高、语种最齐全的机器翻译服务，但是这并不意味着在中英文互译这个领域做得就是最好的。事实上，通过比较研究，笔者发现中国大陆科技公司开发出的百度翻译和搜狗翻译的准确度其实要高于谷歌翻译。我们还是以翻译例 4.1 中的那段英文的准确度作为评判标准。下面的例 5.1、5.2 和 5.3 是百度翻译和搜狗在不同时期对这段英文的翻译。

从这三个例子中，我们可以很清楚地看出来，在几乎同一时期（2017 年 2 月到 4 月），百度翻译对这段英译中的处理要明显优于谷歌翻译。在例 4.2 中我们看出来谷歌翻译至少有 4 个地方翻译得不是很恰当，但是在例 5.1 中，百度度翻译只有两处（"甜腻的食物"、"他是身高 175cm"）。百度翻译的自我纠错能力似乎也要高于谷歌翻译。在例 4.4 中，谷歌仍有一处严重的翻译错误（"损失了 10 公斤"）；而在例 5.2 中，除了一个不太严重的翻译错误外（"他是身高 175cm"），百度翻译已经完全没有任何别的语法、词汇错误了。更值得注意的是，搜狗翻译的准确性从一开始就高于谷歌翻译，甚至也高于百度翻译。例 5.3 是搜狗翻译在 2017 年 4 月对这段话的翻译，在当时这段翻译就已经完全没有任何错误了，可以说是文从字顺，翻译得非常到位。

例 4.1

C. Translation: English to Chinese

In order to lose weight, he wakes up before 5 o'clock every day (不到…) to exercise. In addition, he has changed his dietary habits, and does not eat sweet and greasy foods. As a result, he has lost 10 kilograms in less than one month. He is about 175cm tall and weighs 65 kilograms. Obesity is no longer something he is worried about.（饮食，又…又…，结果，不再…了）

例 5.1

百度翻译（2017 年 4 月）

为了减肥，他每天五点前醒来锻炼身体。此外，他改变了自己的饮食习惯，不吃<u>甜腻的食物</u>。结果，他在不到一个月的时间里瘦了 10 公斤。他<u>是</u>身高 175cm 体重 65 公斤。他不再担心肥胖了。

例 5.2

百度翻译（2018 年 4 月）

为了减肥，他每天五点以前就起床锻炼。此外，他改变了自己的饮食习惯，不吃甜食和油腻食物。结果，他在不到一个月的时间里就瘦了 10 公斤。他<u>是</u>身高 175cm 体重 65 公斤。肥胖不再是他担心的事情了。

例 5.3

搜狗翻译（2017 年 4 月）

为了减肥，他每天早上 5 点起床锻炼。此外，他改变了饮食习惯，不吃甜食和油腻的食物。结果，他在不到一个月的时间里瘦了 10 公斤。他身高约 175 厘米,体重 65 公斤。肥胖不再是他担心的事情。

　　在比较这三种机器翻译在中英文互译方面的准确度时，笔者也发现了另一个有趣的现象，那就是中国大陆公司开发的百度翻译和搜狗翻译，尤其是搜狗翻译，看起来似乎比美国公司开发的谷歌翻译更"了解"中国，他们在翻译一些具有中国特色的词汇时更为准确。我们以对 "comparatively well-off" 这个词汇的翻译为例。这个词出现在上文例 2.1（也就是第 34 课的作业）中。下面的例 6.1、6.2 和 6.2 分别是谷歌翻译、百度翻译和搜狗翻译在同一天对这个英文单词以及这段话的不同翻译。

例 2.1

C. Translation: English to Chinese

1. In the past ten years, it can be said that Chinese people's lives have gradually changed from a state of "comparatively well-off" to a state of "richness." You can tell from the change in people's leisure activities that people's expectations (期望, qīwàng) toward life are getting higher and higher. For example, in the past people seldom travel far, but now traveling abroad is no longer a rare thing. （可以说是；从 A 变到 B；看得出来；远处；出国）

　　　在例 6.2 和 6.3 中，百度翻译和搜狗翻译都将 "comparatively well-off" 这个词正确地翻译成了"小康"。而在例 2.2 和 6.1 里面，谷歌在不同时期对这个词翻译则分别是"相对富裕"和"比较富裕"。我们不能说这两个翻译完全不对，

但这明显是对中国国情缺乏了解的表现。谷歌翻译的这个误译很像一个中英文水平很好但是又不了解中国社会、经济、历史情况的外语学习者所会犯的错误，而百度翻译和搜狗翻译则要更接地气。事实上，搜狗翻译的一大特色就是它可以很准确地翻译出大量的新词、口语词、网络用语，甚至是一些方言词汇。比方说搜狗翻译可以将"扎心了老铁"准确地翻译成"my heart has broken，old fellow"；将"不明觉厉"准确地翻译成"I don't understand what you said, but I think you are awesome"。[7]这是谷歌翻译在本文完成时也没有做到的。

例 6.1

谷歌翻译（2017 年 4 月 1 号）
十年来可以说，中国人民的生活从"**比较富裕**"的状态逐渐转变为"富裕"状态。 从人们的休闲活动的变化可以看出，人们对生活的期望越来越高。 例如，以前人们很少旅行很远，但现在出国旅行已不再是罕见的事情。

例 6.2

百度翻译（2017 年 4 月 1 号）
在过去的十年里，可以说，中国人的生活已经逐渐由"**小康**"状态转变为"富裕"状态。从人们休闲活动的变化可以看出人们对生活的期望越来越高。例如，在过去人们很少旅行很远，但现在出国旅游已不再是一件罕见的事情。

例 6.3

搜狗翻译（2017 年 4 月 1 号）
在过去的十年中，可以说，中国人民的生活已经从"**小康**"状态逐渐转变为"富裕"状态。你可以从人们的休闲活动的变化中看出人们对生活的期望越来越高。例如，过去人们很少旅行，但是现在出国旅行不再是一件稀罕事了。

除此以外，例 6.1、6.2 和 6.3 还可以再一次证明的笔者上述的几个发现，那就是当前机器翻译的翻译水平很可能要高于中级汉语学习者的翻译水平（至少比笔者的学生水平要高）；机器翻译的准确度也会越来越高，因为它的自我学习能力、自我进化能力极强；在中英文互译的准确度方面，目前中国大陆公司开发的搜狗翻译要优于百度翻译，而百度翻译要优于谷歌翻译。

4. 小结及讨论

本文展示了在中文教学领域机器翻译对传统翻译作业的挑战。笔者发现在美国的大学里越来越多的中文学习者开始借助谷歌翻译而不是传统的字典来完成翻译作业。在这个过程中，整句乃至整段修改、照抄机器翻译答案的"作弊"现象时有发

---

[7] http://www.sohu.com/a/133040255_126842

生，而且老师也越来越难对此进行判断。从学生的角度来讲，使用谷歌翻译极其方便，容易得到高分，是应对中文课作业多、压力大的"捷径"。如果教师不特意进行对比研究，则常常会忽视这个现象，导致浪费大量时间来批改机器翻译的答案。因此，教师对学生使用机器翻译完成作业的现象必须有充分的认识和应对措施。

本研究也发现，在中英文互译的准确度方面，就目前来看，中国大陆公司开发的搜狗翻译要优于百度翻译，百度翻译要优于谷歌翻译，而一般来说谷歌翻译的准确度至少要高于美国常青藤大学三年级中文课的学生翻译的准确度。更重要的是，随着科技的发展，可以期待机器翻译的准确度将会进一步提高，有可能会全方面超过外语学习者。同时，外语学习者使用机器翻译来完成学习任务的现象只会越来越普遍，这将给外语教学领域乃至专业翻译领域都带来新的要求（Clifford, et al., 2013）。

这里需要提及的一点是，本研究从研究方法上来说属于实证性质的定性研究，研究对象只是笔者所教的一门中级中文课程的学生以及他们部分的翻译作业，本文的数据分析与结论并不具备统计学上的定量意义。机器翻译技术的发展与应用方兴未艾，该技术在中文教学领域更应该得到足够的重视。未来的教学研究和实践可以着眼于如下三个方面。1）需要更多的研究和数据来分析学生使用机器翻译的情况。本文的研究仅限于翻译作业，在其他类型的作业里，比如完成句子、阅读理解、写作文等等，学生是否也使用了机器翻译？效果如何？2）在新的技术发展环境下需要重新审视传统的教学法和评估标准。比如，既然机器翻译如此方便、准确，那么是不是不应该禁止学生使用机器翻译呢？是不是应该将诸如搜狗翻译、百度翻译等质量相对有保障的工具介绍给学生呢？3）需要有新的课堂教学设计将机器翻译和其他语言处理技术有机地结合在中文教学之中，利用新技术来辅助中文教学。例如，搜狗翻译和百度翻译不但在翻译的质量上高于谷歌翻译，而且它们还提供了额外的教师可以应用在课堂上的功能，如自动划出重点词汇及其翻译的功能、语音合成及发音功能。根据笔者的初步观察，百度翻译的发音在自然度和流利度上要优于搜狗翻译的发音，而搜狗翻译的发音要优于谷歌翻译的发音（Tian, 2018）。已经有学者将语音处理技术应用在了中文教学领域（Da, 2015），而笔者也根据本研究设计了全新的作业系统。这些都是值得未来的教学试验和学术研究进一步探索的方向，本文的讨论只是抛砖引玉，希望能引起大家对这个现象的重视和研究兴趣。

参考文献

Alridge, D. P. (2006). The limits of master narratives in history textbooks: An analysis of representations of Martin Luther King, Jr. *Teachers College Record, 108*(4), 662-686.

Baker, C. (2013). *Student and instructor perceptions of the use of online translators in English Composition.* (Unpublished master's thesis), Mississippi State University, MS, USA.

Becker, H. S. (1996). The epistemology of qualitative research. In R. Jessor, A. Colby, & R. A. Shweder (Eds.), *Ethnography and human development: Context and*

meaning in social inquiry (pp. 53 - 71). Chicago, IL: University of Chicago Press.

Case, M. (2015). Machine translation and the disruption of foreign language learning activities. *Elearning Papers,* 45, 4-16.

Chou, Z., Chiang, J., & Eagar, J. (2011). *A new China: Intermediate reader of modern Chinese*. Princeton, N.J: Princeton University Press.

Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the Landscape: What is the Role of Machine Translation in Language Learning? *The Acquisition of Second Languages and Innovative Pedagogies* (10), 108-122. doi:10.7203/attic.10.2228

Correa, M. (2011). Academic Dishonesty in the Second Language Classroom: Instructors' Perspective. *Modern Journal of Language Teaching Methods, 1*(1), 65-79.

Correa, M. (2014). Leaving the "peer" out of peer-editing: Online translators as a pedagogical tool in the Spanish as a second language classroom. *Latin American journal of content & language integrated learning. 7*(1), 1-20. doi:10.5294/laclil.2014.7.1.1.

Da, J. (2015). The Application of Speech Recognition Technology in Chinese Language Learning: What Can Be Learned from a Pinyin Lab Session. *Journal of Technology and Chinese Language Teaching, 6*(1), 16-24.

Ducar, C., & Schocket, D. H. (2018). *Google Translate: The Good, the Bad and the Ugly*. Paper presented at The IUP Spring Methodology Conference, Indiana, PA.

Garcia, I., & Pena, M. I. (2011). Machine Translation-Assisted Language Learning: Writing for Beginners. *Computer Assisted Language Learning, 24*(5), 471-487. doi:10.1080/09588221.2011.582687.

Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes, 37*, 112-121. doi:10.1016/j.esp.2014.09.001.

Li, Q. (2006). A Review on the Research of Pedagogical Chinese Grammar for Teaching Chinese as a Foreign Language. *Chinese Teaching in the World*, *76*(2), 110-118. [李泉. (2006). 对外汉语教学语法研究述评. 世界汉语教学, *76*(2), 110-119.]

Krashen, S. D. (2003). *Explorations in language acquisition and use: The Taipei lectures*. Portsmouth, NH: Heinemann.

Niño, A. (2009). Machine translation in foreign language learning: language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL, 21*(2), 241-258. doi:10.1017/S0958344009000172

Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 43-72). New York, NY: Academic Press.

Somers, H. L. (2003). *Computers and translation: a translator's guide*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Tian, Y. (2018, April). *Speech recognition and machine translation for Chinese language classroom applications*. Paper presented at the 2018 CLTA-US Annual Conference, Washington, D.C.

Zhao, J. (2010). Review and Re-evaluation of the Methodology in Teaching Chinese as a Foreign Language. *Chinese Teaching in the World*, *24*(2), 243-254. [赵金铭. (2010)．对外汉语教学法回视与再认识. *世界汉语教学，24*(2), 243-254.]

附录

第 35 课作业单（格式有调整，原格式为 4 页。）

CHIN 212 Intermediate Modern Chinese II
Homework: *A New China* Lesson 35

A. Choose the correct word to complete each sentence
1. 他现在（就，只）住在校园里，你什么时候去找他都可以。
2. 他认为美国的老人大多喜欢自己住。我（居然，倒）觉得不见得。
3. 我把自行车给（丢，失去）了，现在每天都得走路上课。
4. 我在美国（上，升）完大学以后，就会立刻回中国。
5. 他现在有钱得很，（由于，至于，无论）他的婚姻情况好不好，我就不知道了。
6. 他从上个月开始就（一直，一向）背生词，到现在已经记住了三千多个词了。
7. 我母亲的生活都是（用，以，当）孩子为中心。
8. 现在已经十二点了，他（大致，大概）不会来了。

B. Complete the dialogues with the expressions provided.
1. A：在你看来，大学生的生活重心应该是什么？
   B：_____。（以…为…）
2. A：你们与其每天吵架，还不如离婚。
   B：_____。（不见得）
3. A：到了晚年，你愿意让孩子照顾你，还是自己照顾自己？
   B：_____。（宁可…也不要…）
4. A：在古代社会中，女人依靠谁来生活？
   B：_____。（一向）

C. Translation: English into Chinese
1. From the individual's perspective, a big household of three generations is not necessarily the ideal form of the modern family. Owing to different living habits, the possibility of recurring friction is greater than that in a small family centered on a husband and wife. （角度；形式；产生；以…为中心）

2. Letting aged parents live in a retirement home can cause problems, too. The expense of living in a retirement home is high, and thus it saddles children with a financial burden. In addition, most of the times the elderly are not satisfied with the care retirement homes provide. （费用；加重；提供；满意）

3. I (myself) am willing to work in a small city because I don't want to have a life of fierce competition and great work pressure. In terms of salary, as long as I can have enough to eat and keep myself warm, making a little less doesn't matter. （愿意；激烈；至于；温饱）

D. Translation: Chinese into English (Please type and attach a separate piece of paper)

Zhōngguó de gǎigé kāifàng budàn shǐ jīngjì   shèhuì chǎnshēng le jùdà de
中国   的 改革 开放 不但 使 经济、 社会   产生   了巨大 的

biànhuà   tóngshí yě gěi Zhōngguó de hūnyīn hé jiātíng jiégòu dàilái le
变化 ， 同时 也 给 中国   的 婚姻 和 家庭 结构 带来 了

qiánsuǒ-wèiyǒu de yǐngxiǎng   Jìn sānshí nián lái   Zhōngguó de jiātíng móshì
前所未有   的 影响 。 近 三十 年 来， 中国   的 家庭 模式

biàn de duōyànghuà   jiātíng guīmó suōxiǎo   hūnyīn gōngnéng yě zhújiàn de
变 得 多样化 ， 家庭 规模 缩小   婚姻   功能 也 逐渐 地

jiǎnruò   Jiātíng guīmó suōxiǎo shì Zhōngguó jiātíng jiégòu chǎnshēng biànhuà
减弱 。 家庭 规模 缩小 是 中国   家庭 结构 产生   变化

de zhòngyào tèzhēng   Zài dàchéngshì lǐ   zuì chángjiàn de shì yóu niánqīng
的 重要   特征 。 在 大城市 里，最 常见 的 是 由 年轻

fūqī hé wèichéngnián zǐnǚ zǔchéng de héxīn jiātíng   ér bùxiǎng shēngyǎng
夫妻 和 未成年 子女 组成 的 核心 家庭 ， 而 不想   生养

háizi de fūqī shùliàng yě zài zēngjiā   zàochéng zhěngtǐ shēngyù lǜ xiàjiàng
孩子 的 夫妻 数量 也 在 增加 ，   造成   整体 生育 率 下降 。

Lìngwài   dúshēn bùhūn yǐjí wèihūn tóngjū de xiànxiàng rìyì pǔbiàn   yě shǐ
另外 ， 独身 不婚 以及 未婚 同居 的 现象 日益 普遍 ， 也 使

jiātíng de chuántǒng xíngshì shòudào tiǎozhàn
家庭 的 传统   形式 受到   挑战 。

Zhèxiē biànhuà zài yídìng de chéngdù shàng yě zàochéng le yìxiē lǎorén
这些 变化 在 一定 的 程度 上 也 造成 了一些 老人

wèntí   Xiàndàirén yóuyú shēnghuó fāngshì de gǎibiàn   běnlái jiù biàn de
问题 。 现代人 由于 生活 方式 的 改变 ， 本来 就 变 得

qīngxiàngyú yǐ jīnqián hé wùzhì de fāngshì lái xiàoshùn fùmǔ   Bù yǔ fùmǔ
倾向于 以 金钱 和 物质 的 方式 来 孝顺 父母 。 不 与 父母

tóng zhù yǐhòu   zǐnǚ gèng yīnggāi guānxīn fùmǔ de rìcháng shēnghuó   cái
同 住 以后 ， 子女 更 应该 关心 父母 的 日常   生活 ， 才

néng bìmiǎn shǐ lǎoniánrén gǎndào gūdú wúzhù
能 避免 使 老年人   感到 孤独 无助 。

眼动技术及其在汉语教学研究中的应用
(Eye-tracking Technology and its Application in Chinese Teaching and Learning Research)

施黎静
(Shi, Lijing)
伦敦政治经济学院
(London School of Economics and Political Sciences)
l.shi3@lse.ac.uk

摘要：眼动研究作为阅读研究和人际交互领域一个重要的研究方法近来也进入广大教育者的视野，促使相关研究激增。本文选取三个与中文教学密切相关的眼动研究领域，回顾近年来的发展，展现其在汉语阅读，阅读策略和实时网络语言教学(synchronous online language teaching)方面的应用和发现，同时显示眼动研究在范式和形式上的多种可能性。本文首先简介眼动研究的历史，随后侧重展现眼动研究在汉语阅读和阅读策略这两个方面已经取得的成果。然后以两个混合式眼动研究(Stickler & Shi, 2015; Shi, Stickler, & Lloyd, 2017)为例来阐述如何创新地使用眼动技术对网络实时汉语教学进行深层次的分析和研究。文章也将为有意开展眼动研究的中文教育者提供一些实用的操作建议。

Abstract: As an important research method in the field of reading research and human-computer interaction, eye-tracking research has recently attracted increasing attention from educators. Thus, more related research on the topic has been published. Selecting three areas closely related to Chinese teaching, this paper reviews the development of eye-tracking research in recent years, and shows its applications and discoveries in Chinese reading, reading strategies, and synchronous online language teaching. It also demonstrates diverse research paradigms and forms of conducting eye-tracking research in language education. This article first introduces a brief history of eye movement research, focusing on the achievements that eye-tracking research has achieved in both Chinese reading and reading strategies. How and what eye-tracking research can contribute to more in-depth analysis of online synchronous Chinese teaching are then exemplified by two innovative eye-tracking studies (Stickler & Shi, 2015; Shi, Stickler, & Lloyd, 2017). The article also provides practical suggestions for Chinese language educators interested in conducting eye movement studies in the future.

关键词：眼动研究，中文阅读，网络实时语言教学

Keywords: Eye-tracking, Chinese reading, synchronous online language teaching

1. 眼动研究及其基本研究范式

随着信息科技在人类生活中的不断深入，人机交互（Human Computer Interaction -HCI）研究增多，眼动研究正在进入普通研究者的视野。简而言之，眼动研究（eye-tracking research)是通过测量眼睛的注视点的位置或者眼球的运动轨迹的一个研究方法。目前在软件设计，应用程序设计，乃至商业网站的设计和使用中，眼动研究都在发挥着越来越醒目的作用。在科技辅助教育的研究领域，使用眼动技术的研究从 2007 年的 1 篇飙升为 2012 年的 39 篇（Lai et al., 2013），范围主要集中在信息处理，教学策略的效果和学习者的个体差异。 这从一个侧面展现了教育研究者对眼动研究的日益青睐。

眼动研究并不是一个时髦的新工具或者新方法，相反，它被用于阅读研究的历史可以追溯到 100 多年前 (Just & Carpenter, 1976)。此后的很长时间里它在阅读研究和信息处理领域有广泛的应用 (Rayner, 1998, 2009)。人机交互的科学家们从 30 多年前开始大量地使用眼动技术来研究电子界面和电子设备的使用效果与设计 (Jacob & Karn, 2003; Nielsen & Pernice, 2010)。

在阅读研究领域，眼动技术主要用于测量阅读、注视静态、动态场景时的眼部活动。Rayner（1998）曾经总结出了一系列有关眼部运动的经典行为。他将新的阅读区域或图像带入视野的眼部运动称为"眼跳 saccades"，是眼睛从已知区域朝向新的未知区域，即在两个注视点之间的跳动。 阅读时眼跳的频率是每秒 3-4 个。而两次眼跳之间的停顿被称为"注视 fixations"， 一般在 200-250ms（毫秒）之间。默读时注视时间大约持续 225ms，朗读时为 275ms。对图像注视的时间大约为 330ms。眼动在文本中从右向左的回看被称为"回视 regression"。由此衍生而出的眼动指标帮助研究者了解人们的阅读习惯，思维方式，理解能力等等。

在人机互动的研究领域，所参考的眼动数据主要有三种： 注视类（如注视总时长、次数、首次注视的时间点等）， 回视类（如次数、振幅等）， 和路径类 (Poole & Ball, 2006) 。通过对这些数据的分析，可以更充分地了解产品设计的效果，人们使用电子产品时的喜好。这个领域的研究者在 30 多年的实践后也发现了眼动数据在一定程度上受制于参与者的体态特征（如，眼球的大小和形状，有没有带眼镜）或具体研究活动设计的限制。因此，Nielsen & Pernice (2010) 建议将眼动研究同其他研究方法想结合，如刺激回忆法（stimulated recall）， 问卷调查，采访和观察，从而进一步提高眼动研究的信度和效度。将量化的眼动数据同质化数据甚至参与者的个体信息相结合不但多层面地展现参与者是如何阅读，如何与内容和他人互动的，并且有利于进一步了解这些行为背后的原因。Nielsen & Pernice 的这些建议

对于从事技术辅助语言教育的研究人员很有价值；近期一些语言教育领域的眼动研究操作与实践也逐渐体现了这样的思路。

本文以下部分将从三个跟技术辅助汉语教学紧密相关的领域来回顾眼动研究的具体使用以及取得的成果：汉语阅读，阅读策略，和在线语言教学。同时展现眼动研究的范式如何从传统心理学的以定量分析为主的方式，逐渐转变为混合式研究中的一项手段。

2. 汉语阅读的眼动研究

关于汉字阅读的眼动研究成果丰富，已在语言学习和认知心理学领域有自己的一席之地。早在 1925 年，Miles & Shen 就发表了关于汉字横读与直读时阅读速度的眼动研究。Journal of Research in Reading 在 2013 年的出版的 Eye movements during Chinese reading 特刊汇集了研究者对于汉语阅读最新的眼动研究成果[1]。其中的六篇论文分别讨论了：成人读者如何用汉语学习新词汇，特别是词语之间插入空格是否有助于学习；他们在阅读句子时如何预先处理位于注视右侧句子中即将出现的信息；字符的哪个组成部分对于不中断的阅读最重要；使用线性混合模型来研究读者对汉语语义的预处理；以及在汉字阅读过程中跳读的特征分析。

白学军等人在 2015 年的回顾文章中展示了过去 20 年的相关研究主要集中于探索哪些因素决定汉语阅读过程中的注视时间和注视位置。针对注视时间的研究已从词边界信息、词频、熟悉性、预测性、语义透明度、具体性、非注视词的特性等方面进行探讨。而对于注视位置，探索的领域包括汉字字号、笔画数、字结构、词长、词结构、词间空格、词频、预测性、合理性、年龄和阅读能力。他们通过比较发现，汉语阅读的眼动特点与拼音文字的眼动特点既有共同之处也存在差异。

比如，Bai 等(2008) 对汉语为母语的成人的研究发现，相对于正常的词间无空格，字间空格和非词空格条件会干扰阅读，而词间空格既不促进阅读，也不干扰阅读。而 Shen 等(2012) 针对美韩日泰四国留学生的同类研究却发现，词间空格有利于留学生阅读汉语文本，且该促进作用不受被试者母语文本呈现方式（拼音文字或表意文字）的影响。

研究人员也探索了汉字笔画数和汉字结构对汉语阅读的注视位置所产生的影响。比如，孟红霞等人 (2014) 发现单次注视条件下首次注视往往落在词的中心，多次注视时首次注视往往落在词的开头位置。当首字为多笔画汉字时，读者的首次注视更多地落在词的首字上。Zang et. al (2013) 发现左右和上下结构的单字词或双字词的眼动模式相似，这似乎表明汉字结构并不会对汉语阅读的注视位置产生影响。

这些以量化统计主导的眼动研究对汉语教学和研究有直接的启发意义，因为在汉语教学中汉字和阅读的重要性不言而喻。而通过以上这些研究的介绍，读者也可

---

[1] https://onlinelibrary.wiley.com/toc/14679817/36/S1

能意识到对于汉字阅读的眼动研究依然是一个值得继续探索的领域， 还有很多研究的空间。就像 Liversedge 等（2013）所指出的，汉语阅读的眼动研究不仅帮助我们了解汉语习得过程中的具体问题，也将推动对人类阅读能力本质的理论诠释。

3. 在线阅读策略的眼动研究

阅读是一个互动的过程，它不仅是读者从字词句的层面开始，由下而上的信息解码，也可以是在阅读策略的指引下，从上而下，提取信息，构建意义。随着科技的广泛使用，阅读也不仅依存于传统的书本之上。与之相反，今天我们阅读的对象大多以超链接文本（hypertext）的形式出现，在线阅读能力成为一种新的技能。所以 Kang（2014) 结合眼动技术和访谈，研究英语母语者和非母语者的在线阅读的策略，继而探索这两者在速度，目的性和理解度上的异同。9 名美国大学生和 9 名中国留美大学生参加了本次研究，他们先阅读在线学习平台上的信息，然后完成 5 个阅读理解问题，他们的整个阅读过程被眼动仪记录下来。在眼动捕捉结束后，被测者还接受了访谈，被邀请参与观看自己在阅读过程中的眼动轨迹并发表自己的看法。Kang 发现英语母语者的阅读速度更快，但他们在阅读过程中的注意力分布和阅读理解的边沿跟非母语者并没有根本上的差异。也就是说，一旦非母语者掌握了核心词汇，对线阅读表现的最有影响的是阅读元策略。在该研究中，质化的数据帮助研究者了解被测者的认知过程和行为背后的具体原因；就研究方法而言，质化数据是量化的眼动数据的有效辅助和补充。这样的眼动研究范式更接近 Nielsen & Pernice (2010)的建议。

对研究汉语阅读策略有兴趣的学者，可在上述研究的基础上继续探索，比如，对比分析母语者和非母语者的汉语阅读策略，研究汉语与其他语言的阅读策略之间的异同，或者调查不同语言水平的汉语学习者的阅读策略差异。眼动技术可以提供直观和量化的数据。

4. "混合式"眼动研究探索实时网络语言的教学

4.1 为什么采用 "混合式"的眼动研究

实时网络语言教学(synchronous online language teaching/learning) 是计算机辅助语言教学（CALL）领域的前沿地带也是其研究重点，因为大量的语言课程在网上依然是以课堂的形式存在的。然而网络语言教学参与者的具体行为及其背后的原因，依然有待更深刻的研究。目前的研究以问卷，前后测试为主， 无法在教学的自然状态下采集参与者的真实行为和第一时间的思维方式。 有些研究注重量化指标（如 p 值），却对教学发生的"情境 context"缺乏考量和反思。但如果以" 社会文化"的研究范式来分析实时网络语言教学，我们不难发现在这个特定环境中的教与学是以各种语言和科技的可供性（affordance) 为媒介的,而传统的研究方法无法深刻地揭示处于多重媒介之中的教学互动。

Stickler 和 Shi 数年来从社会文化理论（Sociocultural theory）的视角持续研究实时网络语言课程中师生的真实互动 (如 Shi, Stickler, & Lloyd, 2017; Stickler & Shi 2015, 2017)。从 2012 年起她们尝试使用眼动研究方法，将它同刺激回忆法，问卷调查，反思等研究工具相结合，探索学生和老师在实时网络语言教学中注意焦点（attention focus) 的特征及其原因。具体的研究问题包括：

- 在实时网络中文教学中的阅读活动中，学生的注意焦点在哪里（拼音，汉字或其他）？为什么？
- 在实时网络中文教学中的对话活动中， 学生的注意焦点在哪里（拼音，汉字或其他）？为什么？
- 在实时网络（中文）教学中，老师的注意焦点在哪里？为什么？

研究的目标是更清晰地揭示自然状态下的网络汉语教学的行为细节，并挖掘这个过程中促进或阻碍教学的因素。从研究伦理的层面来说，参与这两项研究的师生不仅仅是被 研究的对象，也是研究的共同参与者和创造者。通过观看自己或他人的网络教学眼动录像，可以帮助他们反思自己的学习策略或教学策略，从而更有意识地参与网络学习，各自成为更好的网络学习者和教育者。如此一来，行为研究也具备了赋能（empowerment）的作用。

接下来的，将通过两个研究例子来介绍她们是如何开展这种"混合式"的眼动研究及其研究发现。

4.2 研究一：实时网络汉语教学中学习者的注意焦点及其原因

第一个眼动研究（Stickler & Shi 2015）探索实时网络汉语教学中学习者的注意焦点及其原因。研究对象是 10 名参加过英国开放大学远程中文课程的成人汉语学习者，当时他们的汉语水平在 A1-B1 之间。

在眼动研究开始前，他们先填写了关于各自语言能力和信息技术能力的调查量表。然后进入眼动实验室，完成一个简单的阅读理解活动。在学习软件 Elluminate 的界面上，学生可以从上到下看到一段中文文本，对应的拼音文本和英文的阅读理解题。

参与者的眼动数据被 Tobii 60 眼动仪实时记录，在阅读活动结束后，眼动记录马上生成热点图（hotmap），直观地显示参与者在阅读过程中的注意焦点。比如，下图中红色的部分就是一个学习者在阅读活动中注视停留时间最长的地方。
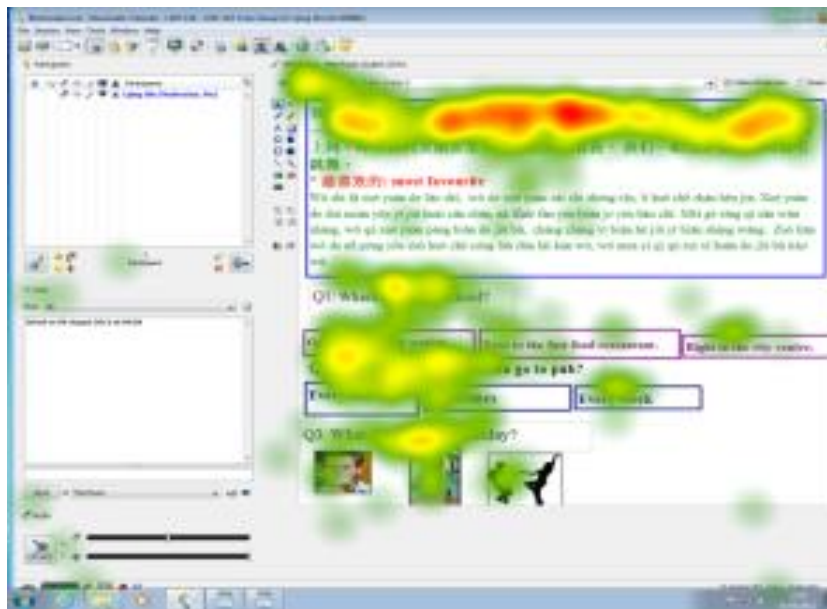
图 1 Activity 1 (reading task), 爱米 (Ai Mi) – excellent（高阅读能力的学生在参与阅读活动后的注视热点图）(Stickler & Shi, 2015, p. 62)

当该学习者完成阅读问题后，研究者就与其共同观看眼动录像，并讨论其阅读过程中的选择，策略和原因。这样，一方面有效收集了学习者对其行为的解释(三角测量法)，另一方面学习者也通过观看自己的眼动录像激发他们对自己的阅读行为进行反思，继而成为更有学习策略意识的语言学习者。

这个研究的第二阶段是对口语互动过程中，学习者的关注焦点进行研究。 每个参与者在眼动实验室中参与十五分钟的网络实时中文课程。这节课程的主题是"交通工具"，有一名中文老师和四名学习者参与；进行朗读，问答，对话等语言互动。而整个过程中，眼动仪全程收集在实验室内那名参与者的眼动数据，也就是他／她的关注焦点。

当这个阶段的眼动研究结束后，参与者稍事休息后，同研究者一同观看互动阶段的眼动录像，回顾这个过程中的感受，解释特定关注焦点的理由，分析自己在当时情景下做出选择的原因。

就阅读任务中学习者的关注焦点而言，通过对比眼动仪生成的注视总时长和学习者的汉语水平，她们发现这十名学习者对拼音的注视比率在 3%－97%之间。水平越高的学生越少注意拼音，而水平低的学习者几乎完全依靠拼音来完成阅读理解题。最有意思的是中间水平的学习者的关注焦点。 他们完成有能力通过汉字阅读来完成任务，可是在阅读过程中，依然关注了拼音。这些学习者在眼动研究后的刺激回忆中详细解释了原因。归纳而言，有的学习者通过拼音来"确认"自己对汉字解码的信息；有的是因为出自母语的影响而更习惯于阅读拼音 （"更方便更快速"）；而有的是通过拼音来再次学习和记忆汉字。

对于口语互动阶段的注视形态，研究者首先将所有参与者在整个过程中的"热点图"叠加，从而清楚地看出他们的关注焦点是在界面所展示的例句和替代词语的拼音上。然后进一步分析关注焦点的特征。 将 Elluminate 的界面分割成三个兴趣区（Area of Interest/AoI）：内容区，社交区和技术区，并计算学习者的注视总时长和分布规律。内容区是指界面上呈现的与学习相关的文字或图像的区域。社交区是指界面左上面显示参与者名字和表情符号的区域， 而技术区是指学习者来发言的"话筒"和文字输入区域。
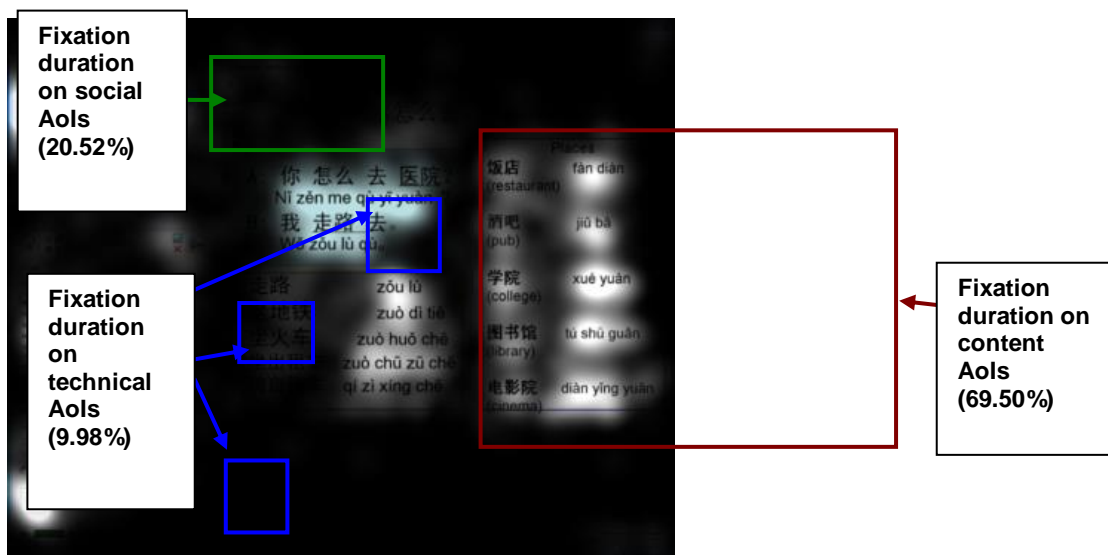

图 2 Three types of AoIs and relative Fixation Duration (学生注视的三个区域及其相对时长) (Stickler & Shi, 2015, p. 68)

她们发现学生近 70%的注视聚焦于学习区，20%在社交区，约 10%是在技术区。这些数据显示了基于网络的语言学习是多媒介的，也从一个侧面说明了社交的需求在此类学习中不可忽视。从某一个层面上来说，网上学习也是一种基于网络的社交活动。这提醒教学者和教育平台的设计应从用户体验的角度去考虑学习界面的使用和设计。

概括而言，这个混合式的眼动研究首次展现了成人汉语学习者在实时网络课程中与内容互动，与他人互动时他们注意焦点的时长，特点，分布，以及与其语言水平之间的关系，并探讨了这些图像数据，数字证据后的社会文化原因。辅之以问卷量表和语言能力测评，眼动数据展现了直接的量化证据，而实时课程结束后进行的刺激回忆法提供了具体的更有深度的质化证据。（更详细的内容，请参阅 Stickler & Shi 2015)

4.3 研究二：实时网络教学中老师的注意焦点及其原因(Shi et al., 2017)

第二个眼动研究在 2013－2014 年进行，探索语言老师在进行网络实时课程中的注意焦点及其原因。研究参与者分别为德语老师 Ella，法语老师 Valérie 和汉语

教师 Xiaomei。这三位老师都是经验丰富的语言教师。但就网络教学而言，他们的经验不同：Ella 熟悉网络教学并熟悉特定网络教学平台；Xiaomei 熟悉网络教学但不熟悉特定网络教学平台；而 Valérie 不熟悉网络教学，也不熟悉特定网络教学平台。

研究方法与前面提到的混合式眼动研究（Stickler & Shi 2015）一脉相承。 每一位老师在眼动研究前接受采访，以收集与他们教育经验，风格，理念相关的数据。然后他们进入眼动实验室，并各自通过 OULive 的界面上了一节 20 分钟的网课，分别和四位同学实时互动。Tobii 全程记录了他们的眼动细节。网课结束后，他们一边观看眼动录像，一边回忆，解释，评论和反思他们的教学过程。

通过分析这三位老师的眼动数据（注视时长），我们发现汉语教师 Xiaomei 在授课过程中，51%的注视聚焦于内容区, 30.6%在社交区，而 18.5%的在技术区域。同上文提到的学习者的注意焦点特征相比，老师更多地注意社交区域，反映出老师更刻意维持学习者在网课过程中的社交连接，积极地通过表情符号来弥补网络交流缺乏面部表情和身体语言的不足，更密切地关注"参与者窗口"以保证不错过学生可能发出的交流信号。之所以 Xiaomei 近五分之一的注视集中在技术区，是由于这位老师对于新的教学界面并不熟悉 ，尽管她有多年网上授课的经验。这一系列的数据也显示出网络实时教学是一项要求高而复杂的（社会）活动；要求老师在短时间内正确有效使用多项（认知和交际）能力。语言教师在网络授课时所面对的不仅仅是技术的挑战。她们还要有能力确保师生之间交流通畅进行，从而完成既定的教学目标。

在眼动试验后的反思环节中，伴随着眼动录像的回放，三位老师回顾了自己的教学体验并评论各自的教学策略，生成了丰富的质化数据。研究者随后对这些反思数据进行了分类和进一步分析。

首先，老师们的反思涵盖以下几个方面：解释自己的眼动特征，反思该研究过程，对网络教学的（整体）反思，对网络教学策略的（特定）反思，和对自己语言教学及其理念的反思。可以看到，汉语教师 Xiaomei 的反思集中在教学策略 （TS）上，占其反思总量的 58.1%，12.9%的反思同 TS- 相关。 TS-是指老师表达了自己的教学理念或者技能因受制于网络环境或者科技的而无法展开。比如，"如果是在面对面的课堂，对于这个语法点我会要求学生小组活动，但在网络教学中，小组活动很难开展，所以我只要邀请学生依次发言"。可见，结合眼动数据和刺激回忆法，老师能够多方面深层次地评测自己在教学过程和特点。

如果把这些反思数据同老师的个人情况对照分析，就会发现"有网络教学经验的，对网络教学平台熟悉"的老师的反思偏"正面"，他们能清楚地解释其特定网络教学行为背后的原因和理念。而"缺乏网络教学经验，对网络教学平台不熟悉"的老师， 其反思偏向"负面"，更多地在于思考网络教学与传统课堂教学的不同，陈述自己无法在网络上实践熟悉的教学理念的无奈。这个发现说明网络语言教学的成功不但依靠教学理念的恰当和先进，还需要老师熟悉网络教学的平台和特定功能。教

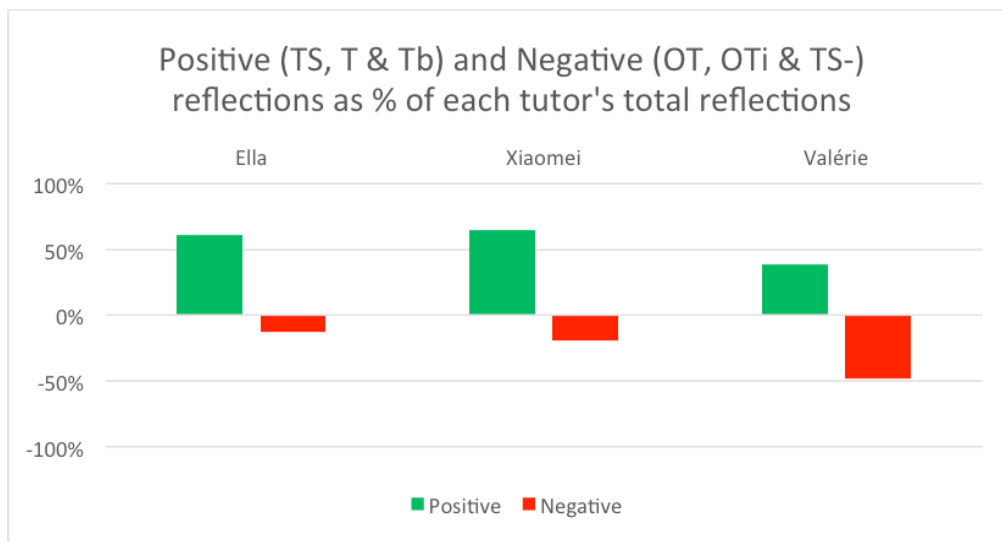育者应当结合学习的特定环境，学习者的个体需求，教学内容的要求，熟练使用地教学科技，发挥其优势，避免其缺陷（更详细的内容，请参阅 Shi et al., 2017）。

图 3 Positive and negative reflections（老师反思的正负面情况）(Shi, Stickler, & Lloyd 2017, p. 226）

　　以上回顾了在过去几年中，采用混合式的眼动研究方法，对实时网络汉语教学的一些探索。由此可见，眼动研究能够提供细致、即时、可靠和丰富的数字和图像数据，能清晰直观地展现在实时线上和线下语言教学过程中学习者和教育者所关注的区域和内容。这些对于了解真实情景下的教学互动具有重要的意义。教育者可以根据这些发现进一步了解网络教学参与者所面对的困难和挑战，从而改进语言教学内容的设计，提升教学的效果和学生的满意度。而将眼动研究与其他研究方法（如刺激回忆法、反思法、问卷调查）相结合，为量化的数据结果提供了详实的语境，增强了研究的信度和效度。更重要的是，将网络学习的过程真实地展现在参与者眼前，帮助他们认识理解自己教学行为的策略或不足，使其成为更有效的学习者和教育者。

5. 开展眼动研究的具体注意事项

　　O'Rourke, Prendergast, Shi, Smith, & Stickler ((2015) 认为眼动研究既可以遵循第二语言习得（SLA）的研究范式关注信息处理和语言习得的过程，也是采取从文化社会观的角度来分析理解人机互动的过程和行为，同样也可以按照行动研究法（Action Research）的范式去关注具体教学法的效果和帮助参与者积极反思。在开始眼动研究之前，需要对眼动设备有一定的了解，但更重要的是考虑眼动研究如何会帮助回答研究者所提出的问题。

　　目前，眼动仪的种类很多。以下是一些使用最广的眼动仪器品牌：Tobii，SMI, EyeLink，ISCAN，LC Technologies，EyeTech，The Eye Tribe，Ergoneers, 等。

这些常见的眼动仪可以分成屏幕式(如 EyeLink 1000 Plus)，眼镜式（如 SMI），便携式（如 Tobii X3）。随着科技的发展，眼动仪也可以跟 VR（虚拟现实）技术相联系结合。

屏幕式眼动仪适合实验室的环境。它的外观跟普通台式电脑无异，但其内置的摄像头可以捕捉屏幕前被测试者的眼动轨迹，并即时通过内置的分析软件记录保存数据。这类眼动仪 适合实验室环境的测试研究，常常用来调查网页或网上课堂的设计和学习效果和师生互动等。一般要求被测人员端坐在屏幕前，尽量不移动头部。由于测试环境可控度较高，眼动数据的准确度也相应更高。有些屏幕式的眼动仪器还配有支架，可以确保被试者的头部问题，更进一步地提高眼动数据的精度。

眼镜式眼动仪适用于研究真实环境中的人类注意力和电子工具的使用习惯 。比如，Facebook 曾使用该设备调查人们每天在家中观看电视的准确信息以及在观看过程中“第二屏”（手机、平板等）对电视观看体验的影响。因为被测试人员可以自由行走，移动，这类眼动仪也被频繁使用对公共场所的标示设计，广告布置，超市的货品陈列等。

便携式眼动仪小巧轻便，可以跟不同型号的电脑，手机或平板电脑连结， 被测的场地也不只局限于实验室的环境。适合多媒体学习资料的研究（如阅读、信息搜索、浏览图像），也可以用以对比课堂教学和网络教学的活动过程等。

这三类眼动仪器中屏幕式的采样率 （sampling rate ）最高，可以达到 1000 - 2000Hz。而便携式的采样率一般为 60-120Hz 之间, 眼镜式眼动仪通常是 50 -120Hz。采样率是指眼动仪每秒采集眼球图像的次数。一台采样率为 60Hz 的眼动仪，其采样间隔为 16.67 毫秒 (ms)，如果采样率为 120Hz，采样率间隔就缩短至 8.33ms。如果眼动仪采样率为 1200Hz，那么采样间隔则分别降低到 0.83ms。也就是说，采样率高的眼动仪能更快速准确地捕捉眼动的细微差别，但同时也会带来一定的数据噪音。

由于精确度和准确度的差别以及配套软件的不同，眼动仪的价格差别很大。需要研究者自己编写分析程序的 The Eye Tribe 的价格不超过 100 美金，而采样率达到 1000Hz 并自带数据分析系统的高端眼动仪则需要 2-3 万美金的预算。

至于选择什么样的眼动仪，要根据具体的研究目的和设计来决定。如果只是通过眼动研究来发现眼动的基本规律，学生在学习过程中大致的注视范围和特点，或者是混合型研究（mixed methods research)中的一部分，低采样率的眼动仪也可以胜任。但需在论文或报告中注明眼动设备的型号。而针对阅读的研究研究对眼动仪的采样率要求更高。

另一个需要考虑的是采用什么样的眼动研究方式，是纯量化的眼动指标数据的分析和模型建立；还是将眼动研究数据同其他研究方法相结合，不但去了解什么时间发生了什么，并探索其背后的原因？类似的问题实则为研究范式的选择。从本质

上说是如何看待研究对象，把他们看作提供信息的被测者还是研究本身的一部分，这也是个值得考虑的问题。

在实验操作方面需要准备和安排。应当提前告知研究对象需要的注意事项和研究的具体步骤。比如，在正式的眼动测试之前，被测者需要端坐在屏幕前，通过眼动软件进行校准（calibration），有些眼镜镜片或眼部的装饰物件可能会影响校准度。在眼动测试中，被测者应该尽量保持头部稳定，不要大幅度晃动；坐姿的改变也有可能影响眼动数据的捕捉。

如果需要用到词典或 worksheet 等辅助资料，最好事先都存为电子版本， 方便通过屏幕获取。因为如果使用的是屏幕式或便携式的眼动仪，被测者的目光一旦离开屏幕，这段时间内的眼动数据会缺失，给后期的数据分析带来挑战。也可以考虑使用额外的摄像头来记录整个实验过程。

此外，一台眼动仪一次只能测试一个被测者，想要通过眼动来研究网络课堂互动的学者需要在研究的设计安排上要需要提前考虑到时间，内容和人数的安排。

眼动研究会生成大量的数据，除了热点图，眼动轨迹图等视觉类数据外，还可以生成大量的数字指标。这也意味着数据的分析变得更有挑战，需要花较多时间去筛选哪些数据能够更好地回答研究问题。首次尝试开展眼动研究的学者，需要一段时间去熟悉特定的眼动仪和配套的分析软件。

6. 结语

科技突飞猛进，计算机辅助汉语教学已是常态。实时网络语言教学也正快速成为重要的发展形式 ，而不再只是少数人的学习方式。随着"汉语热"的持续升温,在未来必将有更多的学习者会乐于通过网上实时学习的平台和软件来学习汉语，来认识了解大中华地区的社会经济动向。眼动研究可以为汉语教学者提升教学质量和探索语言学习的本质提供多种可能性。而这个领域高质量的研究不但有利于汉语教学的研究和发展，也将为全球语言教育的研究和实践贡献新知。

参考文献

Bai, X., Yan, G. , Liversedge, S. , Zang, C. , & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology:Human Perception and Performance*, *34*, 1277-1287.

Bai, X. Li, X. & Yan, G. (2015). Eye Movement Control in Chinese Reading: A Summary over the Past 20 Years of Research. *Acta Meteorologica Sinica*, *31*(1): 85-91. [白学军, 李馨, 闫国利. (2015). 汉语阅读眼动控制: 20 年研究的总结. 心理发展与教育, *31*(1): 85-91.]

Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In R. Radach, J. Hyona, & H. Deubel (Eds.), *The mind's eye, cognitive and applied aspects of eye movement research* (pp. 573-605). Amsterdam: Elsiever Science.

Kang, H. (2014).　Understanding online reading through the eyes of first and second language readers: An exploratory study. *Computers & Education*, *73*, 1-8.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 441-480.

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., . . . Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review, 10*, 90-115. https://doi.org/10.1016/j.edurev.2013.10.001.

Liversedge, S. P., Hyönä, J., & Rayner, K. (2013). Eye movements during Chinese reading. *Journal of Research in Reading, 36*(Suppl 1), S1-S3.

Miles, R.,　& Shen, E. (1925) . Photographic recording of eye movements in the reaidng of Chinese in vertical and horizontal axes: Method and preliminary results. *Journal of Experimental Psychology*, *8*(5), 344-362.

Meng, H., Bai, X., Yan, G., & Yao, Hai. (2014). Number of strokes influences initial landing positions during Chinese reading. *Psychological Science, 37*(4), 809-815. [孟红霞, 白学军, 闫国利, 姚海娟. (2014). 汉字笔画数对注视位置效应的影响. *心理科学*, 37(4), 809-815.]

Nielsen, J., & Pernice, K. (2010). *Eyetracking web usability*. Berkeley: New Riders.

O'Rourke, B., Prendergast,　C., Shi, L., Smith, B., & Stickler, U. (2015) Eyetracking in CALL – present and future. In A. M. Gimeno-Sanz, M. Levy, F. Blin & D. Barr (Eds.), *WorldCALL: Sustainability and computer-assisted language learning* (pp. 285-298). London: Bloomsbury Publishing.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research, *Encyclopedia of human computer interaction*, 211-219.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin, 124*(3), 372.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The quarterly journal of experimental psychology, 62*(8), 1457-1506.

Shen, D., Liversedge, S. P., Tian, J., Zang, C., Cui, L., Bai, X., . . . Rayner K. (2012). Eye movements of second language learners when reading spaced and unspaced Chinese text. *Journal of Experimental Psychology: Applied*, *18*(2), 192-202.

Shi, L., Stickler, U., & Lloyd, M. E. (2017). Language Learning in higher education. *Journal of the European Confederation of Language Centres in Higher Education (CercleS)*, *7*(1) 205-238

Stickler, U., & Shi, L. (2015). Eye movements of online Chinese learners. *CALICO, 32*(1), 52-81.

Stickler, U., & Shi, L. (2017). Eyetracking methodology in SCMC: a tool for empowering learning and teaching. *ReCALL*, *29* (2). 160-177.

Zang, C. , Liang, F. , Bai, X. , Yan, G. , & Liversedge, S. P. (2013). Interword spacing and landing position effects during Chinese reading in children and adults. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 720-734.