科技与中文教学

Journal of Technology and Chinese Language Teaching

Volume 7 Number 1, June 2016
二〇一六年六月　第七卷第一期
ISSN: 1949-260X
http://www.tclt.us/journal

科技与中文教学

Journal of Technology and Chinese Language Teaching

A peer-reviewed online publication with in-print supplement
ISSN: 1949-260Xhttp://www.tclt.us/journal

Volume 7 Number 1, June 2016

Articles

Columns

科技与中文教学

Journal of Technology and Chinese Language Teaching

A peer-reviewed online publication with in-print supplement
ISSN: 1949-260X http://www.tclt.us/journal

Sponsors

Department of East Asian Languages and Literatures, Hamilton College
Department of Foreign Languages and Literatures, Middle Tennessee State University
China Social Sciences Press (中国社会科学出版社)

Contacts

URL: http://www.tclt.us/journal
Email: editor@tclt.us

Big Content in an Educational Engineering Approach (教育学工程方法视角中的大内容制作)

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

Abstract: Language teachers find themselves under increasing pressure to use technology, and to adapt their pedagogical approach accordingly. The first question they are confronted with is: which technology to use, why, how, when and where, and which justification to provide for this choice. A second, more intricate, question is: which content to use, why, how, when and where, and which justification to provide for this choice. This article provides a report on our research on content for language learning in the last thirty years. After defining the concept of content in a more ontological way, we provide an overview of all data and content types that have become available recently. We present educational engineering as a method for making a justifiable choice, but at the same time we point out a number of issues associated with certain content types. The development of content for interactive language courseware is extremely labor-intensive, and it lacks reusability. We explain our attempts to work out a generic model for content structuring, and justify why Open Data seems to become a promising alternative at this point.

Keywords: Content for language learning, educational engineering, Open Data, Big Data

1. Introduction

When we look at the history of education, we can see that content has not really been an issue for many centuries. The classic example in Europe would be Aristotle, who was a *peripatetic* lecturer. He walked about as he taught under the colonnades of the Lyceum in Athens. Oral transfer of knowledge, ambulatory education, combined with illustrated texts on papyrus or parchment, later complemented by bas relief sculptures, stained glass and paintings have been the traditional educational media for centuries. Blackboards and wall maps were typical nineteenth century educational artifacts. The textbook as we know it dates from the early twentieth century, and the sixties saw the appearance of radio, television, tape recorder and the stencil machine in the classroom.

On other continents like Asia, this evolution has been somewhat different, but since the end of the twentieth century, and especially since the emergence of the Internet, a broad revolution has taken place in education worldwide, transforming learning and teaching fundamentally. These changes entail significant challenges for learners,

teachers, content providers (authors), publishers, researchers and policy makers. In this article we want to show why content is currently a challenging issue for language teachers worldwide, and we will explore a strategy based on our own thirty years of experience in the field. This overview entails a certain amount of self-referencing we hope the reader will forgive us.

2. Data, information and content

Language teachers find themselves under increasing pressure to use technology. This pressure is reflected in the pervasive but persuasive terminology they are confronted with: *blended learning*, *digital pedagogy*, *flipped classrooms*, *digital natives*, *21st century skills*, *virtual learning environments* and *serious games*. Each of these terms has been coined some day by a scholar who wanted to give a name to a largely unknown phenomenon, but we all started to use these terms with our own acceptation and purpose in mind. The latest cry is *Big Data*. What does this mean?

Big Data originally had a precise definition. It refers to data sets that are so large or complex that traditional data processing applications become inadequate, and that should be analyzed by powerful computers. The term also refers to the use of predictive analytics or certain other advanced methods to extract value from data in the form of patterns, trends, and associations, regarding human behavior and performance. However, more and more scholars in the field of language learning and teaching started to use the term to indicate the sudden emergence of enormous amounts of content sources which might be useful for language learning and teaching, accompanied by new phenomena such as the Semantic Web, Open Educational Resources (OERs), Massive Online Open Courses (MOOCs), and Learning Analytics.

Let us first come back to our terminology and refine our ontologies. We can define Data as tokens, characters, symbols stored on data carriers, and transmitted as messages that can become Information or Content. Shannon and Weaver (1963) were the first to define Information as a measurable concept, using a mathematical model. The amount of Information in a message is defined as the extent to which the uncertainty (or Entropy) is reduced by the message on the receiver's side. Communication is defined as the exchange of Information. Applied to education in general, and applied to educational technology more specifically, Information in a system reduces uncertainty on the learner's side by providing permanent and just-in-time Information on what to do and how to do it. In the other direction, the system becomes more certain about the learners' needs, achievements, level, attitude etc. by using data retrieved from his/her behavior.

But what is Content exactly in this respect? If we look at it in the context of instructional design and educational technology, specifically Computer Assisted Language Learning (CALL), we can define Content from a pragmatical point of view as Data which can be used in a meaningful way for language learning and teaching, and which can be expected to have an effect on learning. The most traditional content we know are the texts being used in textbooks. Series of characters, symbols, images that we try to remember, transcribe, fill in or translate as tasks or exercise types. This textbook

content has become more interactive in the digital age. The appearance of content can adapt itself to different circumstances, the teacher can edit it, and it can return feedback on the learner's actions.

So Data are series of raw tokens which can turn into Information when they reduce uncertainty, or into Content when they contribute to learning. The Medium (Colpaert et al., 2012) is the carrier of the data. Technology can be defined as any routine or tool based on some kind of specific knowledge to perform an operation or treatment on data on a specific medium.

3. The Big Data tsunami

So *Big Data* is the latest blurred ontology. The term is persuasive as it lets us believe that there is a revolution going on in the world of pedagogy, that we should radically change our behavior as (language) teachers, and that all data are possibly useful. In fact, how much of these data can become Information and how much can become Content?

It is true that the range of data sources has expanded considerably in the last decade. Big Data stands for the overwhelming availability of accessible data worldwide, developed for educational purposes or not, and fans out into several relevant phenomena. They all imply different roles for authors, teachers and learners.

We can classify as Information any data stored as Teaching Information, Learner Information and Research Information. *Teaching Information* groups all documents which are supposed to guide and support the teacher, such as educational programmes (e.g. Common European Framework and national standards), curricula descriptors, course objectives and teacher guidelines. *Learner Information* groups all data that contain Information about the learner, his characteristics, profile, background, preferences, level, achievements and evaluations. Portfolios or e-portfolios are being used by teachers and learners themselves to store all these data. Learning Analytics on the other hand stands for collecting data about learner behavior and performance within a system or environment with a view to improve its design, and indirectly, learner performance. Finally, *Research Information* constitutes a more comprehensive approach in collecting data, often triangulating quantitative and qualitative approaches. Research data are being made more accessible and are being published more and more together with research articles.

But data can also become content when implemented in the learning and teaching process with a view to have a significant effect on learning. This content fans out into a wide range of possibilities:

- Traditional textbooks by publishers (Decoo, 2010).
- Self-authored materials produced by the teacher for use with his/her own students only.

- Authentic documents found on the Web. The Semantic Web (Web 3.0) promotes common data formats and exchange protocols for authors to add meaning to content, to describe the structure of the knowledge about that content, and in so doing to offer promising possibilities for retrieving relevant and meaningful materials for teaching.
- Open Educational Resources are materials that are being shared, reused, improved and re-shared again. They are supposed to reduce workload for teachers and to increase learning effect considerably.
- Interactive materials, also called interactive language courseware or tutorial CALL, have been around since the early eighties (Colpaert & Decoo, 1999).
- Massive Online Open Courses are the most remarkable phenomenon since 2010. Platforms such as *Moodle* and *OpenLearning* allow teachers to create and use language courses accessible worldwide.
- Virtual Worlds (such as *Second Life*) and Serious Games (Cornillie et al., 2012).
- Ambient Intelligence (the adaptation of an electronic device to the presence of the learner) and Augmented Reality (the view of a real-world environment whose elements are augmented by sensory input such as sound, video, graphics or GPS data).
- The Internet of Things stands for real-world objects and artifacts which carry readable data that can be used as content in tasks.

The availability of huge amounts of data that can turn into Information or Content leads to anxiety and choice stress for teachers, next to the pressure they are already experiencing. And all the efforts made to help teachers with this choice, like this article, lead to even more data. The question is: how to make a justifiable choice ?

4. Educational engineering and distributed design

Neither technology nor pedagogy is starting points for designing learning environments (Colpaert, 2015). Nor is content. The specification of the required technology, pedagogical model and content should be the result of a methodological design process (Colpaert & Stockwell, 2016). We will briefly try to explain our approach in the following paragraphs.

Our starting point is that education will never be perfect. Education has always been *l'art du possible*, and this for four reasons. First, by its very nature, education can and will never be perfect. Because we are humans. Secondly, lack of time and resources often prevent us from duly implementing the required changes. Thirdly, any change, even the most justifiable one, entails some kind of resistance, often from stakeholders we misjudge. Last but not least, there is not enough knowledge available in terms of substantiated findings which would enable us to improve education, solve problems or design solutions in a systematic, methodological and justifiable way.

Engineering is 'the strategy for causing the best change in a poorly understood situation within the available resources' (Koen, 2003) or, in other words, the strategy to be used for devising the best possible real-world solutions when not enough knowledge is

available for doing so. It is a way of thinking in the first place, and does not necessarily imply technology.

Engineering is about building knowledge through real-world implementations, in a systematic and verifiable way, using working hypotheses that ate based on theory and practice and that should be empirically and theoretically validated. Hypothesis testing analyzes the effect of modified parameters, taking into account the specificity of the context. Engineering is about formulating and validating working hypotheses regarding the role, order, weight and intensity of these parameters.

Engineering is cyclic, iterative and probabilistic: Engineering seldom leads to proven facts in one project, but it often requires several iterations due to resistance, financial limitations, technological challenges or practical constraints in order to observe significant changes in the effect of the parameters in play (Bayesian epistemology).

In the same vein, *Educational Engineering* is about building the best possible educational artifacts. These educational artifacts can be documents, tools, content, concepts, models and solutions such as textbooks, syllabi, lesson plans, curricula, graded readers, exercises, tests, applications or electronic learning platforms.

Our research in educational engineering focuses on the theoretical and empirical validation of the following hypotheses, which we have grouped under the term *Distributed Design*, referring to the idea that the design process should take into account as many actors and factors as possible.

The four paradigm shifts stand for a radically new way of thinking about ICT in education:

- The *Ecological Paradigm Shift*: No technology has an inherent, measurable and generalizable effect on learning. Only the entire learning environment, seen as an ecology of interacting components, can have this effect.
- The *Process-Oriented Paradigm Shift*: The targeted effect of a learning environment does not depend on product features, but is proportional to its *designedness*. Designedness stands for the extent to which the learning environment has been designed in a methodological and justifiable way. This methodological approach is universally applicable, but leads to polymorphous results.
- The *Psychological Paradigm Shift*: In cases of problematic or lesser motivation, we tend to insist more on our pedagogical goals. This appears to be counterproductive: it is better to focus on personal goals first. Personal goals are defined here as subconscious volitions which hinder or stimulate acceptance and willingness to engage in the learning process. The problem with personal goals is that they are difficult to elicit (Colpaert, 2010).
- The *Demand-driven Paradigm Shift*: Neither technology nor pedagogy is appropriate starting points for design (Colpaert, 2015). The methodological design process creates a need, a strong demand for theoretical knowledge, content and technology.

Distributed Design, our Educational Engineering model based on these four paradigm shifts, can be considered an Instructional Design model of the ADDIE type for guiding the Analysis, Design, Development, Implementation and Evaluation of educational artifacts for learning, testing and teaching.

The Analysis stage is all about understanding the problem in its context. Its output is an accurate description of which aspects can, cannot and/or should change. The Design stage has three substages: Conceptualization, Specification and Prototyping. The goal of the Conceptualization stage is to detect and elicit subconscious goals, to identify the points where these personal goals conflict with the set pedagogical goals and to formulate a hypothesis about the best possible way to find a compromise between these conflicting goals. The output of Conceptualization is the formulation of the expected outcome, which will be compared with the actual outcome during the Evaluation stage.

The Specification stage is nothing more than a detailed (ontological) description of what is needed to realize the formulated construct in terms of pedagogical models (for teaching, learning and evaluation), content, technology and infrastructure.

The purpose of the Distributed Design model is to enable teachers to decide for themselves which pedagogical approach, content and technology to use, when, where, how and why.

5. Big Data in an educational engineering approach

The specification of content in the projects where we have applied the model, has led to a series of interesting observations.

Regarding Information, learners need Information on what to do, on their degrees of freedom, on how to bridge the gap in a TBLT approach, on how to reach the next level in a constructivist approach. They need Just-in-Time Information and Supportive Information in a 4C/ID approach. They want to know where they stand and how well they are doing.

Teachers rely on Learner Analytics and Portfolios in order to keep track of their learners' progress. Learning Analytics can reveal useful Information for diagnostics, remediation, and redesign of the learning environment, which was the main theme of the European project VITAL (http://www.project-vital.eu/en/). In an educational engineering approach, Learning Analytics should in the first place be geared towards validating the design hypothesis by comparing the actual outcome with the expected outcome. Data from Learning Analytics can become Information on how our design process can be improved.

Content is an important part in a Task-based Language Teaching (TBLT) approach. However, in order to make tasks really effective ('optimal' or 'activating'), our research has pointed out that the following aspects should not be neglected:

- Task design is a process (Colpaert et al., 2015). Teachers should follow steps in analyzing the context and specify the best possible tasks in order to realize the set pedagogical goals in the best possible way. Now, tasks 'fall from the sky' too frequently without any methodological approach.
- Any tasks are a hypothesis and should be formulated as such. Task validation consists in comparing the expected outcome with the actual outcome ("I expect my students to invest more time and energy in this paper if their paper will be reused later by other students"). Students even seem to become more interested in a task when teachers actually let them know that it is a hypothesis.
- There is no such thing as a good task or a bad task. Task effectiveness depends on the context. A simple drill-and-practice focus on form exercise may be useful in one context, and totally not in another. It is important that in task design, the specification phase makes the teacher choose from a wide range of skill types (21st century skills, Higher-order Thinking Skills, SAMR model, Digital Bloom Taxonomy).
- Tasks should create acceptance and willingness in the learners' mind. Self-Determination Theory (Deci & Ryan, 2002), Dörnyei's L2 SELF model (Dörnyei & Ushioda, 2009) or the author's Personal Goal Theory (Colpaert, 2010) help in explaining why tasks should be meaningful (what is in there for me?) and useful (what does it mean for others?).

What seems to work particularly well in many contexts is co-construction of knowledge. Our students write the content of our Instructional Design course themselves, adding every year another layer and another focus. Also five-minute knowledge clips as tasks are quite effective in this respect, both on the level of secondary education ("Explain the battle of Hastings", "Who was Confucius?", "Explain uniformly accelerated motion") and on the level of higher education and teacher training (as part of teaching materials they can upload in any digital learning environment).

The already mentioned Open Educational Resources looked very promising, but their success seems to be hampered (Colpaert, 2012) by psychological ('what will others say about my content?'), technological ('what should I use to share my content?'), epistemological ('what does Open exactly mean?') and juridical ('are Creative Commons enough?') challenges.

MOOCs do not seem to break through either, at least not to the extent we had expected. As explained in Colpaert (2014), there is a conceptual problem: MOOCs are not really massive, not really open and they cannot be considered real courses. Their highest potential lies in reaching and bringing together smaller groups of isolated learners on a specific specialized topic.

Learning from all this, the main challenges for the future seem to be personalization (adaptation of difficulty level, task type, … to the learner) and contextualization (adaption of content to the geotemporal location of the learner). In this respect, the role of interactive language courseware becomes more interesting again as long as it is situated in a wider array of content types. But what exactly is the problem with interactive language courseware?

6. Towards a generic structure for interactive content

Since the 1997 CALICO conference with its theme 'Content! Content! Content!', the issue of sustainability, exchangeability and reusability of content has not been out of the public discussion. The emergence of interactive language courseware (Colpaert, 2004), also called Tutorial CALL, made the problem become apparent: due to the complex functionalities needed for the required interactivity, content became less reusable and got lost. The authoring of this content for interactive language courseware was very labor-intensive, hence expensive. This is why, as explained in more detail in Colpaert 2013, we focused our research more and more on how to make content more generic in a first phase.

By using the term learning content, we do not only refer to traditional textbooks, but also to materials such as syllabi and handouts, interactive exercises in applications like *Hot Potatoes* or *QuestionMark*, course content in electronic learning environments like *Blackboard*, video and sound clips (podcasts), presentation slides in *PowerPoint* or *Prezi*, materials for Interactive Whiteboards, web pages, wikis and e-reader content. Writing learning content does not happen in a linear way, but it involves an arduous cyclic process of creating, editing, combining, structuring and formatting materials in several layers. Content should comply with many pedagogical, linguistic and cultural requirements such as to be linguistically correct, adapted to a specific level and context, engaging and attractive, as interactive and relevant as possible and to be politically correct by avoiding any statements or images which could insult or irritate individuals or minorities.

On the other hand, it is not easy to retrieve, select, evaluate and integrate materials developed by others, due to the fact that they are protected by copyright, not accessible or difficult to copy-paste. These materials contain text, images, sound and video, all with or without some levels of tagging, metadata or interactive functionality. Especially in the case of language learning, this functionality can become very complex (Colpaert, 2006). The more 'enriched' or interactive these materials, the higher the cost.

Moreover, language learning content should continuously be updated, adapted, rearranged and rechecked at every change in the learning context. These changes can be due to a curriculum change, a new pedagogical approach such as the 4CD/ID model (Van Merriënboer & Kirschner, 2013) or the Dynamic Systems Approach (Ellis & Larsen-Freeman, 2009), the integration of a new technology such as tablets or Interactive Whiteboards (Van Laer, Beauchamp, & Colpaert, 2012), and to changing learning styles, attitudes and preferences. But existing learning content is not easy to change. This is mainly due to the fact that most learning content has been created in a dedicated format: it is determined by the medium or the technology of the educational artifact as product.

Learning content gets lost far too quickly due to this inability to adapt, to be reused, exported, transferred. In other words, due to its lack of sustainability. In order to remedy this problem, learning content should become more *sustainable*. We define sustainable in this context as the sum of four properties: generic, reusable, interactive and open.

- *Generic*: Content should be authored, structured and accessed independently from any concrete device or medium and should be stored in a separate database. Its structure should not be influenced by any product as possible output.
- *Reusable*: Learning content should be made as transferable or exportable as possible to a wide variety of media, technologies and carriers, such as traditional hard copy textbooks, digital customized printed material on demand, mobile app exercises and materials for Interactive Whiteboard use.
- *Interactive*: Learning content can be 'flat' text, audio or video. There are however several possibilities for offering more information (e.g. enriched materials by semantic tagging afford more accurate selection of suitable learning materials) or more functionality (e.g. interactive exercises containing exercise types, answer possibilities, feedback scenarios, error analysis and remediation, reporting and logging).
- *Open*: Learning content should be as accessible, open and authorable as possible for allowing easier co-construction, updating and adaption.

We have been developing since 1986 a long series of applications and project tools (Colpaert & Decoo, 1999), representing a total of more than 150 man-years in projects for universities, governments, institutions, industrial companies, publishers and Europe. Initially the content of these developed programs was stored in a specific dedicated database structure (every application had its own structure). In 1997 we developed a new platform in Windows that focused on two requirements: generate a wide variety of applications with the same source code, and have the content in a separate database. In fact, there were two databases: the first contained all information for the application to run (identity and appearance, menu systems, behavior and interaction) while the second contained the learning content. Both databases were based on different object models, but they were both open, readable and updatable, at least for authorized people. The learning content was stored and structured in a relational Access database. The advantage of this approach was not only the strong integration with Visual Basic programming, but also the fact that authors could easily make their own interfaces as forms, based on queries, and reports. Some authors even succeeded in writing their own error-checking routines in VBA (Visual Basic for Applications).

In the *Eventail/Arcades Interactive Textbooks* project (with Wilfried DECOO), a longer-term project with publisher Van In, we initially converted language textbooks into interactive applications. Gradually, we started first structuring a database of learning content so that CD-ROM and textbook could be generated at more or less the same time as different output products. We gradually applied this approach in projects where possible: the BIS Online project for the Flemish department of Education, SELOR language tests for the Belgian Civil Service Commission, and a series of European projects. Finally, we ended up working more with the same object models behind the database structure of several different programs. These object models were not technical, but reflected a reasoning for opening, reading, editing and updating a specific database.

In 2004, we developed an object model that complied with all possible requirements and defended this research as doctoral dissertation (Colpaert, 2004; Colpaert, 2006).

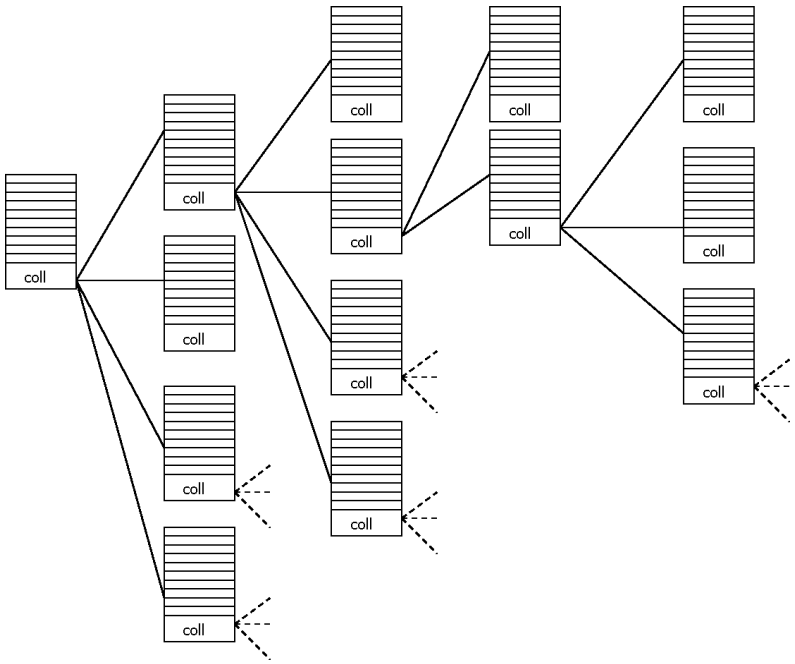


Figure 1. Generic Object Model

While generic structuring (for instance if carried out in a relational database) may appear fairly readable and authorable, surrounding factors in a normal working environment can make things quite complex: large data, many co-authors, complex functionality, author support (queries, forms and reports for content analysis), error checking and prevention (quality control), integration and reuse of existing materials, and generation of content into a wide range of products and services. In order for the object model to be implemented and to lead to significant and sustainable results, we wanted to develop an authoring interface, or at least define an ontology. The adopted methodology was a mixed-method approach consisting of a theoretical, an empirical and an engineering component.

Our research (Colpaert & Cornillie, 2008) has shown that language learning content can and should be structured and stored in a sustainable way, defined as generic, reusable, authorable (open) and allowing interactivity of scalable complexity. Language learning content can be structured as a collection of a collection of a collection of items, with as many collection levels as deemed necessary: from a plain text (only one level with one collection) to complex task-based scenarios. In this collection structure, it is possible to use the same, simple, object model, consisting of properties and methods which govern appearance and behavior of the object content. The proposed structure appears to work well in client/service environments such as web-based applications or mobile apps, but also for generating a wide range of products and services, also non-digital.

But our research has also shown that it remains quite a challenge, if not impossible from a practical point of view, to have this sustainable content structuring implemented in the real life of teachers, authors and publishers. The transition from the actual publishing tools to a generic authoring interface is an enormous undertaking for publishers worldwide, which in the best case would take five to ten years to realize. Teachers and authors need to be informed about the best way to structure their data in a sustainable way. The technological side of the interface was not really the problem, but the psychological side. It will take several other research projects to convince authors of an interface perceived as useful and easy to use at the same time.

7. A special case: Open Data

Open Data is a recent phenomenon which deserves our attention. The adjective 'Open' in Open Data has a slightly different connotation than 'Open' in Open Educational Resources, Open Access (for research articles), Massive Open Online Courses and Open Source. Open Data, according to the open definition (http://opendefinition.org), is any kind of data in any kind of format, that can be accessed, used, modified and shared, by anyone for any purpose. An open license on the container of the data is the only requirement to comply with this definition. More interesting is the goal behind Open Data, which could be interpreted as maximizing the reuse of a dataset. It may be in the best interest of a company that its datasets are being used within other app(lication)s as much as possible. In the same vein, one could consider that it is in a publisher's interest to have learning content reused in as many ways as possible.

While many data sources are already openly licensed (Bizer et al., 2009), as already explained in this article, there are still many hurdles to overcome when trying to publish data for maximized reuse (Colpaert, 2014). Enhancing interoperability with all other published datasets would increase their reuse. Interoperability denotes the ability of diverse systems and organizations to work together (inter-operate). In this case, it is the ability to interoperate - or intermix - different datasets. Interoperability is important because it allows for different components to work together. Measuring this interoperability is difficult (Colpaert et al., 2014), yet dividing the concept into properties may help: technical interoperability, syntactic interoperability, semantic interoperability and querying interoperability. Many of these interoperability problems are being tackled by standardization, yet in a constantly evolving world, standards often lack behind, especially in education.

Open Data, in a simplified view, provides an interface for reading the structure and content of the dataset before actually accessing it. Linked Open Data maximizes the effect of Open Data by linking together large series of datasets. The connection of these datasets to the Semantic Web creates an even bigger enrichment.

The advantages are obvious in science and for government, but the potential for education is huge. Datasets can be made accessible without having to restructure them first. On the level of Information, as defined earlier in this article, datasets about research, education, schools, students, teachers and even parents. On the level of Content: any

existing material could be made reusable again. Provided we add a layer (API or other) which explains how to read the data. In theory, all these interfaces can be different, but the evolution towards some kind of standardization in terms of learning objects or student records may seem attractive and daunting at the same time.

We are currently analyzing to what extent the previously developed Generic Object Model can be reused as an interface object model for any kind of learning content. We are aware that new issues and problems will arise, but the fact that so much lost content may be reactivated again gives us a lot of courage to continue in this endeavor. Open Data can lead to Big Learning Content. There is a promising compatibility with recent paradigms in TBLT (Colpaert et al., 2015), Complex Dynamic Systems (Ellis & Larsen-Freeman, 2009), personalization and contextualization of the learning process, skills (SAMR, Bloom digital taxonomy, 21st century skills, Higher Order Thinking skills…).

8. Conclusions and recommendations

Teachers do not only find themselves under increasing pressure to use technology – and to adopt a new pedagogy – digital or not – , they are also confronted with a tsunami of data, which we have called Big Data. These data belong to two important categories: Information and Content. Teachers need Information to guide and support their learners and to organize their learning environment in the best possible way. Learners need information to know where they stand, where they can go and what their degrees of freedom are. Learning Content on the other hand is fanning out in a wide variety of types, so that teachers are – again – confronted with choice stress. How to decide which content to use, where, how, why, when?

Content can only become useful if it is needed based on a specification of an optimal learning environment. This specification remains a hypothesis within what we have called an educational engineering approach. This angle of attack makes us look in a different way at content. In a more critical way to OERS and MOOCs, but also in a more worried way to interactive language courseware: it is labor-intensive due to the required linguistic-didactic functionalities, and it gets lost all the time due to its non-generic structure.

Open Data, in the same spirit as Open Source, Open Access, Open Educational Resources, but with a slightly different acceptation for ‘Open’, seems like a promising phenomenon to solve this problem. Publishers and authors can continue to produce learning content the way they were, but they will gradually add interfaces to the extent that they want their content to be shared and reused again.

Researchers and developers in CALL will find in Open Data a promising direction which is very compatible with current approaches such as TBLT, personalization and contextualization of the language learning process, Complex Dynamic Systems and digital skills. Teacher and learner roles will change considerably, especially regarding

collaboration in activities such as telecollaboration, course co-construction or knowledge clips.

Again, the most appropriate pedagogical model, technology and content are not starting points for design, but they are the result of good design, and they will always depend on the context. And in that context, the needs, goals and volitions of learners and teachers are important. Like in neuromarketing, the subconscious volitions are more and more considered as decisive starting points for good design. So how can we 'Open' up this source of 'Information'?

References

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data-the story so far. *Semantic Services, Interoperability and Web Applications: Emerging Concepts*, 205-227.

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

Colpaert, J. (2006). Toward an ontological approach in goal-oriented language courseware design and its implications for technology-independent content structuring. *Computer Assisted Language Learning*, *19*, 109-127.

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). Open educational resources for language teachers: A goal-oriented approach. In *Proceedings of the joint CMC and teacher education EuroCall SIGS workshop Bologna*, 29-30 March 2012. 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. Wong, 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.* Mouton de Gruyter (Open Series), 162-173.

Colpaert, J. (2015, April 8). Neither technology nor pedagogy should come first in design. *The Educationist*. Retrieved from https://www.openlearning.com/educationist/NorTechnologyNorPedagogyShouldComeFirstInDesign0.

Colpaert, J., & Cornillie, F. (2008). Between dedicated and generic requirements: The Entreposage Universel Model as a grammar of interactive content. In A. Gimeno, & R. Seiz Ortiz (Eds.), *Aprendizaje de lenguas asistido por ordenador: herramientas de autor para el desarrollo de cursos a través de la Web* (pp.57-92). Valencia, Spain: Editorial Universidad Politécnica de Valencia.

Colpaert, J., & Stockwell, G. (in press). *Technology and the Language Learning Environment*.

Colpaert, J., & Decoo, W. (1999). The role of didactic functions in CALL design. In K. C. Cameron (Ed.), *CALL & the learning community* (pp. 65-74). Exeter, UK: Elm Bank Publications.

Colpaert, J., Aerts, A., Oberhofer, M., Gutiérrez-Colón Plana, M. (Eds.). (2015). *Task design and CALL. Proceedings of the Seventeenth International CALL Research Conference*. Antwerp, Belgium: University of Antwerp.

Colpaert, J., Aerts, A., Wu, W. and Chao, Y. (Eds.). (2012). *The Medium matters: Proceedings of the Fifteenth International CALL Conference*. Antwerp, Belgium: University of Antwerp.

Colpaert, P. (2014). Publishing data for maximized reuse. In *11th Extended Semantic Web Conference*. Retrieved from http://ceur-ws.org/Vol-1165/salad2014-1.pdf

Colpaert, P., Van Compernolle, M., De Vocht, L., Dimou, A., Vander Sande, M., Verborgh, R., & Mannens, E. (2014). Quantifying the interoperability of open government datasets. *Computer*, *10*, 50-56.

Cornillie, F., Clarebout, G., & Desmet, P. (2012). Between learning and playing? Exploring learners' perceptions of corrective feedback in an immersive game for English pragmatics. ReCALL, 24(3), 257–278. doi:10.1017/S0958344012000146.

Deci, E., & Ryan, R. (Eds.). (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

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

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

Ellis, N.C., & Larsen-Freeman, D. (Eds.). (2009). *Language as a complex adaptive system*. Oxford, UK: Wiley-Blackwell.

Koen, B. V. (2003). Discussion of the method: Conducting the engineer's approach to problem solving. NY: Oxford University Press.

Van Laer, S., Beauchamp, G., & Colpaert, J. (2012). Teacher use of the interactive whiteboards in Flemish secondary education—mapping against a transition framework. *Education and Information Technologies*, 1-15.

Van Merrienboer, J. J. G., & Kirschner, P. A. (2013). *Ten steps to complex learning (Second Revised Edition)*. New York, US: Routledge.

Weaver, W., & Shannon, C. (1963). *The mathematical theory of communication*. Champaign, IL: University of Illinois Press.

An Investigation of the Design of a Four-stage Flipped Classroom in Mandarin Chinese (中文翻转教室四个阶段的设计与研究)

Tseng, Miao-fen
(曾妙芬)
University of Virginia
(弗吉尼亚大学)
mt3z@eservices.virginia.edu

Broadstock, Maan
(姜满)
Wright State University
(莱特州立大学)
mjbroadstock@gmail.com

Chen, Henny
(陈妲良)
Moreau Catholic High School
(莫若高中)
hchen@moreaucatholic.org

Abstract: This paper aims to investigate the efficacy of a flipped classroom that was created for twenty novice learners in an intensive summer Chinese language program. The flipped classroom featured four-stage learning: watching videos, taking notes online while watching videos, responding to preview questions, and receiving instant feedback. Three-factor ANOVA, one-way ANOVA, and regression analyses were used to investigate how gender, level of Chinese, instructional topic, and note-taking affected students' learning. Learners also completed a questionnaire to elicit their reactions to the flipped classroom and feedback on the effect it had on their learning. Statistical analysis indicated that the scores differed significantly in instructional topics and that the number of notes taken significantly decreased from the beginning toward the end of the program. Additionally, a significant positive relationship was found between the number of notes taken and test scores. The results of the questionnaire at a 5-point Likert Scale showed that learners strongly recommended the four stages of flipped design despite a slight decrease in the second stage of learning through note-taking while watching videos. Limitations of the study and directions for future research are discussed at the end of the paper.

摘要：本研究旨在探讨暑期初级中文密集项目中翻转教室之有效性以及其对学习之影响。本实验设计之翻转教室分为四个学习阶段，依序为学生看教学视频，进而一边看教学视频一边在网上做笔记，再则回答理解问题，最后收到立即性反馈。本研究采用三因子变异数分析、单因子变异数分析、以及回归分析，以期探索性别，语言水平， 教学主题，网上笔记四个因素对学生学习的影响。除此之外，本研究亦针对翻转教室设计了一个问卷，以期了解学习者对翻转教室的反馈以及其对参与者学习的影响。统计分析显示，学生的学习分数在不同教学主题间有显著的差异，学生笔记次数的多寡自项目开始至终也明显递减。同时也发现，笔记次数多寡与学习分数成正相关。李克量表问卷结果显示，尽管翻转教室在第二个阶段一边看教学视频一边在网上做笔记的数值不如其它三个阶段高，大多数学习者仍然普遍支持并推

荐此翻转教室四阶段的设计以及其未来在中文教学上的应用。本文最后也提出此研究的限制与未来研究方向。

Keywords: Blended learning, flipped classroom, flipped learning, teaching Chinese as a foreign language

关键词: 混成学习，翻转教室，翻转学习，对外汉语教学

1. Introduction

Flipped learning, which originated from research in STEM fields, i.e., science, technology, engineering, and math (Berrett, 2012), has rapidly gained prominence and popularity in innovative higher education E-learning contexts. Disciplines such as physics (Deslauries, Schelew, & Wieman, 2011), engineering electronics and circuits (Biemiller, 2013), statistics (Strayer, 2007), and cinema and TV arts (Enfield, 2013) employ flipped learning. As the Internet continues to explode in the E-learning era, this model has taken hold in secondary education in language arts (Ullman, 2013; Moran, 2014), science (Bergman & Sams, 2012), and math (Fulton, 2013). This evolving pedagogical approach has been presented in several interchangeable terms: Inverted classroom (Lage, Platt, & Treglia, 2000), Just-in-time teaching (Novak, 2011), Flipped classroom (Bergmann & Sams, 2012a), or Inverted learning (Davis, 2013). Baker (2000), who required his communication students to read PowerPoint slides before class through an institutional new computer network, is accredited for conceptualizing the flipped classroom paradigm (Lage & Platt, 2000; Lage, Platt, & Treglia, 2000).

According to Graham's (2006) articulation of three levels of blends on a continuum of technology integration, flip learning enhances blends instead of enabling blends or transforming blends. It aims to provide "enhanced rather than equivalent experiences of traditional classroom teaching" (Hung, 2015) in terms of technology use in hybrid learning settings. It does not simply take place in "brick-and-mortar location" establishments (Staker & Horn, 2012) and is built on extant mobile technological tools that students frequently encounter. The acts of learning that have "traditionally taken place inside the classroom now take place outside the classroom and vice versa" (Lage, Platt, & Treglia, 2000, p. 32). In a flipped classroom, instructional order is reversed in a way that students typically view instructional videos or other types of materials before class and devote in-class time to applied learning activities and interactive higher-order thinking.

1.1 Advantages of flipped classrooms

Researchers have well-recognized the advantages that this option has brought into teaching and learning. Baker (2000) discovered that students had learned a great deal from their peers through collaborative activities. Fulton (2012) identified that instructors were able to allocate more time to individual students and implement the curriculum more effectively. Other positive effects from flipped classrooms include, but are not

limited to, creating inquiry and problem-based learning (Bergmann & Sams, 2012a), eliminating constant homework struggle (Strayer, 2007), and allowing students to catch up with class content in their absence (Herreid & Schiller, 2013). In short, educational proponents have faith in this paradigm shift that empowers them to reach "every student in every class every day" (Bergmann & Sams, 2012a, book title). Although teachers may feel unprepared to "wade into the digital muck of video creation" (Moran, 2014), this technologically reliant idea has unparalleled potential to revamp classroom environments and change how subjects are taught and learned (Bergmann & Sams, 2012a; Bergmann & Sams, 2012b; Fulton, 2012; Tucker, 2012). Its close linkage to theoretical understandings of active learning also "involves students in doing things and thinking about the things they are doing (Bonnell & Eison, 1991, p. 2).

1.2 Flipped classrooms in second language acquisition

Nearly all empirical research on flipped classrooms has been conducted in STEM classrooms. Despite increasing focused attention to flipped learning, surprisingly there have been only a very limited number of studies on flip learning models in second language acquisition. The following literature review offers a small glimpse of what is available in post-secondary, but not secondary, educational settings.

Hung (2015) conducted a posttest-only quasi-experiment to investigate EFL learners' flip learning experiences at a Taiwan university. The experimental design of the study divided students into a flip group with WebQuest learning, a semi-flip group featuring TED-ED viewing, and a non-flip group. The results confirmed the value of active learning and revealed that students in the flip and semi-flip classrooms attained more satisfactory learning outcomes, developed more positive learning attitudes, and devoted themselves more to the learning process than students in the non-flip group.

In his study on developing EFL writing skills through student-created digital videos, Engin (2014) made an attempt to transform passive "sit-and-get" learners into producers and consumers in an active peer learning process. Although student-generated materials and input tended to lower cognitive, linguistic, and cultural load that may be a barrier to language learning (Mercer, 2000), learners who were Arabic native speakers expressed their preference of a teacher's explanation over receiving a peer's input in the flip learning phase.

Mehring (2014) completed a qualitative study in an attempt to investigate how EFL learners who were native speakers of Japanese reacted to the flipped classroom. It was found that students greatly benefitted from "the greater student-centered and active learning environment, the added amount of time participants spent preparing for the face-to-face class, and the significantly enhanced authentic, communicative learning environment" (p. 86).

Hojnacki & Häusler-Gross (2014) compared language gains in four language skills through flip and non-flip sessions in elementary German in a US mid-west liberal arts college. They summarized that learning did take place in each of the two sessions.

Additionally, students who had flip experiences performed slightly better than those who did not in all language skills except for speaking. Despite the fact that no significant difference was reported in change between the two sessions because of small sample size, the survey showed that students who learned through the flipped classroom model were more motivated and held a positive attitude toward the flipped format and therefore recommended continued adoption of the flipped model in the future.

Salazar & Sinclair (2015) reported the findings of their empirical study on the development of lexical competence among elementary learners of Spanish at an American university. In their study, three types of class structures were under investigation: a non-flip class, a flipped class using Quizlet, and a flipped class using vocabulary video tutorials. The differences between the pre-post and post-test indicated that all three designs yielded observable gains that promoted differentiated instruction. Although a combination of Quizlet and videos was recommended, the authors reflected upon potential future improvements to add comprehension checks and a tactile and interactive platform for video tutorials.

The above limited studies on SLA coincidentally were completed at higher education institutions. Whether the findings and pedagogical implications of flip learning design work in foreign language classes at the secondary level remains unknown and deserves further exploration. This pioneering study demonstrates the very first attempt of its kind to investigate the design of an innovative flipped classroom model that involves four stages of learning in Mandarin Chinese among 9^{th}-12^{th} graders at the novice level.

2. Design of the four-stage flipped classroom in this study

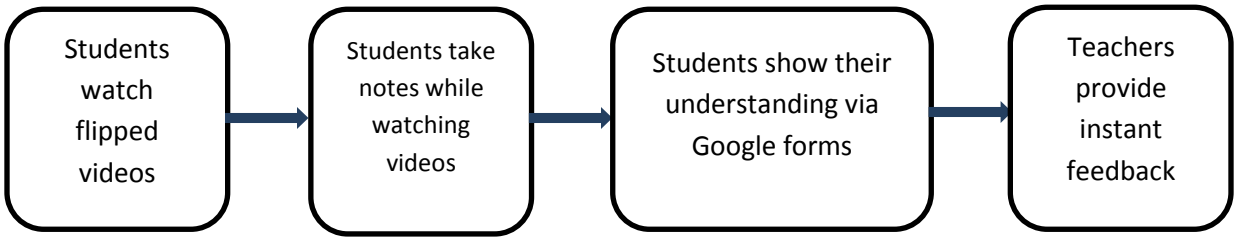
The creation and experiment of the flipped classroom began in the summer of 2014. The flipped classroom launched in the 2014 STARTALK program involved only two stages of learning before face-to-face instruction took place in class, i.e., watching videos and working on Google forms to check for understanding, the latter of which provided valuable insight for teachers to incorporate collective feedback into face-to-face instruction on the following day. It was then re-evaluated, modified, enriched, and expanded to a four-stage innovative model to constantly and systematically monitor learners' progress and ensure that flipped learning achieved its proximal effects before face-to-face learning began. Compared to the traditional flipped classroom model in which learners watch videos at home and then participate in activities in class, the flipped learning model specifically created for this program advanced its existing functionality by adding three critical components: note-taking, comprehension checks, and instant feedback, including scores, error corrections, and explanations for inaccurate items after students completed comprehension questions.

The four stages of the flipped classroom in the 2015 STARTALK program are as follows, described from learners' perspective:

Stage 1: Watching videos

Stage 2: Taking notes online while watching videos
Stage 3: Responding to preview questions
Stage 4: Receiving instant feedback

The four-stage flipped classroom model is highlighted below.



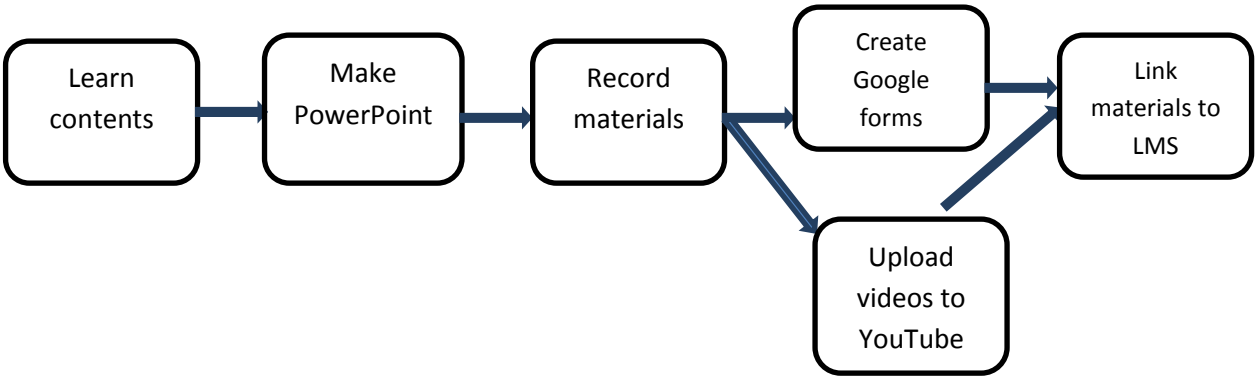
The four-stage model was meticulously laid out at both academic and technological levels. It was intended to be well-sequenced and structured, so each single stage was interconnected with its immediate next stage and the remaining three stages. It was collaboratively created by a technology specialist, four leading teachers of the STARTALK teacher and student programs, and a curriculum developer and key trainer. The pre-class flipped learning components were created to strongly enhance in-class face-to-face learning and support the fulfillment of curricular objectives as the ultimate goal. All materials were uploaded to the program website and well tested before the onset of the teacher and student program. The core instructional team worked tirelessly to ensure that the online operational system remained stable and well-run during the program. The twelve teacher trainees did not become involved in the creation of the flipped learning process. Upon their arrival at the program, the flipped learning online system was already completed and uploaded to the combination program website for the purpose of teacher development.

In total, twenty-two videos were pre-recorded for eight topics to be taught on eight days during the program. Six topics had three videos each, and two topics had two videos each. The videos were created in full alignment with curricular goals and program needs. Specifically, each video included essential learning materials, core sentence structures, and cultural references in a PowerPoint for screen recording. Each video lasted from one to less than three minutes. While a number of free screen recording tools were available online such as Ocam, QuickTime Player Pro, and Jing, the free downloadable version of "Screencast-O-Matic" was chosen to create program videos because of its accessibility on both the Macintosh and Windows operating systems.

After the videos were pre-recorded, daily Google forms and spreadsheets were then created to check for student comprehension through multiple-choice questions, ranging from six to 9 items for each topic. The Google forms and spreadsheets had the capacity to help educators collect, share, analyze, and graph data. A Google spreadsheet add-on was later installed to create a MCQ test, which automatically corrected and scored learners' responses submitted online and sent along explanations for wrong items via email. Compared to other web tools such as Flubaroo, Zaption and EduCanon, the combination of Google forms and the MCQ script is free of cost, and its instant e-mail

feedback and explanatory notes well concluded the four-stage flipped learning process and sufficed for the purposes of this study.

Afterwards, pre-recorded videos were uploaded by the technology specialist to YouTube under the unlisted category and "1know.net," an innovative online learning management system that enables instructors to collect data and monitor students' progress. The 1Know Learning Management System (LMS) recorded all of students' learning data, including the number of notes taken, video and unit learning duration, and video accessing records. The data trail allowed instructors to monitor the note-taking process and outcomes. Instructors were able to know when students took notes while watching the videos, what types of notes and how many notes students typed online, and whether the quantity of notes correlated with the score that students received through Google forms. To put it concisely, LMS generated a database through which instructors could monitor and analyze learners' daily work and progress, and, most importantly, best prepare for face-to-face instruction the next day. The preparation workflow for teachers is as follows.



When students logged into the system, they would be able to see the 8-day assigned online tasks in order, the YouTube videos, and the Google forms on the central 1know platform. They accessed online materials all-at-once at any location that was most convenient for them. An orientation was held on the first day of the program to help students become familiar with the online system and successfully register for the course via the 1Know platform. During this 1Know orientation session, students practiced viewing the videos and taking notes at the same time. Students were also encouraged to post their questions as notes and highlight the questions in red, so their teacher could read their questions more easily. Before attending the program, it was assumed that students had developed note-taking skills in their own high schools. Therefore, what to write for online notes was not the focus of the training session and was open for them to decide. They were advised to write whatever they felt necessary and helpful for their own self-paced learning.

2.1 Research questions and method

The emergence of flipped learning in the STEM area has an increasing impact on foreign language education. More and more foreign language educators have adopted the

concept to create flipped classrooms for students to preview lessons before class, so in-class time can be secured for interactive and engaging discussions. But most flipped classrooms were created without sound research to support the design and instructional strategies (Hamden et al., 2013). More research is therefore needed to inform us of what type of flipped design works and does not work pedagogically.

As indicated earlier, while a very limited number of flipped learning studies in SLA have been situated in post-secondary settings, there is little research to inform us of how flipped learning complements face-to-face instruction to strengthen the learning process and outcomes in K-12 foreign language classes. Unlike the design of most flipped classrooms, this study makes a breakthrough to go beyond one stage of flipped learning and creates a four-stage learning process in order to systematically consolidate and monitor pre-class online learning, which is expected to strengthen face-to-face instruction in teaching Mandarin Chinese as a foreign language to high schoolers in a summer intensive program implemented at an American university.

This study aims to investigate how gender, learners' proficiency level (hereafter level), instructional topics (hereafter day), and note-taking affected students' learning after students went through the four-stage learning process: watching videos, taking notes online while watching videos, responding to preview questions, and receiving instant feedback.

The research questions include the following:

1. While students were watching videos, did gender, level, or topic affect their note- taking behavior?
2. After students went through the four-stage flipped learning, did gender, level, topic, or note-taking affect students' preview question scores?
3. Did students who watched videos score higher than students who did not watch videos?

In addition to the above research questions, the study also intended to explore students' reactions to the four-stage flipped design and their feedback. A questionnaire was developed according to a 5-point Likert Scale to help us understand students' reactions to the four-stage flipped learning and their comments and recommendations for future applications. The questionnaire was intended to gather information about students' input on their motivation level, the effects of flipped learning on their learning outcomes, the friendliness of technology use, and their recommendations for future applications of the four-stage flipped learning model to Chinese language courses or programs (See Appendix I). Learners were also encouraged to share any of their additional comments on the design of the flipped learning model and overall reflective learning experiences.

2.2 An overview of the program

The 2015 program pilot-tested cutting-edge technology tools to create the flipped classroom and achieve proximal learning outcomes in an innovative blended learning

model. The daily routine of teachers' practicum closely followed a comprehensive cycle that was composed of pre-class preview, during-class learning, and after-class review to ensure closely monitored daily progress in alignment with learning objectives.

2.2.1 Student participants

Students who participated in the study learned Mandarin Chinese in the Virginia STARTALK Chinese Student Academy in summer 2015. The program evolved from a half-day Chinese Language and Culture Immersion Program as part of the Virginia STARTALK Chinese Teacher Academy that was launched in 2008 through the STARTALK initiative in the US. The Virginia STARTALK Chinese student Academy (VSCSA) partnered with the Virginia STARTALK Chinese teacher Academy (VSCTA) to increase capacity for learning Mandarin Chinese in K-16 educational settings. The teacher to student ratio was 1:1 in 2015, and students frequently interacted with devoted teachers to satisfy their individual needs on a daily basis. The program strives to create a truly interactive, communicative, and innovative class featuring an integration of a wide array of authentic tasks and cutting-edge instructional technology.

Participants were 20 non-heritage learners who were 9th-12th grade high schoolers recruited from the commonwealth of Virginia and neighboring states in the US. There were 13 female students and 7 male students in the program, including five 9th graders, seven 10th graders, four 11th graders, and four 12th graders. Of the twenty students, seven students had prior learning experiences. One had learned Chinese for three months in a weekend Chinese school, four had completed Level I Chinese, one had completed Level II Chinese, and one had completed a Virtual Chinese course for 3 years. According to the pre-program survey and individual interviews, they were placed into two classes depending on their prior learning experiences, learning styles, age, and gender.

Of the twenty students who participated, two were not able to access the 1Know website. Therefore, only eighteen students watched the videos and completed the online preview questions daily. Among these eighteen students, eleven were female and seven were male. Eleven were novice students, and seven had studied Chinese before. Since the students were not strictly required to complete the assignments, some students missed one or two videos or did not complete the online quizzes. Two students did not watch the videos except on Day One, when all students watched the video in class, but they completed the daily quizzes.

2.2.2 Curriculum

Twenty learners participated in the 10-day non-residential program from July 13th to July 23rd in 2015, excluding Sunday. The curriculum was created by a core instructional team consisting of four leading teachers of the teacher and student program, aided by a technology specialist's expertise in instructional technology and Chinese language instruction through 3-stages of backward design. Morning included three 50-minute slots when students learned through face-to-face instruction in regular classrooms. The target language was exclusively used in class, which allowed learners to gain

enriched immersion experiences. Afternoon sessions were reserved for students to complete technology-related tasks, review learned materials, and interact in cultural activities. English was mingled with Chinese in the afternoon sessions.

The program offered no credits to students. Total instructional time was 65 hours during the day program, and 10 hours for online flip learning and self-review at home. The targeted proficiency level was novice-mid according to ACTFL's proficiency guidelines, but performance level could reach up to novice-high on certain topics. The instructor exclusively used the target language in class, and encouraged learners to do likewise. Learners were enrolled in the virtual US-China Exchange Program to embark on an exciting trip to Beijing. The entire student program lasted for ten days but covered only eight topics on eight days, with the seventh day reserved for review and the last day for an end-of-program performance and closing ceremony.

Instructional topics included the following:

Day 1: Greetings and self-introduction
Day 2: Family
Day 3: Dates and time
Day 4: Hobbies
Day 5: Vising a Chinese family
Day 6: At a Chinese restaurant
Day 7: Review
Day 8: Sightseeing in Beijing
Day 9: Shopping
Day 10: End-of-program performance and closing ceremony

2.2.3 Teachers

Instructors were two leading teachers in the student program and 12 teacher trainees who were recruited nationwide to take a 3-credit graduate level course in teaching Chinese as a foreign language during the program. Of the twelve teacher trainees, ten were in-service teachers who taught in secondary or post-secondary educational settings in the US, and two were graduate students who were about to enter the job market to launch their teaching careers. Three teachers worked together to prepare for team-teaching, and a total of five groups of teachers rotated to complete their teaching practicum. The teachers prepared lesson plans according to STARTALK-endorsed principles for effective teaching and learning as follows, while taking into account program-specific curricular objectives on a daily basis.

- Implementing a standards-based and thematically organized curriculum
- Facilitating a learner-centered classroom
- Using target language and providing comprehensible input for instruction
- Integrating culture, content, and language in a world language classroom
- Adapting and using age-appropriate authentic materials
- Conducting performance-based assessment

The twelve teacher trainees went through a multi-faceted reflective comprehensive cycle that was carefully designed for them to fulfill practicum requirements. They worked closely with leading teachers in the teacher and student program to complete teaching practice and were systematically supervised by the leading teachers, the technology specialist, and the key trainer and program director before and during the program.

3. Discussion and results

Whereas twenty students completed the questionnaire at the end of the program, only eighteen students fully participated in the four-stage flipped learning process from the beginning until the end of the program. Therefore the statistical analysis of the following first section is based on data collected from eighteen students. The discussion is divided into two sections. The first section focuses on statistical analysis of the factors potentially affecting students' preview questions and preview question scores. The second section summarizes the results of the questionnaire pertaining to students' input and comments on the design of the flipped learning classroom.

3.1 Factors potentially affecting students' learning in the flipped classroom

Scores on the daily comprehension questions, i.e., quiz scores, were used to indicate students' understanding of the videos. A statistical analysis was conducted to investigate whether the following factors played a significant role in students' scores and the number of notes they took: gender, level of student's prior Chinese knowledge, and the number of days into the program, i.e., 8 days with 8 different topics. The effect of note- taking on students' quiz scores was also considered.

3.1.1 Score versus gender, Chinese level, and day

A three-factor statistical analysis of variance (ANOVA) model was used to investigate the following research questions:

- Did female students score higher than male students?
- Did students who learned Chinese before score higher than novice students?
- Did the scores differ in eight days, i.e., eight topics?

The three factors analyzed were gender, Chinese level, and day, with day as the repeated measurement.

Table 1 shows the results of the three-factor ANOVA on scores. The results indicated that students' scores did not show a significant difference based on gender, the level of students' prior Chinese knowledge, or the interaction between gender and level. However, the scores showed a significant difference on day ($F(7, 98) = 4.17$, $p < 0.001$).

Table 1: Results of the 3-factor (gender, level, day) ANOVA on scores with day as the repeated measurement

Tests of Hypotheses Using the ANOVA MS for ID (Gender*Level) as an Error Term **Dependent Variable: Score**					
Source	DF	ANOVA SS	Mean Square	F Value	Pr>F
Gender	1	0.00042861	0.00042861	0.00	0.9555
Level	1	0.18101010	0.18101010	1.37	0.2620
Gender*Level	1	0.11533924	0.11533924	0.87	0.3667
Tests of Hypotheses Using the ANOVA MA for ID* day(Gender*Level) as an Error Term **Dependent Variable: Score**					
Source	DF	ANOVA SS	Mean Square	F Value	Pr>F
Day	7	0.44336667	0.06333810	4.17	0.0005
Gender*Day	7	0.07373690	0.01053384	0.69	0.6776
Level*Day	7	0.09677619	0.01382517	0.90	0.5023
Gender* Level*Day	7	0.09165201	0.01309314	0.86	0.5397

A plot of day versus average score by gender is shown in Figure 1, and a plot of day versis average score by level of Chinese is shown in Figure 2. Both figures reveal that the average scores throughout the eight days showed no significant difference based on gender or levels.

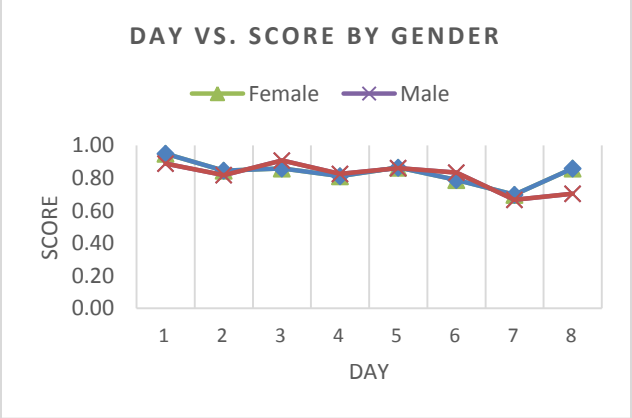


Figure 1: Day versus Score by Gender

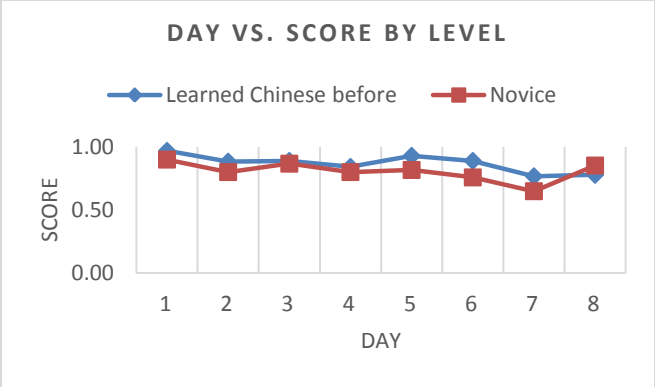


Figure 2: Day versus Score by Level

The results show that day is a significant factor in the scores of student performance. A one-factor repeated measurement ANOVA was then conducted for a detailed comparison of the scores among these eight days. The results were consistent with the 3-factor ANOVA in that there were significant effects for day on the scores ($F(24, 119)= 7.35, p< 0.001$). The average score of Day 1 and the average score of Day 7 were significantly different.

The following figure, i.e., Figure 3, shows the average scores for the eight days.

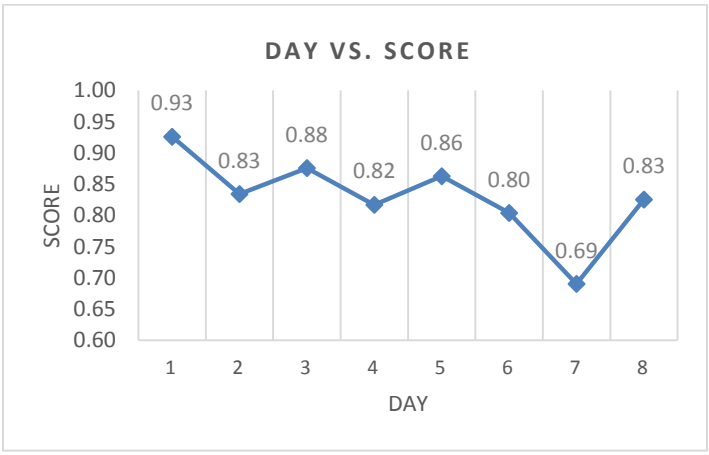


Figure 3: Day versus Score

The Day 1 score was higher than the other days. This might be due to the time that the quiz was taken. On the first day of the program, students received in-class instruction in the morning and then learned how to use the flipped classroom video platform during the afternoon lab time. They watched the videos and took the quiz after they were introduced to the 1Know website. Therefore, the scores reflected their understanding after the in-class instruction and video watching. The score from Day 7 was significantly lower than the scores on other days. This might be due to the higher difficulty level of the content associated with vocabulary and grammatical structures and the design of the quiz questions. For Day 7, students watched two flipped videos and answered six questions after watching the videos. However, two of the questions were answered correctly only 50% of the time. One of the questions was not clearly stated,

and the other question referred to grammar structure that was not explained in the video. Besides Day 7, the rest of the scores show no significant difference among the different days and topics. Therefore, excluding the human error in the question design on Day 7, gender, level of Chinese knowledge, and the day of each topic did not have a significant effect on quiz scores.

3.1.2 Note-taking versus gender, Chinese level, and day

Since gender, level, and day did not have a significant effect on students' scores, the effect of note-taking on students' scores was studied. First, how gender, level of prior Chinese knowledge, and day affected the number of notes students took while watching the videos was studied. A three-factor ANOVA model was used to investigate the following research questions:

- Did female students take more notes than male students?
- Did students who learned Chinese before take fewer notes than novice students?
- Did the number of notes taken change throughout the eight days (topics)?

The three factors analyzed were gender, Chinese level, and day, with day as the repeated measurement.

Among the eighteen students who took the quiz, two students' data were excluded from the data analysis since they watched only the first day's videos and did not watch the rest. Therefore, this statistical model included only the data from sixteen students. In total, there were 1,045 notes taken throughout the eight days of video watching. The notes can be categorized into the following three different types.

1. Repetition of video content
   Ninety percent of the notes (totally 941 notes) that students took repeated what they learned from the videos. This includes listing new vocabulary in Pinyin and English (for example: *da lanqiu = bball),* and summarizing the content they learned in the video (for example*, women jintian qu kan dian ying, hao ma? subject-time-verb-thing*).
2. Questions (94 notes)
   Nine percent of the notes (totally 94 notes) that students took were questions they had about the content. They were encouraged to ask questions and to highlight the questions in red so the teachers could see them more easily. Their questions included queries about language and grammar structure (for example, *In Chinese, do you ever conjugate verbs?*), and questions about culture (for example, *is it disrespectful to ask an adult their age?*).
3. Notes to self
   Only one percent of the notes (totally 10 notes) served as notes to themselves. They included a reminder (for example, *be careful of the pinyin*), and self-reflections (for example, *wo hen xi huan xue xi han yu*).

Although the quality of the notes might be a factor that affected how students learned Chinese language and culture, this study sorely focused on examining whether the quantity of note-taking affected students' learning outcomes. The quality of note-taking is beyond the scope of this study.

Table 2 shows the results of the three-factor ANOVA on note-taking. The results indicated that students' scores did not show a significant difference on gender, the level of students' prior Chinese knowledge, or the interaction between gender and level. However, the number of notes taken shows a significant difference on day ($F(7, 84) = 5.74$, $p < 0.001$); thus there was a significant effect of day on the number of notes taken.

Table 2: Results of the 3-factor (gender, level, day) ANOVA on number of notes taken with day as the repeated measurement

Tests of Hypotheses Using the ANOVA MS for ID(Gender*Level) as an Error Term Dependent Variable: Notes					
Source	DF	ANOVA SS	Mean Square	F Value	Pr>F
Gender	1	646.1519805	646.1519805	2.04	0.1786
Level	1	4.5893630	4.5893630	0.01	0.9063
Gender*Level	1	282.6019129	282.6019129	0.89	0.3641
Tests of Hypotheses Using the ANOVA MA for ID* day(Gender*Level) as an Error Term Dependent Variable: Notes					
Source	DF	ANOVA SS	Mean Square	F Value	Pr>F
Day	7	724.1919430	103.4559919	5.74	<.0001
Gender*Day	7	52.8686549	7.5526650	0.42	0.8880
Level*Day	7	123.3499991	17.6214284	0.98	0.4530
Gender* Level*Day	7	25.9886940	3.7126706	0.21	0.9833

A plot of day versus number of notes taken by gender is shown in Figure 4 below.

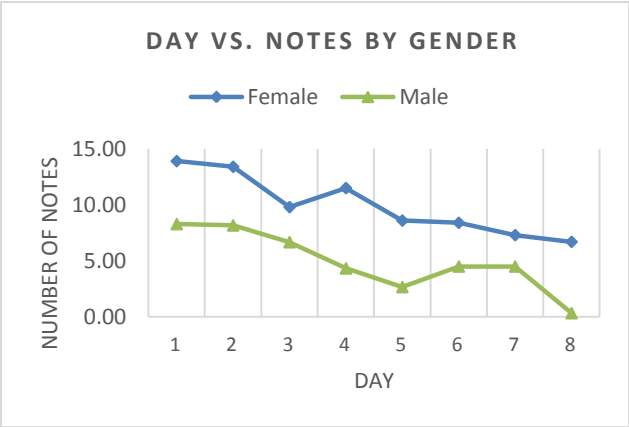


Figure 4: Day versus Number of notes taken by Gender

In Figure 4, the daily average number of notes taken by the female students was consistently higher than that of the male students. However, the results of ANOVA showed no significant difference between genders ($F(1, 12)=2.04$, $p=0.17$). When looking at the averages over the entire eight days for females ($M=9.87$, $SD= 2.17$) and males ($M=5.21$, $SD=2.57$) and taking into account the standard deviations, there was no significant difference between the genders in terms of the average number of notes taken.

Figure 5 is the plot of day versus the number of notes taken by students with different levels of Chinese. The results of the analysis showed that the number of notes taken by the novice students ($M=8.27$, $SD= 2.17$) and the number of notes taken by the students who had previously learned Chinese ($M=7.88$, $SD=2.57$) were not significantly different.

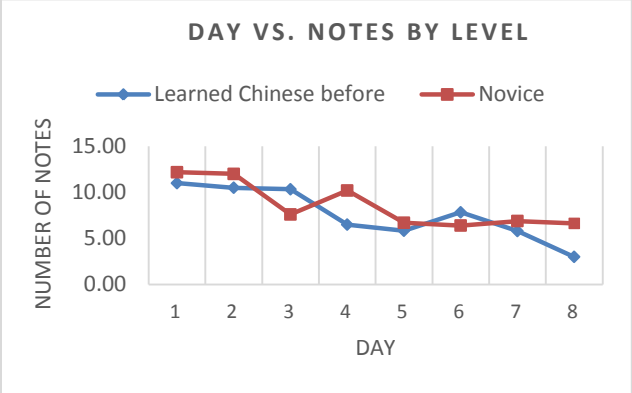


Figure 5: Day versus Number of notes taken by Level of Chinese

Figure 6 shows the average number of notes taken throughout the eight days. The number of notes decreases throughout the eight days. A one-factor (day) repeated measure ANOVA with number of notes as the dependent variable was conducted. The result was consistent with the results from the 3-factor ANOVA in that there was a statistically significant effect of day on the number of notes taken ($F(22,105)=15.21$, $p<.001$). The decrease in the number of notes taken can be explained by the decreasing level of student motivation and energy level to take notes and increasing accumulative workload toward the end of the intensive program. Another factor must have something to do with the design of note-taking that is one-way in nature, lacking the instructor's interaction with students online. Although face-to-face feedback on students' note-taking was given as needed during the next day, daily encouraging feedback and responses to students' comments, regardless of type of notes, were not given online through two-way communication as a program feature.

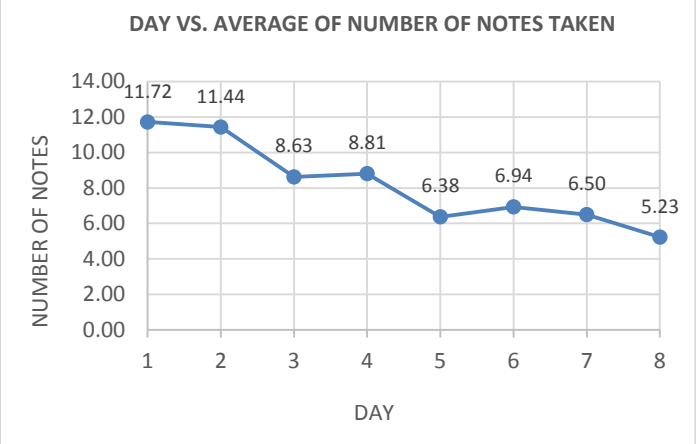


Figure 6: Day vs. Number of notes taken

3.1.3 Number of notes taken vs. score

Gender and level of prior Chinese knowledge did not show significant effects on quiz scores. A regression analysis on the average of individual students' scores and their average number of notes taken was used to investigate whether students who took more notes scored higher than students who took fewer notes, and whether there was a relationship between test scores and the number of notes taken. As indicated in Table 3, the results of the analysis showed that there was a significant positive relationship between the number of notes taken and score ($F(1, 14)=4.69$, $p < .05$). The coefficient of linear fit slope is 0.00737, which showed that there is a positive relationship between the number of notes taken and score, with a significance level of .05. However, the effect on score is very small.

Table 3: Results of the regression analysis of the number of notes taken and quiz score

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	1	0.03228833	0.032288	4.6938
Error	14	0.09630542	0.006879	Prob > F
C. Total	15	0.12859375		0.0480*
Parameter Estimates				
Term	Estimate	Std Error	T Ratio	Prob > |t|
Intercept	0.7944662	0.034563	22.99	<.0001*
AvgNote	0.0073768	0.003405	2.17	0.0480*
Linear Fit AvgScore = 0.7944662 + 0.0073768 x AvgNote				

3.1.4 Watching videos or not versus score

There were two students who did not watch the flipped classroom videos at home but completed the quiz questions, except on Day 2. One of them had learned Chinese for

one year, and the other did not have any prior Chinese learning. To understand how these videos helped novice students learn, the scores of this novice student who did not watch videos were compared with the scores of the novice students who watched the videos. Figure 7 shows that the scores of the novice student who did not watch the videos (M=0.40, SD= 0.15) were significantly lower than the novice students who watched the videos (M=p.83, SD= 0.074).

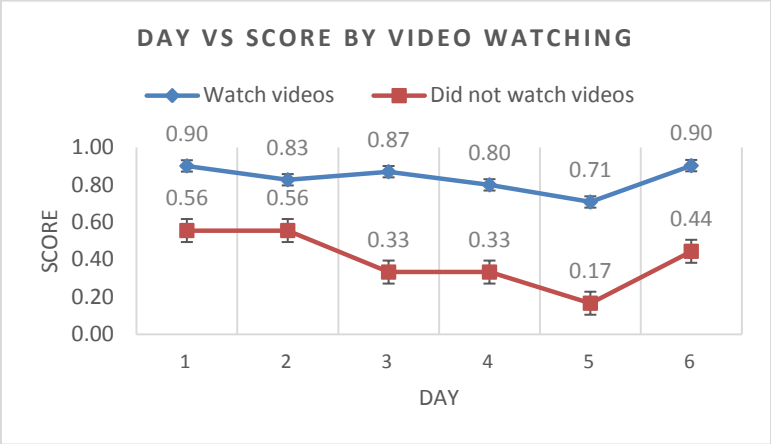


Figure 7: Day versus Score by video watching and no video watching

The results of the statistical analysis showed that gender and level of Chinese knowledge did not have a significant effect on quiz scores or on the number of notes taken. However, day had significant effects on both scores and the number of notes taken. The effect of day on the scores might have been caused by the question design, which may have room for improvement. The effect of day on the number of notes taken was significant, but the reasons for the decrease in the number of notes taken needs further study. The regression test showed that there is a significant positive relationship between the number of notes taken and quiz scores. The students who took more notes scored higher on the quizzes; however, the impact on score was small. Students who did not have prior Chinese knowledge and watched the videos scored significantly higher than the novice student who did not watch the videos. This result indicated that watching flipped classroom videos did help students understand the content and perform better on the quizzes.

3.2 Results of the questionnaire on learners' input and comments

In addition to the above statistical analysis of the impact of gender, level, topic, and note-taking on students' learning, this study also conducted a questionnaire with items created based on a 5-point Likert Scale to gather learner participants' input on the effectiveness of the implementation of the four stages of flipped learning. The questionnaire looked into five categories of each stage of flipped design: time commitment, motivation, effects, technology use, and future recommendations.

As far as time commitment is concerned, when learners watched videos, they watched videos and took notes at the same time, so time was automatically calculated

based on the total time spent watching videos and taking notes. There was no mean indicative of time spent on watching videos only. On average, each learner spent 11.50 minutes watching videos and taking notes, 10.05 minutes responding to preview questions, and 6.10 minutes viewing instant feedback, amounting to 28.1 minutes for daily flipped learning at home. The time spent on completing flipped learning components seemed reasonable and manageable and did not appear overwhelming as learners were fully devoted to Chinese language learning, and no other subjects, in the summer.

The following table summarizes the means of the remaining four categories: motivation, effects, technology, and recommendations.

Table 4: Means of 5-point Likert Scale on four Stages of flipped Learning

Category	Watching videos	Watching videos/taking notes	Responding to preview questions	Viewing instant feedback
Motivation	4.15	3.45	4.15	4.25
Effects	4.55	3.30	4.40	4.45
Technology	4.65	3.95	4.50	4.50
Recommendations	4.65	3.65	4.30	4.60
Average	4.50	3.59	4.34	4.45

As the above table shows, all columns have means higher than 4 except for the second column on watching videos and taking notes, with the mean of each cell ranging from 3 to 4.

Looking at the figures vertically, the means from the highest to the lowest are represented in the following order: watching videos (4.50), viewing instant feedback (4.45), responding to preview questions (4.34), and watching videos and taking notes (3.59). The highest mean lies in watching videos, but surprisingly when it goes with note-taking, the means in four categories become the lowest. Responding to preview questions and viewing instant feedback had their means very close to the highest.

Horizontally, analyzing the means in four categories, motivation overall has slightly lower means, while effects, technology, and recommendations have higher means. Having said that, cells that represent motivation still have means higher than 4, except for the cell that represents watching videos and taking notes, as mentioned earlier. This indicates that learners well-recognized the effects and values of flipped learning and the friendliness of technology use, therefore highly recommending the future use of the four-stage flipped learning model. The only area that did not receive the highest recommendation is online note-taking that accompanies watching videos. Although learners affirmed the value of watching instructional videos as a before-class preview component, they did not hold the same high level of motivation when they were required to do note-taking while watching videos simultaneously. Similarly, learners did not hold the same high level of positive attitude toward note-taking as the effects of flipped learning and the friendliness of technology use.

The relatively low value of taking notes while watching videos is worth exploring and explaining. Apparently, learners' personal preference did not comply with the impact of note-taking on learning outcomes, as indicated in the statistical analysis discussed above.

As discussed in the previous section on statistical analysis, a significant positive relationship was found between the number of notes taken and the learning outcome, i.e., quiz score. This can be interpreted that the more notes taken, the better the score is. The results showed that note-taking does have a positive effect on absorbing video content, therefore enhancing language performance with more satisfactory learning results. It may be possible that 9-12th graders were not yet independent learners in online learning, and were unaware of the value of note-taking and its positive impact on their overall learning. For this group of learners, flipped learning is a brand new concept and experience. They had never watched videos and taken notes online simultaneously, so they were not accustomed to performing similar tasks at the same time. Another factor that may well explain why note-taking is not a welcome option is that the flipped learning process in this study was designed to be a one-way communication; no follow-ups or interactive features were added between the learner and instructor or among learners themselves. This led note-taking to be an individual task with interaction taking place between each learner and the online platform. It is admitted that note-taking has undisputable pedagogical value and its functionality cannot be ignored or under-stressed. It is hypothesized that if learners had received more training on note-taking and understood its positive effect before participating in flipped learning during the study, their input on this aspect might have differed. This hypothesis, however, will have to be verified through future studies.

Aside from the results summarized above based on a questionnaire at the 5-point Likert Scale, it is worthwhile discussing additional comments that some of the learners shared in the questionnaire. Most learners appreciated the opportunity to watch the short videos as the videos helped them prepare for the next class, gain an idea of upcoming content, and get ready for the full immersion class on the next day. One commented that learners can "get a head start by watching the video," adding that these videos are informative and are great resources. One said positively that videos also helped explain grammatical use, and in terms of grammar, another expressed the need to add more grammar to the videos. Of all the added comments, one student who was not used to online learning expressed his frustration because "video content and vocabulary were not on student sheets." The program actually offered a folder of study guides that included daily materials to be learned, and whether it is necessary to create handouts to accompany video content is a matter for flipped classroom instructors to consider further. According to comments collected in this study, this does not stand out as a critical issue.

Unlike mostly positive comments on learners' reactions to watching videos, feedback on note-taking was half positive and half slightly negative.

On the positive side, note-taking helped to "retain content" and "affirm knowledge and keep it fresh in mind." Two learners expressed their appreciation of the

convenience of the technology tool that allowed the video to automatically pause when learners started taking notes. This surprisingly made another student frustrated, revealing completely the opposite response.

One student shared a very interesting observation about grammar covered in the videos. She pointed out that the videos sometimes discussed grammar, whereas in class the instructor rarely discussed it. This point was perceived by that particular student as a drawback, but it is indeed a proof of the successful handling of grammar that was taught through performance-based instruction with no English explanation at all. Apparently, this particular student expected the instructor to use English to explain and analyze grammatical structures, but this was not encouraged in a full immersion setting where only the target language could be used as a best practice.

Two students mentioned that it was unnecessary to take notes during videos without explicating reasons or a justification. One may have experienced technology glitches and could not get the note-taking software to work, and another found it was hard to access the notes online afterwards and preferred paper notes. This leaves the researchers to ponder upon the issue of whether online notes or traditional paper-and-pencil notes bring about more convenience and desirable outcomes. If notes were made readily accessible afterwards, then the objection could have been easily removed. Echoing this point of view, one student requested a physical paper with preview items on it to aid in watching the videos. Several others who provided additional comments unanimously shared positive feedback on preview questions. The preview questions not only showed learners the area where they needed to improve but also helped them check for understanding and retain knowledge. Besides, the tool used to access and respond to preview questions was easy to use, as added by one student.

Additional comments on receiving instant feedback were overwhelmingly positive. Time commitment to instant feedback was reasonable and manageable. The answer keys helped learners "learn grammar," "see mistakes," and "understand what to improve and why." As further noted by three students, the process was "fast, useful, and efficient." While one student agreed that links to email were good, another noted that "email feedback was less convenient; immediate feedback on the site would be better." No one experienced any technology glitches while receiving instant feedback via email, and they just had to check emails.

Overall, the preview section really helped learners "feel prepared and ready for the next day, but without being overwhelming." The value of watching videos is well-acknowledged by learners. Taking notes helped "make material sink in," although several comments demonstrate a need for further thoughts about note-taking. Responding to preview questions helped to "get the best STARTALK experience possible," and students wanted to retain receiving instant feedback in the future program.

3.3 Pedagogical implications and applications

This study contributes to our understanding of effective flipped learning by creating a well-sequenced and structured four-stage flip learning process. Its applicability in secondary and post-secondary classrooms during the regular semester is perceived to have high interest for language teachers, supervisors, and administrators. The four-stage flipped learning model allows instructors to constantly and systematically monitor students' daily progress through a multi-functioned 1know centralized online operational system that was created for the purpose of the study. Undoubtedly, the planning of each single stage is never an easy process. It takes full devotion among teaching and administrative staff on the team.

3.3.1 Creating the first stage of flipped learning

Most teachers tend to have reservations about the required time commitment to create instructional videos. Flipped learning actually engages students to be exposed to a wide variety of learning materials and does not necessarily require teachers to create videos at the expense of their own time. There are different types of materials that can be adopted and used for flipped learning with ease, such as YouTube videos, online films, podcasts, newspapers, TV news, audio files, and a large variety of authentic materials geared toward four language skills across three communicative modes. Watching videos is just one of many effective ways for pre-class preparation, but listening to sound files, reading an article, writing a brief email, or recording a short response are some examples of what students can do to prepare before class. As technology continues to grow and gain popularity, online resources have become more accessible, diversified, and versatile, and teachers do not need to work alone. Collaboration and professional exchanges save teachers time and help achieve mutual goals.

3.3.2 Alternatives to the online platform for flipped learning

Conceivably, creating the multi-functioned 1know central system that was particularly developed for the purpose of this study is beyond what a foreign language teacher can do. Most schools may have limited personnel support, lack of expertise and knowledge, and budget constraints. A possible resolution to this situation is to use any online platform that is available and whatever tools teachers feel familiar with and confident about. In this regard, checking for understanding can be made possible, and it is highly recommended that comprehension be included in whatever format to ensure that learners do what they are expected to do and that learning does take place before class. Providing instant feedback, if not doable through technology, can be replaced by other convenient options that achieve similar evaluative and curricular goals, such as self-checking or a small-scale follow-up evaluation. Alternatives to the adoption of technology tools and for the three stages that follow video watching in order to achieve the same curricular objectives are crucial and of great interest to language teachers. But that is not the focus of this study and will be addressed in a separate paper.

3.3.3 Time allocation during regular semesters

The program was conducted in an intensive summer STARTALK combination program when student participants did not take any course other than learning Chinese, so they were able to fully concentrate on completing required work before and after class. The group of learners was selected from a large group of applicants, and recruited students were highly motivated, well-disciplined, and willing to put in their greatest effort to meet curricular standards. One may argue that while the flipped components work well in such an intensive summer program, they may not work well during regular semesters. As acknowledged by many Chinese language teachers who teach in high schools in the US, they are not allowed to give students too much homework after class and therefore have reservations about assigning preview work in accordance with school policies. This leads them to give little or no work for students to do before entering the classroom, and the result is that all materials have to be taught by the instructor in class, and what can be reasonably and readily self-learned before class is postponed to limited class time.

An allocation of less than 20 minutes for preview work prior to class time is indeed feasible and greatly beneficial to both teachers and students. At any rate, instructional videos are supposed to be as short as 1 to 2 minutes, and no more than 3 minutes. Adding questions for comprehension checks and instant feedback, but excluding note-taking, the preview exercise will not exceed 20 minutes and deplete too much of students' time at home. Wisely arranged, pre-class materials for self-paced flipped learning can be divided into several small segments as needed. To decrease the complexity level on technology, a two-step design that involves videos and comprehension questions may still fulfill pedagogical needs that instructors aim to achieve.

3.4 Conclusion

This study yields practical applications and potential for future research, but like many other studies on pedagogical innovation, it has perceivable limitations. The study had a small sample size at the novice level during a short, intensive summer program. To what extent the results of the study are applicable to other proficiency levels in different types of educational settings is unknown. Future studies should increase the sample size, extend to other proficiency levels, experiment in different courses and programs, and if feasible, include both control and experimental groups to make possible a comparison of students' performance in light of scores of pre-program and post-program tests and other aspects of analysis.

The pedagogical value of note-taking cannot be ignored despite the fact that students did not strongly favor it. Online note-taking accompanying video-watching is an area that deserves in-depth discussion. When learners move to the intermediate and advanced levels, the content of their note-taking may change. In this study, the major note type is the repetition of video content, which accounts for 90% of all notes, leaving 10% to be questions and reminders. The content of elementary Chinese is not composed of a wide range of vocabulary, and the complexity of syntactic structures, functions, and

pragmatic use is not as high as that at the intermediate and advanced levels. It is thus hypothesized that the composition and type of notes may alter in intermediate and advanced Chinese. This poses another interesting research question to study further.

Online note-taking in this study was designed to be a one-way flipped learning component, not a two-way communication. Assuming that learners benefit more from two-way interactive communication, tweaking online note-taking to two-way interactive learning may increase the level of engagement, but this would have become labor intensive and requires more time commitment from instructors or teaching assistants. It is assumed that one-on-one connections between individual students and the instructor will motivate students to work harder, and teachers will be better able to interact with students online. To what extent one-way or two-way note-taking affects learners' input and learning outcomes is another new area to explore in future studies.

Note-taking strategies are another domain to attend to as well. Although student participants attended an orientation provided by the technology specialist and the leading teacher of the student program, how to take notes to best retain content while watching videos was not the focus of the orientation. If students were well guided to learn effective note-taking strategies, such instruction could affect note-taking content. This, however, went beyond the scope of the study but deserves further exploration.

It is hoped that the same or revised flipped learning model in the study be duplicated, modified, and extended to different levels of Chinese language courses and different types of Chinese language programs domestically or abroad. Further studies with such experimental measurement can deepen our understanding of the effective design and implementation of flipped learning in Chinese as a foreign language across different proficiency levels.

References

Baker, J. W. (2000). The "classroom flip": Using web course management tools to become the guide by the side. In J. A. Chambers (Ed.), *Selected papers from the 11th international conference on college teaching and learning* (pp. 9-17). Jacksonville, FL: Florida Community College at Jacksonville.

Bergmann, J., & Sams, A. (2012a). *Flip your classroom: Reach every student in every class every day.* Eugene, OR: International Society for Technology in Education.

Bergmann, J., & Sams, A. (2012b). Before you flip, consider this. *Phi Delta Kappan, 94*(2), 25.

Berrett, D. (2012). How "flipping" the classroom can improve the traditional lecture. *The Chronicle of Higher Education, 58*(25), 16-18.

Biemiller, L. (2013). San Jose State U. says replacing live lectures with videos increased test scores. *The Chronicle of Higher Education.* Retrieved from http://chronicle.com/blogs/wiredcampus/san-jose-state-u-saysreplacing-live-lectures-with-videos-increased-test-scores/40470

Bonnell, C., & Eison, J. (1991). Active learning: Creating excitement in the classroom. Washington, DC: George Washington University Press.

Davis, C. (2013). Flipped or inverted learning: Strategies for course design. In E. G. Smyth & J. X. Volkder (Eds.), *Enhancing instruction with visual media: Utilizing video and lecture caputer* (pp. 241-265). Hershey, PA: IGI Global.

Deslauries, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science, 332*, 862-864. doi:10.1126/science.1201783

Enfield, J. (2013). Looking at the impact of the flipped classroom model of instruction on undergraduate multimedia students at CSUN. *TechTrends, 57* (6), 14-27. doi: 10.1007/s11528-013-0698-1

Fulton, K. P. (2012). 10 Reasons to Flip. *Phi Delta Kappan, 94*(2), 20-24.

Fulton, K. P. (2013). Byron's flipped classrooms. Retrieved from http://www.aasa.org/content.aspx?id=27488

Graham, C. R. (2006). Blended learning systems: Definitions, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *The handbook of blended learning: Global perspectives, local designs* (pp. 3-21). San Francisco, CA: Pfeiffer Publishing.

Hamden, N., McKnight, P., McKnight, K., & Arfstrom, K. (2013). The flipped learning model: A white paper based on the literature review titled *A review of flipped learning*. Retrieved from http://www.flippedlearning.org/cms/lib07/VA01923112/Centricity/Domain/41/WhitePaper_FlippedLearning.pdf

Herreid, C. F., & Schiller, N. A. (2013). Case studies and the flipped classroom. *Journal of College Science Teaching, 42*(5), 62-66.

Hung, H. (2015). Flipping the classroom for English language learners to foster active learning. *Computer Assisted Language Learning, 28*(1), 81-96, doi:10.1080/09588221.2014.967701.

Lage, M. J., & Platt, G. (2000). The internet and the inverted classroom. *Journal of Economic Education, 31*(1), 11.

Lage, M. J., Platt, G., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education, 31*(1), 30-43.

Mehring, J. (2014). *An exploratory study of the lived experiences of Japanese undergraduate EFL students in the flipped classroom* (Unpublished doctoral dissertation). Pepperdine University, Malibu, CA, USA.

Moran, C. (2014). *Changing Paradigms: A Mixed Methods Study of Flipping the English Language Arts Classroom* (Unpublished doctoral dissertation). North Carolina State University, Raleigh, NC, USA.

Novak, G. (2011). Just-in-time teaching. *New Directions for Teaching and learning*, 128, 63-73.

Ryan, B. (2013). Flipping over: Student-centered learning and assessment. *Journal of Perspectives in Applied Academic Practice, 1* (2), 30-39. doi: 10.14297/jpaap.v1i2.64

Strayer, J. F. (2007). *The effects of the classroom flip on the learning environment: A comparison of learning activity in a traditional classroom and a flip classroom that used an intelligent tutoring system* (Unpublished doctoral Dissertation). Retrieved from http://faculty.washington.edu/rvanderp/DLData/FlippingClassDis.pdf

Salazar, H., & Sinclair K. (2015, November). Improving L2 lexical competency through flipping the classroom. Paper presented at the 2015 ACTFL Annual Convention and World Languages Expo. San Diego, CA, USA.

Staker, H., & Horn, M. (2012). Classifying K-12 blended learning. Retrieved from http://www.christenseninstitute.org/wp-content/uploads/2013/04/Classifying-K-12-blended-learning.pdf.

Tucker, B. (2012). The flipped classroom: Online instruction at home frees class time for learning. *Education Next*, 12(1), 82-83. Retrieved from http://educationnext.org/files/ednext_20121_BTucker.pdf

Ullman, E. (2013). Tips and tools for the flipped classroom. Retrieved from http://www.techlearning.com/features/0039/tools-and-tips-for-the-flipped-classroom/53725.

Appendix 1

Questionnaire on Learners' Input and Comments on the Four-Stage Flipped Learning

Instructions: Thank you for participating in the 2015 Virginia STARTALK Chinese Student Academy. This year we have pilot-tested before-class flipped classroom components and homework components. We would like to get your feedback on these. Circle the number that best reflects your learning experiences. Your responses will be kept confidential and anonymous and will have no impact on your position in this program. Thank you very much for your participation and feedback.
Before-class flipped classroom components:

1. Watching videos
2. Watching videos while taking notes
3. Responding to questions via Google Form
4. Receiving instant feedback via email

A. Motivation

1. I am motivated to watch the videos before class.

1 2 3 4 5
Least Agree Strongly agree

Additional comment:

2. I am motivated to take notes while watching videos before class.

1 2 3 4 5
Least Agree Strongly agree

Additional comment:

3. I am motivated to work on comprehension questions in multiple-choice questions before class.

1 2 3 4 5
Least Agree Strongly agree
Additional comment:

4. I am motivated to read the results (score & explanation) of my responses to the multiple-choice questions before class.
1 2 3 4 5
Least Agree Strongly agree
Additional comment:

B. Effects
5. Watching videos before class helps me prepare for the content that I will learn on the following day.
1 2 3 4 5
Least Agree Strongly agree
Additional comment:

6. Taking notes while watching videos helps me learn the content of the videos better.
1 2 3 4 5
Least Agree Strongly agree
Additional comment:

7. Responding to comprehension questions (multiple-choice) after watching the videos helps me understand how well I learn.
1 2 3 4 5
Least Agree Strongly agree
Additional comment:

8. The instant feedback (score and explanation) helps me better understand how well I learn.
1 2 3 4 5
Least Agree Strongly agree
Additional comment:

C. Time commitment
Please write down <u>how much time on average</u> you spend on each of the following components of the flipped classroom.

9. Taking notes while watching videos
Average time spent per day: ( )

10. Responding to comprehension questions
Average time spent per day: ( )

11. Receiving instant feedback on my responses to comprehension questions
Average time spent per day: ( )

D. Use of technology

Comment on the friendliness of the technology tools that the program created for the following component.

12. Watching videos

1 2 3 4 5

Least User Friendly Very User Friendly

Additional comment:

13. Taking notes (iNote) while watching videos

1 2 3 4 5

Least User Friendly Very User Friendly

Additional comment:

14. Responding to comprehension questions

1 2 3 4 5

Least User Friendly Very User Friendly

Additional comment:

15. Receiving instant feedback on my responses to comprehension questions

1 2 3 4 5

Least User Friendly Very User Friendly

Additional comment:

E. Recommendation

Recommendation for the flipped classroom components in next year's STARTALK program

16. Watching videos

1 2 3 4 5

Least recommended Strongly recommended

Additional comment:

17. Taking notes while watching videos

1 2 3 4 5

Least recommended Strongly recommended

Additional comment:

18. Responding to comprehension questions

1 2 3 4 5

Least recommended Strongly recommended

Additional comment:

19. Receiving instant feedback on my responses to comprehension questions

1 2 3 4 5

Least recommended Strongly recommended
Additional comment:

20. Please write 3-5 sentences to provide your overall comments and suggestions on the flipped classroom components.

移动通信终端促成远程讨论参加者间形成合作学习作用分析[1]
(Cooperative Peer Learning Activity of Online Cross-Cultural Communication by Sharing Emoji)

砂岡和子
(Sunaoka, Kazuko)
早稻田大学
(Waseda University)
ksunaoka@waseda.jp

摘要：母语者与非母语者在跨文化讨论中是否能顺利形成互助合作学习，除了语言能力的影响，能够引起参加者情感共鸣的副语言信息也是非常重要的因素。本文介绍了在远程讨论课中由于物理空间间隔造成面对面副语言信息沟通缺失的情况下，不同语言能力的参加者通过在移动终端共享图片信息的方式成功构建了远程同侪合作学习的课堂实践经验，证明了通过传递电子符号可以为远程讨论提供有效的信息共享，从而推进课堂互动，提高教学质量。

Abstract: Peer learning is a pedagogic method which effectively improves activeness of participants during communication. In the present study, we introduce the experience of establishing peer learning activity between distance participants of an international online Cross-cultural communication class by sharing digital information in the group chat. Non-native participants successfully understood the paralinguistic information transmitted by the digital information from native participants and therefore they could effectively participate in the communication. Our experience indicates that paralinguistic information is the essential glue of Cross-cultural distance communication, and these symbolic and rapid electronic information resources played an active role in Cross-cultural discussion.

关键词: 同侪学习, 跨文化远程讨论, 移动终端, 副语言交际

Keywords: Peer learning , cross-cultural distance communication, mobile phone, paralinguistic information

[1] 本文部分内容曾在第九届国际汉语电脑教学研讨会（TCLT9-2016，Macau）上发表。

1. 研究动机与目的

移动数码媒介如智能手机、笔记本电脑、平板电脑的极大普及使全社会迅速进入基于网络平台进行沟通合作的新型交往模式。因此，以往远程教学直接将以教师为主导的面对面教学进行简单的网络转换已经不能适应培养学生在虚拟网络中形成社会交往能力和解决问题能力的教学需求。代之而来的是网络参与式学习模式。在参与式学习模式中，师生界限被打破，同侪启导与合作和授课结成学习伙伴，通过相互协助的方式完成学习任务。与传统的由教师主导的课堂形式相比，拥有能够让参与课程的同学通过相互沟通、协调配合、共同完成学习任务的优点。经典的同侪学习需要参加者共享某个物理空间，如在同一个教室上课，由此形成面对面沟通的伙伴互助关系。但比起真实的面对面合作，在虚拟共同空间里人们容易产生较大的物理空间感（Hall, 1966; Bono, 2009）。身处不同物理空间的参加者能否形成有效的同侪合作学习、如何促进合作学习的形成，目前还鲜有研究关注。

本文通过介绍一次建立在北京、台北、东京、横滨、北九州五地五大高校间共同举办的“亚洲学生会议”的实施情况，来探讨通过移动终端发送表情图片（日语叫做 Emoji）和动画的方式，成功在身处遥远物理空间的、语言文化背景各异的远程课程参加者之间唤起有效互动学习的教学经验。以往在课程实施过程中，学生们只有在需要查找词典或搜索相关资料时才能使用电脑和手机，群组信息只有在课外互动时才偶尔使用。而此次课程中，学生们不仅将在群组互推信息融合成课堂互动的一部分，而且所发送的信息悉数尽为动画和图片，始终没有出现过文字和声音（图8）。

据社会智能理论（Social Intelligence theory），不发言的第三人不单单是受动的旁听者，他还会用眼神和姿势等隐形信号参与讨论，对会话的进行施加影响(Goodwin,1981; Sumi, Ito, Matsuguchi, Fels, & Mase, 2003)。在本次会议中，默默推发动图的学生，虽然表面上并没有以发言的形式参与讨论，但却通过在群组中推送图片信息的方式成功参与了“多通道人机对话（Yang, 2014）”的进行，并有效解决了沟通障碍，激活了课堂气氛。下面通过对网络讨论课的具体数据加以观察和分析，从不同角度探索本次在群组中推送信息的行为对促成同侪学习的作用及其在远程讨论过程中所起的效果。

2.远程讨论课程的设计

2.1 课程理念

随着国际化的发展，生长国家各异、文化背景不同、语言能力各有千秋的学生同聚一堂，单向授课的传统教学方式已难以适应学生间跨文化交流的真实需求。教学重点应转向培养学生社会交际能力、鼓励学生参与课堂互动、激发学生自主学习动机，最终构建以个性化、民主、合作为基础的以学习者为中心的学习环境。通信技术与手段的全球普及为网络互动教学模式打下了基础。（Reinders & Darasawang, 2011，引自 Xu 2016)。

在 21 世纪即将来临的前一年，为了迎接全球国际化时代的到来，早稻田大学发布了教育发展的新战略，提出了基于科技与网络，开放性、流动性、多样性共存的三层教育研究架构方案（Nakano et al., 2014）(图 1)。其中跨文化交流课程（Cross-Cultural Distance Learning，以下简称 CCDL）是该发展战略的中枢。

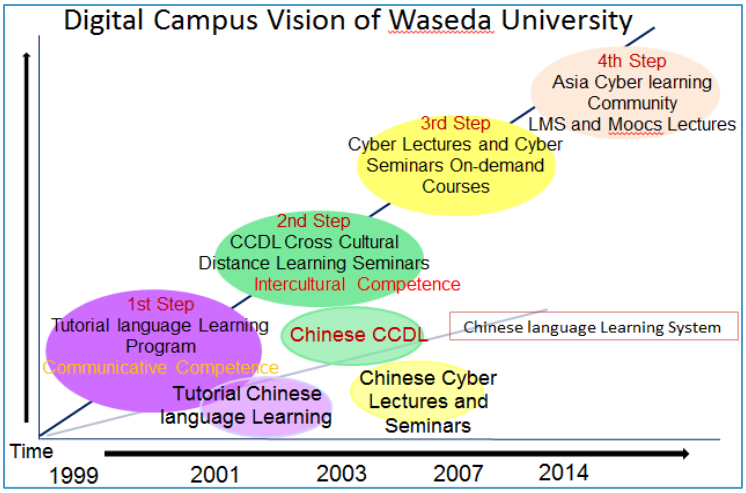


图 1 早稻田大学网络校园发展计划

随着通信手段的发达，已形式化了的技术等显性知识（Explicit knowledge）转瞬间便能传遍全球。与此相对，通过经验积累而获得的信念与价值观、劳动技巧和诀窍等隐形知识（Tacit knowledge），因难以形式化而传播性则较弱（Harvey, 1992）。互联网的发展并没有使世界各地民族、地域之间的对立与纷争减少，反而不同文化之间摩擦与冲突的报道层出不穷。高等教育机构作为培养国际人才的重要基地，培养学生形成对各国文化求同存异、互相理解的能力，以成为促进国际间理解与合作的原动力之一。

基于这个理念，CCDL 以开放性的语言活动与跨文化能力（Intercultural Competence）的培养相结合作为课程定位，在大学单位和合作企业的支持下，致力于打造英德汉日语等多种语言间交融共通的国际网络课程。“亚洲学生远程会议”是其中的汉语分部，在北京、台北、东京、横滨、北九州五地之间通过互联网定期进行视频讨论会。

2.2 课程特点

2.2.1 持久型交流

一般国际之间的视频交流会议，因为组织和技术等多方面的原因，存在举办次数不多，交流时间少，寿命短等缺陷。而“亚洲学生会议”因被早稻田大学和庆应义塾大学列为正式课程，能保证每期课程历时一个学期，期间每周举办一次同步视频会议。在各主办单位的大力支持下，自 2011 年开始到现在，已成功举行了 15 年之久。到 2016 年 6 月为止，共实施过 222 课时。参加学校时有更替，截至目前为

止，曾与北京大学、清华大学、首都师范大学、台湾师范大学、元智大学、淡江大学、韩国高丽大学、日本国内的庆应义塾大学和早稻田大学九州分校等 9 所学校进行过合作交流(Sunaoka, 2016)。早大和庆大的学生每半年或半学期可以选修，包括其他单位平均有 2 名到 10 余名学员，多时共约 30 名，少时也有 10 多名学生参加讨论。

为了避免使用免费视屏软件导致网络联接不稳定而出现课堂交流中断、通话质量不佳等问题，早稻田大学选择与相关企业合作，由其提供专业远程技术和管理系统，负责开发有关组织 CCDL 的各种软件并支持运营维护（Sunaoka, Monden, Morishita, & Ikegami, 2006）。安定的管理系统保证了海内外几所大学的成员能够定期面对面相聚而没有后顾之忧。但刚开始几次，因为参加者身处不同的教室上课，很难拉近距离。由于接触少，语言文化的差异也大，而产生疏远感。但通过一个学期间每周一个多小时的谈话，大家建立起了同学间的信赖感，逐渐形成了一个共同维护公共课堂的无言共识。

2.2.2 多元的学员

课程参加者有包括在北京、台北、日本各地的本国学生（Local Student, LS）和留学生（International Student, IS），因当天参加课程的学校而时有变化。还有来自越南、印度、俄罗斯、波兰、意大利等国的留学生参与其中。参加此课程的参与框架(participation framework)较灵活。从学生的语言文化背景来看，在同一课堂中从最开始只有本地学生参加，逐步演变为拥有留学经历的学生和在当地留学的不同国籍学生增多，变成国际对话的趋势。学生背景的多元化会给学生会议的运作带来一些负担，但总体来说它促成学员对各种语言和文化的包容能力有显著提升，对彼此间的认同感也逐步产生变化。

2.2.3 复语言主义

讨论语言最初原则上用汉语、日语隔周轮换，后来由于参加者大多具有留学经验或在不同文化背景国家生活的经历，近期同学们已经开始自由使用英日汉三种语言进行讨论。CCDL 基于 CEFR(Common European Framework for Reference of Language)的复语言主义（Plurilingualism）理念[2]，鼓励拥有成员们在彼此沟通交流时，运用他们具备的各种语言能力，并允许学生间进行同步翻译。尤其是当他们使用外语时，即使未臻完美也从正面进行积极评价，以达到培养学生社会交际能力、鼓励学生参与学习的目的。在进行文字交流时，尊重当地母语者的习惯使用其常用字体。多语言多文字兼容并蓄的课堂不仅是国际化交流的需要，还有利于促进不同背景的学生平等发言，避免让一些会按照传统外语课堂的惯例认为只能使用指定语

[2] EU 加盟国基于 CEFR(Common European Framework of References)的理念，以促进拥有不同语言和文化背景的人们之间的交流为目标，提倡复数语言教育政策。复数语言(Plurilingual)主义指的是一个人拥有包含母语在内的多数多层次的语言能力，不要求语言使用者以及学习者成为理想的母语话者，而是即使仅获得部分能力依旧积极正面评价，以运用所持能力达到语言使用的目的为目标。来源: http://www.coe.int/t/dg4/linguistic/CADRE1_EN.asp

言的学生自行封杀了发言的机会，也避免让另一些外语能力卓越的学生独占发言权(Sunaoka, 2016)。

2.2.4 学生主体的运作

“亚洲学生会议”最重要的特点是从课程议题的选择到会议的组织进行等一连串的活动流程，皆以学生为主体，由每个学校轮流负责。大家事先通过电邮商榷当天的讨论议题，开会时一般由主持人指名发言，也可大家举手自由发言。为达成以上教学目标，并使课程能够顺利举行，任课教师和各地助教（TA）紧密联系，与技术部门合作维护平台运营。如议进行中的声音画像等问题，由任课教师通过助教与技术部门进行沟通协调；若是其它人为事故，则课后由一方助教以书面文件形式提出建议，以免类似问题再次发生。

这样每周 80 分钟同海外高校进行网络互联，在信息空间中与成员伙伴积累了解，构建信任，用语言表达各自的信念与价值观，渐渐形成了相互包容、相互理解、相互关心的同侪意识。基于以培养社会交际能力、解决问题为目的的教学理念，该课程的学生成绩评价（Assessment）以参与这个会议的贡献度，比如主持会议，互助翻译等行为的参与程度进行评价（Formative evaluation）。再者参考讨论发言次数和质量，包括上课前后向 BBS 提交的发言草稿与总结报告等语言习得成果进行总括评价（Summative evaluation），最后认定学分（Sunaoka, 2016）。

2.2.5 讨论题目

许多跨文化交流活动喜欢选择有关国际民生(Global issues)的大议题。起初“亚洲学生会议”也曾基于教师的提议就时事或是社会问题进行讨论，但由于不同地区的学生对当事国的社会背景及文化传统的理解有明显差别，且对外语能力要求甚高，讨论本身易流于表层的资讯交换，无法进一步深入探讨。因此在后期的课程实施中，我们把选定论题的权利交给了学生，发现他们确实能够提出与自己切身利益相关，同时大家都感兴趣的话题[3]。

在大学里，学生的外语学习多依赖于课本和教师讲解，非常缺乏实际运用该外语和母语话者沟通的机会。“亚洲学生会议”提供了在轻松愉快的交流氛围中，让非母语者定期与母语者沟通交流、阐述观点的机会。这种将自身体验语言化的方式，不仅帮助学生习得外语的综合能力，而且促使其学会如何将自身观点传递给与自己文化背景以至思维方式不同的人，从而使得各自所属社会的隐性知识得以传递。

[3] Discussion topics of the “Asian Students’ Conference” http://www.f.waseda.jp/ksunaoka/

3.观察数据简介

3.1 课程记录

12 月 10 日的讨论语言以中文为主，必要时可以使用日语和英语。课程从当日日本时间上午 10 点 45 分到 12 点 15 分共进行了 90 分钟（发言时间共 85 分 6 秒，空白时间共 4 分 4 秒）。参加本次讨论的同学包括 9 名汉语母语者（Native Speaker, NS）和 9 名非汉语母语者（Nonnative Speaker, NNS）（表 1）。他们的母语背景、外语水平和文化背景差异较大。除了早大东京校区、庆应和北大曾经合作过之外，包括主持人（早大北九州的中国留学生）在内的其他同学都是首次参与。尽管如此，在本次讨论的一个半小时里，没有出现过讨论中断的尴尬状况。5 地学生参与程度相对均衡，共计 64 次发言被记录了下来。将发言内容转换成文本后，如表 1 所示共计 8701 个汉字、558 个英文单词。按 4 个讨论议题划分后，NS 和 NNS 的发言次数以及发言时长分别如图 2、图 3 所示。

3.2 分析方法

根据(Hayashi, 1991; Enomoto, 2007）的研究，由多人组成的会话层级（Conversational floor）以参加者之间的关系与分享信息量的不同，可以有单层（Single conversational floor）到多层（Mutiple conversational floor）的多种划分方法。参加者之间的互动与信息共享越多，可划分的会话层越多，会话层级越深。基于以上理论，本次研究首先将讨论课当日的讨论课由影音资料转换为文本资料，然后从三层会话层级—参加者个体、共享物理空间的各讨论组、共享线上空间的参加者全体—的角度出发，通过统计发言次数、通信次数等相关信息，来分析讨论进展过程和合作学习程度。最后我们由所得到的结果来探讨移动终端共享图片信息对跨跃语言文化和物理空间的同侪合作学习的促进作用。

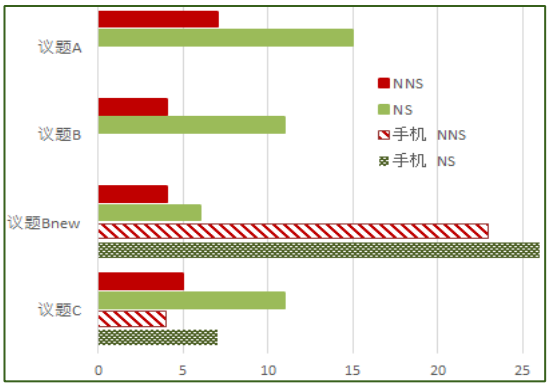


图 2 各议题中成员的发言次数和手机通信次数

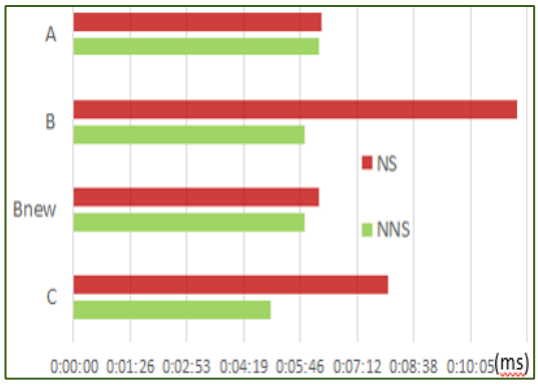


图 3 各议题中发言的平均时长

表 1 . 2015 年 12 月 10 日参加讨论人员的语言背景与发言量

学校	语言背景	人数	发言次数	发言字数	国籍	发言所用语言
庆应义塾大学（K）	NS	2	7	1164	中国	汉语
	NNS	2	6	848	日本	汉语
北京大学（P）	NS	2	6	1137	中国	汉语
	NNS	2	7	1044	日本	汉语
台湾师范大学（T）	NS	2	5	819	台湾	汉语
早稻田大学东京校区（W）	NNS	3	6	1010	日本	汉语
早稻田大学北九州校区（Wips）	NS	3	23	2679	中国	汉语
	NNS	2	4	558	波兰	英语
合计	**NS 9 人，NNS 9 人**		**64**	**9259**		

4. 分析結果

4.1 讨论议题

2.2.5 所示，“亚洲学生会议”的讨论议题每次由学生根据自身的兴趣与关注点来选择，12 月 10 日的讨论题目为“传统语言与颜文字（emoji）”。课堂讨论为 ABC3 个议题顺序进行，资料整理过程中发现 B 议题后半部分的讨论内容和之前的明显不同，因此在本次分析中将 B 议题切分成 B 和 B-new 两个部分，加上原本的 A、C 议题，共形成 4 个议题。

A ----- 介绍自己的年度词汇；
B ---- 讨论 emoji 的流行原因，使用场合与频率；
B-new --- 讨论 emoji 给日常交流带来的改变；
C ---- 讨论 emoji 与传统词汇相比对沟通的作用；

各地参加者在 4 个议题的发言次数及发言时间的统计结果如**图 2、图 3** 所示。

4.2 NS 的话语组织策略

进一步分析发现 B 议题发展成 B-new 和 NS 与 NNS 在讨论时所使用的话语组

织策略有关。如**图 6** 所示，NS 在讨论时，不仅陈述个人的经验与想法，更愿意从描述社会或群体的现象出发，引起参与者的共鸣。比如以下是 Wips 一名 NS 男生的发言；

> 我认为 emoji 会流行的一个根本原因，实际上是它在慢慢的趋近于人性化，也就是满足人类本质的沟通习惯和需求的事情。不断的贴近我们人与人之间，面对面的交流。举个例子，大家都尝试过在去追女朋友的时候，更多的是眼神，肢体语言和表情之类的东西，而不是用嘴去说出来的，这个我觉得大家都深有体会。其实也可以看看我们的老祖宗有几句话说的特别有道理“含情脉脉”“暗送秋波”“勾魂摄魄”“满面春风”“欲拒还迎”，这些成语都特别的生动贴切。比如说我们举个例子，如果一个姑娘说“我在含情脉脉的看着你”，和她真正的用含情脉脉的眼神看着你，这两个完全是不同的效果。所以我觉得，通过表情或者表情贴纸会使得我们的聊天更加方便，含义更加丰富，以及有更强的表达能力。

与此相对，大多数情况下 NNS 只能讲述自身经历和感想，引起参加者共鸣的效果不如广为人知的例子；另一方面，NS 所使用的例子中很多由于文化差异并不为 NNS 所熟知，如上例中所列出的成语，再加之语言能力的差距，共同导致了 A、B 议题中 NNS 难以和 NS 形成共鸣，越来越难以理解议题讨论动向的问题。从 B-new 议题开始共享图片信息之后，NNS 排除了语言能力的阻隔，通过图片所传递的副语言信息迅速和 NS 形成了共鸣，成功重新在讨论中活跃起来，且发言内容的深度不输 NS。于是在议题 B-new 进行的过程中，开始了合作学习。

4.3 合作模式的启动

之前在议题 A、B 中持续的个体发言方式，非汉语母语者的发言次数整体少于汉语母语者发言次数的一半。这使得言语能力更胜一筹的 NS 掌握了话语领导地位，NNS 参加者明显处于劣势，尤其是话题 B 的差异最大（图 2、图 3）。

议题 B-new 开始，以 NS 向参加者全员所在的微信群发信息为契机，参加者同时使用移动终端互推信息，迅速由议题 A、B 的单层会话变成议题 B-new、C 的多层关系，形成了深层会话模式的合作学习。图像信息能够有效帮助 NNS 参与论题，促成了远距离单位之间互动合作的情感，使得进行议题 C 时，讨论重新活跃起来。

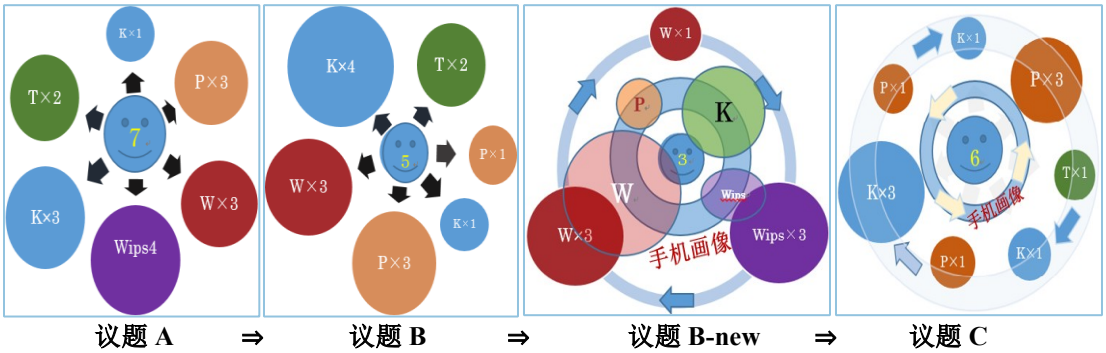


图 4 各议题参与者的互动关系

说明：K、P、T 等为各大学名称的缩写（参见**表 1**）。“×2”、“×3”等是每个地点发言者的人数（＝发言次数）。椭圆的大小与发言次数成正比，次数越多面积越大。椭圆间的“⇒”表示合作互动的方向。每个图中央的笑脸和数字表示主持人的发言次数。

4.4 讨论中的资源共享

4.4.1 移动通信终端

由于议题 B 的讨论题目“emoji 的流行原因”相对抽象，发言者集中在了同一所高校的 NS。在早稻田大学的 NNS 出现参与困难情况时，先由早大北九州男生（NS），之后北大女生（NS）分别在微信群中推发了（图 8）所示的几条贴图，然后大家在主持人的提示引导下，参加者全员开始积极响应微信群组信息，陆续推发共 68 条动图或图片，形成了各地间一边参照图片信息一边讨论议题的合作学习链。以下是当时的一连串发言摘录；

早大男生 NNS1：(看到了一条由早大北九州一名 NS 同学推发来的贴图后)我很喜欢那个笑哭的 emoji。我觉得那个 emoji 很有用。我想表达我失败的时候，或者我有错的时候，我可能用那个表情，我觉得那个很有用。我觉得微信和 Line 的 emoji 不太一样，（中略）微信有比较讽刺的 emoji，Line 有比较可爱的 emoji (此时北大女生又推发了贴图，然后主持人告诉大家参看) 。所以我觉得 Line 和微信的 emoji 用法不大一样。emoji 给我带了很多乐趣，看 emoji 很开心，所以我们经常用 emoji。就这样，谢谢! (大家陆续推发大量的动图片) 。

早大到北大留学的男生 NNS2：我觉得我们对 emoji 的印象，最近可能有了一点点变化。我高中的时候，emoji 是一种比较温柔的表达方式。最近我觉得，用 line 的时候，朋友只发 stamp 的话，会比较冷淡，所以 emoji 用的太多的话，我们的语言表达能力会减弱。

早大北九州校男生 NS1：我觉得 emoji 可以表达一种比较模糊的感情，起到活跃聊天气氛的作用。但是我觉得如果太过于依赖 emoji 的话，会降低我们自己的语言表达能力。相信大家应该也知道有些朋友，他们在微信中聊天的时候显得很能说，很活跃，但是现实聊天的时候却发现他并没有通过工具聊天这么能说会道，其实很大程度是因为 emoji 弥补了他很大一部分语言表达能力的缺陷，但是对于一些比较正式的场合的时候，就会突出自己语言表达能力的不足，所以我觉得 emoji 用时可以经常用，但是同时也要多思考，自己语言表达能力方面有没有什么可以提升的部分。谢谢。

早大北九州波兰的男生 2：So I think the reason why emoji is getting popular those days is that it is a easy way to express your true feelings your true emotions and it also can save a lot of time when you try to write a long sentences in words or express what you are thinking about, you can just chose a face and send to other people. I think emoji changed my communication with my friends. It can make other people feel that I am friendly and I have a intension to keep this conversation going , but I think if you use emoji too many times it can change the way we are seen by other people, some of the people may even lose respect to what you said. （后略）

学生都是 90 后的数码新一代，很熟悉这种网络参与式的交流方式。Line 和微信是“亚洲学生会议”参加者平时用于课外交流的手机软件（Sunaoka, 2016）。而本次课程中，这些软件提供的表情和动图，不仅对于课程讨论难题的解决起到了巨大的作用，而且通过这些图片信息，参加者迅速构建了共同的讨论平台。图 5 展示了当时的课堂情景。保证讨论持续进行的方法后，议题进到 B-new 时，进一步促进

了同侪对等学习的进行。

4.4.2 动图符号

在本次讨论过程中，移动终端的推送没有出现过文字，都是图片或动图。如图 6 所示，在讨论群中发信最多的是早稻田大学的 NNS，其次是北京大学的 NS，相比二者，其他学校的发信数量较少。通过共享图片、动画等能够体现发言者态度与情绪的内容，使对语言一知半解的 NNS 更加精确的理解了讨论内容（Kubota, 2012），跟上了整个讨论的节奏。在 NNS 最终理解了议题 B 的设问后，成功激活了相关的背景知识，进到议题 B-new 时提出了较有深度的问题。其他参加者通过对这些提问的思考，从新的角度开始审视与理解议题 C。

由此可见，在不能实现真实面对面眼神或姿态共享的远程参加者之间，通过分享相关的图片和动画能够使参加者迅速实现“共情”，实现相互理解与相互帮助的学习模式，推动讨论交流的活跃与顺畅 （图 4）。学生们不仅将在群组互推图片融合成课堂互动的一部分，同时分享动图符号所含的象征性信息，还能迅速准确地理解异文化。反映了在当今国际交流的多语种及多元文化的背景下,“象征能力（symbolic competence）”将更为完善人与人之间的沟通能力（Kramsch, 2006）。

图 5. “亚洲学生会议”10/12/2015 课堂情景

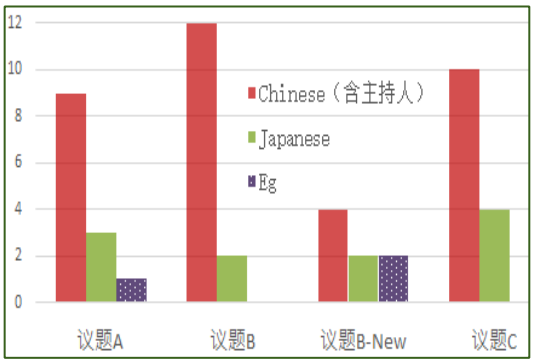


图 6. 不仅陈述自己的发言

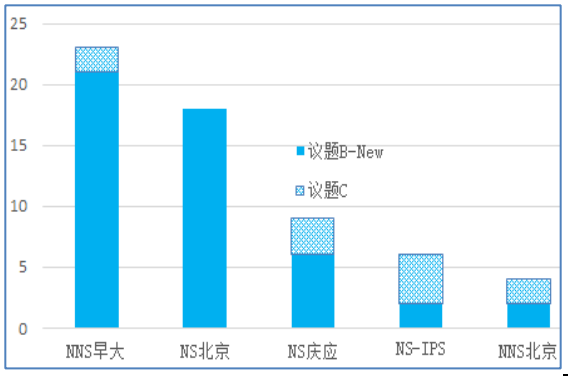


图 7. 手机通信次数

图 8. 互推动画和图片群

5. 结语

随着电子信息科技的发展和智能移动终端的广泛应用，课堂教学也不必再局限在某时某地某教室。但在物理空间间隔使参加者失去了通过眼神、表情、姿势等辅助语言交流的条件下，形成同侪合作学习的跨语言跨文化讨论困难重重。本文通过使用定量分析的方法，介绍了在东京、北京、横滨、北九州、台北五地共同举行的远程讨论课程中，参加者通过在移动终端共享图片信息，迅速形成对课题的共鸣，有效弥补了非母语参加者对议题理解困难的问题，成功实现了异地远程同侪合作学习的课程实践经验。

从本次课程经过推及至日常生活，可以窥见在 Emoji、贴图和动图等流行的背后，存在着共同的社会心理原因。发送文字信息避免了打电话那样直接沟通所可能带来的不便，却因只用文字未免显得态度冰冷。于是卡通人物为主体表情包应运而生，不仅恰当地传递了话者的情绪信息，还因适度的夸张而带来更加幽默风趣的效果，作为辅助文字的工具，确实能有效地解决沟通障碍。

本次视屏会议中，学生推发大量表情动图，起初仅是为了帮助无话可说的非母语同侪参与讨论，后来却引起了全体学生的共鸣，激活了所有参加者的学习兴趣。其使用动机与私人短信交流时发送 emoji 向对方展示当下情绪的心态相近。在缺失副语言信息（Paralanguage information）的情况下，贴图像背景音乐那样，将参加者带入共同的情绪环境，有效地缓解了沟通障碍，激活了课堂的气氛。与此同时，无声的贴图不会影响课堂的语言讨论，而且能随手下载，可迅速收发，用起来很方便。

由此本研究为在虚拟空间中，如何使用移动终端推送副语言信息，营造出面对面交流的效果，使学生间生成自然的合作情感，克服物理空间造成的障碍，实现同侪合作学习提供了一些证据。有人曾担心在课堂上让学生随便使用智能手机、平板电脑等会不会影响正常的教学秩序。而学生以往课堂上很少使用移动终端，在群组中推送信息就更少，只有在需要查找词典或搜索相关资料时才能使用。本课程成员都认为因自己要承担保证会议进行的各种责任，所以不会发生用手机窃窃私语，甚至侵害隐私权，或过激言论等问题。

科技大爆炸、大普及、大提高催生的数码新一代学习者希望可以通过数码信息和交流技术、通过网络与外部真实世界相联系，最后希望交互使用虚拟课堂和实地课堂，希望以学习者为中心的自主（因人而异）学习，特别是动态学习（Xu 2016）。教师也应该接受他们的要求，尽可能地尝试将数码媒介结合到语言教学中去。

网络开放式教学与传统的由教师主导的课堂形式相比，能有效地形成学习者的社会交往能力和解决问题能力。能够让参与课程的同学通过相互沟通、协调配合、共同完成学习任务的优点（Nakano et al., 2014; Nakano, 2015）。当然，由学生主宰运作的课堂，未免存在讨论范围缺乏系统性，讨论深度参差不齐的问题。因此教师和助教在课后会将各种相关材料上传至 BBS 平台进行补充，同时基于学科的衔接

设计（Curriculum Articulation），鼓励学生并行选修高级讨论课和留学，全面提高综合能力（Sunaoka, 2016）。

包括“亚洲学生会议”，早稻田大学 CCDL 正赶上了 CALL（Computer Assiste Language Learning，电脑辅助语言学习）的全方位开拓期（2000-2005）和社会 CALL 发展的新趋势（2006-2015）（Xu, 2016）。参与式学习模式和 Social CALL 的兴起为 CCDL 的国际交际合作打下了坚实的基础。据（Nakano et al., 2014; Nakano, 2015）的研究，参加 CCDL 以培养社会交际能力为主学习环境中操练英语的学生，他们一直保持足够的学习动机，与其他上一般英语课的学生相比，他们的英语水平提高得更快。在开设课程的十数年中，包括汉语 CCDL 在内，有关 Tutorial Chinese Language Course 及 Chinese Cyber lecture Course 的建设与教学资源的累积，都取得了一定的成果（Sunaoka, 2016）。

另外，本文将使用电子信息辅助远程跨文化讨论形成合作学习作为分析的重点，对语言讨论等方面没有进行充分的说明，今后我们将在其他文章中对此进行讨论。

致谢：本论文是在 2016 年 5 月 28 日澳门大学举办的 TCLT9 会议上，口头发表时，受到在场各位老师的指教以及来自北京语言大学到早大留学的孙悦同学的翻译校对下完成的。至此致谢！

参考文献

Bono, M., & Takanashi, K. (2009). *Methodology for analyzing multi-party interaction*. Tokyo: Ohmsha. [坊農真弓, & 高梨克也. (2009). *多人数インタラクションの分析手法*. 东京: オーム社.]

Hall, E. (1966). *The hidden dimension*. Garden City, N.Y.: Doubleday & Company, Inc.

Enomoto, M. (2007). Variety and dynamism of multi-party interaction: What is increased in multi-party interaction (Workshop in the 18th Congress). *The Japanese Journal of Language in Society, 9*(2), 154-158. [榎本美香. (2007). 多人数インタラクションの多様性とダイナミズム：多人数インタラクションでは何が多くなるのか? (第 18 回研究大会ワークショップ). *社会言語科学*, *9*(2), 154-158.]

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Harvey , D. (1992). *The condition of postmodernity: An Enquiry into the Origins of Cultural Change*. Cambridge, MA: Wiley-Blackwell

Hayashi, R. (1991). Floor structure of English and Japanese conversation. *Journal of Pragmatics,16*(1),1-30.

Kramsch, C. (2006). From communicative competence to symbolic competence. *Modern Language. 90*(2), 249-252.

Kubota, H. (2012). What does emoji tell? In M. Yamanashi, et al (Eds.), *Cognitive Linguistics 10* (pp. 143-192). Tokyo: Hituzi Shobo , Ltd . [久保田ひろい. (2012).

絵文字は何を伝えるか. 山梨正明他編，*認知言語学論考 (10)*, 143-192. 東京:ひつじ書房.]

Nakano, M., Bonham, M., Owada, K., Ueda, N., Oya, M., Negishi, J., & Kondo, Y. (2014) . Networked English language learning from English tutorials to cyber interactions at Waseda University. In R. Tsai, & G. Redmer (Eds.), *Language, Culture, and Information Technology* (pp. 1-31). Taiwan: Bookman Books, Ltd.

Nakano, M. (2015). Capitalising on internet and computing technologies for improving learning and teaching: The experience of Waseda University. In C. Ng, R. Fox, & M. Nakano (Eds.), *Reforming Learning and Teaching in Asia-Pacific Universities: Influences of Globalised Processes in Japan, Hong Kong and Australia* (pp. 213-230). Singapore: Springer.

Sumi, Y., Ito, S., Matsuguchi, T., Fels, S., & Mase, K. (2003). Collaborative capturing and interpretation of interactions. *Journal of Information Processing, 44*(11), 2628-2637. [角康之, 伊藤禎宣, 松口哲也, シドニー・フェルス, 間瀬健二. (2003). 協調的なインタラクションの記録と解釈. *『情報処理学会』論文, 44*(11), 2628-2637.]

Sunaoka, K. (2016). Conversion from promotion for overseas study to learning in campus with a co-educational environment. In K. Sunaoka, & Y. Muroi (Eds.), (2016). *Voices from Japan in a Multilingual World: Prospects for Fostering Human Resources and Global Communication* (pp. 183-203). Tokyo: Asahi Press. Retrieved from http://www.waseda.jp/fpse/winpec/assets/uploads/2016/03/No.E1521Kazuko-SUNAOKA.pdf. [砂岡和子. (2016). 留学振興から国際共学へ. 砂岡和子、室井禎之编『日本発多言語国際情報発信の現状と課題-ヒューマンリソースとグローバルコミュニケーションのゆくえ』183-203. 東京:朝日出版社.]

Sunaoka, K., Ma, Y., & Shi, X. (2015). Practice and challenges in the international distance learning—A case study of the “Asian Students’ Conference”. *The Journal of Modernization of Chinese Language Education, 4*(2), 70-79. Retrieved from http://xuebao.eblcu.com/xuebao/essay/eightterm/09.pdf. [砂冈和子, 马燕, 施信余. (2015). 国际远程教育的实践与经验——以“亚洲学生远程会议”课程为例. *中文教学现代化学报, 4*(2),70-79.]

Sunaoka, K., Monden, Y., Morishita, Y., Ikegami, D. (2006). The construction of foreign language learning environment in international distance conference., In M. Nakano (Ed.) *CCDL [Cross Cultural Distance Learning]: Teachers’ Manual. Vol. 1* (pp. 195-122) . Tokyo: Media Mix, Waseda University.

Xu, D. (2016). Issues in CALL studies. *Journal of Technology and Chinese Language Teaching, 6*(2), 1-16. Retrieved from http://www.tclt.us/journal/2015v6n2/xu.pdf.

Yang, M., Tao, J., Li, H., & Chao., L. (2014). Multimodal natural human-computer interaction dialog system. *Computer Science, 41*(10), 12-18,35. [杨明浩, 陶建华, 李昊, 巢林林. (2014). 面向自然交互的多通道人机对话系统. *计算机科学, 41*(10), 12-18,35.]

Journal of Technology and Chinese Language Teaching Volume 7 Number 1, June 2016
http://www.tclt.us/journal/2016v7n1/li.pdf pp. 56-77

Collocation Analysis Tools for Chinese Collocation Studies (可用于汉语搭配研究的搭配分析工具)

Li, Shouji
(李守纪)
Massey University
(梅西大学)
s.li.1@massey.ac.nz

Guo, Shulun
(郭曙纶)
Shanghai Jiao Tong University
(上海交通大学)
gshulun@163.com

Abstract: The importance of the teaching and learning of collocation has been widely acknowledged among researchers in the field of second language acquisition and teaching. In recent years, a number of lexical tools that can be used for collocation analysis such as WordSmith, AntConc, PowerConc, BNCweb, and CQPweb have become available for researchers. Despite their availability, the understanding and mastery of these tools and related techniques remain challenging for many researchers who are technically less competent. This is particularly the case in the studies of Chinese collocation within the field of Teaching Chinese as a Foreign Language (TCFL). Some of the abovementioned tools cannot be used for processing and analyzing Chinese texts, and even if they could, extra steps would be needed before they could be employed. This study aims to provide an overview of the collocation tools and corpora for Chinese language learning, with a focus on tools that can be used for Chinese language. Using the corpus analysis toolkit AntConc as an example, the study explains the procedures involved in the analysis of Chinese collocations, the configurations of the toolkit, and the statistical measures that are relevant to Chinese text analysis. It is hoped that this introduction provides some useful information for researchers who are interested in using collocation analytic tools in their studies.

摘要：搭配教学的重要性近年来已经得到研究第二语言习得和第二语言教学界的普遍认可。相应地，可以用于搭配研究的一些词汇学工具例如 WordSmith, AntConc, PowerConc, BNCweb, CQPweb 等也已经开发出来并为诸多学者所使用。然而，对于很多缺乏这方面技能的研究者来说，理解并掌握如何利用这些工具进行搭配研究却成为一个挑战。这种情况在汉语搭配研究当中更为突出，原因是上述很多工具最初是针对分析英语文本而设计的，无法直接用于汉语搭配的分析，或者是有些工具虽然可以用于汉语文本分析，但使用前需要一些额外的步骤对汉语文本进行处理。本文首先概述了可以用于汉语搭配研究的词汇工具和语料库的基本情况，然后以语料库分析工具 AntConc 为例，详细说明了使用它对汉语语料中的搭配进行提取和分析时的操作步骤、软件设置、跨距和统计测量值的设定等，以期对那些有兴趣使用这些搭配分析工具进行汉语搭配研究的学者提供一些有用的信息。

Keywords: Collocation studies, Collocation analysis tools, Chinese corpus, TCFL

关键词：搭配研究，搭配分析工具，汉语语料库，对外汉语教学

1. Introduction

In recent years, the importance of the teaching and learning of collocation has been widely acknowledged among researchers in the field of second language acquisition and teaching. As a subset of formulaic sequence, collocation is regarded by some scholars as one of the most important aspects of language teaching and also one of the most significant challenges faced by second language learners. For example, Palmer (1981) regards collocation as a crucial key to language learning. In his opinion, "language are learned collocation by collocation rather than word by word" (p. 21). Likewise, Lewis (2000) believes that "collocation is the most powerful force in the creation and comprehension of all naturally-occurring text" (p. 45). Nesselhauf (2003) stresses that collocations are especially important for learners who want to achieve a high level of proficiency in their target language because they help to enhance both accuracy and fluency.

Despite the realization of the importance of collocation in language acquisition and learning, the definition and identification of collocations remain a challenge for many researchers. Over the years, there have been various taxonomies proposed by researchers with regards to how to define, identify, and distinguish collocations based on multiple criteria. These views, however, can be largely summarized as belonging to two major groups: frequency-based views and phraseological views. The frequency-based view, represented by Sinclair (1991), defines collocation as "the occurrence of two or more words within a short space of each other in a text" (p. 170). This definition is based on the probability of co-occurrence of two or more lexical items and is supported by statistical data obtained from the analysis of large language corpora. The second approach to defining collocation, namely the phraseological view, is based on the analysis of syntactic structure and semantic motivation of co-occurred lexical combinations using certain criteria. Each of these two approaches has its own pros and cons. The data-driven approach runs the risk of identifying some recurring lexical clusters (such as *and the*, *of a*) that have little psycholinguistic validity to native speakers, and the phraseological approach lacks the statistical evidence of the frequency of their actual use in language communication (Henriksen, 2013, p. 31). To overcome these disadvantages, researchers usually apply both approaches in their studies on collocations; first identifying high-frequent lexical bundles in large scale language corpus using a set of statistical measures and subsequently excluding those combinations that do not meet the collocational criteria according to the analysis of their syntactic structures and semantic motivations. It is, therefore, safe to say that the use of collocation analytical tools and language corpus has become increasingly indispensable for collocation studies.

2. **An overview of collocation analysis tools**

With the development of technology in the last few decades, a number of corpus analysis tools that can be used for collocation analysis have become available to researchers. These tools include WordSmith, MonoConc, AntConc, PowerConc, BNCweb, JustTheWord, COCA, TANGO, the Gutenberg Collocation Tool, Wmatrix, SketchEngine, Phrase in English and CQPweb. According to Hardie (2012), these concordance tools can be classified as four generations according to their power, flexibility, and usability. The first-generation tools are mainframe-based software such as the CLOC (by Reed, 1978) concordancer used at the University of Birmingham, and the second-generation are mainly PC software such as the Kaye concordancer (developed by Kaye, 1990), the Longman Mini Concordancer (by Chandler, 1989) and Micro-OCP (by Hockey, 1988). The third generation tools, represented by WordSmith (developed by Scott, 1996), MonoConc (by Barlow, 2000) and AntConc (by Anthony, 2005), were designed partially to meet the need for tools that can be used by the majority of linguists who are less technically competent and lack programming skills (Hardie, 2012, p. 382). The second- and third-generation corpus tools are usually desktop computer-based programmes, and are usually under a Windows, Macintosh or Linux operating system.

These corpus tools, however, as Hardie (2012) argues, have some limiting factors in terms of power and usability. The term 'power' refers to the capability of achieving high query speed on very large corpus datasets. The second- and third generation corpus tools often lack the power to handle large-scale corpus data that researchers need to achieve their goals due to the processing ability of computer hardware. Usability refers to user-friendliness, which has been a critical factor that constrains non-technical linguists from using these analysis tools. Hardie (2012) points out that a general relationship between power and user-friendliness can be summarized as that the more powerful a corpus analysis is, the less user-friendliness it has. For example, although WordSmith is a usability-oriented tool, one of its most powerful features is actually the indexing process, which is often less accessible for most non-technically-savvy corpus analysts. To address these issues, the fourth generation corpus tools use a different approach to improving both their power and usability by using a client/server model. According to Hardie (2012), CQPweb is one of the best web-based fourth generation corpus analysis tools, because it can enable non-technical corpus linguists to carry out corpus-based data analysis much like browsing web pages. It not only offers all the functionalities that most second- and third generation tools offer such as word list, concordancing, collocation, and keyword analysis, but also allows users to search for metadata of texts such as pre-defined genres, speakers' language proficiency and gender information. Furthermore, it provides corpus specialists with more advanced search options so that they can make more sophisticated queries. In terms of its language dependency, CQPweb has two additional features that are more attractive to corpus linguists whose research objects are languages other than English. According to Hardie (2012), CQPweb is "both language-independent and writing-system-independent" and "corpora in the Arabic, Bengali, Chinese, Cyrillic, Devanagari and Latin writing systems have been analyzed using CQPweb" (p. 398). Other highly regarded fourth generation tools include Wmatrix (developed by Rayson, 2008), SketchEngine (by Kilgarriff et al., 2010), and the corpus.byu.edu system (by Davies 2005, 2009, 2010). All the aforementioned analysis

tools can be used for collocation analysis based on English corpora, with a few specifically designed for processing and analyzing collocations in English, such as JustTheWord, TANGO, and the Gutenberg Collocation Tool.

With regards to whether they handle Chinese text analysis, only WordSmith, MonoConc, AntConc, SketchEngine, and CQPweb have the capability to analyze collocations in Chinese corpora. Some later versions of WordSmith, MonoConc, AntConc are capable of processing and analysing Chinese texts if they are encoded in Unicode and tokenized/segmented in advance. The fourth generation tools Sketch Engine and CQPweb can also be used for collocation analysis based on corpora that have already been provided on their servers. For example, Sketch Engine provides four Chinese corpora: Chinese GigaWord 2 Corpus (mainland, simplified, 250,124,230 tokens); Chinese Giga Word 2 Corpus (Taiwan, Traditional, 455,526,209 tokens), ChineseTaiwanWac (Traditional, 349,198,060 tokens) and ChineseTaiwanWac (Universal Sketch Grammar, 465,102,710 tokens). Users can also create their own corpora by uploading files from their own computer or download corpus datasets from FTP sites. Since CQPweb is an open source corpus query system, users can create their own corpora to use. For example, Xu (2014) created a BFSU CQPweb, which has 35 corpora in seven languages including Chinese and can be accessed at http://111.200.194.212/cqp/. Thus far, BFSU CQPweb has four original Chinese corpora: Lancaster Corpus of Mandarin Chinese version 1 (LCMCv1, Brown family, 1991), Lancaster Corpus of Mandarin Chinese version 2 (LCMCv2), TORCH2009 (Texts of Recent Chinese, Brown family, 2009, 2013 summer edition), and the UCLA Corpus of Written Chinese (2nd edition).

When choosing the most suitable tools for their collocation studies, researchers need to consider their research goals, available datasets or corpora, as well as the limits of available corpus analysis tools. According to Xu & Jia (2013), the third- and fourth generation of corpus analysis tools will co-exist for a period of time due to their advantages and limitations. They suggest that, in consideration of the researchers' actual needs, the third generation tools are probably more suitable for carrying out individual corpus-based studies. As a third generation tool, WordSmith has a complex interface and does not support regular expressions; in comparison, AntConc is easier to use and supports regular expressions, but it has less functionality and computational efficiency. It also tends to freeze or quit unexpectedly when processing larger corpus data. The fourth generation tools, as discussed above, are Internet-based network applications. Such tools are based on data and index technology, having faster retrieval response time and offering better user experience. When supported by the necessary hardware such as powerful servers, they are more capable of handling large corpus such as BNC. However, as the flexibility of these tools is not sufficient, users usually have difficulties with processing and analyzing corpus data stored on local desktop computers rather than a web server.Due to the limits of the index format and the scale of the data, their retrieve grammar is relatively simple, and they do not support complex searches. Considering all the factors, the authors choose AntConc to demonstrate the process and procedure involved in conducting Chinese collocation studies.

It is worth noting that the list of collocation analysis tools discussed here is not exhaustive. There are many other tools that might be more suitable for an individual

researcher's purposes. The aim here is to provide some useful information for researchers who are less technically competent but would like to gain a basic understand of such tools.

3. An overview of Chinese corpora and interlanguage corpora

3.1 List of corpora

In recent years, there have been major advancements in natural language processing. This enables more language corpora to be established for linguistic studies. To the best of the authors' knowledge, there are a few large scale Chinese corpora and Chinese learner corpora (or interlanguage corpora) freely available for language researchers. Below is a list of corpora that can be used for Chinese collocation studies:

(1) The BCC Corpus

The BCC Corpus (Beijing Language and Culture University Corpus Center) is a contemporary Chinese corpus developed by the Institute of Big Data and Education Technology of Beijing Language and Culture University. It has a 15 billion character collection of text samples of present-day written language from various sources including microblogging, science and technology, literature, and the press. The BCC Corpus provides an online concordancer which can be accessed at http://202.112.195.249/bcc/. Users can use search query to extract the use of words, grammatical markers, and other unique units such as separable words (离合词). It offers a statistical function, which not only allows users to search collocations using some formulaic expressions, but also provides statistical data of frequencies of different collocates. This is particularly useful if researchers want to find the frequency of certain collocates used in a specific text genre. For example, if one wants to know how the word 良好 collocates with other nouns in news texts, he or she can use the formulaic expression 良好* n to search. The following screenshot shows the search result of all the possible collocates of 良好。

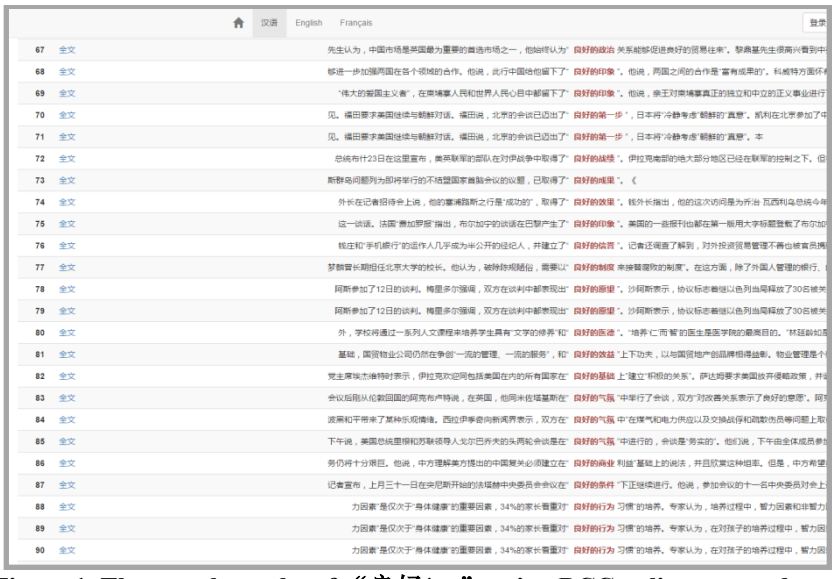


Figure 1. The search results of “良好* n” using BCC online concordance

Figure 2. The statistical data of nouns that in “良好* n”

The list is useful as it gives the frequency of each noun that can enter into the collocational phrase “良好的 n”. The list can also be downloaded for further analysis. However, one point that needs to be noted here is that some of these figures have to be manually checked for accuracy. For example, the list shows that 良好的社会 is the most frequent phrase used in the news texts with a frequency of 2581. But after a careful observation of the contexts 良好的社会, one will find that 社会 is actually not the collocate of 良好; it is only part of the nominal modifier. See the following examples:

(a)…在进行结构改革的同时还必须建立良好的社会保障。
(b) …没有良好的政府信用，就绝不可能建立起良好的社会信用。

(c) …提高服务质量，赢得了良好的社会信誉。

It is clear that in (a), the collocate of 良好 is 保障，in (b) is 信用 and (c) is 信誉。

(2) The PKU-CCL Corpus

The PKU-CCL Corpus is an online Chinese language corpus with 477 million characters, a collection of Chinese written texts of different genres. The corpus can be accessed at http://ccl.pku.edu.cn:8080/ccl_corpus/. This online query system supports word search, formulaic expression queries as well as queries of some unique patterns in Chinese, such as 高高兴兴(usually summarized as AABB pattern of duplicated adjectives), therefore it can also be used for Chinese collocation studies. However, it does not provide statistical functions similar to the BCC corpus provides.

(3) General Contemporary Chinese Corpus

This corpus is also an online Chinese language corpus with a token size of 19455328 and can be accessed at http://www.cncorpus.org/. It is sponsored by the State Language Affairs Committee of the Ministry of Education of China. Although this corpus only provides users with word or words online queries, it offers some very useful and free corpus analysis resources for users to download. For example, it offers the segmentation and POS tagging tool CorpusWordParser.ext and the word frequency tool CorpusWordFrequencyApp.ext, which are very useful in collocation studies. In Section 4, the authors describe how CorpusWordParser.exe can be used to prepare Chinese text for further collocation analysis using AntConc.

(4) The LIVAC (Linguistic Variations in Chinese Speech Communities)

The LIVAC corpus is a synchronous corpus developed by Hong Kong City University. It aims to offer a system that can be used to store the data and to analyze the linguistic development of printed Chinese texts in difference Chinese communities. The corpus also provides an online query system that can be accessed at http://www.livac.org/search.php.

(5) Academia Sinica Balanced Corpus of Modern Chinese

Sinica Corpus is a balanced corpus developed by the Academic Sinica in Taiwan with 10 million words (character token size 17,554,089). All the texts in this corpus are collected from different genres and categorized according to five criteria: genre, style, mode, topic, and source. Every text is segmented and every word is POS tagged. It also offers a web-interface which can be accessed at http://www.sinica.edu.tw/SinicaCorpus/. The online query system is designed for statistical comparison according to users' specification of topics, genres, etc. It is worth noting that this the characters in this corpus is traditional Chinese characters rather than simplified ones.

(6) ToRCH2009: Texts of Recent Chinese

The Corpus is also known as the 2009 Brown family Chinese corpus ToRCH 2009. It is the acronym of 'Texts of Recent CHinese'. The name ToRCH was proposed by Xu Jiajin and the corpus was released in summer 2013. ToRCH 2009 contains texts of 15 types (Press: Reportage, Press: Editorial, Press: Reviews, Religion, Skill and hobbies, Popular lore, Belles-lettres, Miscellaneous: Government & house organs, Learned, Fiction: General, Fiction: Mystery, Fiction: Science, Fiction: Adventure, Fiction: Romance, and Humour). The corpus size in tokenised words is 1,066,347 (1,670,356 Chinese characters) and one of the editions of ToRCH 2009 was already tokenised/segmented using ICTCLAS2012. It can be downloaded at http://www.bfsu-corpus.org/channels/corpus as a corpus stored at a local device such as desktop computers; therefore, it is a good dataset for researchers who want to perform collocations studies according to their own research goals.

(7) The HSK Dynamic Composition Corpus

The HSK Corpus is a collection of 11,600 essays (approximately 4.3 million Chinese characters) written by learners of Chinese for the HSK test. It was developed by the Research Center for Studies of Chinese as a Second Language at Beijing Language and Culture University. The HSK Corpus not only annotates the error information of characters, words, sentences, and text features such as cohesive devices, but also provides some possible corrections to these errors marked with tags. Some important text attributes such as nationality, gender, and age of CFL learners who took the test are also provided for researchers to consider when conducting studies. Statistical data pertaining to the use of characters, vocabulary, sentences, and discourse contained in compositions are also available for researchers. The corpus has an online query system which can be accessed at http://202.112.195.192:8060. Researchers can extract erroneous use of characters, words, collocations, and cohesive devices by using this online query interface.

(8) Chinese Learners Corpus

Developed by the Advanced Center for the Study of Learning Science of National Taiwan Normal University, the Chinese Learners Corpus collects written text samples from learners of Chinese at different levels with 40 L1 backgrounds and contains 3 million characters. It separates the texts written for assignments and for exams. This is particularly useful for studies on contributing factors of second language writing quality and writing strategies. Information for this corpus can be found at http://advancedcenter.top.ntnu.edu.tw/achievement5_1.html.

Due to the limit of space, the authors can only list these corpora which are currently available. This list of corpora is by no means exhaustive. For more information about Chinese corpora and Chinese interlanguage corpora, please refer to Wu & Li (2009).

3.2 The pedagogical value of the corpora

Among these corpora, the HSK Dynamic Composition Corpus, the BCC corpus, and the Chinese Learner Corpus probably have much more pedagogical value than the rest of the corpora listed above, although the last one contains only Traditional Chinese characters, which may not be suitable for learners who learn Simplified ones. In a TCFL classroom where the access to such corpora is available, teachers can first encourage learners to search for the use of some confusing words in the HSK Dynamic Composition Corpus, or teachers can show learners a few erroneous instances found in the HSK corpus. The learners can then use the BCC corpus as a reference corpus to search forthese words' high-frequency collocates. As described in Section (1), BCC corpus has the statistical function of listing all the high-frequency collocates for users. By studying the erroneous use of the words and their high-frequency collocates, learners are given opportunities to observe and interpret the collocational patterns of these words. On the basis of such language experience, learners are able to develop their competence in identifying errors, classifying various types of collocates, and generalize to a more abstract pattern. Such an approach to teaching vocabulary has been regarded by some scholars as beneficial and valuable because it allows learners "to observe *what* is typically said in given circumstances, and *how* it is typically said, and to relate the two" (Sinclair, 2004, p. 18). These three corpora also serve as very valuable datasets for TCFL teachers to conduct classroom-based studies. For example, Li (2016) investigates the collocational errors in compositions written for the HSK by American learners of Chinese on uses the HSK Dynamic Composition Corpus and the BCC corpus, and provides some useful strategies to improve the outcomes of the teaching of vocabulary. Due to the design principles and aims of data collection, other corpora listed in this study are probably more suitable for corpus linguistics studies with more specific research goals.

4. A demonstration of the use of AntConc for Chinese collocation study

As discussed in Section 2, most of the corpus analysis tools currently available are originally designed for English text analysis. However, some of them such as AntConc can handle Chinese text analysis if the text is prepared carefully in advance. In this section, the authors use AntConc as an example to demonstrate how this freeware concordance program can be used for Chinese collocation studies.

AntConc is a corpus analysis toolkit that can be used as a concordancer to extract multiple examples of words or phrases and their common collocates from a corpus for analysis. It is developed by Prof. Laurence Anthony from Waseda University in Japan and can be downloaded at http://www.laurenceanthony.net/software/antconc/. More information about this software can also be found at this website.

To use it for Chinese collocation studies, the encoding of Chinese characters needs to be UTF-8. This can be set up under the "characters encoding" category in the "global settings" tab of AntConc.

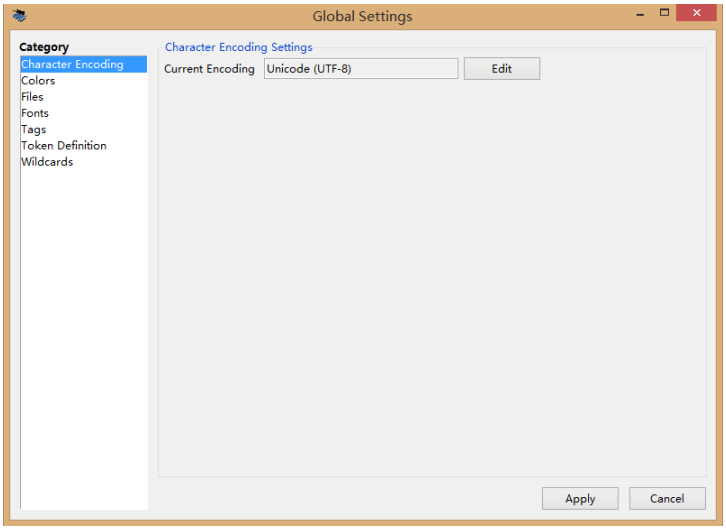


Figure 3. The encoding settings of AntConc

The next step is to set the statistical measures such as MI or T-Score. MI (mutual information) value and T-Score are two most frequently used statistical measures in collocation analysis. They are used to decide whether two co-occurred lexical items are by chance or whether their association is significant and has psychological validity. The Mutual Information value can be defined as the extent to which observed frequency of words occur in collocation differs from what we would expect. The T-Score can be defined as a measure that calculates the absolute frequency of co-occurrences of words. In his discussion of these measures, Stubbs (1995) points out that the issue of the MI value is that low-frequency collocates are more prominent in the MI based lists, while T-Score puts more emphasis on the number of joint frequencies. Consequently, more function words are likely to be included in a collocation list based on a T-Score than on MI, whereas lexical collocates that infrequently co-occur with the searched word are more likely to be found in a collocation list based on MI value(pp. 9-12). Most current corpus analysis tools provide either MI or T-Score, or both. Users are able to identify the collocates according to these two values if the MI value $\geqslant$3 and T-Score $\geqslant$2). It is worth mentioning that, according to Stubbs (1995), although the statistics can be generated mechanically by software or tools, users need to be aware of the differences of these measures and make their decisions accordingly.

According to Bai & Zheng (2004), to identify potential collocations in Chinese texts, the two measures needs to be set as MI $\geqslant$3, T$\geqslant$2.33, so co-occurrences with higher collocational strength can be highlighted. This can be set under the category of "collocates" in the "Tool Preferences" tab. Under the heading "Other Options" next to "Selected Collocate Measure", there are two options available: MI and T-Score. Choose one of them and click on "Apply" to complete the setting up process.

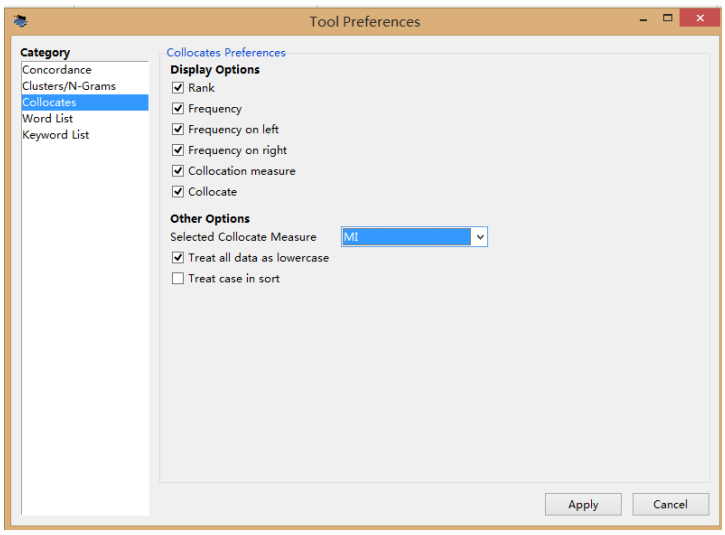


Figure 4. Setting the statistical measures

Once the set up is done, run AntConc and click on "File" on the menu bar. Choose "Open Dir…" , which will lead to the corpus data directory. See the following Figure 5.

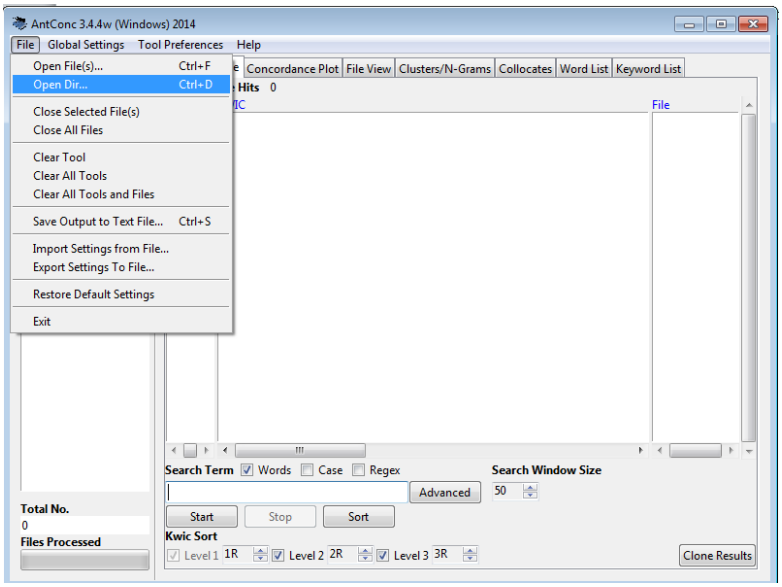


Figure 5. Importing the UTF-8 plain text files

Take the ToRCH 2009 corpus as an example. There are two sets of data available, one is ToRCH2009_ANSI—20140720, the other ToRCH2009_UTF-8—20140720. To search for collocates of a certain lexical item, just navigate to this folder that contains ToRCH2009_UTF-8—20140720 and click on "OK". The plain text files in UTF-8 encoding will be imported to the toolkit. If the plain text files in ANSI encoding are imported, the result will be "No Hit" or just a blank page after the files are processed. Once the right texts are loaded, the following settings need to done before performing the collocation query:

(1) Click on the "collocates" tab, and the relevant collocation settings will appear.
(2) Input the word to be investigated into the search box under the "search term" heading.
(3) Set the span of the words to the left and right of the searched term.

The default setting is 5L and 5R, which means 5 words to the left and 5 words to the right of the searched word. However, since the span is a critical concept of the statistical-based approach to studying collocation, the setting of span has much impact on the validity of the research findings. For example, Sun and Huang (1998) investigated the collocational distribution of noun (能力), verb (培训), and adjective (广泛) based on large scale Chinese corpus, and proposed that the best window spans for these three types of part of speech are: noun (-2, +1), verb (-3, +4), and adjective (-1, +2), here the figures in the parentheses mean the words to the left and to the right of the searched word, for example, (-2,+1) means two words to the left and one word to the right of the searched word. Setting these figures right may lead to more accurate results and retrieve rates. Bai (2004) and You (2005) also did similar research, and determined that it was a more suitable span for verbs and their collocates in Chinese text is (0,+5) as it covers most of the high-frequency collocations. It is, therefore, the researchers' decision to define the window spans to suit their research goals. For the purpose of demonstration, here the span is set as (0, +5) and search term (also known as *node* in some other studies) is a verb 提高。

(4) Decide how to sort the result. There are six methods to choose from: Freq, Freq(L), Freq(R), Stat, Word, Word End. Among them, the first four are more useful for the collocation analysis. "Freq" refers to the number of times the searched term and its potential collocates co-occur in the corpus, "Freq (R/L)" refers to the number of times the potential collocates occurs to the right or left of the searched term, and "Stat" refers to either MI or T-Score value which measures the collocation strength of the co-occurrence in the corpus. For this example, the result is sorted by Freq (R).

(5) The minimum collocate frequency can be set as the default setting "1".

After these settings are done, click on "Start", a warning message will pop up saying "AntConc needs to Jump to the word list tool to generate a word list". To calculate the collocation strength, it needs to know the frequencies of all the words in the corpus. Click on "ok" and the toolkit will start processing the data. A collocation report will be generated. See Figure 6.

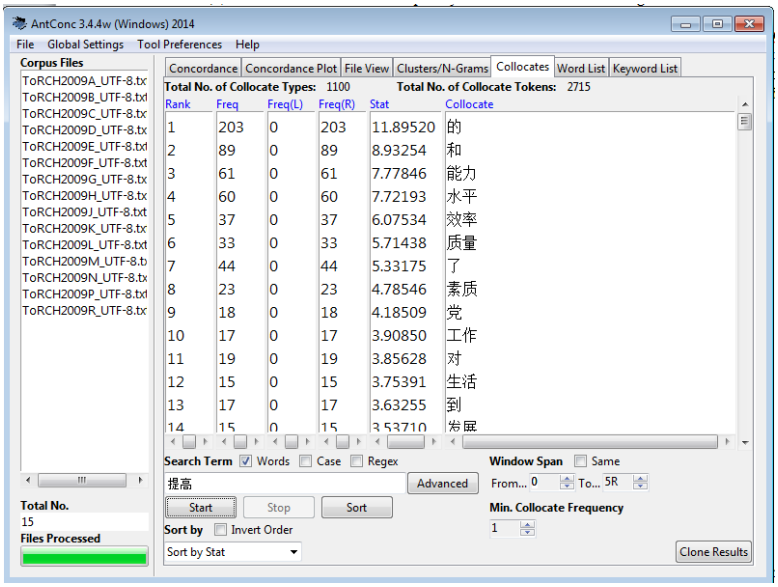


Figure 6. Collocation report of the verb 提高

It can be seen that the highest frequency and T-score value is 203 and 11.89520 respectively, and the word that collocates with 提高 is 的。However, this is just a high frequency recurring lexical bundle with little psycholinguistic validity, so words similar to this can be excluded from the high frequency list. We can see that the word 能力 has a high frequency at 61 and a high T-score at 7.77846. To view it in context, we can just click on the word 能力 and produce a concordance set of result. This step is essential as it will allow researchers to manually check if the co-occurrence has psycholinguistic validity and whether the word is a real collocate of the searched term. Figure 7 shows the concordance examples of the word 能力：

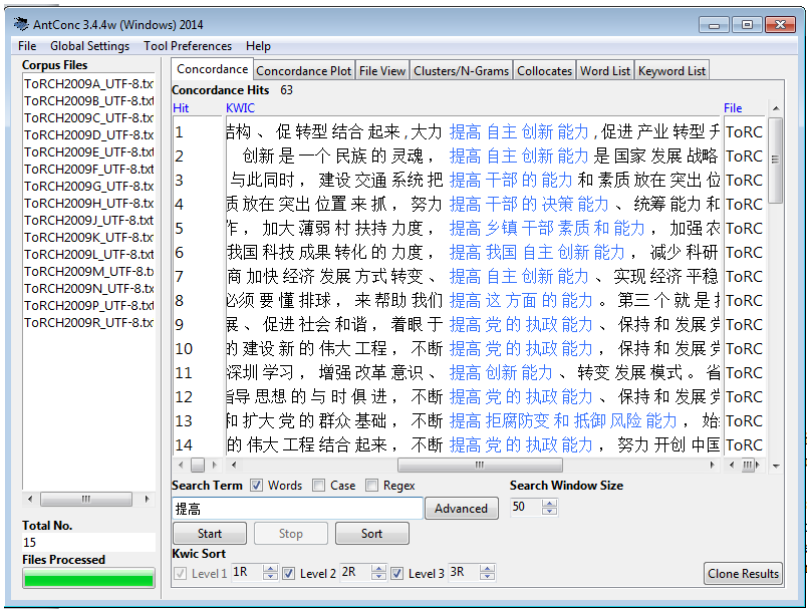


Figure 7. The concordance examples of 提高

One may also find that the word 党 does not look like a real collocate of 提高, although it has a high frequency at 18 and high T-Score at 4.18509. By checking its concordance result, we can see that the word 党 is actually part of the modifier for the nominal head words such as 水平, 能力 etc., see Figure 8. The word 党 therefore should be excluded from the high frequency list of 提高。

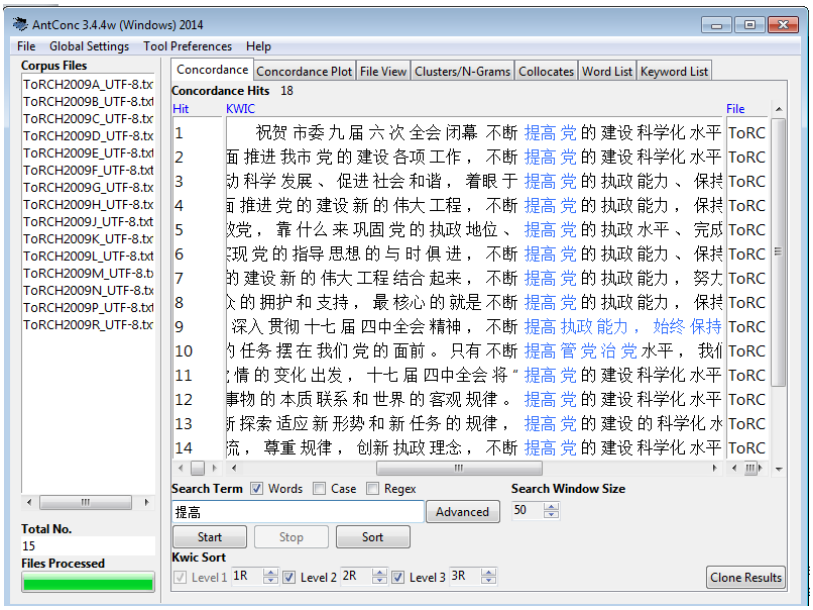


Figure 8. The concordance result of 党

To export the statistical data of the high frequency list of 提高 for further analysis, one can just click the File tab and choose "Save Output to Text File". The data will then be saved as a txt file.

Due to the relatively smaller size of the ToRCH 2009, the high frequency list of collocates is rather limited in terms of the insight it may provide of the features of a particular lexical item. For example, if we search for the high-frequency collocates of 丰富 from ToRCH 2009 and set the two measures as MI ⩾3, T⩾2.33, we only get the following list:

Table 1. High frequency collocates of 丰富 in ToRCH 2009

Collocate	MI	Frequency	Collocate	T-Score	Frequency
营养	8.88	6	营养	2.446	6
经验	8.38	6	经验	2.439	6
			内容	2.434	6
资源	6.76	5			
维生素	8.38	3			
资料	6.833	2			

To address this limitation, we may need to resort to larger scale corpus such as

BCC corpus mentioned in Section 3. However, due to the fact that BCC does not offer downloadable corpus raw data, it is impossible to use AntConc to perform collocation analysis. Even if it did, the sheer size of the corpus would impose a great difficulty for corpus analysis toolkits such as AntConc to process simply because of its limited power. One way to resolve this issue is to use BCC's online query system to generate concordance result of a particular lexical item, then use AntConc to perform further collocation analysis. Still using 丰富 as an example, we can search for the word on its online query page at http://bcc.blcu.edu.cn/. To limit the number of retrieved concordance results, it would be much more viable to choose one of BCC's sub-corpus such as literature corpus as the dataset (the system returned 31386 instances of 丰富), rather than its comprehensive corpus (returned 203210 instances of 丰富). The online query system allows registered members to download 10000 instances of the total results, which should be large enough for researchers to perform analysis and generate some valid high-frequency lists.

Once the concordance result of 丰富 is downloaded and saved as plain text file, the next step is to remove the <u> and </u> marks surrounding the word 丰富. The file is then parsed using ICTCLAS (Institute of Computing Technology, Chinese lexical Analysis System) or the POS tagging tool CorpusWordParser.ext downloaded from the General Contemporary Chinese Corpus. Figure 9 shows the screenshot of the interface of CorpusWordParser.exe when parsing the downloaded concordance result of 丰富。

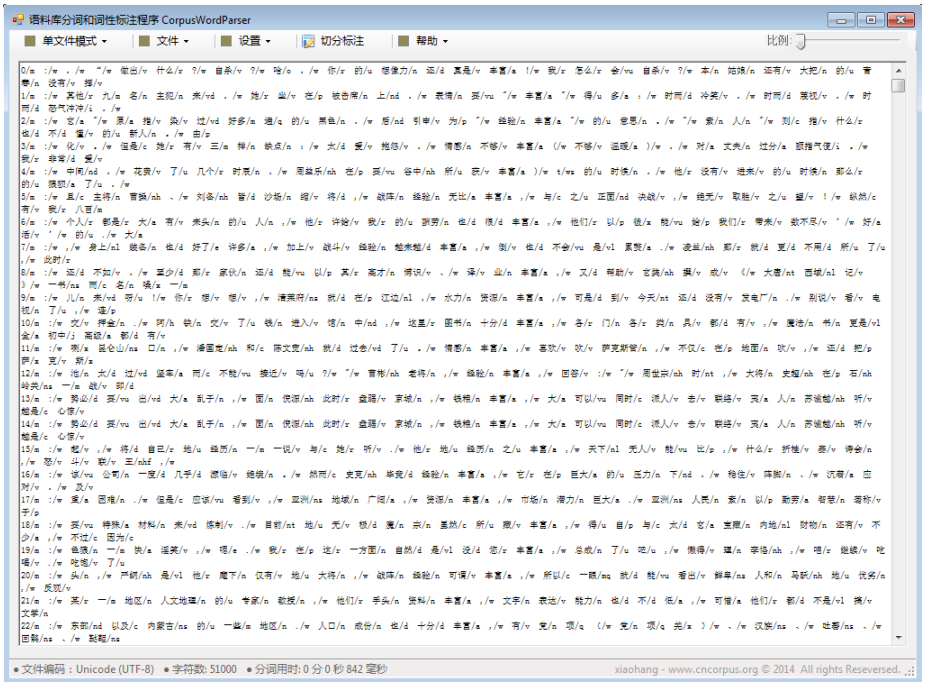


Figure 9. Using CorpusWordParser.exe to parse Chinese text

Once the text is complete, make sure to save it as a plain text file with UTF-8 encoding. See Figure 10.

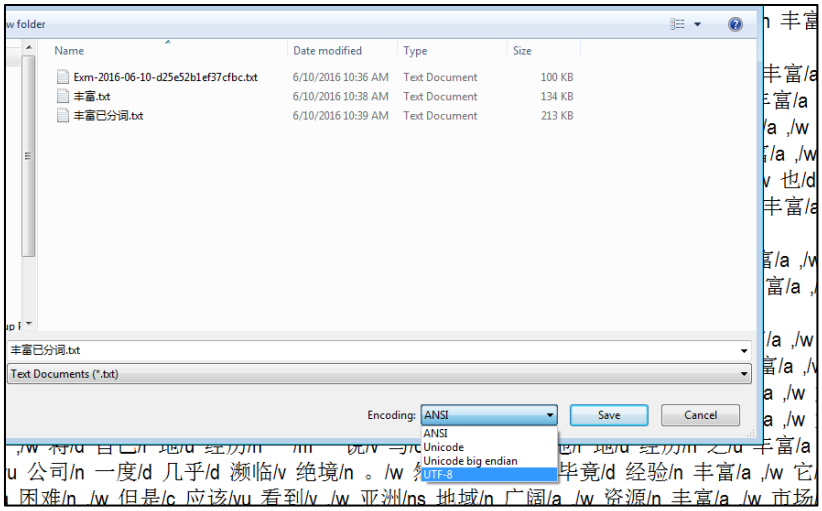
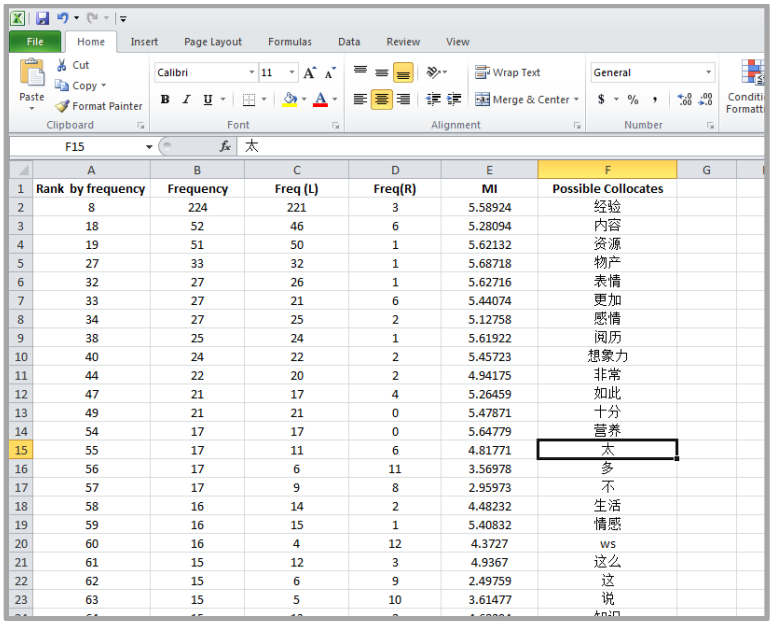


Figure 10. The parsed text needs to be saved as plain text file with UTF-8 encoding

Once the file is saved correctly, it can then be analyzed using AntCont. When setting the window span, the number of words next to the node word has to be doubled. AntConc treats the POS tag as a character. For example, if a researcher decides the window span for noun, verb, and adjective are noun (-2, +1), verb (-3, +4), and adjective (-1, +2) respectively, when he or she searches their collocates in parsed text, the window span should set as noun (-4, +2), verb (-6, +8), and adjective (-2, +4) respectively.

Once these are set, the text can be imported to AntConc and a collocation analysis performed. The analysis result can be exported as a plain text file. To conduct a further analysis, the result file can be imported into Microsoft Excel using the Data>From Text tool bar function. Figure 11 shows the process of analysis of the possible collocates of 丰富。



	A	B	C	D	E	F
1	Rank by frequency	Frequency	Freq (L)	Freq(R)	MI	Possible Collocates
2	8	224	221	3	5.58924	经验
3	18	52	46	6	5.28094	内容
4	19	51	50	1	5.62132	资源
5	27	33	32	1	5.68718	物产
6	32	27	26	1	5.62716	表情
7	33	27	21	6	5.44074	更加
8	34	27	25	2	5.12758	感情
9	38	25	24	1	5.61922	阅历
10	40	24	22	2	5.45723	想象力
11	44	22	20	2	4.94175	非常
12	47	21	17	4	5.26459	如此
13	49	21	21	0	5.47871	十分
14	54	17	17	0	5.64779	营养
15	55	17	11	6	4.81771	太
16	56	17	6	11	3.56978	多
17	57	17	9	8	2.95973	不
18	58	16	14	2	4.48232	生活
19	59	16	15	1	5.40832	情感
20	60	16	4	12	4.3727	ws
21	61	15	12	3	4.9367	这么
22	62	15	6	9	2.49759	这
23	63	15	5	10	3.61477	说

Figure 11. Using Microsoft Excel to further analyze possible collocates

As mentioned before, a manual check process is essential to ensure those lexical bundles which have little psycholinguistic validity are excluded. The final high-frequency collocate list of 丰富 (MI ⩾3) can be found in Table 2:

Table 2. High-frequency collocates of 丰富 in BCC literature corpus

Frequency	Freq (L)	Freq(R)	MI	Possible Collocates
224	221	3	5.58924	经验
52	46	6	5.28094	内容
51	50	1	5.62132	资源
33	32	1	5.68718	物产
27	26	1	5.62716	表情
27	21	6	5.44074	更加
27	25	2	5.12758	感情
25	24	1	5.61922	阅历
24	22	2	5.45723	想象力
22	20	2	4.94175	非常
21	17	4	5.26459	如此
21	21	0	5.47871	十分
17	17	0	5.64779	营养
17	11	6	4.81771	太
16	15	1	5.40832	情感
15	12	3	4.9367	这么
15	13	2	4.68294	知识
14	13	1	5.73025	极为
13	13	0	5.43069	收获
11	5	6	5.73025	肥沃
11	5	6	4.48924	思想
10	8	2	4.59275	那么
10	9	1	5.35174	极其
10	7	3	5.73025	学识
10	10	0	5.14529	不断
9	7	2	5.09282	物质
8	5	3	4.40832	经历
8	7	1	5.27082	材料
8	7	1	5.40832	内涵
7	5	2	5.36768	形式

Table 2 can be further analyzed according to the statistics. For example, one can clearly see that most of the nouns (经验, 内容, 资源, 物产, 表情, etc.) and adverbs (非常, 如此, 十分, 太, 这么, 那么, etc.) tend to appear to the left of the word 丰富, and the number of their Freq (L) are much higher than that of their Freq (R). Such information is useful for both TCFL teachers and learners of Chinese as this can give them a more

accurate and specific collocational pattern (Li, 2016).

If we compare Table 1 and Table 2, it is obvious that the size of a corpus is important for generating more valid research findings. Researchers need to consider carefully which corpus and corpus analysis toolkit best suit their research goals and make decisions accordingly.

A last note for the study is that due to the limit of space, this study does not provide introduction of the use of some fourth generation corpus analysis tools such as Sketch Engine (https://www.sketchengine.co.uk/) and BFSU CQPweb (http://111.200.194.212/cqp/), both of which are capable of performing Chinese collocations studies online. This may be offered in our future studies.

5. Implications for the teaching of vocabulary

In recent years a number of studies have been conducted on the use of corpus-based approach in L2 classrooms (Chan, 2002; Souza Hodne, 2009; Jafarpour, Hashemian, & Alipour, 2013). Some scholars also call this the data-driven learning approach. In this section, the authors use a few confusing words in Chinese to demonstrate how the collocation analysis can be helpful for learners not only to improve their collocational competence, reduce their erroneous use of collocations, but also to increase their awareness of collocations in the target language and therefore to produce more natural utterances.

In the HSK Dynamic Composition Corpus, some of the deviant use of 营造，造成，and 达成 are found:

(d) …男女分班制度由于缺少另一方而造成{CC 营造}了不平衡的生活圈子。
(e) …首先喜欢流行歌曲的人可以造成他们的同样的感情
(f) … 我认为流行歌曲可能会造成不好的文化，可流行歌曲的作用也不能忽视
(g) …可是对某些人类却造成{CC 达成}了很多负面的影响{CC 效果}。

It can be seen that these words are somewhat confusing to learners, as these words either share some constituents such as both 营造 and 造成 have a constituent 造，or they share a meaning that expresses “something is made to happen”.

In a TCFL classroom, teachers can ask learners to first search for these words’ high-frequency collocates using BCC corpus online concordancer, and then ask them to use the statistical function to list them on paper. Because main errors in the above sentences are the writers have used wrong nouns that collocates with verbs, teachers can ask learners to use regular expressions 营造*N，造成*N，and 达成*N to search for their instances in the corpus. Due to the space limit, only one of the screenshots of the statistical data is demonstrated here:

Figure 12. The high-frequency collocate list of 营造

Figure 12 shows that the high-frequency collocates include 环境，氛围，气氛，舆论，文化，空间，城市，优势，效果，景观， 市场，防护林等。 Among these words, most of them can be said to have an [ABSTRCT] semantic feature except 城市，景观，防护林。Teachers could then ask learners to check the modifiers of these nouns by click on the words listed in this statistical page. See the following screenshot:

Figure 13. Screenshot of the context of 营造*环境

A careful observation of the context can tell us that most of the abstract nouns are modified by words that have positive meanings such as 安全，公平，和谐，自然，良好，优美，健全 etc. Therefore, we can add one more semantic feature [POSTIVE] to the use of 营造*环境. The collocation pattern of 营造 therefore can be summarized as: Following such procedure, teachers can guide learners to work as groups to discuss and try to figure out this collocation pattern.

It is If TCFL teachers want the collocation pattern to be more accurate, they can use collocation analysis tools such as AntConc to conduct an analysis and generate a more reliable and comprehensive collocate list using the method we have demonstrated in the previous section.

Using the same procedure, teachers and learners should be able to work out the following collocation patters for the three words:

(1) 营造＋collocates [ABSTRACT/POSITIVE] except a few words that have concrete meaning such as 城市，景观，防护林 etc.

High-frequency collates: 环境，氛围，气氛，舆论，文化，空间，优势，效果，市场

(2) 造成 + collocates [NEGATIVE]

High-frequency collocates: 影响，死亡，损害，后果，浪费，危害，污染，伤害，伤亡，困难，问题，事故 etc.

(3) Two or more parties + 达成 + collocates [NEUTRAL]

High-frequency collocates: 协议，共识，一致，意向，意见， 妥协， 目的，交易，和解，默契，愿望，合作，效果, 契约 etc.

Using such lists, learners can work together to identify the reasons why sentences (d)-(g) are wrong, and then correct the sentences by themselves. Such process will help leaners to learn more productively about collocations and improve their collocation competence.

6. Conclusion

Corpus analysis tools and large scale corpora are becoming increasingly indispensable for studies of applied linguistics and second language acquisition. However, there are some challenges researchers face when employing these tools and resources. This is particularly the case for researchers who are less technically savvy. This study offers an overview of corpus analysis tools available for analyzing both English and Chinese text, and some of the most well-designed and well-compiled Chinese corpora that can be used for studies on Chinese collocation. The study then introduces the steps and procedures that are involved in using AntConc, a corpus analysis toolkit, to perform collocation analysis based on Chinese corpora such as ToRCH 2009 and BCC corpus. It is hoped that this demonstration may provide some useful information for those who are keen to employ similar tools in their own Chinese collocation studies.

References

Anthony, L. (2004). AntConc: A learner and classroom friendly, multi-platform corpus analysis toolkit. In L. Anthony, S. Fujita, & Y. Harada (Eds.). *Proceedings of IWLeL 2004: An Interactive Workshop on Language e-Learning.* Paper presented at the IWLeL 2004: An interactive Workshop on Language e-Learning, Waseda

University, Tokyo, Japan, 10th December (pp. 7-13). Tokyo: Waseda University.
Bai, M., & Zheng, J. (2004). Study on ways of verb-verb collocation. *Computer Engineering and applications, 27*, 70-72. [白妙青, 郑家恒, (2004). 动词与动词搭配方法的研究. 计算机工程与应用, *27*,70-72.]
Barlow, M. (2000*). MonoConc Pro*. Houston, TX: Athelstan.
Chan, M. K. M. (2002). Concordancers and concordances:Tools for Chinese language teaching and research. *Journal of the Chinese language Teachers Association, 37*(2), 1-58.
Chandler, B. 1989. *Longman Mini Concordancer*. Harlow: Longman.
Davies, M. (2005). The advantage of using relational databases for large corpora: Speed, advanced queries and unlimited annotation. *International Journal of Corpus Linguistics*, *10*(3), 307-334.
Davies, M. (2009). The 385+ million word corpus of contemporary American English (1990–2008+): Design, architecture and linguistic insights. *International Journal of Corpus Linguistics*, *14*(2), 159-190.
Davies, M. (2010). More than a peephole: Using large and diverse online corpora. *International Journal of Corpus Linguistics*, *15*(3), 412-418.
Hardie, A. (2012). CQPweb—combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, *17*(3), 380-409.
Henriksen, B. (2013). Research on L2 learners' collocational competence and development–a progress report. Retrieved from http://www.eurosla.org/monographs/EM02/Henriksen.pdf
Hockey, S. (1988). *Micro-OCP* (*OCP Version 2*). Oxford: Oxford University Press.
Jafarpour, A. A., Hashemian, M., & Alipour, S. (2013). A corpus-based Approach toward teaching collocation of synonyms. *Theory and Practice in Language Studies*, *3*(1), 51.
Kaye, G. (1990). A corpus-builder and real time concordance browser for an IBM PC. In J.Aarts & W. Meijs (Eds.), *Theory and Practice in Corpus Linguistics* (pp.137-162). Amsterdam: Rodopi.
Kilgarriff, A., Rychly, P., Smrz, P. & Tugwell, D. (2004). The Sketch Engine. In G. Williams & S. Vessier (Eds.), *Proceedings of Euralex 2004* (pp.105-116). Bretagne, France: Université de Bretagne-Sud.
Lewis, M. (Ed.). (2000). *Teaching collocations. Further developments in the lexical approach*. Hove, England: Language Teaching Publications.
Li, S. (2016). A corpus-based analysis of collocation errors by American learners of Chinese and its implication for the teaching of vocabulary. *Chinese as a Second Language -The Journal of the Chinese Language Teachers Association-US (CSL)*, *51*(1), 62-78.
Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied linguistics*, *24*(2), 223-242.
Palmer, F. R. (1981). *Semantics – A New Outline*. Cambridge: Cambridge University Press.
Rayson, P. (2008). From Key words to Key semantic domains. *International Journal of Corpus Linguistics*, *13*(4), 519-549.
Reed, A. (1978). *CLOC User Manual*. Birmingham: University of Birmingham.
Scott, M. (1996). *WordSmith Tools.* Oxford: Oxford University Press.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, J. M. (Ed.). (2004). *How to use corpora in language teaching* (Vol. 12). Amsterdam: John Benjamins Publishing.

Souza Hodne, L. C. (2009). *Collocations and teaching: Investigating word combinations in two English textbooks for Norwegian upper secondary school students* (Unpublished master's thesis). University of Bergen, Bergen, Norway.

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of language*, *2*(1), 23-55.

Sun, H., & Huang, C. (1998).The distribution characteristics of collocations in texts. *Proceedings of 1998 International conference on Chinese Information Processing*, Beijing, China (pp. 230-236). [孙宏林, 黄昌宁. (1998). 词语搭配在文本中的分布特征. *1998 中文信息处理国际会议论文集(pp. 230-236)*.]

Wu, W. & Li, S. (Eds.). (2009). *Linguistics and Chinese as a Second Language*. Hong Kong University Press, Hong Kong.

Xu, J., & Jia, Y. (2013). The design and development of the R-Gram-based corpus analysis software PowerConc. *Computer-assisted Foreign Language Education*, *1*, 57-62. [许家金, 贾云龙. (2013). 基于 R-gram 的语料库分析软件 PowerConc 的设计与开发. *外语电化教学, 1*, 57-62.]

Xu, J., & Wu, L. (2014). Web-based fourth generation corpus analysis tools and the BFSU CQPweb case. *Computer-assisted Foreign Language Education*, 5, 10-15,56. (许家金, 吴良平. (2014). [基于网络的第四代语料库分析工具 CQPweb 及应用实例. *外语电化教学, 5,* 10-15, 56.]

You, L., & Wang, S. (2005). Rules and distributions of Chinese verb-verb collocations. *Computer Engineering and applications, 23,* 179-181. [由丽萍, 王素格. (2005). 汉语动词——动词搭配规则与分布特征. *计算机工程与应用, 23,* 179-181.]

URL (网址): http://www.tclt.us/journal
Email (电子邮件): editor@tclt.us