Journal of Technology and Chinese Language Teaching

Volume 9 Number 2, December 2018
二〇一八年十二月　第九卷第二期

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: 1949-260X http://www.tclt.us/journal

Volume 9 Number 2, December 2018

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 Chinese Language and Literature, University of Macau
Department of World Languages, Literatures, and Cultures, Middle Tennessee State University
China Social Sciences Press (中国社会科学出版社)

Editorial board

Editorial staff

Contacts

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

Five Principles of a Relational Pedagogy: Integrating Social, Individual, and Material Dimensions of Language Use

Richard Kern
University of California, Berkeley
(加利福尼亚大学伯克利分校)
rkern@berkeley.edu

Abstract: This article views technology broadly in terms of its greatest contribution: textual communication. It argues that well-grounded pedagogy should not just focus on what is 'new' in digitally-mediated literacies, but should also focus on what digitally-mediated practices share with all forms of literacy. The article discusses material, social, and individual dimensions of language and literacy, and proposes five pedagogical principles that underlie literacy technologies from the origins of writing to the digital era. These five principles are the basis for a 'relational pedagogy' that aims to foster in students an ability to reflect on meaning-making practices broadly, with particular emphasis on how materials and technologies interact with social worlds and individual creativity. Pedagogical implications of these principles are presented to encourage a goal of critical symbolic awareness in twenty-first century language education.

Keywords: Literacy, technology, mediation, relational pedagogy, reading, writing, critical thinking, strategies

1. Introduction

Technology's greatest contribution to education has been the development of writing and textual communication. Textual communication is the subset of communication that gets passed on, that provides wide access to human knowledge and creativity, that becomes representative of people and cultures, that is archived and becomes historical record. For centuries, textual communication has taken the form of handwritten or printed texts. However, with the advent of film and digital media that allow speech to be easily transmitted and recorded, the scope of textual communication has broadened considerably. In the current era of intense technological and social change, educators need to think carefully about how they approach the 'new literacies' born of digital technology. In this paper I will argue that rather than attempting to distinguish between 'new' and 'old' literacies, educators should focus on what they share. That is to say, rather than attempting to predict what skills today's students might need at some future time, educators should focus on fundamental principles that have always underlain language, literacy, and communication broadly—principles that potentially help people see connections across modes of expression and between past and present practices, giving them a critical

perspective that will prepare them to understand and shape whatever future practices develop with technologies that have yet to be invented.

This of course begs the question: how might one identify such principles? In *Language, Literacy, and Technology* (Kern, 2015) I approached the question by studying past periods of technological and social change (e.g., the origins of writing in the third millennium BCE, the development of paper at the dawn of the Common Era, movable type and the printing press starting in the thirteenth century, the telephone in the late nineteenth century, and the computer in the twentieth and twenty-first centuries). My premise was that technological artifacts always interact with *what* we communicate, *how* we communicate it, and, broadly, how we interpret meanings. What I found was that how we use language on the internet, for example, relates in interesting ways to earlier technologies of language, and that the material infrastructures of our communication technologies have always been tied to particular communicative cultures that shape how we read, how we write, how we construe and share knowledge, and ultimately how we understand ourselves in relation to the world. I also found that these communicative cultures are always informed by relationships among material, social, and individual factors. In this paper, then, I will attempt to demonstrate the importance of these material-social-individual relationships and present five principles of what I call a "relational pedagogy" that aims to foster a critical symbolic awareness in students.

I should begin by explaining what I mean by literacy. Textual communication always relies on some form of material technology. Literacy is the know-how people need to deal with that technological mediation—not only the know-how to produce texts, but also the know-how to interpret them. The language teaching profession typically uses the terms "reading skills" and "writing skills" to talk about these abilities. I prefer to use the term "literacy" because it is more holistic and less oriented toward discrete skills. Literacy allows for a more unified discussion of relationships between readers, writers, texts, culture, and language learning. Literacy frames reading and writing as complementary dimensions of textual communication—and so we are led to focus on their *interrelatedness* rather than on their separateness as distinct skills. Literacy also highlights the importance of socialization and social practices, which are often less visible when reading and writing are treated primarily as internal, cognitive processes. Finally, literacy also lends itself to today's digital technologies that afford the possibility of creating texts that are not just linguistic but that integrate images, graphic layout design, color, font variation, and sometimes audio and video. These technologies of *textualization* are not easily assimilated under the rubric of "writing," and they require interpretative skills that extend beyond those of "reading." Although texts have never been purely linguistic in nature, they have never had as wide a range of potential forms as they have today—which is one reason why people often talk about *literacies* in the plural.

2. Material, Social, and Individual Dimensions of Language and Literacy

Making meaning with texts involves three primary kinds of resources:

- Material resources, such as paper, writing instruments, computers, smartphones, recorders, film, cameras and video cameras.
- Social resources, such as language and other semiotic systems, social practices, norms, conventions, cultural values, and ideologies.
- Individual resources, such as creativity, imagination, intuition, emotion, aesthetic sensibility, along with circumstantial factors such as available time, energy, and motivation.

These various resources shape our language use and literacy practices. Historically, the surfaces on which we have written (e.g., clay, bark, stone, palm leaves, papyrus, paper) have influenced the shapes of the scripts we use (Kern, 2015), and the devices we use today (e.g., smartphones, computers) continue to affect how we read and produce texts.

Of course, social resources, and especially language, are absolutely key to all forms of communication. But it is important to remember as well that all of our uses of material resources are socially constructed, and social resources influence the ways that we use technologies. Consider, for example, how conventions for beginning and closing conversations on the telephone differ by language and culture, but also by age and gender. Shaffer and Clinton (2006) point out that people do not *use* new technology objects so much as they *interact with and through* them (p. 289). Moreover, today's globalized information networks have their own 'sense-making' agency in the form of relevancy filtering, which determines which advertisements appear on our screens, what results we get from search engines, and what recommendations we get from online stores. All this is made possible by a social contract that makes internet users into providers of internet content

If language and literacy practices are always socially embedded, they also involve *personal* choices, drawing on individual resources to create an original voice. A writer's tone, rhythm, style, rhetorical flair, irony, and wit all contribute to the individuality of his or her voice. One excellent example of multilingual language play that contributes to personal rapport is Eva Lam's (2009) study of an Instant Messaging exchange between Kaiyee, a Mandarin-speaking seventeen-year-old immigrant who lived in a mostly Cantonese-speaking community in the United States, and Dawei, a male schoolmate whose family was originally from Taishan. Kaiyee impresses Dawei by writing in Cantonese, using both standard and non-standard characters. But she also aligns herself more personally with Dawei by writing in Taishanese, using a combination of alphabetic pinyin and Chinese characters. They both sporadically use English and Romanized spellings to write about things relevant to their American school context, as well as emoticons. In Lam's study, participants' communication is not organized and enacted within the bounds of a single language, but rather in terms of the teenagers' strategic uses of linguistic resources from Mandarin, Cantonese, Taishanese, and English; and graphic resources in the form of traditional and simplified characters, pinyin, Roman script, emoticons, and punctuation.

In today's digital environments, even when people are communicating in just one language, writing technology often makes them operate *multi-symbolically*. Consider, for example, the exchange between A and B below:

A: wuz^
B: nmhu?

A's utterance uses two distinct strategies: a phonetic representation of a youth speech variety of English ("what's" with a deleted /t/ and a voiced /s/) and a graphic representation of an upward pointing arrow, cueing an association between this icon and the word *up*. There is no question mark, so we must infer that this is a question from our prior familiarity with the expression "what's up?" B's response also mixes strategies. The first three letters are an initialism representing the first letters of the written words "not much, how about…" (or possibly "not much here"), but the u is a rebus, *phonologically* representing the whole word "you." Again, the reader has to switch processing conventions mid-stream from graphic to phonological in order to correctly parse the utterance. This can make reading difficult initially, especially because the strategy switching points are never marked but must be discovered by trial and error. But if the reader has been socialized into reading utterances like nmhu on a recurring basis, he or she will recognize the whole string *as a unit,* without added cognitive load.

Although many people talk about digital writing like this as simplified or reduced language, it is actually, at least initially, a *complexification* in terms of processing, on the part of *both* the writer and the reader.

These kinds of multi-symbolic complexities, resulting from interactions among material, social, and individual factors, have been with us from the very earliest days of writing. In Mesopotamia, for example, early clay tablets were bookkeeping devices and did not have any discernable syntax. While these tablets might seem extremely ambiguous to modern readers, they were not entirely unlike today's pared-down text messages, tweets, and emoticons. To further complicate things, a cuneiform sign for one word could be used to represent other words with the same or similar sound. But a single sign could also represent semantically related words that nevertheless had vastly different sounds. As a result of this homophony and polyphony, the intended meaning of cuneiform signs could often only be determined by context.

This context effect is of course still important today. When you see 10 on a chalkboard, your default assumption may be to read it as "ten," but if you were in a computer science class and working in binary code, you would read it as "two." Does 1/4 indicate a fraction? Yes, but it can also be a date, and depending on the language/culture, it might designate January 4, or it might designate April 1. The word *coin* may make you think of money, but it makes a Frenchman think of a corner, or the quack of a duck. How do you pronounce the word *entrance*? That depends on whether it is a noun or a verb, and we can only know that from context. In order to interpret meaning, we need to know what the relevant symbolic system is and what contextual relationships are operating.

So, what history teaches us is that at a very fundamental level, literacy is still the same as it has always been. It is about designing meaning from graphic signs with the resources and constraints of a particular medium and culture. It requires understanding relationships among forms, contexts, and the meanings they mediate. It is about expressing identity and affiliation through writing and sharing texts. And it is about becoming

socialized (and socializing others) into particular cultural practices related to texts—practices that both require and confer social power.

Of course, against the backdrop of these fundamental characteristics, the specifics of how people do things with texts have always varied widely and always will. Both across and within different cultures, people read, write, analyze, and think for myriad purposes, in myriad contexts, with different tools, and in different mediums.

In terms of preparing our students for the future, and whatever new technologies may be developed, it is hopeless to speculate about what kinds of future skills will be needed. Our best hope, in my opinion, is to expose learners to as broad a range of purposes, contexts, modes, and mediums of language use as we can—while focusing on fundamental principles that underlay language, literacy, and communication—to help learners see relationships across modes of expression and to foster a critical perspective that will prepare them to understand and shape whatever future practices develop with technologies that have yet to be invented.

3. The Five Principles of a Relational Pedagogy

The balance of the paper will describe the five principles I developed in *Language, Literacy, and Technology* and present some ideas for how they might be implemented in language teaching.

Principle 1. Meanings are situated and relational, not autonomous.

It is a basic fact that understanding language always involves contexts of interpretation. No text can signify in the absence of some context of interpretation. We saw this above in the examples of how context influences our interpretation of signs. There is nothing new about this. We have always known that signs mean different things in different cultures, for example. But today, the Internet brings multiple cultures into what we experience as *one place*, with participants who are often from different cultures, and we generally have fewer contextual clues to guide us in our interpretation. In fact, the Internet overlays *its own* culture on top of all the cultures represented by people interacting with one another online.

To give a trivial but interesting example, the internet trend of abbreviating words with numbers (like l8r in English) is ubiquitous across languages. But the systems underlying those abbreviations are multiple and are never explicitly marked, making interpretation difficult unless one can be socialized by practitioners. 143 in English designates “I love you” (based on the number of letters in each word). In Chinese, 88 is ba ba (which sounds similar to English bye bye, meaning two separate systems come into play) but in German, 88 is a neo-Nazi symbol (H is the eighth letter of the alphabet, forming "HH” or Heil Hitler). In Japanese, 39 is pronounced “sankyuu” (just like 3Q in Chinese); but in Dari and Pashto, 39 is a swear word. 555 designates laughter in Thai (ha ha ha), but in Chinese, it indicates crying (wu wu wu). Then, in Korean as well as some other Asian languages, OTL or Orz operates within an iconic system, providing the viewer with a “side view” of a stick figure kneeling out of disappointment or bowing to show deference. OZ is

the simplified version. The basic principle here is that you have to know the relevant system behind the signs. The signs don't have intrinsic meanings in and of themselves. From this perspective, we need to think of computer-mediated exchanges as what Mary Louise Pratt (1991) calls contact zones: "social spaces where cultures meet, clash, and grapple with each other" (p. 34).

A less arcane example is the fact that words and phrases often activate mental representations of particular contexts when we hear or read them. Consider the contexts of situation you imagine when you hear utterances like "Catch a tiger by the toe," or "All hands on deck!" or "This one's on me." When these contexts are not understood, problems can arise both at the level of interpretation (e.g., "This one's on me" interpreted as "It's my fault") and at the level of production (e.g., B responding "What a nuisance" when A says "Terry's father died") (Richards, 1990, p. 75). Strategic pedagogy that focuses on guessing word meaning from context is important, but it addresses only one side of the issue—teachers also need to recognize the importance of the *inter*textual resonances of texts, developing a sense (over time) of the contexts associated with certain words, expressions, genres, styles and so forth.

One pedagogical goal for applying Principle 1 is to develop learners' awareness of how changes in framing and context affect meaning. One tool we've developed at Berkeley that is helpful in this regard is the Berkeley Language Center's Library of Foreign Language Film Clips (Lumière), which holds 4,200 films in 54 languages, with 19,000 clips tagged for vocabulary, grammatical structures, speech acts, and cultural notions. This allow instructors to search for a particular speech function (e.g., apologizing) and get multiple examples involving different situations, different historical periods, and different characters from different age groups, social classes, and regions. Instructors can then ask students to analyze certain features as they view the clips. Information about this database (which is available to other institutions) can be found at blcvideoclips.berkeley.edu.

Another way of approaching this pedagogical goal is to ask questions that show different dimensions of "meaning." That is, distinguishing among *referential* meaning (i.e., what do the words refer to, according to their framing?), *metaphorical* meaning (i.e., are words being used literally or metaphorically? What is the referent being compared to, and what is the effect of that comparison?), *structural* meaning (i.e., how does the particular ordering of words and clauses contribute to your interpretation? how might changing the order affect meaning?), *intertextual* meaning (i.e., how do elements of this text echo elements of other texts you have encountered? What is the effect of that echo?), *social* meaning (i.e., what kind of relationship do the words of the text establish between the writer and reader? or between characters in the text? what conventional meanings are attached to the genre or form of the text?), *personal* meaning (i.e., what feelings do the words of the text evoke in you? What personal experiences are called to mind?), *symbolic* meaning (i.e., beyond metaphors, are textual elements used in allegorical or emblematic ways? Can the text itself be considered a symbolic act?), and *ideological* meaning (i.e., whose interests are served by the text? are meanings consistent with or do they challenge dominant discourses?). It is important for students to realize that these levels or frames of meaning operate in many different kinds of texts, and not just literary texts.

Principle 2. Language, literacy, and communication rely on both convention and invention.

When we produce an utterance or a text we are creating something new and unique in relation to a particular context, and yet to do so we must rely on resources and practices that are well established. In other words, we are recycling old materials in fresh ways, establishing new relationships among stock elements. This means that all our acts of communication are social and historical as well as individual. That is, they involve both convention and invention.

Emergent grammar is one example of this principle. An utterance like "my red new car" is typically considered ungrammatical, with English grammar telling us that "my new red car" is correct. But suppose I have recently bought *two* new cars, one red and one yellow. Now when I tell you about the red one, as opposed to the yellow one, it is suddenly okay to say "my red new car." Here the structure has been adapted to fit the particular needs of people in a specific communicative situation. Paul Hopper (1998) explains that grammar is not immutably fixed in a speaker's head, but is negotiated in particular situations. Language forms are derived from prior practice (sedimented in spoken and written texts) but rearranged in sometimes novel ways.

We see convention interacting with invention all the time on the internet. A recent article in Wired Magazine (Anderson, 2018) describes how feminists in China have used emoji to outwit government censors of sexual harassment #MeToo forums. Using juxtaposed emoji of a bowl of rice and a bunny rabbit (spoken aloud as "mi tu" in Mandarin), they have succeeded in avoiding censorship.

One pedagogical goal for applying Principle 2, then, is to show the importance of social conventions in discourse, but also show how people adapt conventions and resources for individual and collective purposes. Getting students to reflect on conventions can be approached in terms of both process and product. Learners can think about how social conventions develop in response to material limitations (e.g., the keys of a cell phone or limits on characters in Twitter) and how in their own experience they have "worked around the system" when they have encountered obstacles in accomplishing a communicative goal. They can also do stylistic analyses of others' texts in terms of the graphological, phonological, lexical, syntactic, semantic, and pragmatic features, paying special attention to the aesthetic or persuasive *effects* of the writers' choices on their interpretations of those texts (see Fowler, 1986; Simpson, 2014; Widdowson, 1992 for more on stylistics in language learning).

A second pedagogical goal for Principle 2 is to make learners aware of their own agency in choosing and configuring the semiotic resources they use. Students can record and transcribe segments of their own or others' face-to-face communication and compare the features they find with those of various forms of online communication (emails, text messages, chats, videoconferences, etc.), reflecting on how the forms chosen relate to the material and social dimensions of the situation. This kind of activity dovetails with the educational goals of the third principle.

Principle 3. The medium matters.

Writing means different things in different mediums. For example, emails, text messages, and handwritten letters are all forms of written correspondence but they are produced and read differently from one another. Similarly, digital storytelling involves a different constellation of knowledge, skills, and practices than composing a prose narrative. Consequently, people can be quite literate in certain mediums but not in others, and no one is literate in all possible mediums.

The pedagogical goals proposed for Principle 3 center on developing learners' ability to reflect on how communication is shaped at least partly by its material context. Teachers can ask how performing some act (such as persuading, apologizing, inviting, breaking up) would be different in one medium versus another, how it would be influenced by the affordances of each medium. For example, a phone call allows the speaker to gauge an interlocutor's immediate reaction to what is said; a letter doesn't, but it can be drafted and rewritten to help the writer find just the right tone. Comparing book and film mediums is another way to address Principle 3. Students can analyze a scene of a novel that has been made into a film, and compare the textual and filmic scenes feature by feature. Whose point of view is expressed? How is it expressed? Does it change during the scene? If so, how is the shift marked? How does the director express in film what the author expresses in an interior monologue written in free indirect discourse? Or students can even make their own filmic transformations of written texts (Porter, 2009).

Another dimension of Principle 3 is to analyze mediums critically for ideological or commercial underpinnings. The goal here is to get students thinking about how writing systems and other technologies of literacy have histories and ideologies embedded in them. For example, how simplified characters were developed in the People's Republic of China to facilitate the spread of literacy among the masses, but traditional characters were retained by Taiwan, Hong Kong, and Macau—and how learners of Chinese are making a *political choice* when they learn to write in one system versus another; how speakers of Hindi/Urdu cannot avoid revealing their religious affiliation when writing, since the Nastaliq and Devanagari scripts mark Muslims and Hindus respectively; how Arabic and Chinese scripts present the appearance of linguistic unity whereas spoken varieties are often mutually unintelligible; or how most formerly colonized peoples of the world are permanently marked as such by using the writing system (if not the language) of their colonizers. With respect to digital technology, how spelling/grammar checkers militate against anything unusual (e.g., spelling to imitate speech, incorporating foreign words, using neologisms), and how the autocorrect feature changes non-normative forms often without the writer's awareness; how PowerPoint's default settings may influence how we organize information and how much information we organize; how Facebook profiles are constrained by the categories of information authorized by Facebook. Teachers should emphasize students' critical reflection about the Internet, which collects their online habits, their search queries, their purchases, the music they listen to, the information they put on Facebook and other social media sites—all of which is then used by companies to personalize the information and advertisements they see on their screens.

Principle 4. Texts and communication are always multimodal.

We naturally communicate multi-modally. When we talk, the sounds, rhythms, and pacing of our speech play an important role in constructing meaning. When we write, our handwriting, the layout of our text, the font, the physical surface on which we have written, and how we mark that surface all contribute something to the meaning. Consider this haiku by Anita Virgil (2004), which highlights the iconicity of the letter i:

> speeding along the awning's edge
> raiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiin

Today we are faced with more choices than ever about how to communicate. What should we express through pictures or video? What should be expressed in speech? What in writing? And what will it 'say' to express it in one medium versus another? As we think about modes of expression like music, theater, dance, and visual art, what are the specific dimensions of each that contribute to the construction of meaning?

One pedagogical goal for Principle 4 is to get students to reflect on how linguistic and non-linguistic elements interact in texts, as well as in face-to-face communication. Starting with writing, have students interpret the effects of typefaces (styles and sizes) and page layout design in a variety of text types. How are visuals or colors used? How do all these visual elements complement or detract from the verbal message? In face-to-face communication, have learners analyze recorded interactions to see how gesture, body language, facial expressions, tone of voice, and other nonlinguistic features affect the participants' co-construction of meaning.

Another goal is to get students to develop rhetorical skills in a range of modes. For example, have students analyze how language interacts with music, sound, and visual design in theater, film, and other multimodal forms of expression. Have them listen to the musicality of Martin Luther King's delivery of his "I Have a Dream" speech and discuss what effect that musicality has on their reception of the speech as compared to reading the script of the speech silently or listening to someone else reading it aloud.

A third goal is to develop learners' awareness of codes and programming in digital multimodal environments. Students should be exposed to the human actions that underlie the surface of what they see and hear on digital devices—and have an opportunity to discuss the values implicit in programming decisions. They should also be required to think about the ethical responsibilities that go along with that power and how the design of an interface may encourage certain kinds of behavior, both positive and negative (for example, an interface based on user anonymity could promote honesty or it could foster flaming or cyberbullying).

Principle 5. Language, technologies, and texts mediate between the social and the individual; between ourselves and real and imagined worlds.

Finally, the fifth principle frames language, technologies, and texts as kinds of filters through which we interact with the world. Reading allows us to extend our knowledge far beyond what we experience directly in the world. The internet has been

tremendously beneficial in that we now have access to information that was previously impossible or difficult to obtain, making our "world" much larger. But at the same time, personalization algorithms on the internet give rise to what Pariser (2011) calls "identity loops, in which what the code knows about you constructs your media environment, and your media environment helps to shape your future preferences" (p. 233). Relevancy filtering effectively *narrows* one's exposure to new information, as echoed by Manjoo (2016): "when confronted with diverse information choices, …we gorge on information that confirms our ideas, and we shun what does not" (p. B1). In this way, technologies have the potential to negatively shape our behavior and our thinking when we interact with them often.

Just as language, technology, and texts mediate 'outside in' between the world and ourselves, they also mediate 'inside out' between ourselves and the world around us. Sometimes this is overt and structural (as in the case of a Facebook profile). Other times it is more subtle, as in lexical and grammatical choices, handwriting style, language play, pseudonym choice, emoticons, and other affect-laden visual dimensions of writing. These examples remind us that affective and aesthetic dimensions of written forms play an important role in making meaning.

One pedagogical goal associated with Principle 5 has to do with developing learners' awareness of mediation and its role in learning. This can be approached by getting students to understand the affordances of different kinds of mediation in their own learning. To do this, they can reflect on how different learning mediums (e.g., books, lectures, online forums) position learners differently, how they embed different ideas about how power circulates among participants, how jointly (or not) knowledge is constructed, how much (or how little) dialogue can occur, and how controlled or free-flowing the interaction is. Books afford slow, careful reading and critical reflection, whereas lectures allow listeners to hear the speaker's voice, to witness a performance, and to interact directly with the speaker. Online forums afford highly interactive discussion of a topic in a structurally flat power hierarchy, and they are "textual" so interactions can be saved, re-read, and analyzed. With each learning medium, learners are encouraged to think and act in certain ways.

A second pedagogical goal related to Principle 5 is to develop learners' awareness of how people create social identities through their use of language and technologies. They can consider how people's speech patterns, accent, diction, pace and rhythm, all contribute to an identity image they project when they talk. In writing, students can consider how handwriting and expressive style are interpreted by others as revealing something about themselves. To do this, students can pass around samples of their handwriting anonymously and comment on what they infer about the writer's personality from the size, style, or color of the writing itself. The point is not to show that there is uniform consensus about what a particular style of handwriting reveals about the writer, but rather to demonstrate that handwriting style is *susceptible to interpretation*. Students can then do research on cultures where employers traditionally ask for handwritten, rather than typewritten, cover letters.

A third pedagogical goal tied to Principle 5 is to evaluate the authenticity and validity of information. In today's media environment, fake documents can be made to

look authentic, and the ability to critically evaluate the source and quality of information is particularly important. Leu et al. (2007) reported on seventh grade readers' ability to critically evaluate information online by creating a spoof website, *Save the Pacific Northwest Tree Octopus*, which gave information about the tree octopus with the support of photos, graphics, and links to external resources. They found that even the most proficient young readers could be fooled about the reliability of information they found on the Internet, even when they were well aware of how unreliable online information can be. In this case, the majority of students not only believed the fabricated information, but many also persisted in their belief that the tree octopus existed, even after researchers explained that the information had been made up. Students can be given practice in discriminating between authentic and illegitimate materials by focusing on subtle details of language, layout, and provenance of the text, and they might even try designing their own false texts to see if they can fool their classmates, discussing the results afterwards.

Another, even more subtle kind of authenticity that is harder to even be aware of, much less reflect critically on, is that of cultural materials typically found in textbooks. Vinall (2012) argues that a central problem in foreign language teaching is dealing with historical events, which are too often presented in textbooks as lists of facts attached to a specific time and place, without critical reflection about how those seemingly neutral facts have been constructed, how they have been transformed over time, and what feelings they evoke in people. Vinall takes as her example the history of the Conquest of the Americas as represented in an intermediate-level Spanish textbook. She proposes a three-phase approach in which students reflect critically on discourse, explore issues of power, and reframe the discourse world of the textbook.

Finally, Principle 5 acknowledges the importance of imagined worlds, suggesting a pedagogical goal tied to aesthetic dimensions of communication. Students' attention can be focused on their visceral responses to texts based on material considerations (for example, the smell and feel of a leather binding, the weight and texture of the paper, the feel and flow of a pen, the brightness and resolution of a screen). Roland Barthes (1977) posited the phrase "the grain of the voice" to describe the pleasure or displeasure that one gets just from the sound of a performer's voice without taking into account the words uttered or the notes sung. His point was that raw sensory impressions cannot be separated from the ultimate meaning we derive from producing or interpreting a text. Roman Jakobson (1960) made a somewhat similar point with respect to form in language when he described the poetic function of language. Students can be asked to attend to sounds, sound patterns, visual forms, intonation, and language play from an aesthetic standpoint. As Cook (2000) has argued, language use is not always rational and transactional, and language play allows learners to see how contexts of interaction motivate and shape language forms—and how language forms themselves can sometimes motivate and shape interaction.

The appendix recapitulates these five principles of a relational pedagogy with sets of broad questions to guide teachers' planning. It is important to recognize that these principles and questions are heuristic in nature, designed to get teachers and students thinking, discussing, and learning together. They are not intended to be definitive or programmatic, but rather as starting points for teachers and learners to explore and extend in their local contexts.

4. Conclusion

Technology plays an important role in the educational goals outlined above. As Dourish and Bell (2011) point out, "The technologically mediated world does not stand apart from the physical one within which it is embedded; rather, it provides a new set of ways for that physical world to be understood and appropriated" (p. 132). It is in the spirit of fostering reflection on relationships between physical worlds, cultural worlds, and technologically mediated worlds that I have proposed the five principles outlined in this paper.

Whereas language teaching tends to emphasize the "thingness" of language and literacy by focusing on vocabulary, grammar rules, styles and genres as so many *items,* or static facts, what I'm calling a relational pedagogy attempts to draw learners' attention to the crucial "in between" relationships that bring those items to life. It shifts emphasis from unvarying *a priori* rules to appreciation of how mediums, cultural practices, situational circumstances, and individual creativity interact when people make meaning—and how those interactions are reflected in language forms such as orthography, grammar, and genres.

In other words, a relational pedagogy aims to foster an ability to reflect on meaning-making practices broadly, but with particular emphasis on how materials and technologies interact with social worlds and individual creativity in those practices. A relational pedagogy deals with fundamentals, teaching language and literacy in the broadest sense. It aims for communicative competence but also for critical competence and symbolic competence. It aims to get students to see that language is not just a normative system, but also an adaptive practice that interacts with its cultural and technological mediations.

By focusing on the fundamental dimensions of meaning-making, by showing learners how material, social, and individual factors interact (and do so differently in different contexts of communication, across different situations of technological mediation, and across different moments in time), we stand the best chance of preparing students for the future while simultaneously connecting them to the past—and, crucially, helping them to see relations between the two.

References

Anderson, M. (2018, March 30). How feminists in China are using emoji to avoid censorship. *Wired*. Retrieved from https://www.wired.com/story/china-feminism-emoji-censorship/

Barthes, R. (1977). *Image music, text* (S. Heath, Trans.). London: Fontana Press.

Cook, G. (2000). *Language play, language learning*. Oxford: Oxford University Press.

Dourish, P., & Bell, G. (2011). *Divining a digital future: Mess and mythology in ubiquitous computing*. Cambridge, MA: MIT Press.

Fowler, R. (1986). *Linguistic criticism*. Oxford: Oxford University Press.

Hopper, P. J. (1998). Emergent grammar. In M. Tomasello (Ed.), *The new psychology of language* (pp. 155-175). Mahwah, NJ: Erlbaum.

Jakobson, R. (1960). Closing statement: Linguistics and poetics. In T. Sebeok (Ed.), *Style in Language* (pp. 350-377). Cambridge, MA: MIT Press.
Kern, R. (2015). *Language, literacy, and technology*. Cambridge: Cambridge University Press.
Lam, W. S. E. (2009). Multiliteracies on instant messaging in negotiating local, translocal, and transnational affiliations: A case of an adolescent immigrant. *Reading Research Quarterly, 44*(4), 377-397. doi:10.1598/rrq.44.4.5
Leu, D. J., Zawilinski, L., Castek, J., Banerjee, M., Housand, B. C., Liu, Y., & O'Neil, M. (2007). What is new about the new literacies of online reading comprehension? In L. S. Rush, A. J. Eakle, & A. Berger (Eds.), *Secondary school literacy: What research reveals for classroom practice* (pp. 37-68). Urbana, IL: National Council of Teachers of English.
Manjoo, F. (2016, November 2). How the Internet is loosening our grip on the truth. *New York Times.* Retrieved from https://www.nytimes.com/2016/11/03/technology/how-the-internet-is-loosening-our-grip-on-the-truth.html
Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. New York: Penguin.
Porter, J. (2009). Filmmaking and foreign language instruction. *Berkeley Language Center Newsletter, 24*(2), 10-11.
Pratt, M. L. (1991). Arts of the Contact Zone. *Profession*, 33-40.
Richards, J. C. (1990). *The language teaching matrix*. Cambridge: Cambridge University Press.
Shaffer, D. W., & Clinton, K. A. (2006). Toolforthoughts: Reexamining thinking in the digital age. *Mind, Culture, and Activity, 13*(4), 283-300.
Simpson, P. (2014). *Stylistics: A resource book for students* (2nd ed.). New York: Routledge.
Vinall, K. (2012). ¿Un legado histórico?: Symbolic competence and the construction of multiple histories. *L2 Journal, 4*(1). Retrieved from https://escholarship.org/uc/item/99b712hv
Virgil, A. (2004). *Summer thunder*. Forest, VA: Peak's Press.
Widdowson, H. G. (1992). *Practical stylistics*. Oxford: Oxford University Press.

Appendix

Heuristic Questions for the Five Principles (Kern, 2015)

Principle 1: Meanings are situated and relational

- What are the contexts relevant to the interpretation of this text (e.g., material, situational, social, historical, ideological etc.)? How might the text and context inform one another?
- How does this text allude to, contest, build on other texts, even in other mediums?

Principle 2: Language, literacy, and communication rely on both convention and invention

- How have conventional semiotic resources been appropriated, adapted, or recontextualized for individual or collective purposes in this text? To what effect?

Principle 3: The medium matters

- How does the text's medium affect language *form*? Are words written/spoken differently? Is syntax modified? Is text length affected? Are cohesion and coherence devices the same and if so, are they used in the same ways?
- How do such differences affect listening, reading and writing?
- How could the meanings expressed in this medium be expressed/re-mediated in a different medium? (to achieve a similar or a different effect)

Principle 4: Texts are multimodal

- How do linguistic elements interact with nonlinguistic textual design elements to produce particular meanings?
- How are time (e.g., rhythm, timing) and space (e.g., visual layout, movement) used to create particular meanings or effects?
- What communicative acts (e.g., establishing rapport, sharing ideas, persuading, negotiating, expressing feelings) are facilitated or rendered more difficult by the medium?
- What are the social consequences (in terms of who is included or excluded, how participant interactions might be reconfigured, how cultural processes and products might be affected) of using one medium and technology versus another?

Principle 5: Language, technologies, and texts mediate

- How does our use of language, technologies, and texts affect how we think about, produce, and use knowledge?
- How are traces of the communicator's identity or persona signified?
- How do aesthetic qualities contribute to meaning?
- Whose interests are at stake, and how are those interests identifiable? Are beliefs, attitudes, myths, and assumptions marked as such or can they be mistaken for facts?

Mapping Stylistic Variation with Correspondence Analysis (以对应分析统计法图示语体变异维度)

Zhang, Zheng-sheng
(张正生)

San Diego State University
(圣地亚哥州立大学)
zzhang@sdsu.edu

Abstract: This article demonstrates the use of an alternative method for carrying out multi-dimensional research on stylistic (or register) variation. Correspondence Analysis (CA) is an alternative statistical procedure to Factor Analysis, which has been employed in most multi-dimensional studies on stylistic variation. The advantages of CA are ease of use and intuitive visualization in the form of "stylistic maps." The method will be illustrated with the author's work on stylistic variation in written Chinese as well as a pilot study using English.

提要：本文介绍一种多面向语体研究的另类统计方法。对应分析（Correspondence Analysis）是普遍使用的因子分析的一种变异形式。此法的优点是容易使用并可提供直观的语体分布图。方法的演示将以中文书面语体研究为例，以及英语语体的一个初步研究。

Keywords: Stylistic variation, factor analysis, correspondence analysis, stylistic map

关键词：语体变异、因子分析、对应分析、语体分布图

1. Multi-feature, Multi-dimensional Study of Stylistic/Register Variation

Even though the focus of this paper is on illustrating the use of an alternative methodology, a brief introduction is nonetheless in order on stylistic/register variation and the multi-dimensional framework for the study of such variation. For more in-depth discussion of these matters, the reader is advised to refer to a series of publications by the author listed in the references, especially Zhang (2017).

1.1 Stylistic/register variation

It may be safe to say that stylistic/register variation exists in all languages. Stylistic/register characteristics, such as formality and literariness, can be seen in different

areas of language, most notably in the lexical domain. In English, a contrast can be seen between the colloquial *eat*, *buy* and their formal counterparts *dine*, *purchase*; in Chinese, there are also lexical doublets, such as 买 [buy] and 购 [purchase], 在 [at] and 于 [at]. One commonly evoked stylistic/register distinction is that of spoken vs. written or formal vs. informal. More subtly, as demonstrated by Biber, Douglas, Johansson, Leech, Conrad, & Finegan (1999), even basic syntactic categories, such as parts of speech, can be shown to have certain stylistic affinities. They show that a greater frequency of nouns and nominalized elements seems to be a hallmark of formal written texts.

1.2 Problem with simple binary distinctions

A single dichotomous distinction such as spoken vs. written simply cannot accommodate the complexity of stylistic variation. For example, this distinction may not align with another commonly evoked distinction of formal vs. informal. What is written in style may not always be formal, and vice versa. The same can be said about a distinction based on formality. In the case of Chinese, even though classical Chinese elements dating back to the pre-Qin era, such as 与 ("and," which corresponds to 和 in modern Chinese), have a close affinity with the modern written style, they may not always be the most formal. As we will see later, the most formal registers actually do not have the most classical Chinese elements. In order to accommodate the intricacies of stylistic/register variation, obviously more than one distinction is required.

Past work on stylistics was mostly on features in isolation, such as word length and lexical and syntactic choices, in the absence of a broader account of how such features may figure in the overall scheme of stylistic variation. As an example of specific problems with single features, categorization based on single features can be problematic, as different features may lead to contradictory classifications. For example, word length as feature cannot be used to classify texts into written vs. non-written types. Monosyllabicity, a distinct characteristic of classical Chinese often associated with the written style, cannot be used to characterize modern written Chinese as a whole, which tends to favor disyllabic words.

A third problem is methodological. Much of previous work on stylistics was based on introspection and anecdotal evidence, without empirical and quantitative support. For example, the stylistic markings in dictionaries, mostly based on the compilers' intuition, are neither complete nor consistent (Zhang, 2017). Introspection is also limiting. The aforementioned stylistic affinity of syntactic categories like nouns and morphological processes like nominalization will be hard to ascertain by introspection alone, without using corpora and statistical methods. Finally, without the support of quantitative information, it will be hard to go beyond simple dichotomies and entertain the possibilities of continuums, which are gradient in nature.

1.3 Multi-feature and multi-dimensional (MM) framework

Biber (1988) was perhaps the most influential work in introducing a multi-feature and multi-dimensional (MM) paradigm. MM-style research is different in three major ways:

a. multiple features, instead of single features, are examined simultaneously
b. multiple dimensions, instead of a single distinction, are entertained
c. more empirical, being corpus-based and quantitative in methodology

Using 67 mostly grammatical features, 23 registers from spoken and written corpora (such as press reports, official documents, and speeches) and the statistical method of Factor Analysis, Biber (1988) extracted six dimensions of register variation for English, driving home the fact that no single feature/dimension can characterize register variation:

1. Informational vs. involved production
2. Narrative vs. non-narrative
3. Explicit vs. situation-dependent reference
4. Overt expression of persuasion
5. Abstractness/nonspecific
6. Online informational elaboration

Since Biber's initial study on English, a number of MM studies of register variation have been carried out for other languages, including one on Taiwanese (Jang, 1998) and a number of them on Mandarin Chinese by the present author (Zhang, 2012, 2013, 2016, 2017, & forthcoming).

According to Zhang (forthcoming), two dimensions, rather than a single dichotomous distinction, characterize the style in modern written Chinese. In addition to the literate dimension, which is related to the common spoken vs. written (or formal vs. informal) distinction, there is a separate dimension, dubbed Alternative Diction, concerning the manner of expression. This will be explained in Section 4, where the two dimensions are presented in detail.

Although not couched in the same terms, the idea of two dimensions also seems to be implicit in Feng's (2010) analysis, which is mostly derived conceptually rather than using corpus data. His analysis separates situation (正式 [formal], associated with 现代书面语 [modern written Chinese]) from diction (庄典 [dignified and elegant], associated with 古代词语 [classical diction). Texts using more classical diction are not necessarily more formal, nor vice versa. To show the disassociation of diction from formality, Feng uses the examples of the *Yellow Emperor Epitaph* [黄帝祭文] and *Romance of the Western Chamber* [西厢记]. Although the *Yellow Emperor Epitaph* is both formal and classical, the classically-worded *Romance of the Western Chamber* is not formal at all.

1.4 Gathering data in MM-style variation research

Research on stylistic variation requires frequency data of the relevant linguistic features in various contexts, such as registers. Therefore, the initial steps in a MM-style study are:

a. Selecting stylistically relevant linguistic features
b. Selecting a balanced corpus with multiple types

c. Collecting frequency data of features in various types

In what follows, each of these steps will be illustrated.

1.4.1 Feature selection

To avoid any *a priori* assumption, in principle any feature that is potentially relevant to stylistic variation should be included. It can be a lexical item, a phrase, a syntactic category, such as noun and verb, and even a whole construction. However, there are practical considerations that limit the number and type of features that can be included. First of all, only those that can be searched for with little or no manual work can be included. Searching for discontinuous strings such as 除了… 以外 [in addition; except for] will require extra work. To ensure reliability, only features with high enough occurrences should be selected. This consideration favors whole parts of speech and function words. In order for the features to be visually legible on the bi-plots generated by SPSS, the number of features should not be too large either.

1.4.2 Balanced corpora

As the most important information in stylistic variation concerns how style varies in different contexts, the selected corpus needs to provide such variant contexts, namely multiple registers. In other words, the corpus needs to be balanced. One example of balanced corpora is the Lancaster Corpus of Mandarin Chinese (LCMC), which includes as many as 15 registers, despite its modest size of one million words (see Table 1).

Table 1. LCMC registers

Register	Abbreviated label
News reportage	NewsRep
News editorials	NewsEd
News reviews	NewsRev
Religion	Religion
Skills, trades and hobbies	Skill
Popular lore	PopLore
Essays and biographies	Biography
Reports and official documents	Official
Science (academic prose)	Academic
General fiction	FicGen
Mystery and detective fiction	FicDec
Science fiction	FicSci
Adventure/martial arts fiction	FicMart
Romantic fiction	FicRom
Humor	Humor

As can be seen, LCMC is quite finely differentiated. Some major register categories have subcategories. Fiction alone has five different subtypes and the journalistic type have three subtypes. Such fine differentiation allows us to gain insight into intraregister as well as interregister variation.

Another well-known balanced corpus is the much larger BCC corpus from Beijing Language and Cultural University, albeit with only four types: 文学 [literature], 报刊

[press], 微博 [tweets], and 科技 [science and technology]. Its large size will prove vital in studying lower frequency lexical items, such as near synonyms.

1.4.3 Collecting frequency data

Frequency data of features in registers can typically be obtained by searching the chosen corpus through an interface. For example, LCMC and several other similarly structured corpora (UCLA, ZCTC) are available online at the Beijing Foreign Studies University CQP website (http://111.200.194.212/cqp/), which hosts altogether 41 corpora, both Chinese and English. The BCC corpus has its own online interface, which provides rather extensive search options. The frequency information is given both in raw counts and normalized counts (typically per million words), which facilitates comparisons between registers.

The retrieved frequency data can be first organized in an Excel file, before they are entered into SPSS. Seen in Figure 1 is a partial screen capture using the LCMC corpus with its 15 registers.

	ADJ	V	N	ADV	adj	adv	n	v	p	deN	deV	Vde
NewsRep	2.9	11.7	18.58	4.9	0.14	0.018	0.84	0.225	3.1	4.405	0.303	0.12
NewsEd	3.1	13.3	18.23	5.824	0.14	0.031	0.68	0.283	3.31	5.216	0.173	0.128
NewsRev	3.4	12.6	20.03	4.655	0.07	0.014	0.36	0.091	3.62	6.352	0.238	0.053
Religion	2.8	12	17.69	5.515	0.33	0.035	2.03	0.84	3.65	5.713	0.13	0.046
SkillHob	4.3	14.7	18.9	5.518	0.26	0.037	2.01	0.453	3.79	4.45	0.166	0.113
PopLore	3.5	13	17.31	6.136	0.14	0.029	1.21	0.305	3.18	4.975	0.271	0.144
BioEssay	3.1	12.9	14.93	6.293	0.21	0.03	1.13	0.347	3.2	4.376	0.362	0.166
Official	2.2	13	24.93	2.53	0.06	0.006	0.35	0.119	3.4	4.481	0.18	0.017
Academic	3.3	12.1	21.46	4.958	0.13	0.017	1.06	0.271	3.75	6.536	0.229	0.054
FicGen	3.6	13.7	13.24	7.166	0.19	0.021	1.43	0.252	2.56	3.975	0.512	0.303
FicDec	2.8	12.9	14.42	6.461	0.38	0.043	1.31	0.453	2.83	4.169	0.4	0.201
FicScifi	3.8	13.1	13.3	6.787	0.15	0.062	0.86	0.154	2.69	5.401	0.678	0.131
FicMart	3.2	13.1	12.75	7.866	0.46	0.138	2.57	0.877	2.34	2.737	0.269	0.323
FicRom	3.6	12.9	11.83	7.485	0.23	0.053	1.07	0.26	2.73	4.39	0.663	0.308
Humor	2.4	15	14.28	6.433	0.07	0.02	0.8	0.21	2.07	2.608	0.399	0.261

Figure 1. Data Matrix in Excel file (columns= features; rows=registers)

2. Two Statistical Procedures

To make sense of the large amount of frequency data gathered, statistical methods have to be employed to extract a smaller number of underlying factors or dimensions, which may reveal some meaningful patterns. This kind of procedure is generally referred to as dimension reduction. In this section, two kinds of dimension reduction procedures will be briefly described, namely Factor Analysis and Correspondence Analysis.

2.1 Factor Analysis

A method such as Factor Analysis is commonly used to reveal underlying factors hidden in the bewildering amount of variability. According to Wikipedia

(https://en.wikipedia.org/wiki/Factor_analysis), "Factor analysis is a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors**"** (see Gorsuch, 1983, for a classical exposition of the method).

In the field of teaching Chinese as a foreign language, Factor Analysis has been used in the study of learning strategies, such as those for learning Chinese characters (Shen, 2005) and in the study of rating criteria used in assessment of learners of Chinese (Chen, 2016).

For the non-statisticians though, Factor Analysis presents quite a steep learning curve. Correspondence Analysis (CA), which also can accomplish the reduction of large amount of data into smaller number of dimensions, provides a more user-friendly alternative.

2.2 Correspondence Analysis (CA)

Even though Factor Analysis has been the preferred method for dimension extraction in MM-oriented research, the present author's studies have all employed Correspondence Analysis. According to an online publication:

> In a nutshell, correspondence analysis (CA) may be defined as a special case of principal components analysis (PCA) of the rows and columns of a table, especially applicable to a cross-tabulation. …... Its primary goal is to transform a table of numerical information into a graphical display, in which each row and each column is depicted as a point (https://www.mimuw.edu.pl/~pokar/StatystykaII/EKSPLORACJA/CorrespondenceAnalysis/UNESCO_IDAMS_CorrespAnal.pdf; for details of this method, refer to Greenacre, 1984).

CA has been commonly used in market research, such as brand preference, by different demographic groups. It has also been used in literary studies to uncover hidden patterns of linguistic usage, such as a writer's preferences in lexical choices. According to Gries (2015), CA is only occasionally used in corpus linguistics work. As far as the author is aware, it has not been adopted for the study of register variation using the MM model.

Correspondence Analysis is much easier to use than Factor Analysis. It is highly flexible with data requirements, the only strict data requirement being a rectangular data matrix made up of columns and rows with no negative entries. According to Tabata (2007), which also employs Correspondence Analysis instead of Principle Component Analysis (PCA) and Factor Analysis (FA) in his study of English literary authors, "one advantage CA has over PCA and FA is that PCA and FA cannot be computed on a rectangular matrix where the number of columns exceeds the number of rows." As the number of columns is most likely many times the number of rows, the data do not readily lend themselves readily to Factor Analysis without extensive re-organization. There is also no need to deal with

the choice of rotation methods, which do produce different results. Like Factor Analysis, Correspondence Analysis is also available in common statistical packages such as SPSS.

The greatest appeal of Correspondence Analysis lies in its intuitive bi-plot visualization: "Categories that are similar to each other appear close to each other in the plots. In this way, it is easy to see which categories of a variable are similar to each other or which categories of the two variables are related" (SPSS help). In other words, the closer things are together, the more alike they tend to be, as in the Chinese saying 物以类聚 [birds of a feather flock together]. This kind of intuitive visualization can help detect relationships among categories and also aid in the interpretation of dimensions. Correspondence Analysis is therefore better suited for exploration and practical applications. An example of a bi-plot is given in Figure 2, which shows the distribution of the 5 text types from COCA (Corpus of Contemporary American English, https://corpus.byu.edu/coca/) along the horizontal and vertical dimensions.

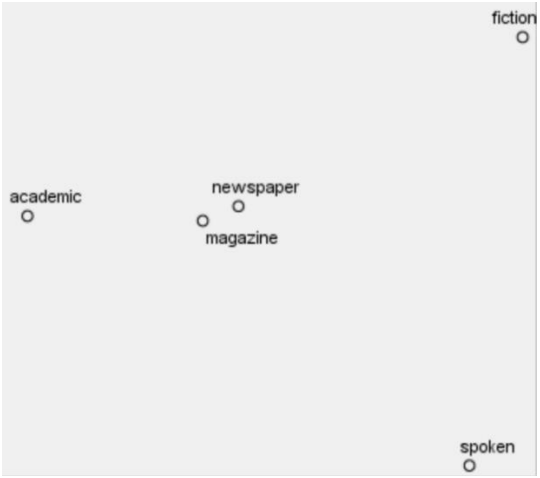


Figure 2. Distribution of COCA text types

Newspaper and *magazine* are understandably closer together than anything else, as both are journalistic registers. On the other hand, *academic* writing, *fiction* and *spoken* are farthest apart possible, academic being most different from fiction and spoken on the horizontal dimension and fiction and spoken are most distinct from each other on the vertical dimension. Details of the pilot study using COCA will be given in section 5, where the two dimensions will be interpreted.

3. Running Correspondence Analysis in SPSS

In this section, the technical details on applying Correspondence Analysis in SPSS will be briefly described. The major steps for extracting dimensions and generating bi-plots include the following:

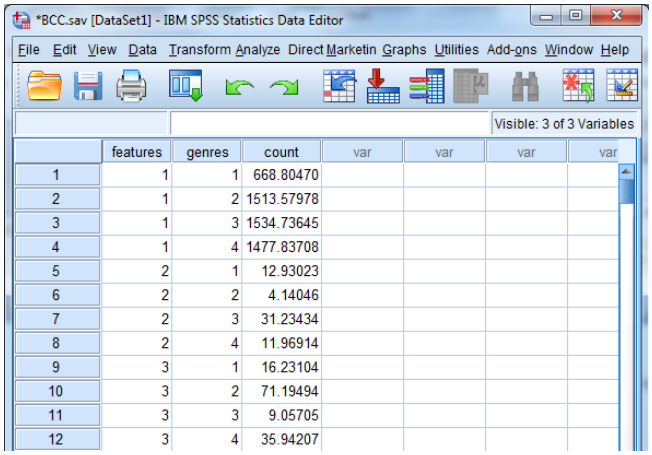
1. Coding frequency data (as shown in Figure 1) in SPSS
2. Selecting options for dimension extraction

3. Selecting options for bi-plot display

Interpreting the dimensions and plots will be done in Section 4, where sample results of the author's recent studies will be presented.

3.1 Data coding in SPSS

After gathering data in the form of frequency of occurrence of different features in various contexts, the next step is to create the data file in SPSS. A special note is in order on data coding, as this step is potentially the most unintuitive part of the whole process. For researchers familiar with Factor Analysis, a notable difference exists in how data are coded in the SPSS implementation of Correspondence Analysis. Instead of using variables directly to represent linguistic features (columns in the earlier Excel file) and registers (rows in the Excel file), all the linguistic features are represented with one single variable (named, for example, as Feature), which is then divided into the same number of values as the number of original variables; in the same manner, all the register features are also represented with one single variable (possibly named Register), which is likewise divided into the same number of values as the number of original registers. The values of both features (frequencies of features in registers) are coded in the third feature called Count. A screen capture showing this scheme is given in the Figure 3 below. The rows represent the values of the features *Features*, *Genres* (registers) and *Count*:



	features	genres	count	var	var	var	var
1	1	1	668.80470				
2	1	2	1513.57978				
3	1	3	1534.73645				
4	1	4	1477.83708				
5	2	1	12.93023				
6	2	2	4.14046				
7	2	3	31.23434				
8	2	4	11.96914				
9	3	1	16.23104				
10	3	2	71.19494				
11	3	3	9.05705				
12	3	4	35.94207				

Figure 3. Example of coding scheme in SPSS

The current version of SPSS (v25) supports both English and Chinese text, as seen in Figure 4, which uses a mixture of Chinese and English in the labels for the text types:

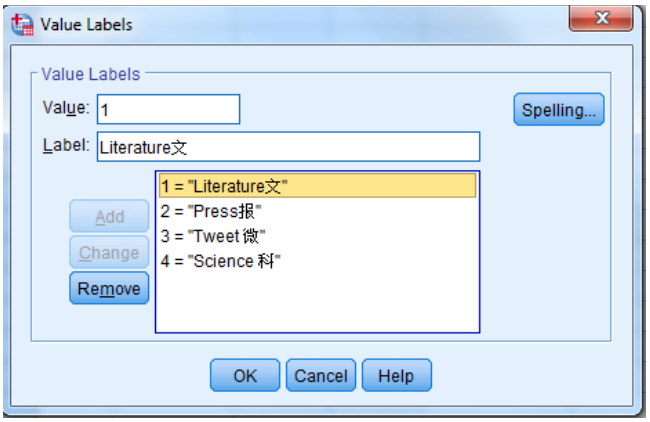


Figure 4. Text in either English or Chinese

3.2 Applying Correspondence Analysis (SPSS)

To run the Correspondence Analysis procedure, choose the *Dimension Reduction* option from the pulldown menu *Analyze* and then choose the *Correspondence Analysis* option, as seen in Figure 5:

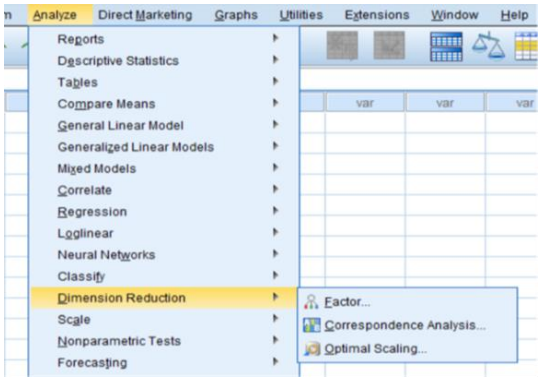


Figure 5. Correspondence Analysis in SPSS

You will then be asked to choose from the various options associated with the extraction of dimensions, as shown in Figure 6:

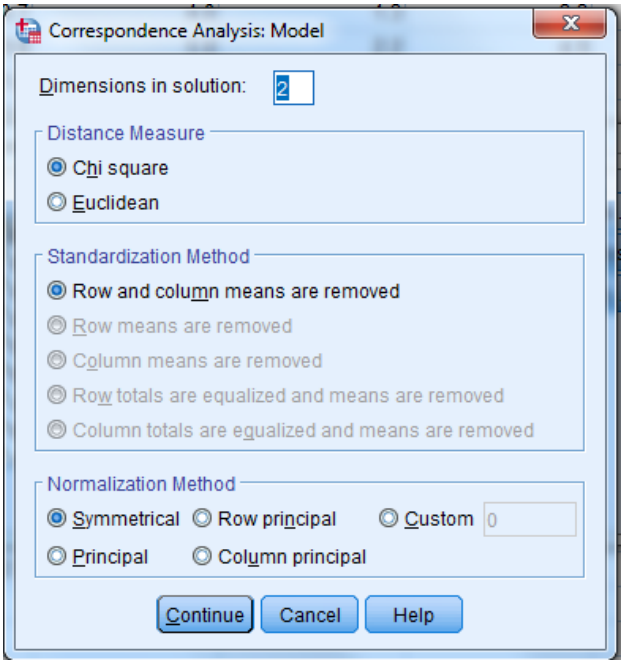


Figure 6. Options for extraction of dimensions

The number of dimensions in the solution can be specified by the user. The procedure can automatically generate up to N dimensions (N=number of registers minus 1). So for LCMC, 14 dimensions (15-1) will be generated and for BCC three dimensions (4-1) will be generated. While the present author has found the first 2 dimensions to be most interpretable, it is up to the reader to explore greater number of dimensions.

The default choice for Distance Measure is "Chi Square," the default for Standardization Method is "Row and column means removed," and for Normalization Method the default choice is "Symmetrical." The choices "Row principal" and "Column principal" have the effect of stretching the horizontal or the vertical dimension.

There are also various options for the display of Plots, as seen in Figure 7.

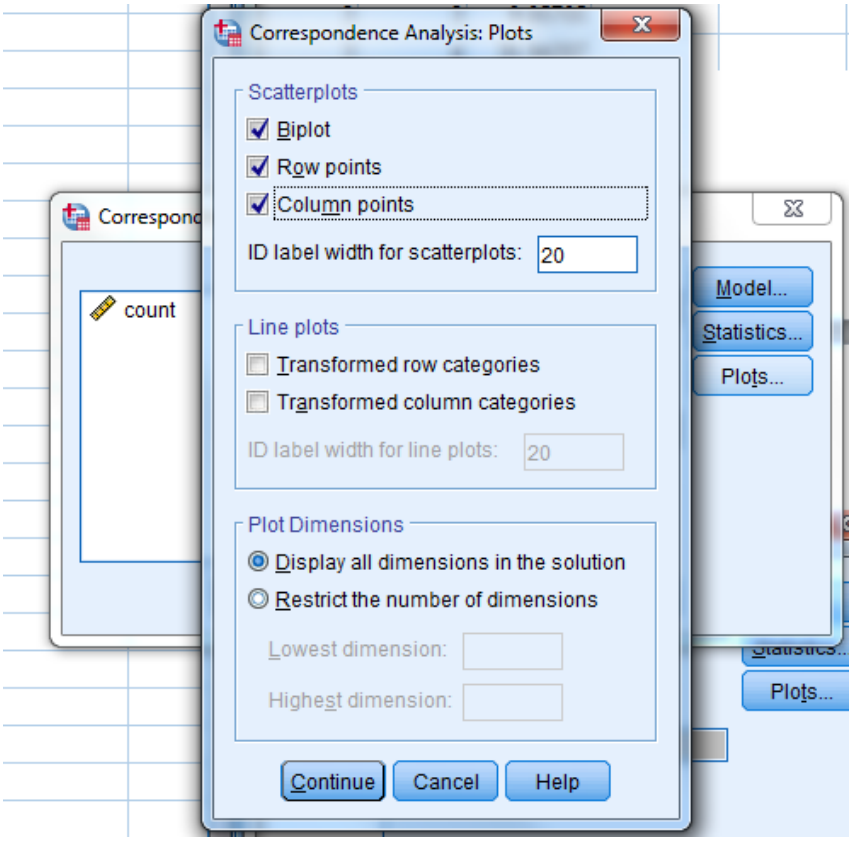


Figure 7. Options for plots display

When the "Row points" and "Column points" options are checked under Scatterplots options, separate bi-plots for features and registers are generated.

Finally, when running the procedure, it is possible to select a subset of contiguously numbered features or registers with the option of "Define Range," as is shown in Figure 8. This can be used to explore the effects of excluding certain features or registers from computation.

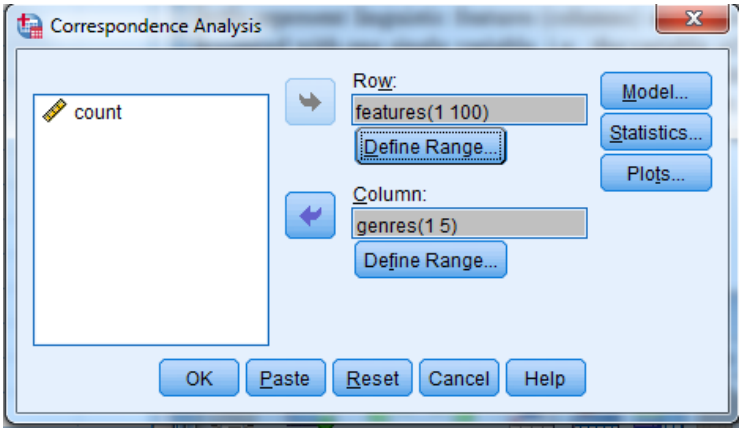


Figure 8. Option for using subset of features & genres (registers)

4. Illustrative Results: Mapping Stylistic Variation in Written Chinese

In this section, some sample results from two studies in Zhang (2017) will be used as an illustration. They respectively demonstrate the two stylistic dimensions in written Chinese and a more close-up look at two sets of near synonyms mapped along these two dimensions.

4.1 Two dimensions of written Chinese (LCMC)

To demonstrate the overall dimensions in written Chinese, the LCMC corpus with more finely differentiated registers (albeit with a small size) will be used. Most of the 50 features used are structurally related items, as they occur more frequently.

In this section, it will be shown that contrary to the commonly assumed one single formal/written vs. informal/spoken distinction, at least two dimensions of stylistic variation are found for written Chinese, which are respectively the "Literate" dimension and the "Alternative Diction" dimension. As mentioned in section 3, even though the number of dimensions can be as many as the number of registers minus 1 (15-1 for LCMC), it has been the author's experience that only the first two dimensions have sufficiently clear interpretations.

As Correspondence Analysis is mostly used for exploratory studies, measures of statistical significance such as eigen values and scree plots will not concern us further. Suffice it to say although small in number, the total variation accounted for by the two dimensions is very high (around 80%). Furthermore, the correlation between the two dimensions is small, which means the dimensions are sufficiently independent of each other.

4.1.1 Dimension 1: Literate

Dimension 1 (the horizontal dimension) is a very strong dimension, accounting for two thirds of all variation. On the bi-plot in Figure 9, where 50 linguistic features are distributed in a two dimensional space, several clear patterns can be noted:

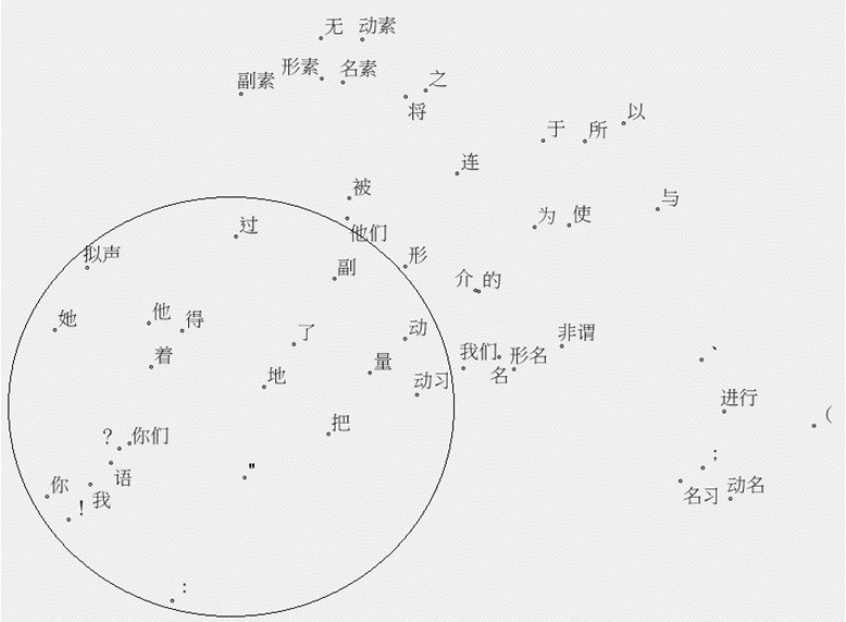


Figure 9. Clustering of interactive/narrative features on Dimension 1

1) There is a clear clustering of interactive/narrative features on the left (encircled), such as pronouns (我 [I]、你 [you]、她 [she]、他 [he]、你们[you plural]), particles (语), aspectual features (了、着、过), the two verbal *de* (地、得) and measure words (量).

2) There are clear contrasts between verbal and nominal features: verbs (动) and their associated features, such as the two verbal *de* (地、得) and the aspectual markers (了、着、过) are on the left, whereas nouns (名) and associated features, such as nominalized verbs (动名) and adjectives (形名) and attributive adjectives (非谓), are on the right.

3) There is a clear contrast between verbs in general (动) and light verbs: verbs in general are very centrally located while light verbs such as *jinxing* (进行) are very peripheral, being almost at the extreme right of the plot. There is a similar contrast between adjectives in general and attributive adjectives: adjectives (形) are fairly centrally located while attributive adjectives (非谓) are to the right of it. The contrast between the three *de* (得、地、的) is also quite astounding. The nominal 的 is clearly to the right of the verbal 得 and 地, as highlighted in Figure 10:

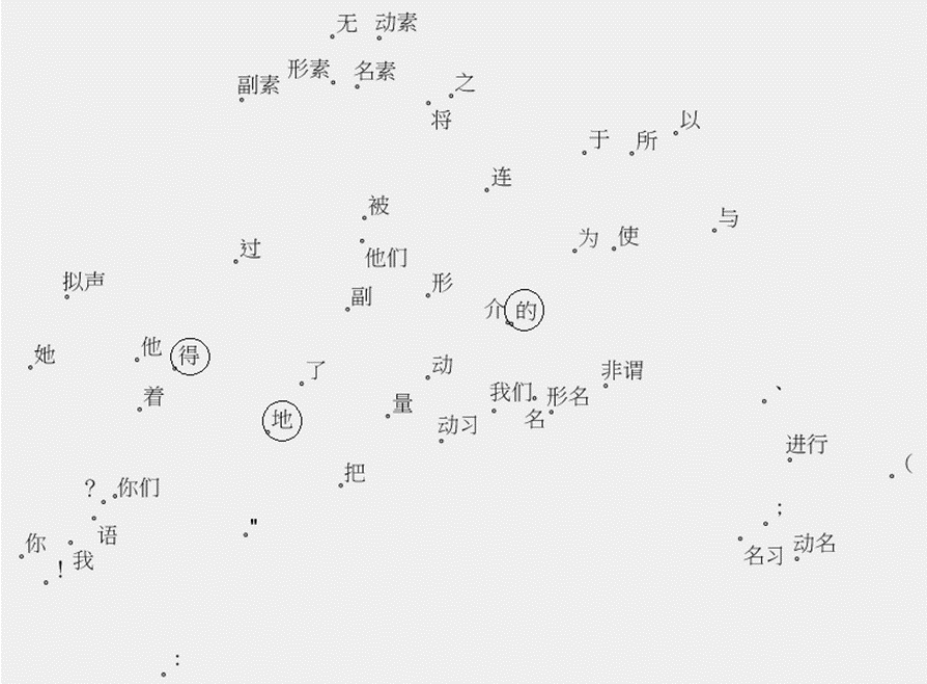


Figure 10. 得、地、的 on Dimension 1

4) The contrast between the two kinds of punctuation marks, shown in Figure 11, is no less astounding: question, exclamation, colon and quotation marks are all on the left, whereas parenthesis, semi-colon, and Chinese-style pause marks (、) are all on the right.

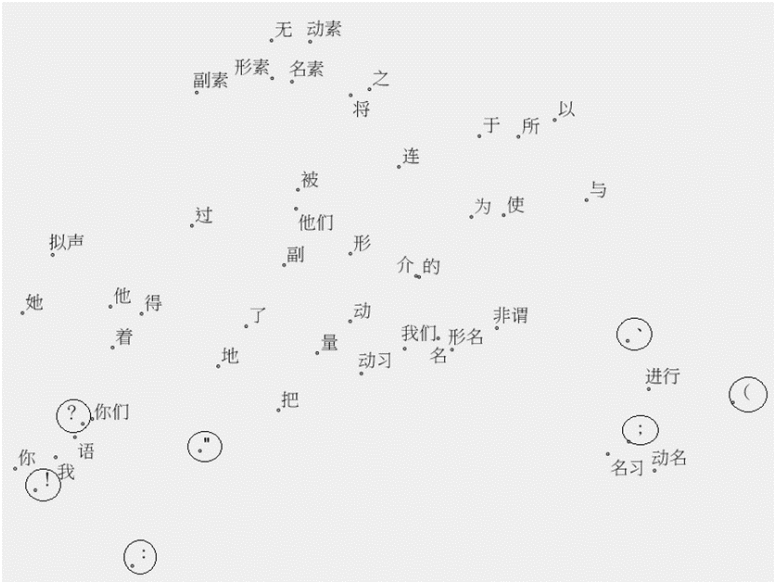


Figure 11. Two kinds of punctuation marks on Dimension 1

The distributional patterns are strikingly reminiscent of the commonly evoked "spoken vs. written" distinction, even though the LCMC corpus is exclusively written. It seems the same parameters that distinguish spoken and written styles are also at work here, such as the degree of interactivity, narrativity, nominal vs. verbal and unplanned vs. pre-planned, and so on. We will dub Dimension 1 the "Literate" dimension.

4.1.2 Dimension 2: Alternative Diction

While Dimension 1 seems to evoke the commonly used "spoken vs. written" distinction, the existence of the second dimension should be particularly worthy of note. Obviously, this is only possible with the multi-dimensional framework. But substantively, what can the second dimension be?

On the plot in Figure 12, one immediately notices the clear clustering of classical Chinese elements (encircled) on the upper half of the bi-plot.

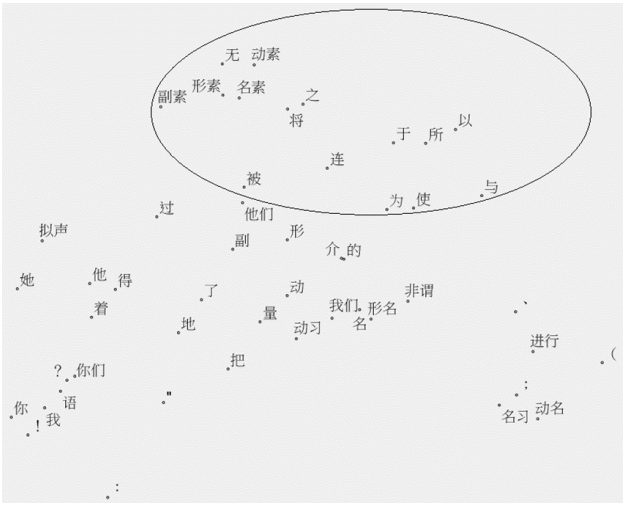


Figure 12. Clustering of classical elements on Dimension 2

These classical elements are of two kinds: individual items (为、以、所、与、于、之、将、无、使), located towards the bottom right of the encircled area, and four class features (名素、动素、形素、副素), located towards the top left part of the encircled area. These class features may require some explanation. The 素 in their labels means that these are bound morphemes that can only be part of compound words in modern Chinese but can still occur as standalone words in classical and classical-flavored texts. An example can be given from the LCMC corpus: 堤 下 鸡 鸣 ， 鸟 叫 ， 犬 吠 [under the dike, roosters crow, birds chirp, and dogs bark]. The three underlined syllables are such morphemes. 鸣 [chirp] is verbal (动素); 犬 [dog] is a nominal morpheme (名素), and; 吠 [bark] is again verbal.

It is worth noting that of the two kinds, the classes of bound morphemes are more positive on this dimension than the individual words, although they are less literate on the "Literate" dimension. This may mean that they are less integrated into modern written Chinese. They also seem to be more likely to be content words (实词) rather than function words (虚词).

The contrast between classical words and their non-classical counterparts can also be seen in the two minimal contrastive pairs of 将 vs. 把 and 之 vs. 的, given in Figure

13. The classical 将 and 之 are both north of their non-classical counterparts 把 and 的 respectively.

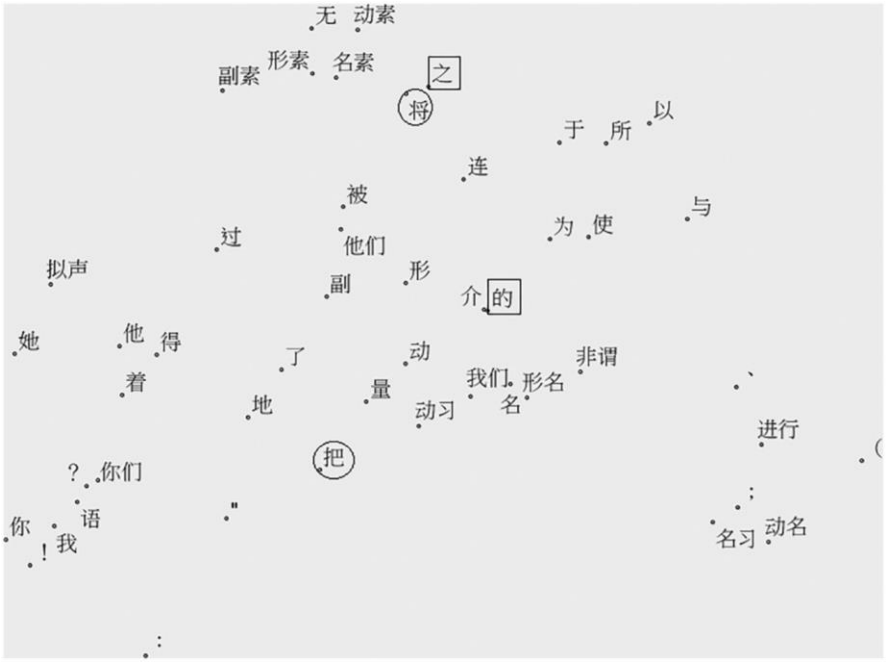


Figure 13. 将 vs. 把 and 之 vs. 的 on Dimension 2

Even though "classical" seems to suggest itself readily as the interpretation for Dimension 2, "Alternative Diction" may in fact be more accurate (Zhang, 2017). This is because classical elements are not the only ones with this distribution; some non-classical elements also are similarly distributed. These non-classical items include literary elements, dialectal elements, internet neologisms and other non-canonical forms, whose presence can be seen in larger corpora such as BCC.

The term Alternative Diction may require some explanation, as it is quite unlike the familiar notions of formality and writtenness. The best examples to illustrate Alternative Diction are the minimal pairs presented in Figure 13, which are identical in meaning but distinct in word choice resulting from a stylistic contrast unrelated to the dimension of Literateness. 将 is an alternative word choice to the synonymous 把, and 之 is an alternative word choice to the synonymous 的. The stylistic difference between 将 vs. 把 and 之vs. 的 cannot be attributed to formality or literateness on the basis of distributional evidence shown in Figure 13.

"Alternative Diction" is even better motivated when we go beyond Chinese. In Section 5, a two-dimensional analysis will also be presented for English, which shows remarkable similarity to Chinese. But needless to say, the second dimension in English, also interpretable as "Alternative Diction," cannot be classical for obvious reasons.

Compared with Dimension 1, the second dimension is much weaker in that it only accounts for less than one-sixth of the total variation (two-thirds accounted for by the first

dimension). This statistical information may well be indicative of the relative importance of the two dimensions.

4.1.3 Distribution of the 15 registers in LCMC

The distribution of the 15 registers of LCMC is given in Figure 14. The distribution pattern provides additional support for our interpretations of the two dimensions.

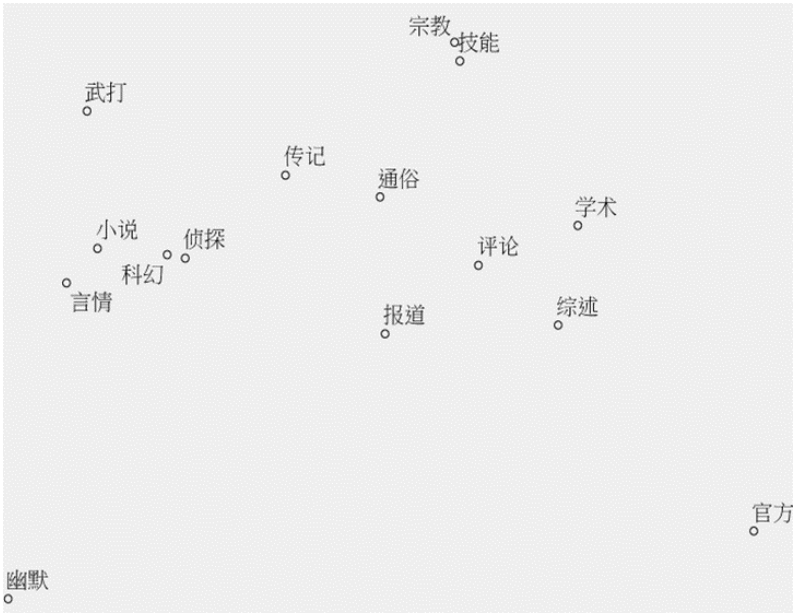


Figure 14. Distribution of *LCMC* registers

On Dimension 1, official documents (官方), academic writing (学术), and one of the three news types (综述) are the right most on this dimension, whereas all five types of fictional writing (小说、言情、科幻、侦探、武打) and humor (幽默) are on the opposite side, with news reports (报道), hobbies and skills (技能), religion (宗教), popular lore (通俗), and biography (传记) in between. This seems to concur with both our intuition and the distribution of features seen earlier.

On Dimension 2, the text types most alternative in diction are hobbies and skills, religion, and martial art novels. The least alternative are official documents and humor. In the middle are all types of fiction, all news types, and academic writing. While this may be initially surprising, it is in fact independently collaborated. Tao (1999) observed that in hobbies/skills texts, such as recipes, the classical 将 is more frequent than the synonymous modern 把, which is found more often in formal texts such as political commentaries. It is also fairly reasonable to accept that highly conventionalized official documents and academic writing are not as alternative in diction as martial arts novels.

Having two dimensions allows a text type to have different stylistic values on the two dimensions. For example, official documents rank the highest on Dimension 1 but almost the lowest on Dimension 2; official documents and humor, which are very close to

each other on Dimension 2, actually occur at the opposite ends of Dimension 1. Finally, martial arts fiction, which is sandwiched between general and romantic fiction on Dimension 1, is actually quite far from the other fiction subtypes on Dimension 2. The addition of the second dimension allows us to avoid some of the quandaries that the single spoken vs. written distinction forces on us. For example, the separation of the two dimensions makes it possible to account for the fact that text types having more classical Chinese elements, such as hobbies and martial arts fiction, are not necessarily more literate, and vice versa.

4.2 Mapping near synonyms

One of the practical applications of Correspondence Analysis is clearer explication of near synonyms. Near synonyms are a major source of difficulty in learning Chinese. It is also difficult to describe their differences in a clear and objective manner. Reference works such as dictionaries can be vague and equivocal. Correspondence Analysis can provide a more fine-grained and empirically-based picture of the differences between them.

Presented in this section is a study based on Zhang (forthcoming). For this study, the same methodological procedure is used as in Section 4 above. Naturally, as the objective here is to show the differences between near synonyms, sets of near synonyms will need to be added to the feature set. Another difference is that the larger BCC corpus is used, which is more suitable for the study of lexical features due to its greater size.

One set of near synonyms, 妇女、女性、女子、女士 and 女人, will first be used for illustration. As shown in Figure 15, these five near synonyms for "women" are distributed along the horizontal dimension in a fairly astounding manner, from most literate to least: 妇女→女性→女子→女士→女人 (the orientation of the bi-plot is flipped from earlier bi-plots using the LCMC corpus). While the plot offers a more gradient picture, which introspection cannot hope to do, it seems in fact quite intuitive. Another striking fact is that 女子 differs from the other four in being more neutral on the literate dimension but stronger on the alternative diction dimension. While this is hard to come by from intuition alone, it too seems to agree with our intuition, as it is more commonly found in classically-inflected texts, such as martial arts novels or theatrical literature, in the classical style.

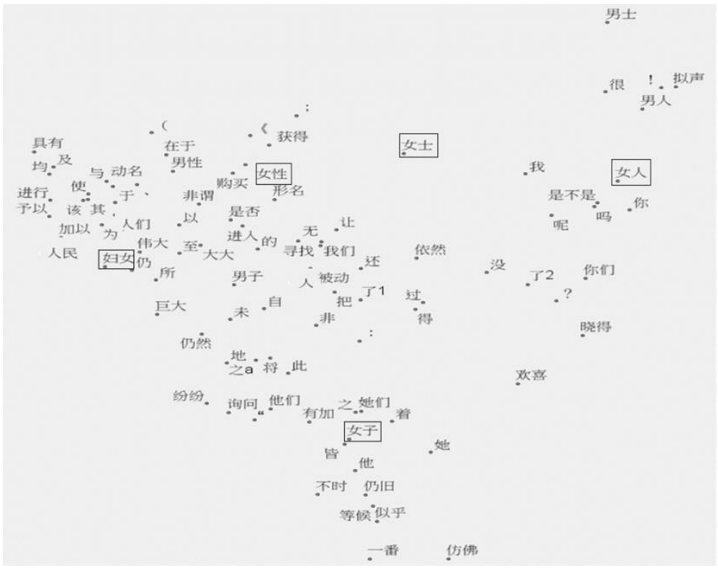


Figure 15. Five near synonyms (left: + Literate; bottom: +Alternative Diction)

The second set of synonyms for illustration is the homophonous pair 作 and 做 [do/make]. This pair has been the bane of Chinese language users, as they seem very similar and are sometimes used interchangeably. Although Lü (1980), Teng (1996), Wang (2005), Yang and Jia (2003), and G. Zhang (2010) all agree that the two are different in collocation, 作 being more abstract (better rendered as "doing") but 做 more specific and concrete (better rendered as "making"), their judgements nevertheless differ. While Lü (1980) considers 作 to be more classical in flavor, G. Zhang (2010) considers both 作 and 做 to be neutral stylistically.

The use of corpus data and Correspondence Analysis helps resolve the difference in subjective judgement. The difference between the two shows up quite clearly on the bi-plot in Figure 16, as 作 lies at the more literate end of the horizontal dimension, consistent with Lü's observation that it is more abstract than 做. As noted by Biber (1998), abstractness is also associated with literateness.

Figure 16. Partial bi-plot contrasting 作 and 做

5. An English Example

It will be shown in this section (based on Zhang, 2017) that the same methodology can be applied to English, too. Furthermore, the two dimensional analysis of Chinese gets cross-linguistic support.

For this pilot study, the large COCA corpus (The Corpus of Contemporary American English) was used (https://corpus.byu.edu/coca/). According to the developer of the corpus, Mark Davis of Brigham Young University: "The Corpus of Contemporary American English (COCA) is the largest freely-available corpus of English and the only large and balanced corpus of American English." It includes 520 million words with 5 types of texts: Spoken, Fiction, Magazine, Newspaper, and Academic. For this study, 88 lexical and grammatical features were chosen based on their potential effect on stylistic variation.

Given the 5 types of texts, 4 dimensions (5 minus 1) can be automatically generated. But as in the case of Chinese, only the first two dimensions seem clearly interpretable. Striking similarities between Chinese and English are found.

5.1 Dimension 1: Literate

Figure 17 shows the distribution of the 88 features. Given the distribution patterns, it seems reasonable to assume that this primary, horizontal dimension can be interpreted the same way as in Chinese, i.e., as one of literateness. A number of literate features are found on the left (encircled). Passives, both the by-variety and the one without *by*, are left of center. Nominal suffixes, such as *-ity*, *-tion* and *-ness*, also lean toward the left. Also found here are literate and formal words and expressions like *upon*, *whom*, *thus*, *hitherto*, *demise*, *due to*, *of the opinion*, *e.g.*, *i.e.*, and *etc.* Parenthesis, hyphen and semi-colon, all associated with carefully crafted texts, also appear in this region. In contrast, the right side of the plot is populated by features such as personal pronouns, colloquial expressions (*absolutely*, *kind of*, *a couple of*), contracted forms, interjections, and punctuation marks such as quotation marks, colons, questions, and exclamations.

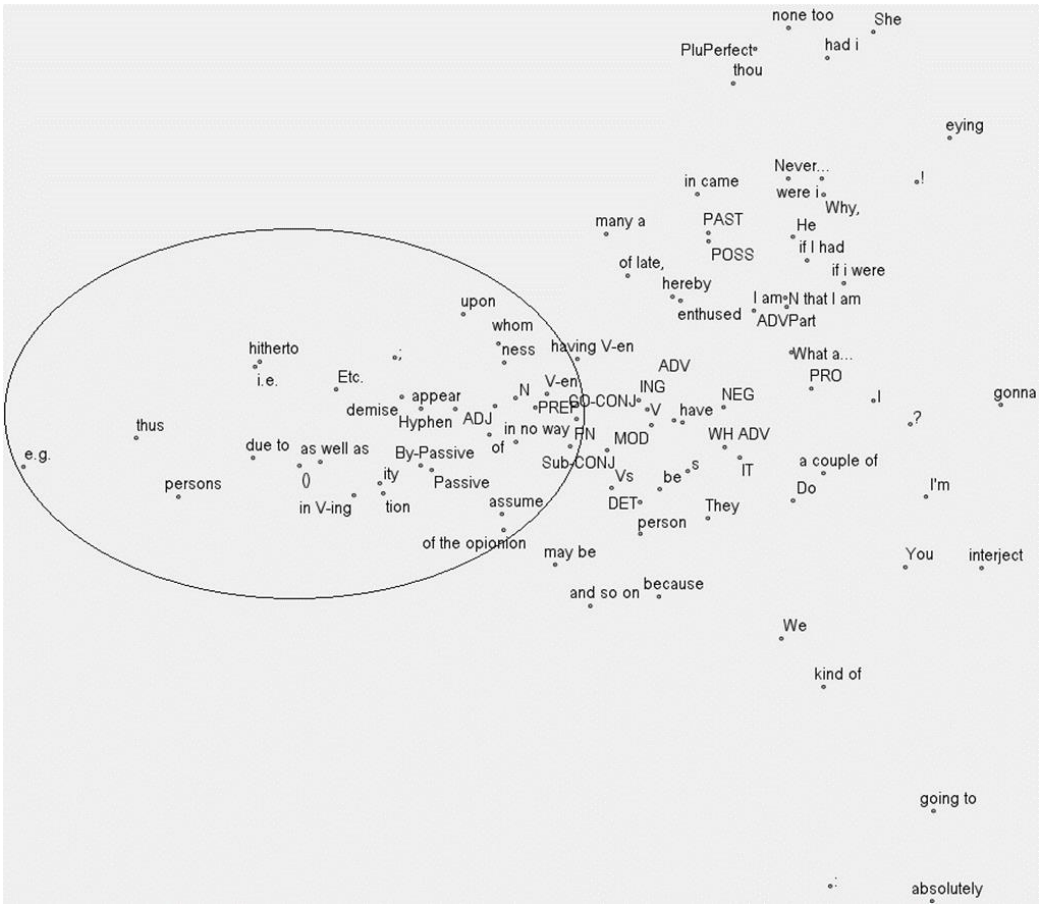


Figure 17. Clustering of Literate features (encircled)

There are some clear minimal contrasts between synonymous items such as *etc*. versus *and so on*, *I'm* versus *I am,* and *because* versus *due to*. Most noteworthy is the distinction between the singular *person* and the plural *persons*, located far apart on this dimension: while the singular is stylistically more neutral, the plural evokes the flavor of legalese, as in: "persons of known heart conditions should refrain from using the spa." The stylistically equivalent plural counterpart of *person* is thus not *persons* but more likely *people*.

5.2 Dimension 2: Alternative Diction

On the secondary, vertical dimension, there are some distributional facts that clearly support the interpretation of this dimension as "Alternative Diction." A number of words, expressions, and constructions are found at the top (encircled) of Figure 18. In addition to literary-sounding lexical items, such as *thou*, *hereby*, *enthuse*, *eying*, there are also constructions such as *none too+adj*. (e.g. *none too pleased about the prospects of meeting the family*), *of late*, *what a + noun* (e.g. *what a wonderful morning*), *many a + noun* (e.g. *many a thing you know you'd like to tell her*), *noun + that I am* (e.g. *fool that I am*). A feature that assumes extreme positive value on this dimension is the pluperfect construction, as used in *when I got there, he had already left*. While the simpler *when I got there, he already left* may now be more frequently used in speech, the pluperfect can still be found in literary texts.

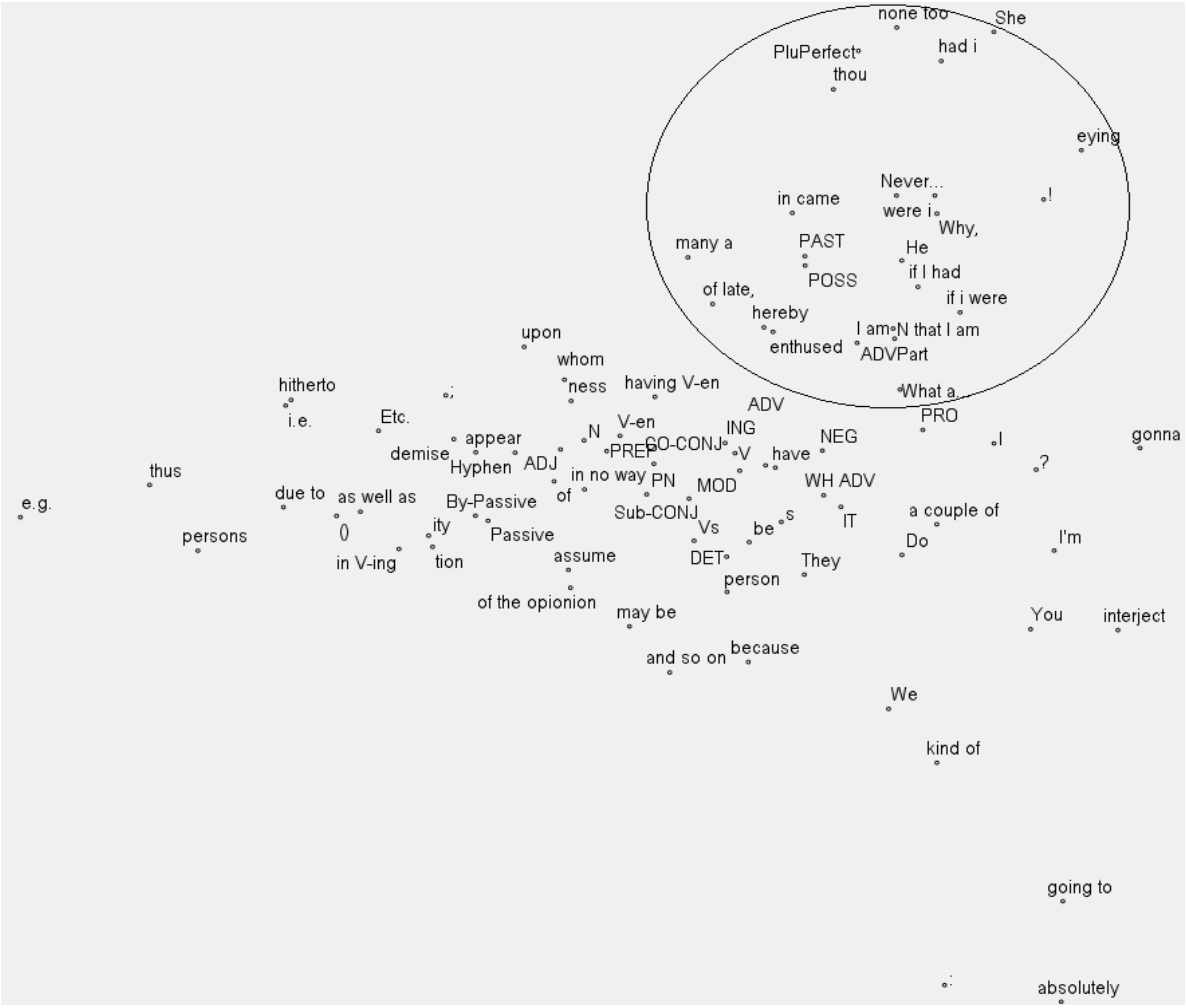


Figure 18. Clustering of Alternative Diction features (encircled)

Also found in this area is the non-canonical, inverted construction that is associated with literary usages (Green, 1982). They include *never + inverted clause (e.g. never have I been so insulted!)*, *were I* (alternative to *if I were*), *had I* (alternative to *if I had*), *in came* (alternative to *came in*).

The distribution of the 5 COCA text types, already seen in Figure 2 and repeated in Figure 19, provides clear support for the interpretations of the two dimensions: academic writing is the most literate while fiction is the most alternative in diction.

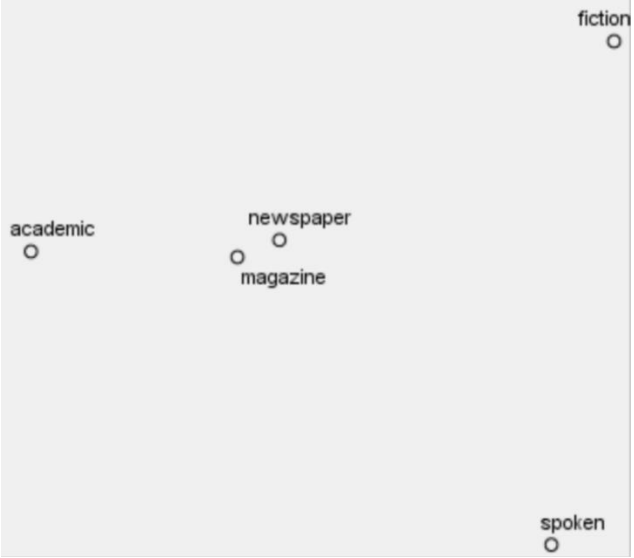


Figure 19. Distribution of 5 *COCA* text types

6. Theoretical and Practical Implications

As has been demonstrated, Correspondence Analysis is a more intuitive and easier to use alternative for carrying out multi-dimensional research on stylistic variation. But the wider use of this tool should have theoretical and practical significance beyond the obvious methodological advantages.

Theoretically, the easier application of Correspondence Analysis may encourage more researchers to conduct corpus-based multi-dimensional studies of stylistic variation. On stylistic matters, one no longer has to rely solely on introspection, which tends to be grossly simplistic. Not only can the multi-dimensional framework present a broader and more fine-grained picture of stylistic variation than simple dichotomies, it also allows for the possibility of gradient continuums, which binary distinctions do not.

Pedagogically, two-dimensional "stylistic maps" can obviously be helpful as well. Instead of relying solely on the admittedly useful heuristic notion of spoken vs. written, learners can see in a visually intuitive manner much finer shades of stylistic differences. As seen in section 4.2, near synonyms such as 妇女、女性、女子、女士 and 女人 can be shown to be clearly differentiated along not just one but two dimensions.

Another important application, especially in upper-level Chinese classes, is in the area of classical Chinese elements. A better understanding of the role of these elements in modern written Chinese will no doubt lead to more judicious instruction. We have shown that contrary to a common assumption, they are actually not exclusively associated with formal written texts. This has important pedagogical implications for issues concerning priority-setting, selection, and sequencing, which have not been sufficiently addressed. For example, how important are classical elements to overall proficiency in written Chinese? What kind of classical elements should be prioritized and introduced first? What kind of

texts should be chosen that contain such priority elements? Corpus-based, multi-dimensional research using Correspondence Analysis should have a direct impact in this area.

References

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, UK: Pearson Education.

Chen, G. Y. (2016). Examining rating criteria used to assess U.S. college students' Chinese oral performance. *Journal of Chinese Language Teachers' Association*, 51(3), 286-311.

Feng, S. (2010). The mechanism of register and its grammatical properties. *Zhongguo Yuwen*, 5, 400-412. [冯胜利. (2010). 论语体的机制及其语法属性. *中国语文*, 5, 400-412.]

Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Green, G. (1982). Colloquial and literary uses of inversions. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 119-154). Norwood, NJ: ABLEX Publishing Corporation.

Greenacre, M. (1984). *Theory and applications of correspondence analysis*. London: Academic Press.

Gries, S. T. (2015). Quantitative designs and statistical techniques. In D. Biber, & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 50-71). Cambridge: Cambridge University Press.

Jang, S. C. (1998). *Dimensions of spoken and written Taiwanese: A corpus-based register study* (unpublished doctoral dissertation). University of Hawaii, Hawaii.

Lü, S. (1980). *Eight hundred words in contemporary Chinese*. Beijing, China: Commercial Press. [吕叔湘. (1980). *现代汉语八百词*. 北京: 商务印书馆.]

Shen, H. (2005). An investigation of Chinese-character learning strategies among non-native speakers of Chinese. *System*, 33(1), 49-68.

Tabata, T. (2007). A statistical study of superlatives in Dickens and Smollett: A case study in corpus stylistics. *Digital Humanities 2007 Conference Abstracts* (pp. 1-5). Retrieved from https://pdfs.semanticscholar.org/23f7/e3a6ca295129bf18a4a3ca619038957a6c35.pdf

Tao, H. Y. (1999). On the grammatical significance of register distinctions. *Contemporary Linguistics*, 3, 15-24. [陶红印. (1999). 试论语体分类的语法学意义. *当代语言学*, 3, 15-24.]

Teng, S. H. (1996). *Chinese synonyms usage dictionary*. Beijing, China: Beijing Language Institute Press. [鄧守信. (1996). *汉英汉语常用近义词用法词典*. 北京: 北京语言学院出版社.]

Wang, H. (2005). *A dictionary of Chinese synonyms*. Beijing, China: Beijing Language and Culture University Press. [王还. (2005). *汉语近义词典*. 北京: 北京语言大学出版社.]

Yang, J., & Jia, Y. (2005). *1700 groups of frequently used Chinese synonyms*. Beijing, China: Beijing Language and Culture University Press. [杨寄洲, &贾永芬. (2005). *1700对近义词语用法对比*. 北京: 北京语言大学出版社.]

Zhang, G. Q. (2010). *Using Chinese synonyms*. Cambridge: Cambridge University Press.

Zhang, Z. S. (forthcoming). Visualizing stylistic differences in Chinese synonyms. In X. Lu, & B. Chen (Eds.). *Computational and corpus approaches to Chinese language learning*. Singapore: Springer.

Zhang, Z. S. (2017). *Dimensions of variation in written Chinese*. Oxford: Routledge.

Zhang, Z. S. (2016). A multi-dimensional corpus study of mixed compounds in Chinese. In H. Tao (Ed). *Integrating Chinese linguistic research and language teaching and learning* (pp. 215-238). Amsterdam, Holland: John Benjamins Publishing Company.

Zhang, Z. S. (2013). The classical elements in written Chinese: A multidimensional quantitative study. *Chinese Language and Discourse*, *4*(2), 157-180.

Zhang, Z. S. (2012). A corpus study of variation in written Chinese. In E. Csomay (Ed.), *Corpus linguistics and linguistic theory: Contemporary perspectives on discourse and corpora* 8–1 (2012) (pp. 209-240). Berlin, Germany: Walter de Gruyter.

應用「中文階層式語法庫」於華語文寫作教學的效果評估
(Applying Chinese Hierarchical Grammar Bank to the Evaluation of Chinese Writing Instruction)

洪嘉馡
(Hong, Jia-Fei)
國立臺灣師範大學
(National Taiwan Normal University)
jiafeihong@ntnu.edu.tw

邱詩雯
(Chyu, Shihwen)
國立成功大學
(National Cheng Kung University)
swchyu@mail.ncku.edu.tw

宋曜廷
(Sung, Yao-Ting)
國立臺灣師範大學
(National Taiwan Normal University)
sungtc@ntnu.edu.tw

張道行
(Chang, Tao-Hsing)
國立高雄應用科技大學
(National Kaohsiung University of Applied Sciences)
changth@kuas.edu.tw

摘要：教學語法(pedagogical grammar)一直是語言教學中，高度被重視及討論的重點。Nasaji and Fotos(2011)提到，若二語學習者在學習目標語的過程中，可以發展溝通技能並能夠準確流利的使用語言達到溝通目的。對於第二語教學而言，釐清教學語法的內容是相當重要的，歷來已有眾多研究確切分析了兩者間的差異(史存直 1986；Odlin1994；呂文華 2008；鄧守信 2009；Nasaji and Fotos 2011)等。本研究藉由語料庫語言學，以使用頻率做為語法排序基底，以學習者為導向的高頻偏誤語法，呈現語法點教學上較為適當的排序。首先，從多本語法書籍彙整、分析之語法點中，根據實用性(practicality)與頻率(frequency)，篩選出 339 個語法點，包含句子種類 61 個、句子架構 110 個和句子成份 168 個。其次，根據母語與二語的語法點對比，驗證其正確使用頻率及偏誤使用頻率，並依照學習語法點時的四種現象:「常出現、常錯誤、不常出現、不常錯誤」，建立「中文階層式語法庫」，提供更具理論性、全面性與實用性的資源，應用於華語文寫作教學與學習，強化教學效能及提升學習成效。最後，以臺灣師範大學僑生先修部之華語學習者(卓越華語班學習者)為研究對象，探討使用「中文階層式語法庫」於「自傳」類之華語文寫作教學的成效。教學設計將透過寫作任務結構化，採用中文階層式語法點之教學法，強化文本結構內的基礎語法和段落結構間的連貫語法，一步步引導學生完成「自傳」。本研究藉由「中文階層式語法庫」在華語寫作數位課程的教與學，預期提供教師在教學上具有更多元的數位教材，提升學習者在寫作時整體寫作表達及思維能力的輔助工具，並具體提升寫作課程的整體教學品質。

Abstract: Grammar teaching and Pedagogical grammar have been a highly discussed issue. Nasaji and Fotos (2011) indicate that in the process which second language learner develop communication skill for precise and fluent use of the target language; the "pedagogical grammar" must play an important role in language teaching designs and activities. Therefore, the importance of course designs and activities in grammar teaching are apparent. However, it is quite a dilemma to strike a balance between theories and real classrooms. For one, the researches on the sequencing of Chinese grammar teaching are important yet unfound. For another, although pedagogical grammar and theoretical grammar are closely associated, they are inherently different. Past studies have distinguished the differences between the two (e.g., Shi, 1986；Odlin, 1994；Deng, 2009；Nasaji and Fotos, 2011). Therefore, it calls a need for CSL teaching to clarify the structure of pedagogical grammar. For this reason, the corpus-based study comprehensively incorporates CSL teaching and learning into a Chinese hierarchical Grammar bank. This Chinese hierarchical Grammar bank (total 339 grammars: 61 grammars for sentence type, 110 grammars for sentence structure and 168 grammars for the main components of the sentence) would provide teachers with teaching assisted resources and learners with individualized learning, accommodating all the aspects of both teaching and learning. Above all, the present study looks forward to achieving adaptive and elaborative Chinese pedagogical grammar teaching and CSL writing instruction with theoretical and empirical perspectives. Eventually, the study would implement the Chinese hierarchical grammar bank in real classrooms. Through the teaching experiments, the project would evaluate the scientificity, objectivity, and practicality of the applying this system on Chinese writing courses. This study takes teachers and students from Division of Preparatory Programs for Overseas Chinese Students in National Taiwan Normal University as research targets, observing the writing performance after they use the Automated Paper Scoring for Han Aid (AES-Han Aid). The study aims to figure out how Chinese Hierarchical Grammar Bank could apply to AES-Han Aid and help teachers and students during the Chinese writing course.

關鍵詞：華語教學、教學語法、漢語句法學、中文階層式語法庫、華語文寫作教學

Keywords: Teaching Chinese as an Second Language, Pedagogical Grammar, Chinese Syntax, Chinese Hierarchical Grammar Bank, Chinese Writing Teaching

1. 引言

學習者語言能力發展基本上圍繞著四大學習項目，分別是：聽、說、讀、寫之語言能力。在第二語言學習中，研究普遍發現學習者對口說及寫作的能力較聽力和閱讀更難掌握，因為寫作包含了完整句構與書面語使用等等，學生需投入更多的心力及時間，而教師也需用上較多時間批改及修正學生之語法和詞彙。寫作是語言能力的綜合表現，透過寫作能力可以反映出學習者的基本辭彙和語法能力(Hong et al., 2014)，其中寫作最具實用性但卻是不容易掌握得很好的寫作文體，主要是因為掌握了語法的使用就能夠準確流利的使用語言達到溝通目的。

由於外語學習者一般很難在短時間內增進寫作能力(Buckingham & Pech, 1976)，倘若在寫作教學的階段，教學者將適合教學內容的語法點放入課程設計，不但可以教導學習者所需之語法點知識，亦可強化學習者在寫作內容的豐富度。有鑑於此需求以及難點，如何增進學生寫作能力是語言教學上甚為重要的課題。在寫作教學實務上，已有相關研究證實文體教學法(Genre-based Approach; GBA)之效益，並廣泛應用(Macken-Horarik, 1999; Kleeman, 1999)。本研究將以「文體教學法」為本，並結合「中文階層式語法庫」，進行理論與科技應用所結合的華語寫作數位課程，並評估其教學效果。

因此，「教學語法(pedagogical grammar)」是華語文寫作教學中的重要議題。Odlin (1994)定義了「教學語法」，指的是關心整體語法體系中不同語法的教學順序，結合理論的語法學根基與課堂的語法教學實務所發展而出的系統性語法。倘若在學習過程中，可以習得目的語的完整訊息且有邏輯的語法結構，對於學生達到溝通及掌握該目的語的能力非常重要。

此外，運用科技於語言教學中普遍被認為是一種有效的教學方法，而國外在英文線上寫作學習方面已有十多年的歷史，如英文託福考試或 ETS Criterion 教學系統，但若要直接應用於華語寫作上，困難重重。本研究以臺師大「華語文寫作語料庫」團隊所開發之「華語文寫作自動評分與教學回饋平臺(Automated Essay Scoring for Han，AES-Han)」為本，並結合以「中文階層式語法庫」為概念所建置的句法庫系統(Hong and Chen, 2016)，冀望能運用在以華語為二語學習者的寫作數位課程的教學上，提供教學者具有更全面性的語法訊息，協助學習者具有更完整性的語法知識。

其中，「華語文寫作自動評分與教學回饋平臺」的建置，是結合了「中文斷詞與詞性標記系統(WECAn)」(Chang et al., 2012)、「可容錯式中文文法剖析器(HanParser)」(Chang et al., 2013)、「中文錯別字偵測與校正工具(HanChecker)」(Chang et al., 2018)、「句法偏誤偵測工具(KNGED)」(Chang et al., 2014)等核心基礎技術，以及搜集低、中、高階能力學習者的寫作文本，並由華語文教學專家依照 ACTFL 的寫作評分準則進行人工評分及偏誤標記的「華語為二語寫作偏誤標記語料庫」(陳浩然 2017；Hong et al., 2014)。而「中文階層式語法庫，以科學性與系統性的數據為研究基礎，兼以具有多年經驗的華語文教師之實際教學經驗，主要在於發展出一套簡明且嚴謹

的中文語法點結構。由於考量到語法點的複雜度及多樣性，也為避免分類太瑣碎造成查找的困擾，因此，語法點建置為可開合的階層性表格，最上層主層是對於語法成分大略的劃分，分別是句子種類、句子架構及句子主要成分，每一類底下再展開細分，分為次層、語法點、語法點次層、語法點細類共四項。

在「華語文寫作自動評分與教學回饋平臺」裡，依照使用者的寫作程度，分成兩個子系統：(1)初級學習者(AES-Han Aid)、(2)中高級華語學習者(AES-Han)。由於使用 AES-Han 的中高級學習者可根據老師所設定的寫作題目及提示寫作，完成文章後能立即得到系統全面性的寫作回饋，包含：分數與評語、語法回饋、文章範例等。而考慮學習者學習目的與學習階段性的不同，本研究於寫作課程中，以 AES-Han Aid 為寫作數位平臺，引導、培養學習者練習寫作時，對於詞語表達、語法使用、句子陳述等相關寫作的能力。

本研究以臺灣師範大學僑生先修部卓越華語班 36 位學習者為對象[1]，將以「AES-Han Aid」為教學平台、應用「中文階層式語法庫」於華語寫作課程，設計教學實驗，探討「中文階層式語法庫」的特點與應用，驗證、評估其在教學上的效果。

因此，本研究將循序說明「中文階層式語法庫」的設計、功能及應用和 AES-Han Aids，分為三大部分：一、說明 AES-Han Aid 和「中文階層式語法點」建置與設計理念；二，實際運用寫作數位平臺於華語學習者的華語文寫作課程內並進行實驗；三，呈現應用「中文階層式語法庫」於華語文寫作教學實驗的結果與探討。

2. 研究目的

由於正規課程的華語文教材將華語視為一門學科，其主要目的是希望按部就班，讓學習者穩紮穩打習得華語，因此正規式華語文教材規劃的學習課時較長，也會加重語法點的說明及練習篇幅，也因為是正規式華語文教材，因此在本研究所蒐集到的各種華語文教材中的第一冊，皆可以看出各編寫者對於語法點的安排有相當默契，大部分都以華語教師的教學經驗為參考依據，篩選認為對學生來說簡單、必要的語法點優先排序，並設想一連串的情境去串聯語法點。這樣的作法是目前較為廣泛的編寫方法，但仔細觀察可以發現，不管是短期式教材或正規式教材，對於語法點的安排及挑選這一塊幾乎都是經驗掛帥。如此教材編寫的作法當然是可行且具有教學意義的，但若每一套教材都以個人教學經驗為主要規準，無法提出更有數據性的參考依據，如此一來，雖然經驗教學內容、順序、流程大同小異，但對第一線的教師與學習者而言，難免有水準參差不齊之虞。

如同上段內容所提，本研究觀察到各種華語文教材的編寫團隊除了對於語法點

[1] 臺灣師範大學僑生先修部，是海外華裔子女來臺灣升讀大學的先修教育園地。由於各僑居地文化背景、華語程度的差異，僑先部針對有華文學習需求的學生，增設有特別輔導班(即卓越華語班)，提供一年期的華語文密集訓練課程。

的安排有相當默契，在整體規劃上不會跳脫該層級難度太多，也不會在初級教材中隨意安排頻率低或實用性低的詞彙及語法點，但還是有些語法點存在著頗大的順序差，這些差距即是我們需要探討的問題，倘若語法點能以數據歸納排序，定能提供教材編寫者與華語教師更有力的教學參考依據。有鑑於此，本研究以具科學性與系統性之語料庫數據為本，結合語言學的語法理論原則與語言教學的語法點實務應用的理念，建構一套具有規則性、系統性的「中文階層式語法庫」，冀望能豐富華語語法教學的內容，增進華語文寫作教學的成效。

3. 教學語法

在語言教學領域中，教學語法(pedagogical grammar)與理論語法(theoretical grammar)本身具有密不可分的關聯性，教學語法需要以理論語法為基礎，再依據第二外語學生的學習需求建立語法架構。但兩者卻又相當不同，相較於理論語言學，教學語法更專注於個別語法點的結構與系統分析，並且不能拘泥於系統性和理論性地呈現語法內容。而 Odlin (1994)亦進一步解釋了教學語法的重要性，來自學生從課堂教學中培養出自學能力，並透過語料的驗證方式來釐清自己母語負遷移所造成的系統性偏誤，讓學生建立自己的二語學習的模式，在教學驗證中所獲得的實際經驗的證據(empirical evidence)，更可以在二語學習過程中提供準確的訊息(Nasaji and Fotos，2011；Xing，2006；鄧守信，2009；Odlin，1994)。然而目前的語法教學還是較著重於語法點的說明，而非從語法架構來進行教學，亦沒有研究學習者在真實對話環境中所需語法的狀況。

近幾年來在關注華語語法教學的議題上，周小兵(2002)曾提到，教學語法(Pedagogical Grammar)的最大特點，是從漢語第二語言的教學實際需要出發，促使學生更快地習得漢語語法，而不拘泥於理論語法的理論性和系統性。Nasaji 和 Fotos (2011)也曾於書中提到，語法是語言的根本，若少了語法，語言也就不存在。

陸儉明(2000)提出語法教學的內容要根據三方面因素來考慮，一是漢語本身。漢語中哪些語法點是必須而且最急需教給學生的？二是漢語和母語在語法上的異同。語法上的共同點在哪裡？最主要的差異在哪裡？哪些差異特別會影響學生對漢語的學習？三是學生在學習漢語過程中出現的語法毛病。學生最容易、最經常犯錯的語法是什麼？陸儉明(2000)也認為語法教學在初級階段必須堅持隨機教學，同時到一定階段需進行總結性的及一定針對性的“鞏固基礎語法”(consolidating basic grammar)教學。針對這樣的教學方式，陸儉明(2000)提出兩個方向：一是課文的選用和編排不能片面地決定於課文要教的內容，而是應該按計劃中所需給學生的字、詞、語法點來編寫和安排課文，也是說，課文中語法點的安排要有講究。二是到一定階段有必要進行帶總結性的、並有一定針對性的「鞏固基礎語法」教學。

楊寄洲(2000)則提出對外漢語教學語法具有程式性，因為語法教學不可能是隨意的、無序的，而是根據教學規律按照一定程式來進行的；也是按照難易順序輸入的，具有強力的可操作性。課堂上講授的語法必須按照簡繁和難易排出一個教學的

先後次序。

呂文華(2008)曾針對華語文教材中語法點的選擇與編排做了說明，他在書中提到教材中語法點的選擇應以兩個重點為依據：一是應通過科學的頻率統計，篩選出最基本常用，即語言交際中必不可缺少的形式，剔除某些不常用，在基礎表達階段不需出現的語言形式，式教材更加簡明實用；二是隨著近年來教學法傾向的變更，教材中選用最適宜表達交際功能的題材、詞彙與語言形式，大量的口語形式進入教材中，如無主句、獨詞句等都應該歸納說明，相反，某些多用於書面，而少用於口語的語言形式，則應於以剔除。

由理論語法系統細分出來的語法點，可視為一個語法的資料庫，因此，該如何從裡面抓取需要的資訊，則是考驗了教學者與教材編寫人員的能力。鄧守信(2009)提出教學語法建構在理論語法的基礎上，是一種被安排過的規範性語言，而這些安排與規範性的訂定理據，則可以從使用頻率、難易度、第二語言習得順序、情境等決定。此外，鄧守信(2009)也提到情境、詞彙、結構為三個影響教材裡面語法點排序的主要因素，這三個主要因素都具有結論性的否定權，因此，在教學上，若以訂定具有層次關係的中文語法庫為教學材料，即可強化華語文寫作教學內容的多樣化與豐富度。

4. 中文階層式語法點

本研究採用語料庫以應用於語法教學研究，根據語料庫語言學與教學的資料驅動學習(data driven learning, DDL)，以語料庫中的大量真實語料為基礎，發展從構詞逐步到篇章所有語言層面的課堂活動與練習，對比於傳統機械式的學習(Hadley，2002; Tribble 和 Jones，1990)，是為一種更適合透過探索逐步上手的學習方式(Johns 1991)，相關研究如董子昀等人(2015)、王意婷等人(2013)。另一方面，語料庫亦可以結合電腦輔助教學(computer-assisted instruction)，以真實語料使用頻率計算為本，應用於教學之中，加上多樣化的行動載具，能夠突破時間及空間限制(Liu and Hwang，2010)，現今也有許多單位利用語料庫研發各種以學習者為主的系列語法書籍、電子書、學習軟體與教學平台等[2]，結合了學習、測驗、評量及互動回饋等機制以及管理功能，內容、架構看起來都非常完整，但無論是書籍或是網站，幾乎是以編寫者經驗為依歸，少有以學習者為中心的學習實證數據佐證，部分雖已經以 CEFR 或 HSK 的規準做分級查詢，但在編寫教材及平台時，仍較缺乏科學性的數據支持，對於中文語法點的結構，未有確切分級的依據及說明，也尚未針對教學者在教學現場所需而提供適當的訊息。

目前在市場上，除《漢語水準等級標準與語法等級大綱》(劉英林 1997)外，不論是教材、教學理論抑或是相關工具書，尚未能提供一份具組織性的語法點編排，

[2] 相關教材如《看圖學中文語法：基礎篇》、《看圖學中文語法：進階篇》，數位語法平台如「Chinses Boost Grammar」及「Chinese Grammar Wiki」等；網站如「華語 e 起來學習網」(http://cel.wzu.edu.tw/)，教學平台如臺師大 COOL Chinese Plus(http://coolchinese.org/plus/)等。

特別是有數據支持，並排出學習順序的語法點架構。本研究為優化華語語法教學及服務教學現場的教學者，冀以微薄之力提供一個相對完善的語法點階層架構。構建以學習為導向的語法點架構，最重要的原則即是以學習者的角度出發，在正確的結構群中安排學習的順序，就如鄧守信(2009)、呂文華(1999)的研究提到，語法點排序的方法有很多種，也都各有其依據及篩選標準，但基本上難度與頻率是目前篩選的兩大趨勢。

本研究採用頻率排序取代難度或情境排序，因為若以難度作為排序篩選具有一定的限制性，根據《Principles of Language Learning and Teaching(第二語教學最高指導原則)》(Brown 2007)所提到的對照分析假說(Contrastive Analysis Hypothesis, CAH)，其中指出語音系統的相互對照提出困難層級制(Stockwell et al. 1965)，以及將語音系統替換為文法對照，經刪減後，提出六個關於文法的困難層級，分別為零級到六級(Prator 1976)。但後期即有另一派學者(Wardhaug 1970)就認為 CAH 透過預測判斷語言學習困難之理論太過於簡單武斷，即使困難層級有六級，但在細微差異上沒有詳細說明，也很難決定哪種對比屬於哪種等級。或許 6 級難度可以用來解釋學習第二外語時可以觀察得到的學習難點，卻不能拿來當作絕對標準。同時根據鄧守信(2009)提到：情境、詞彙、結構為三個影響排序的主要因素，如圖 1 所示，這三個主要因素都具有結論性的否定權。在教學上或編寫教材時，若已經訂定了一個層次關係，在篩選時即可以成為主要依據。而周小兵(2002)所說某些語法在理論語法中的出現順序跟外籍生習得順序不同。不少教科書按理論語法的次序安排語法，教師和學生既定的順序教與學，但學習者實際的習得順序卻不盡相同。而鄭懿德(1991)在語法教學對漢語語法研究的需求和推動以及教學語法研究的特點上主張急用先學，盧福波(2008)也對語法教學提出應注重實用性原則，同時須注意語義及組裝規則。

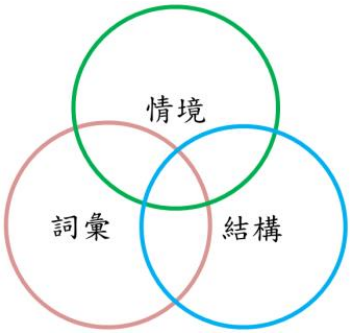


圖 1 語法排序的三個要素

上述證明了概括性、描述性的原則其實無法涵蓋所有語法屬性，雖然六級難度已經能將大部分的語法做分類，卻無法描述兩種語言中細微的差異。因此如使用難度分級語法點，可以想見將會有許多無法比對的情形，導致過多的特殊句型產生，進而增加學習者的負擔。故因以困難度或情境的不客觀性，在排除經驗法則與減少特殊句前提下，頻率是以目前情況而言較為客觀且合理的排序方式。

5.「中文階層式語法庫」建置與設計理念

5.1 建立中文階層式語法架構

本研究所構建的語法架構將同時考量到教學語法及理論語法的概念，在實踐語法點架構的建置時，整合第二語習得理論及實用為先(盧福波 2008)的主要原則，一來為讓語法架構有其科學數據可參考，二來確保語法架構在教與學兩方面的實用性及可行性。除了以學習者為導向而構建外，也希望各個語法點能各歸其所，而不是將無法歸類抑或多重語法功能的語法結構都列為特殊句型，因此中文階層式語法架構建立之際，參照《實用現代漢語語法》(1996)、《對外漢語教學語法》(2009)及《Mandarin Chinese (漢語語法)》(Li and Thompson, 1981)等工具叢書，將整個語法點架構建置成一個階層式的總表如下圖 2 所示。

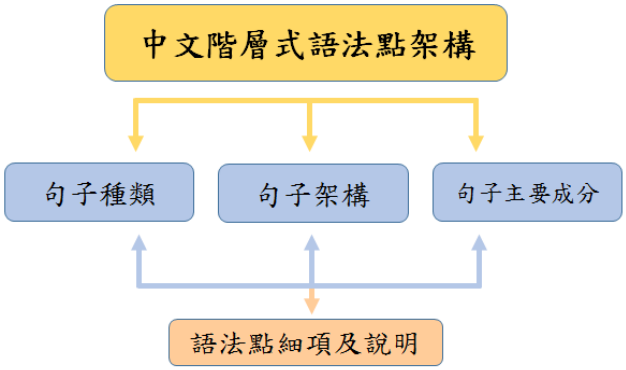


圖 2 中文階層式語法點架構

本研究將語法點建置為可開合的階層性表格，最上層主層是對於語法成分大略的劃分；分別是句子種類、句子架構及句子主要成分，如下圖 3～圖 5 所示。

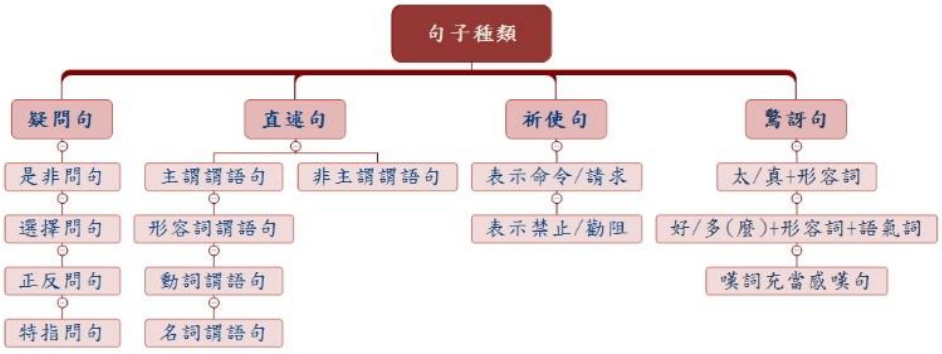


圖 3 中文階層語法點系統之「句子種類」階層圖

圖 4 中文階層語法點系統之「句子架構」階層圖

圖 5 中文階層語法點系統之「句子主要成分」階層圖

如上圖 3～圖 5 所示，每一類底下再展開細分，分為次層、語法點、語法點次層、語法點細類共四項。次層為主要分類下的子種類；語法點才代表各個種類下不同的結構；而語法點次層及細類，則是詳細描述了上一階層組成要素與結構安排，如上圖 2 所示，學習者最主要需要學習的層次在細類中的項目，但每個項目後都具有其意涵與來源，進階學習者與教學者可以從表格當中溯其原本，也能整合之前所學，相同類型的語法點，提升學習效能。

5.2 以語料庫為本的語料收集、分析與統整

本研究所建置的中文階層式語法庫，為了可以實際應用於華語文寫作教學課程中，故先由母語者語料庫－中文十四億字語料庫(Chinese Gigaword Corpus)(Lexical Data Consortium, 2009)、中介語語料庫－華語文寫作語料庫(Chinese Written Corpus)(陳浩然 2017；Hong et al., 2014)與華語文教材(當代中文課程、成功之路、新版實用視聽華語、新實用漢語課本)等三個面向進行語料收集，探討語法架構所彙整之語法點的客觀性及實用性，用以確認母語者、二語學習者及教材研發者三種角度發展之語料觀察語法架構是否符合「常出現常錯誤」以及「常出現不常錯誤」的

高優先教學順序。

5.2.1 中文十四億字語料庫(Chinese Gigaword Corpus)

中文十四億字語料庫(Chinese Gigaword Corpus) 2.0 版是收錄 1990 年至 2004 年的語料，語料庫的內容包含了十四億中文字，分別來自大陸、臺灣、新加坡的大量語料，包括八億多字來自臺灣中央社的資料、近五億字來自大陸新華社的資料，以及大約三千萬字來自新加坡早報的資料。中文十四億字語料庫藉由中文詞彙特性素描 (Chinese Word Sketch Engine)進行搜尋，其首頁介面如下圖 6 所示：

圖 6 「中文詞彙特性素描」系統首頁介面

本研究使用從教材中所蒐集的語法點歸入已建置完成的語法架構，並從母語者使用語法點的情況觀察其母語者使用語法點頻率的高低，因此，以中文十四億字語料庫(Chinese Gigaword Corpus)(LDC 2009)做為母語者語料使用依據，以中文詞彙特性素描系統(Hong and Huang 2006)做為篩選工具。

以「中文十四億字語料庫」做為母語者語料使用依據，這是為了確認母語者在進行溝通時，最常使用的語法及語言結構的情況。此外，若單就語法教學而言，「中文十四億字語料庫」所提供的語法相關訊息，可依照「常用先學」的原則，教學者考量的是，若以母語者使用習慣且高頻的語法點優先教起，那麼學習者將能更快進入狀況，提高學習效率，更早達到有效的溝通效果。

5.2.2 華語文寫作語料庫(Chinese Written Corpus)

除了母語語料庫之外，本研究以華語文寫作語料庫(Chinese Written Corpus, CWC)(陳浩然 2017；Hong et al., 2014)中的外籍生寫作語料進行錯誤分析，從語料當中獲取學習者易犯錯的語法點資訊檢視整個階層式語法架構建置適切性。

華語文寫作語料庫(Chinese Written Corpus)是華語文為外語之寫作語料庫，搜集單一題目下之大量華語為外語學習者寫作文本，搜集的樣本為外籍生經一段時間學習華語後所寫的作文(陳浩然 2017；Hong et al., 2014)。寫作文本搜集時間從 2010 年 9 月至 2014 年 6 月，蒐集來自共 37 個不同母語的外籍生，完整記錄了每一篇作文的資訊，包含題目、書寫者之中英文姓名、國籍、母語及在台學習華文之教學單位等，並轉為文字檔及圖檔儲存。寫作文本目前有四個題目，共 2,802 篇，句子共有 112,080 句。

本研究以「華語文寫作語料庫」做為二語學習者語料使用依據，這是為了可以具體呈現二語學習者在使用華語過程中，實際出現的偏誤例子，為了提供給教學習者在教學時，可以有效給予真實語料當作佐證，提醒學習者在使用上避免犯錯，達到正確使用目的語的目標。

5.2.3 華語文教材(Chinese Textbooks)

近年來，隨著華語教學理論與教材編寫理論的發展，加以教學經驗的累積，教材種類也越來越豐富，曾有學者將教材分為初中高階系列教材、學歷教材及非學歷教材、通用教材及專業技能教材(聽、說、讀、寫)、國別性教材、特殊用途教材(如：文化類、醫學類)等等分類(李泉 2005)。每一套教材對於課文、生字、生詞、語法及教學活動皆自有其編寫原則及規範。本研究以目前較常被使用的正規、通用型系列華語教材：《當代中文課程》(鄧守信，2015)、《成功之路》(張輝，2008)、《新版實用視聽華語》(葉德明，2007)、《新實用漢語課本》(劉珣，2002)四套教材特別針對語法點編排部分進行對比，其比較結果，如下表 1 所示。

表 1 四套華語文教材第一冊課文數及語法點數量

教材名稱	當代中文課程	新實用漢語課本	新版實用視聽華語	成功之路起步篇
課文數	15	14	12	12
語法點數量	137	107	92	57

四套教材因著重性質與教學目的略有不同，因此，在課時、單元、語法點出現的排序上也略有差別。在單元上，教材從 6 課至 15 課不等；在語法點出現的順序上，也發現到其相異之處，如：「時貌：了」，在《遠東生活華語》(1970)第 2 課就出現，在《新版實用視聽華語》(2007)跟《當代中文課程》(2015)皆直到第 10 課後才陸續出現，而《成功之路》(2002)與《新實用漢語課本》(2002)第一冊皆尚未將其列入教學語法點中。從上述表 1 得知，各教材第一冊的語法點安排，雖相去不遠，但仍有差異存在，這些差異也是本研究欲探討課題之一。

5.3 結合數位、教材、教學法的架構設計

在基於「科技內容教學知識理論(TPACK；Technological Pedagogical Content Knowledge)」(Mishra 和 Koehler, 2006)的架構，如下圖 7，統合了教學教材與大量學習者語料的「內容知識(Content Knowledge)」、第二語言習得理論的「教學知識

(Pedagogical Knowledge)」，及運用自然語言處理技術與科技輔助教學系統的「科技知識(Technological Knowledge)」，建置了本研究的「中文階層式語法庫」。

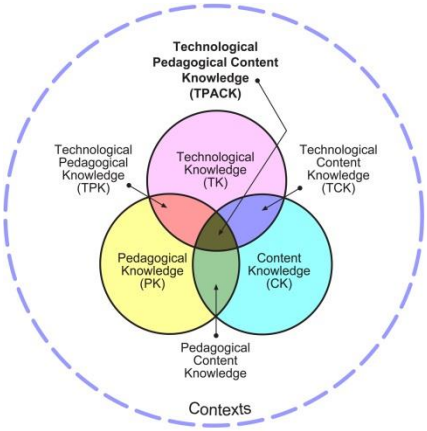


圖 7 科技內容教學知識理論(TPACK)模型

科技內容教學知識理論，主要探究如何透過教師具有的學科知識，連結教學與學習理論的基礎，並結合數位科技的運用，進行有效益的科技融入課堂教學模式。而「中文階層式語法庫」的設計，亦奠基於此三方面的知識之上：透過「科技知識」與語料庫技術，將教師、教材具有的「內容知識」數位化，即是把已排序及分析後的語法點及相關說明、真實語料例句系統化、數位化的建置於「中文階層式語法庫」中。

本研究所建置的「中文階層式語法庫」是一個兼具理論、科技、系統與應用屬性的系統，結合語言學語法理論、科技數位化教學知識，以應用在數位教學平台上。在華語教學當中，尤以寫作教學對教學者的負擔最重，因為需要提供足夠的訊息給予學習者，此外，學習者在寫作方面若可以有充分、豐富的語法訊息，則可大大提升其學習成效及寫作成果。因此，為了秉持著科技數位化的教學理念，結合本研究所建置的「中文階層式語法庫」於華語文寫作數位平台(Automated Essay Scoring for Han Aid，AES-Han Aid)，提供教學者於寫作教學時，對於中文語法相關資訊之所需，以達到充實教學者的教學內容，及有效改善、提升學習者的學習成效。

6. 「華語文寫作數位平台」(AES-Han Aid)建置與設計理念

針對初級華語學習者設計的華語文寫作數位平台(Automated Essay Scoring for Han Aid，AES-Han Aid) (http://www.aeshan.net/)，如下圖 8，以寫作自動評分及語法自動偵錯為核心技術。AES-Han Aid 具有獨有的設計理念及特色：

第一、AES-Han Aid 是以寫作任務為導向，包含卡片、信、電子郵件、表格、便條、留言及一般記敘文等不同寫作類型，輔以三大主題(theme)為教學核心，內容涵蓋「個人與社會」、「休閒與工作」、「都市與郊區環境」三個主題，每個主題包含

多種話題(topic)，如下表 2。

圖 8 華語文寫作數位平台首頁

表 2 寫作主題及話題清單

主題 (theme)	話題 (topic)
個人與社會	日常起居
	教育
	飲食
	個人資料、外貌、性格
	身體健康
	關係
	購物
休閒與工作	就業
	娛樂
	假期
	媒體
	體育活動
	科技
	交通
都市與郊區環境	環境方面的顧慮
	全球問題
	鄰里
	自然地理
	城鎮及其設施
	天氣

第二、透過教師設定的引導問題及關鍵字，讓學生理解該主題下的核心概念並進行創作。教師設定「個人與社會」主題下的「飲食」話題，並設定具體寫作題目為「你最喜歡的食物」，引導題目分別為「你最喜歡的食物是什麼？」、「它吃起來怎麼樣？」，關鍵字為學生可能答出的關鍵字，如「飯」、「面」、「冬粉」、「珍珠奶茶」，「軟」、「香」、「熱」等，如下圖 9、10。

圖 9 設定寫作引導問題 1 及關鍵字之範例(你最喜歡的食物是什麼？)

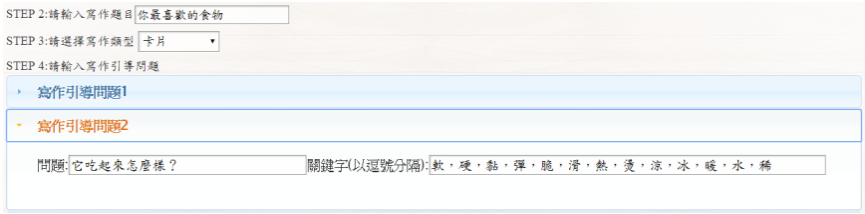


圖 10 設定寫作引導問題 2 及關鍵字之範例(它吃起來怎麼樣？)

第三、學生撰寫過程中，教師設定之引導問題以語音互動提示的方式呈現如圖 11，引導學生回答問題並完成作文，完成後系統即提供分數與語法回饋，如圖 12。

圖 11 「AES-Han Aid」之語音方式引導寫作

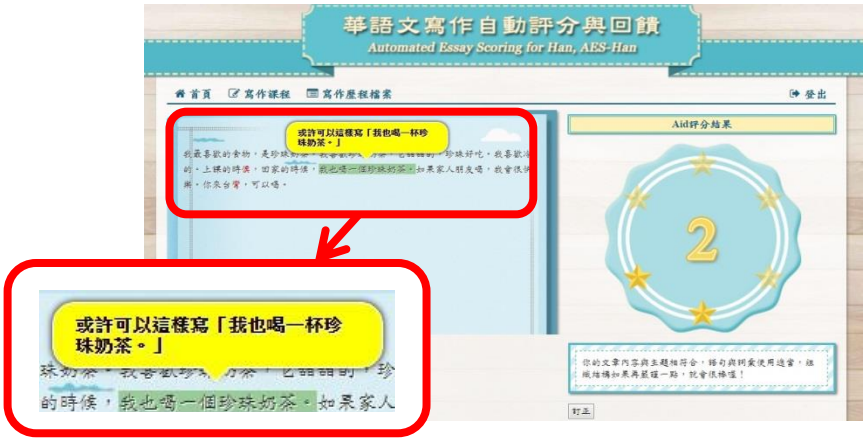


圖 12 「AES-Han Aid」之寫作回饋

7. 「華語文寫作數位平台」的寫作課程設計與實驗

7.1 實驗對象

本研究以臺師大僑先部卓越華語班36位元學習者為實驗對象，包括印尼、越南、巴拉圭、墨西哥、日本等地，國籍分佈如表 3，皆在測驗之前已於僑生先修部接受 18 周、每週 20 小時的密集華語文訓練。換言之，即已在台學習華語的時數達 360 小時，具備基本語法及 500-1000 個基礎詞彙量，符合華測會(國家華語測驗推動工作委員會，TOCFL)進階級(B1)程度。

表 3 學生國籍暨人數分佈情形

國籍	印尼	越南	巴拉圭	墨西哥	日本
人數	30	3	1	1	1

7.2 實驗設計

實驗設計分為前後兩次測驗，前測以傳統紙筆方式進行，後測則帶領學生使用 AES-Han Aid 進行寫作練習。在學生分別完成兩次作文之後，再將原先的前測紙筆寫作成果，輸入系統，對比兩次寫作成績，觀察其進步狀況，以掌握寫作數位平臺、「階層式語法點」，對於二語學習者練習華文寫作的實際幫助。

7.2.1 前測

在前測方面，教師以「說明你的漢字設計」為題，讓學生自選某個漢字，進行漢字特色敘述。教師選定「說明你的漢字設計」為前測的題目，是因為書寫漢字是臺師大僑先部卓越華語班學生每天都必須進行的學習，所以「漢字」已經是這些學生的生活一部份，他們對於「漢字」並不陌生。教師在出題時，告知學生可以從以下幾個方向拆解說明，包括：漢字部件、造詞、句法、意義，以及你選擇的理由等等，教師提供幾個寫作時可能會使用到的關鍵詞與語法點，然後給予學生一段時間，

讓學生搜集相關資料，進行文本寫作。學生在經過思考與資料搜集過後，使用紙筆方式書寫，在 50 分鐘內，完成平均 80 字左右的短文寫作[3]。

7.2.2 正式施測

在學生完成傳統紙本寫作測驗之後，教師再帶領受試者使用平臺。正式課堂開始後，教師先針對平臺的功能予以說明，確認學生瞭解目的以及操作過程後，回歸到華文寫作教學課程原有流程，由教師指定題目「介紹你自己」作為本次題目，教師指定「介紹你自己」為正式施測題目，是因為這些學生在完成卓越華語班的課程之後，將會申請臺灣的大學，都是需要寫「自傳」，而「介紹你自己」就是「自傳」的一部份，這是學生們日常學習的一部份，他們對於「自傳」類的寫作題目也不陌生。

介紹題目後，依照課程進度，首先帶領學生鑒賞佳作，並分析文章作法，接著讓學生透過平臺，開始進行線上寫作。此次寫作，教師利用「AES-Han Aid」的引導寫作功能，以「介紹你自己」設計了六道提示題目與提示關鍵詞，如表 4，這些題目的設計，其所使用的語法點，皆參考自本研究所建置的「中文階層式語法庫」的訊息，以確保教學之所需，學習之程度等級。

表 4 AES-Han Aid 中正式施測題目「介紹你自己」之提示項目

題號	提示題目	參考訊息
1	你是哪國人？	印尼人、越南人、日本人、墨西哥人、巴拉圭人……
2	你家裡有誰？	爸爸、媽媽、哥哥、姊姊、弟弟、妹妹
3	他們在做什麼？	上班族？做生意？家庭主婦？學生？
4	他們和你的關係怎麼樣？	民主？自由？嚴格？期待高？
5	你的個性怎麼樣？	樂觀？活潑？嚴謹？認真？
6	有沒有其他想說的？	任何想說的

學生熟悉平臺系統後，即開始自由寫作。在系統中，學生依照系統提示，每完成一個提示題目後，點選下一道提示，透過引導問題、關鍵字及語音互動提示來完成文章。

[3] 36 位受試者，字數最多者 120 字，最少者 38 字，平均 77.33 字。

8. 華語文寫作教學實驗結果與評估

以「AES-Han Aid」為華語文寫作的教學數位平臺，在受試者完成前測題目「說明你的漢字設計」後，我們將結果以純文字檔方式輸入系統，發現多數受試者的成績落點集中於 2 分，傳統紙筆寫作成績不盡理想，儘管教師給予學習者準備時間，亦給予寫作方向提示，讓學生收集資料，但實際寫作時學生仍然無法完全掌握題目，在篇幅或結構上的成果皆有待加強。如表 5 所示。

表 5 AES-Han Aid 測試－前測題目「說明你的漢字設計」測試結果

分數	人數
3 分	8 人
2 分	28 人

而在正式施測題目「介紹你自己」中，教師在教學過程裡引入 Aes-Han Aid、「階層式語法點」協助課程，藉由系統提示題目與輸入關鍵字的功能，無須重複說明、暗示、提醒學生寫作方向，僅需回覆個別學生的問題，在教學上確能事半功倍；而學生亦能藉由系統中題目的語音互動提示，反覆思考與構築自己的文章結構。在 50 分鐘的寫作課程後，學生皆能在平臺的協助下順利完成本次寫作。

本研究的實驗結果，可從篇幅以及分數兩個方面探討：

第一、從學生寫作的篇幅及字數來看，統計此次學生完成的短文，字數最少 60 字，最多 811 字，平均字數為 245 字。與傳統紙筆寫作方式相比，增加了 3 倍以上，顯示若在即時的引導之下，學生寫作的空間與彈性明顯的提高，寫作內容的豐富度與可能性也提高許多。

第二、從兩次測驗所顯示的分數可以發現，36 位受試者在 Aes-Han Aid、「階層式語法點」的引導寫作之下，除了書寫的篇幅變長之外，在內容的豐富、語法點的使用恰當與增加及整體結構上，多數學習者皆有長足進步，以致在兩次題目整體分數有顯著差異，如圖 13 所示。

(1) 分數從 2 分進步到 6 分者有 4 位。
(2) 進步 3 分者有 12 位：1 位從 3 分進步到 6 分，11 位從 2 分進步到 5 分。
(3) 進步 2 分者有 6 位：4 位從 3 分進步到 5 分，2 位從 2 分進步到 4 分。
(4) 進步 1 分者有 7 位： 2 分進步到 3 分有 7 位。

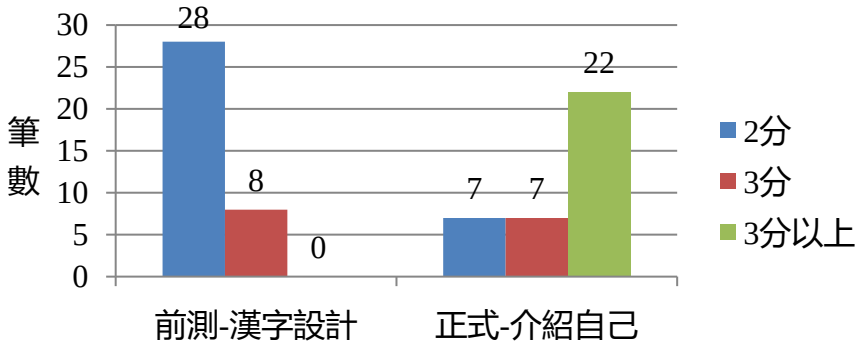


圖 13 兩個題目分數對比結果

除了上述進步的情形外，另有分數持平者 4 位，退步者 3 位。觀察成績退步的學習者，比較兩次寫作內容，第二次使用平臺寫作時，總字數亦不超過 80 字，內容也多為提示題目的部分簡答，由此可知退步的情形，主要受到學習者寫作態度及意願幹擾。而就多數學習者而言，使用數位寫作平臺、句法庫系統等輔助華文學習與寫作訓練，確實能夠經由系統提示來申述題旨，增大寫作篇幅並豐富內容，提升寫作成效。

9. 結論

本研究透過教學語法、語料庫數據及數位科技所建置之「中文階層式語法庫」，是一個具有理論支持且經過實證的研究成果，透過「科技知識」提供的平台與語料庫技術，將教師、教材具有的「內容知識」及「教學知識」數位化，讓使用者更易於查詢及使用「中文階層式語法庫」中的語法點及相關說明，並藉由「中文階層式語法庫」中的大量真實語料例句了解目標語法點的使用方法及情境。

在數位學習科技的輔助下，教師可以更有效益的安排及進行教學，而學習者亦能同時兼顧教師所給予的教學重點與練習，也能在「中文階層式語法庫」所提供的母語者語料及二語學習者的語料中，進行自我學習與檢測，學習正確的中文語法用法，亦可了解偏誤語法的使用狀況，如此一來，可以讓教學更具彈性，讓學習更具真實性，對於實際應用「中文階層式語法庫」於華語文教學課程的設計，可以應用於華語文句子教學與寫作教學上，搭配教學實驗，以評估和檢驗系統對於電腦輔助教學的成效。

致谢： 1) 感謝科技部計畫(MOST 107-2410-H-003-053 -)的補助支持，感謝教育部「高等教育深耕計畫」(107J1D0202)，以及國立臺灣師範大學「華語文與科技研究中心」支持。2) 教學實驗進行當時，共同作者邱詩雯老師任職於國立臺灣師範大學僑生先修部。

參考文獻

Buckingham, T., & Pech, W. C. (1976). An experience approach to teaching composition. *TESOL Quarterly*, *10*(1), 55-65.

Brown, H. D. (2007). *Principles of Language Learning and Teaching* (5th Ed.). Harlow, UK: Pearson Longman.

Chang, T. H., Sung, Y. T., & Hong, J. F. (2018). Introduction to *HanChecker*: A proofreading tool for Chinese misspelling words. In *Proceedings of the 10th International Conference and Workshops on Technology and Chinese Language Teaching* (pp. 20-25). Taipei, Taiwan.

Chang, T. H., Sung, Y. T., Hong, J. F., & Chang, J. I. (2014). KNGED: A tool for grammatical error diagnosis of Chinese sentences. In *Workshop Proceedings of the 22nd International Conference on Computers in Education* (pp. 48-55). Nara, Japan: Social Sciences Education.

Chang, T. H., Liu, C. L., Chen, B., Sung, Y. T., & Su, S. Y. (2013). A grammar parser for automated essay scoring of Chinese as a second language: Development and implementation. Paper presented at *the 21st Annual Meeting of the International Association of Chinese Linguistics*. Taipei, Taiwan.

Chang, T. H., Sung, Y. T., & Lee, Y. T. (2012). A Chinese word segmentation and POS tagging system for readability research. Paper presented at *of 42nd Annual Meeting of the Society for Computers in Psychology*. Minneapolis, MN, USA.

Chen, H. J. (Ed.). (2017). *Corpus and Teaching Chinese as a Second Language*. Taipei: Higher Education Press. [陳浩然(主編). (2017). *語料庫與華語教學*. 台北：高等教育出版社.]

Hadley, G. 2002. An introduction to data-driven learning. *RELC Journal, 33*(2), 99-124.

Hong, J. F., & Chen, Z. Y. (2016, November). *A preliminary study of Chinese hierarchical grammar development.* Paper presented at International Conference on Asian Language Processing (IALP 2016), Tainan, Taiwan.

Hong, J. F., Chang, J. I., Chang, T. H., & Sung, Y. T. (2014, June). *Chinese written corpus development and analysis.* Paper presented at the 8th International Conference Technology and Chinese Language Teaching (TCLT 8), Boston, MA.

Hong, J. F., & Huang, C. R. (2006). Using Chinese Gigaword corpus and Chinese Word Sketch in linguistic research. In *Proceedings of the 20th Pacific Asia Conference on Language, Information and Computation (PACLIC-20)* (pp. 183-190). Wuhan: Huazhong Normal University, China.

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning. *ELR Journal*, 4(s), 1-16.

Kleeman, G. (1999). '*TEP 283 Unit Resources' in Curriculum and Instruction in the Secondary School II General Information and Study Guide*. School of Education, Macquarie University, Sydney.

Lexical Data Consortium. (2009). Tagged Chinese Gigaword Corpus 2.0. Retrieved from https://catalog.ldc.upenn.edu/LDC2009T14

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. Taipei: The Crane Publishing Co., Ltd.

Liu, G. Z., & Hwang, G. J. (2010). A key step to understanding paradigm shifts in e-learning: towards context-aware ubiquitous learning. *British Journal of Educational*

Technology, *41*(2), E1-E9.

Li, Q. (2005). *Reflections on TCFL theories*. Beijing: Educational Science Publishing House. [李泉. (2005). *對外漢語教學理論思考*. 北京：教育科學出版社.]

Liu, X. (Ed.). (2002). *New practical Chinese reader* (Volume 1). Beijing: Beijing Language and Culture University Press. [劉珣(主編). (2002). *新實用漢語課本*(第一冊). 北京：北京語言文化大學出版社.]

Liu, Y. H., Pan, W. Y. & Gu, T. (1996). *Modern Chinese grammar for teachers of Chinese as a second language & advanced learners of Chinese*. Taipei: Shida Shuyuan. [劉月華, 潘文娛, & 故韡. (1996). *實用現代漢語語法*. 臺北：師大書苑.]

Liu, Y. L. (Ed.). (1997). *The outline of HSK standards and grammatical levels.* Beijing: Higher Education Press. [劉英林(主編). 1997. *漢語水準等級標準與語法等級大綱*. 北京：高等教育出版社.]

Lu, J. M. (2000). Duiwai hanyu jiaozhong de yufa jiaoxue. *Language Teaching and Linguistic Studies*, 3, 1-8. [陸儉明. (2000). 對外漢語教中的語法教學. *語言教學與研究*, 3, 1-8.]

Lu, F. B. (2008). Yufa jiaoxue de jiben yuanze yu caozuo fangfa. *Language Teaching and Linguistic Studies,* 2, 25-31. [盧福波. (2008). 語法教學的基本原則與操作方法. *語言教學與研究,* 2, 25-31.]

Lv, W. H. (2008). *Duiwai hanyu jiaoxue yufa tansuo.* Beijing: Beijing Language and Culture University Press. [呂文華. (2008). *對外漢語教學語法探索*. 北京：北京語言大學出版社.]

Lv, W. H. (1999). *Duiwai hanyu jiaoxue yufa tixi yanjiu*. Beijing: Beijing Language and Culture University Press. [呂文華. (1999). *對外漢語教學語法體系研究*. 北京：北京語言大學出版社.]

Macken-Horarik, M. (1999). *Learning Guide-Language in Social Context*, Faculty of Education, University of Technology, Sydney.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A new framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017-1054.

Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context.* New York, NY: Routledge.

Odlin, T. (1994). *Perspectives on pedagogical grammar*. Cambridge University Press.

Prator, C. (1967). *Guidelines for planning classes and teaching materials. Work-papers in English as a Second Language: Matter, Methods, Materials.* Los Angeles: Department of English, University of California.

Shi, C. (1986). *An Outline of Chinese Grammar History.* Shanghai: East China Normal University Press. [史存直. (1986). *漢語語法史綱要*. 上海：華東師範大學出版社.]

Stockwell, R. P., Bowen, D. J., & Martin, J. W. (1965). *The grammatical structures of English and Spanish*. Chicago: The University of Chicago Press Books.

Yang, J. Z. (2000). Duiwai hanyu jiaoxue chuji jieduan yufa xiangmu de paixu wenti. *Language Teaching and Linguistic Studies*, *3*, 9-14. [楊寄洲. (2000). 對外漢語教

學初級階段語法項目的排序問題. 語言教學與研究, *3*, 9-14.]

Yeh, T. M. (Ed.). (2007). *Practical audio-visual Chinese* (Volume 1). Taipei: Cheng Chung Book Co., Ltd. [葉德明(主編). (2007). *新版實用視聽華語*(第一冊). 台北：正中書局.]

Yeh, T. M. (Ed.). (1970). *Far East everyday Chinese* (Volume 1). Taipei: The Far East Book Co Ltd. [葉德明(主編). (1970). *遠東生活華語*(第一冊). 台北：遠東圖書股份有限公司.]

Teng, S. H. (ed). (2015). *A course in contemporary Chinese (Volume 1).* Taipei: Linking Publishing. [鄧守信(主編). (2015). *當代中文課程課本*(第一冊). 台北：聯經出版事業公司.]

Teng, S. H. (Ed). (2009). *A pedagogical grammar of Chinese* (Revised Version*)*. Taipei: The Crane Publishing Co., Ltd. [鄧守信. (2009). *對外漢語教學語法*(修訂版). 台北：文鶴出版有限公司.]

Tribble, C., & Jones, G. (1990). *Concordances in the classroom: A resource book for teachers*. London: Longman.

Tung, T. Y., Chen, H. J., & Yang, H. M. (2000). Error analysis of the *le* construction based on *Chinese Learner Written Corpus*. *Computational Linguistics and Chinese Language Processing, 20*(1), 76-95. [董子昀, 陳浩然, &楊惠媚. (2015). 以「華語學習者語料庫」為本的「了」字句偏誤分析. *中文計算語言學期刊*, 20(1), 76-95.]

Xing, J. Z. (2006). *Teaching and learning Chinese as a foreign language: A pedagogical grammar.* Hong Kong: Hong Kong University Press.

Wang, Y. T., Chen, H. J., & Pan, I. T. (2013). Investigation and Analysis of Chinese synonymous verbs based on the *Chinese Learner Corpus*: Example of “bang”, “bang-zhu”, “bang-mang” and “bian”, “bian-de”, “bi-an-cheng”. *Journal of Chinese Language Teaching, 10*(3), 41-64. [王意婷, 陳浩然, 潘依婷. 2013. 基於中介語語料庫之近義動詞混用情形調查與分析--以「幫」、「幫助」、「幫忙」及「變」、「變得」、「變成」為例. *華語文教學研究*, 3, 41-64.]

Wardhaug, R. (1970). The Contrastive analysis hypothesis. *TESOL Quarterly, 4*(2), 123-130.

Zhang, H. (Ed.). (2008). *Road to success: Threshold*. Beijing: Beijing Language and Culture University Press. [張輝(主編). (2008). *成功之路：入門篇*. 北京：北京語言大學出版社.]

Zheng, Y. D. (1991). Duiwai hanyu jiaoxue dui yufa yanjiu de xuqiu yu tuidong. *Chinese Teaching in the World, 4*, 217-222. [鄭懿德. (1991). 對外漢語教學對語法研究的需求與推動. *世界漢語教學*, *4*, 217-222.]

Zhou, X. B. (2002). Chinese as a foreign language: Characteristics of instructional grammar. *Journal of Sun Yat-sen University (Social Science Edition), 42*(6), 137-142. [周小兵. (2002). 漢語第二語言教學語法的特點. *中山大學學報(社會科學版), 42*(6), 137-142.]

以语音识别技术逆向分析汉语远场群体讨论中非母语者的交互策略

(Using Automatic Speech Recognition Technology to Reverse Analyze Communication Strategies between Non-Native Speakers in a Chinese Long Distance Group Discussion)

砂冈和子
(Sunaoka, Kazuko)
早稻田大学
(Waseda University)
ksunaoka@gmail.com

摘要：近几年，语音识别技术（Automatic Speech Recognition: ASR）的精度大幅提升，已突破了从技术走向实用的门槛。本文对非汉语母语者（NNS）与母语者（NS）的远场群体讨论语料用 ASR 技术进行识别精度。逆向验证了最新 ASR 对单一发言人、母语者、标准口语的识别精度非常高，已达到现场应用的水平。但不管对 NS 还是 NNS，对含有情感及多声源等干扰等语音，ASR 识别率都出现大幅度下降。因此很难应用到识别具有远场(far-field）,多通道(multi-channel）,多模态（multi-modal）特征的语音。相比之下，参加群体讨论会的 NNS 存在汉语口音不够标准，语句碎片化等问题，但充分利用多个信息通道、多种沟通模态来与 NS 进行互动共享信息。最后简介未来 ASR 技术的趋势，同时显示几种“ASR＋汉语教学”模式，从而探讨如何更好地与智能语言工具互补共存。

Abstract: In recent years, the accuracy of Automatic Speech Recognition (ASR) has greatly improved in terms of its practical applications. An ASR test on a real corpus of a multi-person long-distance group discussion between non-native speakers (NNS) and native speakers (NS) of Chinese was used to compare the accuracy of ASR with NNS speakers. It was found that the latest ASR has very high recognition accuracy for single speakers, native speakers, and the standard spoken language. ASR has now reached a new level in field applications. However, for both NS and NNS, the recognition rate of ASR significantly decreased while capturing emotional and multi-channel speech. Therefore, it is difficult to apply to far-field, multi-channel, or multi-modal speech. In contrast, NNS participants who made full use of multiple information channels and modalities were able to successfully communicate and interact, although their Chinese pronunciation was not standard and included fragmented statements. This paper also discusses trends in future ASR technology and introduces several

"ASR + Chinese teaching" methods to explore how they may better coexist with smart language tools.

关键词：语音识别技术，远场群体讨论，多通道，多模态，互动信息共享

Key words: Automatic Speech Recognition, Long Distance Group Discussion, Multi-channel, Multi-modal, Interaction by information sharing

1. 研究动机和目的

近几年，基于大规模数据库的深度神经网络（Deep Neural Networks，DNN）等声学模型在 ASR 研究上均获得了巨大成功(Graves et al., 2016；刘洋，2017)。谷歌、微软、苹果、亚马逊等全球智能语言技术大巨头研发的 ASR 识别率声称已堪比人类听力水平[1]。百度、科大讯飞、搜狗等中国公司对汉语普通话的语音识别率都已达到 97％，识别速度为每分钟 400 字之快(陈鹏，2017;戴礼荣等，2017）。ASR 是人机交互的基础，也是推动机器翻译，自然语言理解等技术发展的前提条件。语音输入法、语音助手、车载语音交互系统等 ASR 智能语言技术已渗入到人们生活的方方面面，并开始应用于会议演讲等实时互译的技术上(中村等人，2017；河原等人，2018）。日本政府研究机构开发的同声传译 Voice Tra，其日英文口译与笔译水平分别与 TOEIC 800 分到 900 相当[2]。

ASR 与机器翻译（MT）的发展既能低成本、高效率地消除不同语言之间的沟通障碍，又能为视听障碍者带来极大的方便。与此同时，自动翻译的普及给外语教学产业带来了强烈的冲击[3]。AI 语言技术会不会改变外语学习的方法？如有正面效果，它对第二语言学得有如何影响？实际上一些公司宣传的性能评价过高, 目前市面上的 ASR 错误率超过 15%甚至 30%[4]。它的难题主要在远场的精确识别以及对于口音、多人语音、多语种、大词汇等场景数据的获取上。当前的人机对话，预先定义好的特定域(domain)内才能实行，可扩展性仍然存在很大的问题（中村等人，2017; 篠田，2017）。

二语教学致力于在模拟真实的语言环境中，注重于物理、生理、心理方面的互动与情感，培养学生的语用能力和沟通能力(Clifford et al., 2013; 罗华珍等人，2017; 徐琦璐,2017;)。ASR 可作为一个很强的听写机，既可以增加自己可操作的教材资源，还可以辅助汉语发音练习。而 ASR 技术还不成熟，直接应用于会话学习或课堂教

[1] 《电子发烧友网》, http://www.elecfans.com/video/yinpinjishu/20161019441265.html [2016-10-19].
[2] VoiceTra. 7.0 (iOS 版). http://voicetra.nict.go.jp/ [2018-10-3].
[3] 《科技视界》, 2016 年 20 期. http://bianke.cnki.net/web/article/F085_1/KJSJ201620081.html.
《人工智能》, http://www.stdaily.com/rgzn/tuijianq/2017-09/04/content_574332.shtml [2017-09-04].
[4] 《瞭望》, https://www.iyiou.com/p/78283 [2018-08-04]

学活动弊大于利。如何设计“AI＋汉语教学”模式是我们面临的重要课题。本文基于这些研究动机和目的，使用远场群体讨论的语料，对此应用 ASR 技术的测验，从而了解 AI 与人类对语义理解之优劣之处。

以后分四个部分进行论述：首先将介绍本次研究的语料与分析框架;其次将对比 ASR 对 NS 与 NNS 两种语音的识别结果;接着浅析非母语者参加外语群体讨论时的交互策略与当前和未来 ASR 技术的趋势；最后提示几种“ASR＋汉语教学”模式作为本文的小结。

2. 分析方法与语料特征

2.1 分析框架

本文使用远场群体讨论课（后文称“讨论课”）的课堂录影材料，以信息共享互动理论（Theory of interaction by information sharing）作为分析框架(安西，2017)，用民俗学 (Ethnomethodology)的话语分析方法，对“讨论课”学生的交互行为进行观察。由于“讨论课”的语料具有多通道、多模态特征的自然口语。所以我们对收集的语料以文字和符号，尽可能地详细转写。一方面试用 ASR 对此小片语料进行听写测验转写成文字，从而比较 ASR 的精度与 NNS 学生交互策略的差异（详见第 4 节）。

信息共享互动理论基于认知神经科学和信息处理最新研究的成果，将构建一种包含所有生物与物体的交互机制与共享信息的模型（后述）。本文将 NNS 和 NS 之间交互方式用以民俗学话语分析方法进行细致描述（Goffman,1981；Sacks, Schegloff, & Jefferson ,1974; Schegloff, 1996），再用信息共享互动理论的交互概念给它试加标记，从而获取群体讨论成员交互过程的认知功能与它的机制。这样既有利于 NNS 所发挥的多种认知功能与互动策略，还可以互补印证与目前 ASR 功能的差异，帮助二语教学对智能语言工具更好地理解进行和取舍。

2.2 术语的定义

由于本文中所使用术语的含义与该术语在它领域的技术类定义有些许出入，特将这些术语在本文中的意思解释如下：

- 信息共享互动理论（Theory of interaction by information sharing）
- 通过认知神经科学和信息处理科学相结合，统一解释物体交互的机制，从而构建一种跨媒体的信息处理框架(Anzai, 1992; Anzai, 2013；安西, 2017; 潘煜等人, 2018；Tomasello et al., 2005)。该理论不仅对人人交互，还对人机交互的认知过程都有较强的理论和应用价值。该理论认为交互者的内部由叫做

GRAMES 的系统与机制构成[5]，并相互衔接和连贯进行信息处理。如目标导向机制（G：Goal-directed mechanism），奖励系统（R：Reward systems），注意力系统（A：Attention systems），动机机制（M：Motivational mechanisms），情绪机制（E：mechanisms for Emotion），社会信息处理机制（S：Social information processing mechanisms）等。本文将GRAMES的概念扩充解释为讨论者所运用的谈话策略**表 4**，初步探索对群体讨论语料作标记。

- 远场（far-field）：指的是发言人距离麦克风较远。本文用的是含有发言人之间的空间距离与网速等干扰因素的广义“远场”(Clark & Brennan， 1991)，以便验证 ASR 技术对远程群体讨论录音语料的识别精度。
- 多模态（multi-modal）：除了用以语音传递信息外，还通过感知觉、意识、记忆、注意、联想、动机和情绪等非语言模式进行交互。
- 多通道/多道（multi-channel）：在多数交互对象之间，通过动作、手势、表情、视线、姿势等进行交流。相对于单一发音人、单通道/单道（single-channel）的语料，对 ASR 难度较高。
- 智能语言工具：包括 ASR 与 MT 以及对话系统等使用 AI 技术开发的语音产品。

2.3 语料特征

笔者自 2011 年以来，通过互联网在北京、台北、东京、横滨、北九州五地之间实施远程汉语讨论课(砂冈,2016)。这堂课基于早稻田大学 Cross-Cultural Distance Learning（CCDL）的建构主义教学理念，针对 CEFR 进阶 B 级以上汉语水平的非母语学生，以培养 Intercultural Competence 能力为学习目标，通过让国内外学生同步互动的平台之上，展开非导向性对话（Non task oriented dialog）。一般每轮对话由一名主持人与一名指定发言人进行一对一问答，但通常由非指定听者参与讨论，有时旁听者之间以视线或姿势进行交互(上同)。

本文使用的是于 2016 年 5 月 19 日实施的一次主题为“介绍自己养宠物经历”（后文称“养宠物”）的对话片段。此次 4 校共 13 名学生参加，具体属性如**表 1** 所示。尽管言语能力占优势的 NS（8 名）和 NS-EC（2 名）掌握了话语领导地位，3 名 NNS 也积极参与讨论，在整个讨论时间 73 分钟之中，很少出现讨论中断等尴尬状况，如**表 6** 自 2B-1 至 2B-6 这一段话总共花 18 分钟，其中发言时间占 90%，而话轮空白（静场）只占 2%。

表 2 所示（A）朗读，（B）演讲等距麦克风单独、近讲、无噪音、中立情绪、单模态等语音类易于 ASR 识别。与此相比（C）现场直播，（D）自由对话等多通道、噪音大、带情绪的语音，为 ASR 带来难题。“讨论课”的语料属于（E）群体讨论语料，既有多人数（共有 13 名 Speaker）、超远场、强噪声等特点外，还含

[5] 基于当前脑神经科学研究的水平，与其他神经基础的结构相对清楚分离的功能称为“系统”，而还未分离的功能叫“机制”。由每项机制的头字母命名为 GRAMES（安西,2017）。

有不同汉语口音、以及二语学习者的非正规语音干扰。因为自然口语的讨论，所以还有大量副语言（如韵律、语气词）、非语音信息（如停顿、笑容）等因素（砂冈，2016；Sunaoka,K.2018 a）。第 4 节**图 3** 所示，目前 ASR 还未支持多模态、多通道信息的处理与理解。

表 1 参加 2016/5/19 讨论课学生的属性

单位	NS（MorF）	NS-EC	NNS	Total
WT 校	2（M1/F1）	1（F）	2（M1/F1）	5（M2/F3）
WB 校	4（M2/F2）			4（M2/F2）
B 校		1（M）	1（M）	2（M）
T 校	2（M2）			2（M）
Total	8（M5/F3）	2（M1/F1）	3（M2/F1）	13（M8/F5）

说明：WT：早稻田大学东京校园；WB：早稻田大学北九州校园；B：北京大学； T：台湾师范大学；M:Man；F：Female; NS:以汉语为母语者； NS：非汉语母语者； NS-EC： Ethnic Chinese 华裔学生
在论文引用时使用这些标号。

表 2 ASR 对语料难度类型(Feature comparison of speech types)

	语音类型	发言人数	通道	话速 mora/s word/min	距 话筒	噪音 水平	情绪强弱	模态	识别难度（%）
A	朗读	单人	单一	7.26(189)	近	安静	平坦	单	易
B	演讲	单人	单一	7.31(191)	近	较安静	较平坦	单/多	较易 約 90%
C	现场直播	单/多人	多道	8.51(222)	近	嘈杂	大起大落	单/多	较难 約 60-80%
D	自由对话	多人	多道	每人不一	较远	嘈杂	每人不一	多	较难
E	群体讨论	多人	多道	每人不一	远	嘈杂	每人不一	多模态	极难

（改编自 有木康雄，2003）

说明

1）Mora 是日语发音的基本单位,以“元音”或“辅音+元音”构成，用平假名表示一个音拍。日语词汇一般为汉字和假名，有时用片假名或罗马字混合书写。日文没有严格的正书法，

因此常用词汇平均由几个字构成就说法不一。话速(word/min)是笔者将原英文的词数换成日文字数（約 2.3 倍）后计算。

2）识别难度据（河原等，2018)。

2.4 测试方法

本研究使用 Google Translate 将语料从语音转换成文字。该平台具有识别精度高，支持多种语言，全球免费在线语音识别与同步翻译等特点[6]。因为随着 Google 不断改进其产品所用的内部语音识别技术，所以对同一样本进行了两次测试。一次为 2018 年 3 月 31 日，第二次是 2018 年 11 月 10 日。结果发现第二次识别精度显著提高。主要表现在 1）强化抗噪音干扰，增强了远场识别功能[7]。但如果多个声源在相近的方位还是难以识别**图 3**。2）情感语音识别能力提升（陈师哲等人，2018; Do et al., 2015） 3）同步翻译的结果比以前自然得多。感觉替换错误减少，而换误和漏误有所增加。

到第二次测试时期为止，Google Translate 主要支持单声道语音文件、单一话者韵律的识别，不支持同一音频文件中话者的转换[8]。因此我们先把音频文件转为单声道，并按照话者进行切分，然后把切分后的文件再输入到 ASR 系统上进行识别。识别操作在安静的室内进行，先将谷歌平台设置成需要的语言，点击图标开始收音，再将需要被识别的音频对着电脑片刻之后，谷歌便会在左侧框中输出识别结果**图 1**。另外，由于 Google Translate 平台每一次翻译的结果都有些许差异[9]，本研究采用同一天（2018 年 11 月 10 日）的多次识别结果中错误较少、翻译质量较好一次的识别结果进行分析比较。

3. ASR 语音识别结果

3.1 ASR 对母语者对话语音的识别结果

下面是将一名来自北京的 WT 校留学生（WT/NS1，汉族）与台湾的主持学生（T/NS1，对外华语教育专业研究生）的一段对话用 Google Translate 转换成文字信息的结果。同原发言相比，共 163 个字中只有 6 个单词被识别错，其识别精度竟达 96%[10]。测试用的样本即便是远程讨论课现场采录的 MP4 压缩格式，音频质量较差，对 Google Translate 并没有造成影响。不过同步自动翻译出了人脑不可能犯的错译，部分英译令人无法理解（后述）。

[6] 其他还有 Google Cloud Speech-to-Text 等更先进的技术，但需要付费，不适合教学用。至于科大讯飞、百度等中国制作的 ASR 精度更适于汉语口语的识别及合成（笪骏，2018），因为登记需要中国大陆手机号，所以不方便国外用户使用。免费的语音开放平台中国手机号

[7] 百度云 Far field speech recognition，https://cloud.baidu.com/product/speech/fsr [2018-10-02]

[8] 还可供选择视频转录模型，

[9] Google 在云端有不同数据中心处理，另外每次搜索算法各自不同，因而产生差异。

[10] 如果把 1A-2 段漏听的 7-8 个词算在一块，识别度不会讲到 90%以下。

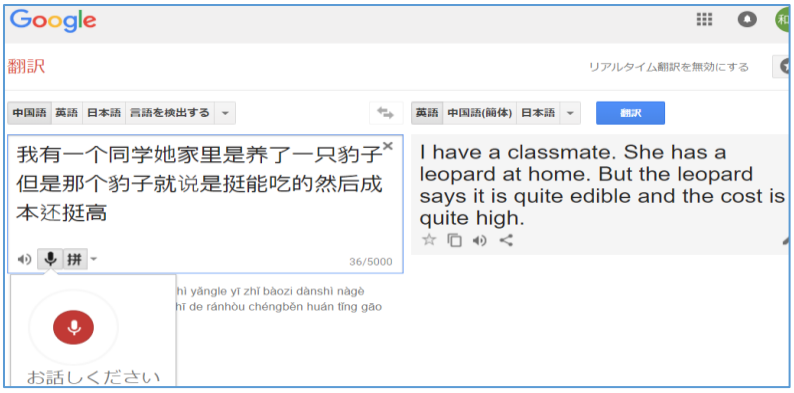


图 1 Google 对 NS 话语的识别

有关 6 个 ASR 听错的词，主要因为发言人把那些词说得较重而造成的。比如 1A-3 段，WT/NS1 将句中“大”和“养”的语调稍微重又长些**表 4**、1B-3，结果 ASR 分别听错为“打”“讲”等字，甚至发生同一个语音换成几种不同文字的情况**表 3**。其实“养”是当天讨论题目的核心动词，大家在讨论中都反复听和说过好多次，并在此段前面 1A-1 段里也出现。所以参加讨论的同学不会听错的。果然不少 NS 与 NS-EC 同学已经笑出来**表 4**、1B-c。

既然 WT/NS1 的发音都符合 ASR 的语音要求，如汉语母语者标准语音，手拿话筒**图 2** 近讲，约 200 多字/分的说话速度[11]等，为何 ASR 识别失败？原因在于 ASR 跨句交叉匹配功能远不如人类。尽管长短时记忆神经网络（Long Short Memory Networks:LSTM）及它的提升版为目前在 ASR 声学模型中广泛应用的算法，可以有效的对时序信号的长时相关性进行建模。理论上可以看到无穷远的上下文，长时间可以保存记忆[12]，实际上，程序运行速度和精度之间此升彼降的矛盾关系，成为制约 ASR 性能提升的瓶颈（中村，2017;篠田，2017）（第 4 节再述）。

另外 ASR 把两处“豹子”分别听错为“报纸”和“能报”（**表 3**，1B-2、**表 4**，1A-2、1A-5，WT/NS1 和 T/NS1 的发言），不仅与原发言内容相去甚远，整句的意思差强人意。ASR 识别错误的原因除了 WT/NS1 对这个词语气放得重些外，还有 T/NS1 说话里夹杂重叠和笑声**表 4**，1A-2，导致话速减慢（约 140 多字/分），造成 ASR 识别不清。后面再有 T/NS1 和 T/NS2 两名台湾学生同时说出“花豹”等词汇，结果发话重复（1B-2），造成 ASR 停掉了识别工作，出现漏听现象。尽管最新 ASR 会支持一个一音频文件中包含不同话者的识别，如果多个声源在相近的方位，ASR 还是难以识别。包括 T/NS1，两名台湾学生都是汉语教学专业。NS1 的华语稍带有南方口音，但其他地方的发言被识别得很清楚，说明他说的汉语足够标准。可知即便是 NS 的自然标准口语，ASR 还难于识别多通道的语料。

[11] 中国人普通语速是 200 到 250 字左右/分，新闻播音速度在约 300 字左右/分（孟国 2006）。
[12] 《瞭望》采访科大讯飞鄢志杰，https://www.iyiou.com/p/78283 [2018-08-04]

ASR 同步汉英翻译输出了不可思议的译文**图 1，表 3** 1A-3，**表 4** 1B-3，这是因为 MT 对文章语境缺乏整体把控能力。此处没能抓好中文的主谓结构，结果英译失败。

- 但是[据]说 [它]的成长速度非常的快→[机器翻译]＊The leopard says it is quite edible and the cost is quite high.）
- 过了半年就已经比我那同学还要打[大]了→[机器翻译]＊After half a year, he has already played more than my classmate.

上面重点指出目前 ASR 在语音识别方面的盲点。智能语言工具还未能支持的背后，往往含有人类特有的交互方式。先将发言内容与非发言及互动行都转写出来，其后在第 4 节将综合分析群体讨论会上，NNS 如何与 NS 交互并共享信息。

表 3（会话片段 1A）Google ASR 识别结果

（下划线处是识别错误）

（ASR 转换后的文本一律无标点符号，去除了元发言的语气词和重复等口语成分。下划线红色是识别错处）

话轮记号	Speaker 记号	ASR 转换后的文本
1A-1	WT/NS1	我有一个同学她家里是养了一只豹子但是那个豹子就说是挺能吃的然后成本还挺高
1A-2	T/NS1	我想确认一下是报纸是就是一种大型的猫科动物有就（之后不识别）
1A-3	WT/NS1	是的他开始进的时候那个豹子还很小但是说他的成长速度非常的快然后过了半年就已经比我那同学还要打了
1A-4	T/NS1	你的同学现在还好吗
1A-5	WT/NS1	家里人呀我也同学还不错能报这是在笼子里的所以现在可能还没有问题呢
1A-6	T/NS1	那行谢谢谢同学分享

表 4（会话片段 1 B）人工转写

话轮记号	Speaker 记号	交互功能标记 (大约按出现顺序标记)	Time(sec) Total 65 sec.		发言转写（含交互行为的描述）
1B-1	WT/NS1	(A)(G)	9	12	*啊：：我有一个同学,然后他家里是养了一只豹子↑,但是那个豹子就是说很能吃的，然后成本还挺高.*
1B-a	交互行为	(A)(E)(S)			包括 T 校其他学校 NS 与 NS(EC) 表示惊讶或发笑。NNS 倒没有表情反应，恐怕没听懂“豹子”这种词汇
	静场		3		
1B-2	T/NS1	(A)(G)(M)(E)(S)	13	14	*ah…〈我想确认↑一下是..豹子↑是..就是那种..大型↑的..猫科动物..有..就是 hh..有耍师..那一种 hh 花豹的那一种 hh 吗.〉*

	T/NS2	(A)(G)(R)	重复		*（同时说）ng 是花豹*
1B-b	交互行为	(S)(M)(A)			13 个同学当中有 5-6 个 NS 与 NS(EC)有笑脸，但 NNS 还没有笑脸，一名 NNS 开始查电子词典。
	静场		1		
1B-3	WT/NS1	(G)(A)	12	15	*啊：是的，但是他开始养↑的时候那个豹子还很小，但是据说它的成长速度非常地快，然后过了半年就已经比我那同学还要大↑了.*
1B-c	交互行为	(A)(M)(S)			WT 校一名 NS 拿出自己的手机给旁边坐的两名 NNS 看，他们侧身观看之后才露出笑脸(图 3)
	静场		3		
1B-4	T/NS1	(M)(A)(E)(R)	3	8	*〈..你的..同学..现在..还..好 hh..吗.〉*
1B-d	交互行为	(S)(E)			几乎所有同学笑逐颜开，有的发出笑声
	静场		5		
1B-5	WT/NS1	(G)(E)(A)	8	9	*是他家里人养的，我同学还不错，豹子是在笼子里的，所以现在可能还没有问题的.*
1B-e	交互行为	(S)(E)(R)			所有同学发笑，幕后的教员也发出了笑声
	静场		1		
1B-6	T/NS1	(E)(S)(R)	6	*7*	*hh hh [好..谢谢..谢谢 hh..这位同学的分享] .*
1B-f	交互行为	(S)(M)			在大家的爆笑中主持人开始指定下一个发言人

3.2 ASR 对非母语者语音的识别结果

会话片段 2A 是一名非汉语学生（B 校 NNS1）与其他发言人的一轮对话。下面是用 Google Translate 转换成文字信息的结果。

表 5 （会话片段 2A）Google ASR 识别结果（下划线红色处是识别错误）

话轮记号	Speaker 记号	ASR 转换后的文本
2A-1	B/NNS1	我以前养过的苍雪的话跟家里的衣柜去宠物店然后一起商量来的但是我现在也现在在家里养的兔子的话我爸爸就你一个人去买的所以引起了妈妈的反感谢谢
2A-2	B/NNS1	其实我妈妈基本上讨厌宠物兔子的话比较一下规模比较大的优点就座椅沙发没有那么做所以基本上没有问题的但是就是因为那个爸爸去工作吧所以那个时候养要养的是妈妈吧所以就让买买醉的发自然后引起了很大的反感谢谢

2A-3	T/NS1	请问那只兔子还在吗
2A-5	B/NNS1	还在已经碎了

据此 ASR 输出的文字，很难理解这轮会话的内容（2A-1，2A-2，2A-5）。与下面对同一轮会话内容的人工转写结果**表 6** 相比较，得知 B/NNS1 发不准汉语发音造成 ASR 识别失败。一是他发[u]时,圆唇的程度不够（如“鼠、突、臭、五”）中，二是送气不够（如“突、臭”），另外后鼻音与四声不到家（如 “然、反、感”），导致 ASR 都不能正确识别，甚至输出的文字一团乱麻。

表 6 （会话片段 2B）人工转写

话轮记号	Speaker 记号	交互功能标记 (大约按出现顺序标记)	Time(sec) Total≒17.7min.		发言转写（含交互行为的描述）
2B-1	B/NNS1	(G) (S)	12		((确认自己手机上的发言稿))大家好,ah: :我,我在日本,我现在,养的是兔子,一个兔子. ah: :之前养过的是仓鼠.
2B-a			≒14 min.		此后 4 校 13 名学生都轮流发言，共有 26 个话轮
2B-2	B/NNS1	(S)(G)	23		好，我以前养过的仓鼠的话 ah：：跟家里人一起去 ah 宠物店=然后一起商量买的.但是 eng 我现在我现在在家里养的..ah：兔子的话..我爸爸突然一个人去买的..ah：所以 eng：引起了妈妈的反感..谢谢.
2B-b	交互行为	(S)(R)	≒2 min.		B NNS1 说完后和旁边的同学一起露出了笑颜。其他同学，除了主持人点头表示理解外，并没有笑。此后 3 校 5 名学生轮流发言，共有 12 个话轮.
2B-3	WT/NNS3	(S)(M)(G)(A)(R)(G)	13	16	((WT/NNS3 抢着话说))［不好意思］可以问一下北京同学的刚刚 ah：他说有愿意说引起了妈妈的反感 ah hh 我想问一下这个故事 hh 可以吗 hh hh.
2B-c	交互行为	(S)(A)(R)			WT 校所有同学笑容鼓励。主持人送视线同意，指定 WT/NNS1 发言。
	静场		3		
2B-4	B/NNS1	(S)(E)(A)（G）	43	45	((笑声应邀，比手划脚地))ah：aha ha 其实我妈妈基本上讨厌：宠物 eng.兔子的话，ah 比较比较规模比较大，而且有点臭，所以 ah：：仓鼠的话 ah：：没有那么臭.所以基本上是没有问题的.但是兔子的话，eng：：因为那个 ng：爸爸去工作吧，所以那个时候 ah 要养的是妈妈吧 hh.所以突然买买兔子的

					话 eng：突然买兔子然后引起了很大的反感，谢谢.
2B-d	交互行为	(S)(E)(R)			B 说完后满开笑脸。还有 3-4 名同学笑颜鼓励，但 S 主持人没有笑。
	静场		2		
2B-5	T/NS2	(A)(S)(G)(R)	3	5	*请问那只兔子还在吗*↑
2B-e	交互行为	(S)(E)(A)			旁坐的主持人 S 发出嗤笑声，其他同学也随着发出笑声或笑脸。
	静场		2		
2B-6	B/NNS1	(G)(E)(S)	3	3	*兔子还在，已经五岁了.*
2B-f	交互行为	(S)(E)(R)			所有同学都有笑容或发笑声

表 7 本文中语料文字转写及交互行为的标记

(3 sec.)	话语计时	↑	升调
,	话语继续	ah,hh,eng	重叠，笑声，吸呼气等副语言
.	话语结束	(( ))	笔者说明
:	表示拖音 (冒号越多表示拖音越长)	（G）	目标导向策略:如用长期记忆，语言功能等达成目标
..	停顿	（R）	奖励策略:如用动作，情感等表示鼓励
=	紧随话语	（A）	注意力策略:如感知联想，语言提醒等
--	话语截断	（M）	动机转移策略:调整目标优先级
［］	同步话语	（E）	情绪调节策略:如快乐，焦虑等
＜＞	符号内是语速明显较慢的话语	（S）	社会共存策略:如礼仪，同情，笑容等

图 2 NS 与 NNS 互动镜头(2B-3)

4. 非母语者交互能力与未来 ASR 技术

基于前节测验的结果，下面对 NNS 的多模态，多通道交互行为进行分析。然后整理当前 ASR 的功能范围及其未来应用技术的发展趋势。

4. 1 NNS 采取多模态交互策略

参加“讨论课”要求 NNS 学生 CEFR 的 B 级以上语言水平，不过大多数 NNS 离 B 级应有的独立交际能力还有一段距离，有时 NNS 汉语发音欠佳影响口语交际。ASR 更不善于识别非标准语音，结果输出让人费解的文字**表 5**。实际上，其他同学以记忆或类推与核实等手段，都理解 B/NNS1 想要说什么。对话句法理论指出，人类自然交流中，重复现象可以帮助构建话语的衔接和连贯。此时，感知、意识、记忆、注意、联想、动机和情绪等多模态信息为整体话语的语义连贯和深层理解提供充足的资源(Tannen, D. 2007; Susan M.et al., 2008）。**表 4** 与**表 6** 所示，多处可以发现 NNS 用多模态交互策略，试图补偿二语能力的差距与 NS 进行沟通。下面基于 GRAMES 改为谈话策略的标记，重新整理 NNS 与 NS 的多模态交互行为。

（表 6 会话片段，2B-3 到 2B-f，共约 69 秒） 划下线[__]指谈话策略，(S)(M)等指策略分类

B/NNS1 说完后，他好友 WT/NNS3 拿过话筒[(S)，(M)动机转移]，引起 B/NNS1 说过的一句[(G)目标导向]，向他要求澄清一遍 [(A)注意力，(R)(G)]。然后 WT 校所有同学满脸笑容鼓励(R)。主持人送视线同意，指定 WT/NNS 发言 [(S)(R)]。B/NNS1 就发出笑声应邀[(S)，(E)情绪调节]，比手划脚地[(S)]重述了前段的内容(2B-4)[(A)(G)]。此间其他同学们以笑颜鼓励（2B-c，2B-d）[(S)(E)(R)]。此后另外一位 NS 颇有风趣地插话提问(2B-5)[(A) (S)(G)(R)]，旁坐的主持人发出笑声，其他同学也随着发出笑声或笑脸[(S)(E)(A)]。对此 B/NNS1 斩钉截铁地答复(2B-6)[(G)(E)(S)]使得大家都笑起来（2B-f）[(S)(E)(R)]。

初步统计所用 GRAMES 策略的结果是;(S)社会共存策略 9 次＞(R)奖励策略 6 次＞(G)目标导向、(E)情绪调节、(A)注意力等策略，均 5 次＞(M)动机转移策略 1 次(次数不含重复出现的，下同)。其中 NNS 运行的策略从多到少排序后有，(S)7 次＞(E)5 次＞(R)4 次＞(G)(A)均 3 次＞(M)1 次。可见不仅是 NS 还是 NNS，运用社会共存策略的占多数，其次是奖励、目标导向、情绪调节、注意力等策略受 NS 的重用。而 NNS 较多用情绪调节策略的倾向。尽管随意选来的短时交互语料，这种结果值得我们重新思考群体讨论的组织框架与条件。

4 .2 NNS 运用社会共存策略

NNS 与 NS 共同维护课堂是“讨论课”的无言共识。各单位的教师、助手与技术人员在后方支持课程活动的顺利运行，保证大家平等参与讨论（砂冈，2016）。如上所示同学最多用**(S)**社会共存策略，可以认为跟这堂课的组织原则有关。课堂需要自我管理，使大家有意识和无意识地注重社会共存策略。其中被 NNS 采取的(S)策略有如，用社会礼仪打招呼（说谢谢）、笑颜笑声达成感情共鸣（DuBois，2014）、用手势或加强语气示意（比手划脚或斩钉截地答复）或主动帮助别人（拿过话筒），平衡分布话语权等行动。其他策略与(S)衔接，相互连贯加深交互（安西，2017）。可知 NNS 虽受外语能力的制约，对语言管理意识并不损于 NS。NNS

常用笑容礼仪等(S)策略，要缓解群体讨论的紧张气氛，成功扮演一个共生情感的中介人角色，与同侪合作共享信息(砂冈, 2018b)。

NNS 能够参与讨论离不开与 NS 的互动。其中与 NNS 接触经验丰富的 NS-E(Native Speaker-Experienced)，无意识地监测双方信息传输是否顺利，还可以具体提出支援方案等语言管理能力(栁田，2015; Sunaoka, 2018a)。如上面 WT/NNS1 经常注意全体会员有无举手要发言，将视线放到发言人的方向，使用体态语表示“了解”。另外一名 NS-E（W/NS2）主动为 NNS 提示生词，帮助他们语义理解，让他们能赶上会话进度有贡献**表 4**，自 1B-b 至 1B-d，e。譬如：

W/NS1 话中的“豹子”这个词汇，起初只有部分 NS 表示惊讶或发笑声，而 NNS 他们似乎没听懂，没有表情反应（1B-a）。此时主持人 T/NS1 向 W/NS1 确认是否真的“豹子”（1B-2）。这句话因发自于 T/NS1 的内心疑惑和惊奇的感情，富有情感，话速变慢。正因为这个语音特色造成 ASR 听别失败，而浓厚情感的音色与音调引起了大家对“豹子”这个生词的选择性注意（王萍丽等人, 2015)。显然 T/NS1 发言后一名 W/NNS3 开始查字典，但还是弄不明白。于是旁坐的 W/NS2 拿出自己的手机（看来帮他们查字典或网络检索）给旁边坐的两名 NNS 看（1B-c）**图 3**。他们侧身观看之后才露出笑脸。此后两名 NNS 都能赶上会话的进度，带着笑容参与讨论（1B-d，e），最后达成了全体成员对 W/NS1 发言内容的理解。NS-E 的无言社会共存行为促进了 NNS 交互与共享信息（Sunaoka, 2018a）。

4 .3 NNS 长时间运用情节记忆

参加网络视频交流活动，由空间距离和时间距离，导致身体动作（姿势、视线交集等）、感官知觉（如听觉、视觉、嗅觉、触觉）的使用都受到一定限制，从而提高参与者之间的交互成本(Clark & Brennan，1991)。正由此限制，群体讨论成员多用情节记忆（Episodic memory），注重推理和联想,以提高与同侪共享信息的效果。如自从 WT/NNS 第一次介绍自己养宠物经验后，第二次重述，竟花了 18 分钟**表 6**。中间还夹有几段其他同学的发言，之后因他好友 WT/NNS2 的澄清要求，WT/NNS1 有机会重复自己的发言，通过核实、重述、交叉匹配等手段，最终消解了大家的歧义。此间大家保持运用长时间情节记忆，尽管要花不少时间，加强互动共享信息。现在 ASR 既有长短时记忆功能（第 3 节已述），最新又出现了装有情节记忆能力的机器人[13]，但目前只能对物体的情节记忆，未能支持对抽象情节的记忆。智能语言工具还难以突破 NNS 学生长时间运用情节记忆的能力。

4.4 ASR 的现在与未来发展趋势

ASR 即使是人工智能目前落地最成功的技术，不过如上所示，ASR 与人脑的认知相比还存在诸多缺陷，但总是在朝着精度不断提高、功能更强大的方向发展。下面引自据最新文献和资料（中村等人,2017;篠田,2017;陈师哲等人，2018），总

[13] 麻省理工学院开发的「ComText」http://news.mit.edu/2017/robot-learns-to-follow-orders-like-alexa-0830. [2017-08-30]

结现阶段 ASR 的基本原理与功能范围，在此基础上介绍 ASR 及其应用技术的未来发展趋势。

ASR 整个系统由语音识别（听写功能），机器翻译（MT），语音合成（TTS）3 种技术组合而构成**图 3**，下同。其中统计匹配模式是整个系统的核心部分。近年来，基于海量语音、语言数据库，通过云端加强 DNN 训练的搜索算法模型，包括 TTS，ASR 与 MT 的精度亦获得了迅速提高。不过这些模式要求的数据都不是原始“自然口语”的语料，而是“标准声音信号”。由于目前 ASR 在嘈杂环境中获取特定声音较为困难，尤其在语言切换时，系统容易出现混乱。因此先去除副语言和非语言成分等“冗余信息”，之后对“声学波形”进行端点检测（除多余的静音和非说话声音）、降噪（去掉杂音与噪音）、特征提取等筛选处理（除掉口音，性别及年龄等个人语音特点），保留语音的“关键信息”，再按照一定规则对数据加以整理构成模式库，最后实行模式匹配。因此在应用 ASR 时，环境要安静、单人（Single Speaker）、说话要离麦克风近、发音要标准、不能持续对话、不能打断[14]。最近 ASR 对情感语音的识别精度有所提升（陈师哲等人，2018），但分开语音所含的韵律特征与说话者特征是语音工学的长年课题，至今尚未解决（篠田,2017）。

有关话者的姿势、视线交集等身体动作以及如视觉、嗅觉、触觉等非语言感官知觉信息的研究本属于开发仿生机器人（Human robot）的范畴（浅田，2010）。近几年 ASR 朝着同非语言技术并存与互补方向发展，如情感语音特征提取研究与图像识别技术的兼容不断取得进展（石黒,2015; 石田等人,2018），已开始应用到汉语的实用业务服务（潘忠德等人,2015; 韩伟等人,2016; 戴礼荣等人,2017; 李银河等人,2017）。但如何提取不同模态下（比如面部表情）最优的情感特征还没有定论。不同模型只能用分流、分层进行运算，计算复杂度极高，机器负载亦过大（上同）。

至于 ASR 的同步翻译功能，DNN 系统往往将词汇表限制为高频词，并将其他所有低频词视为未登录词（刘洋，2017）。词汇量的受限导致译文中出现原文单词漏译和重复翻译、少见的专有名词翻错或 unknown words 等现象（同上）。最新 ASR 提升了长时间记忆性能，但如上所示，还远不如人的记忆功能。尽管 DNN 对 MT 与 ASR 具有里程碑意义，但整个对话系统还是不完美的。因此目前对话系统只能面向旅游查询、订票、数据库检索等一个狭窄领域。包括情感语音识别，语言符号与情境、社会之间的语义协商能力，仍然是 ASR 的挑战课题（陈师哲等人, 2018; 刘振焘等人, 2017; 邵兵等人, 2016）。

[14] CSDN 博客 https://blog.csdn.net/ffmpeg4976/article/details/52348412 [2016-08-28]

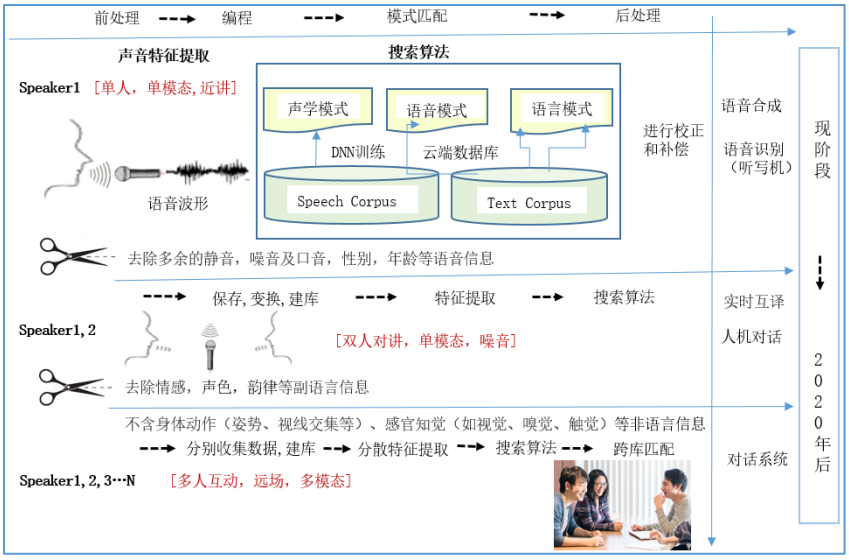


图 3 ASR 到对话系统的基本原理图（改编自中村哲，2016）

5. ASR 在汉语教学应用小结及展望

虽然人工智能语言技术还不太靠谱，但用起来省时省力，已成为一个能支持 NNS 二语习得的重要辅助工具之一。下面介绍几项 ASR＋汉语教学方案。

5.1 ASR＋汉语教学方案（由易到难的排列）

1) 中文听写机＋TTS（初中级）

ASR 可以作为一个很能干的“听写机”，它将逐字逐句听懂汉语并转化成书面中文字，既能省力又能增加自己可操作的教材。譬如为讲演或电影自动生成字幕（张雪梅等人，2012），口述听写成书面教材等，不仅对 L2 学生有用，还为听障人带来福音（河原等人，2018）。另外将 ASR 技术与(Text To Speech: TTS)相结合的教学模式更能提高二语教学的效果。TTS 是已达到成熟的一门技术，它精度比 ASR 还要高，并有如男生、女生、年长者、年轻者等可以选择有不同声音的语音朗读服务功能。如果用 TTS 听不同情感的朗读语音，可以提高汉语的听力能力，又能帮助学生自学能力。

2) 发音练习（初中级）

ASR 可作为一个辅助汉语发音练习的工具（笪骏，2015; Da,2018），如上述 NNS 说有几句发音不准确的话语，ARS 及时识别出了错误的汉语。借以 ARS 这种听写机的功能, 二语学生可以发现自己说汉语中的一些问题。还可以把教材文章朗读出来，用 ARS 来评估他/她朗读结果是否与原内容一致。缺点是 ARS 没有明确提示何处错、为何错。甚至还会出非类人或令人无法理解的识别错误。学生只好把输入用的文本和输出后的字符串相对照才知道自己哪里错。从语言教学的角度来讲，

这种靠感知和摸索的学习法，效率不佳。也可先让学生用 TTS 做复述练习（shadowing），然后再用 ASR 复述会减少二语学生的学习负担。

现在出现即时对汉语学习者的发音给出打分和正音反馈的产品。比如“尔雅中文 app”是一款以学生为主体进行发音练习，同时帮助老师监测学生的发音情况，为汉语教师提供客观数据反馈，以提高课堂语音教学效率，更科学的发音教学模式（魏巍等人，2018）。上面已述在发音时要注意环境安静、单人、标准发音、靠麦克风近讲、不能打断且不要太长，以免 ASR 识别失败。

3) 口语类机考（中高级）

ASR 技术已作为对英文学生进行口语类考试计算机考（机考）的实现手段。采用自动评分的方法，能够根据评估和反馈的结果，对口语考试实行进一步的优化，从而提高口语教学的质量，来满足学生口语学习的需要和口语能力的培养需求（陈卫兵,2015）。在完成人机对话模拟测试之后，教师应要对学生的测试结果进行评价总结，给予学生指导自己汉语能力缺陷的原因，以此弥补 AI 语言技术的局限性(上同)。

4) 对话练习（高级）

ASR 已开始应用到实用业务的问答系统（Question Answering）和信息抽取（Information Extraction）等领域中。但它对人机交互功能还有局限，因此对二语教学的应用目前限于较简单的指令控制，以及对口译课堂教学的部分应用（李霄垅等人, 2018）目前人机对话对二语习得的应用以英文教学占多数（陈卫兵,2015;杜建萍,2018;刘菁菁,2002），而中文教学的却少见。英文教学的例如 Microsoft Research 利用最新的人工智能聊天机器人，将华语为母语人士建立机器学习模型，可直接语音和软件沟通对话。日本一家私企开发的学英文 APP「Tera Talk」，通过与 AI 的模拟对话训练引导学生掌握口语表达技巧，低廉有效地自习英语，产品已推广到 136 个国家[15]。

随着 AI 语言技术的不断完善与成熟,基于云端和大数据的 ASR 引擎可以提供关于真实世界的信息反馈，还可以部分模拟真实世界的人际交互。如果未来 ARS 技术成功应用到互译系统上，它会成为支持交互式口语教学的重要辅助技术（朱坤鸿,2018）。

5.2 展望和课题

通过以上 ASR 与远程讨论课学生之间的语义协商的对比分析，得知 NNS 以跨段、多通道、多模态作为沟通渠道，与 NS 学生合作进行消解歧义。情感语音促成互动共鸣，自律协助行为课堂管理以及语义协商取得了很大的成效，使得全体最终

[15] https://itunes.apple.com/jp/app/terratalk/id1114037031?l=en&mt=8（iOS 版）
https://play.google.com/store/apps/details?id=jp.co.joyz.terratalk（Android 版）

达到了理解。与此相比，当前 AI 工具还严重缺乏社会性，无法主动选择目标，灵活转移注意力、调节情绪而达成目标。自我管理是外语教育的关键能力之一，它决定知识与技能的发挥效率[16]。人的语言无论是选择词汇，还是语气声调，都是自我管理的彰显，这才是“自然”语言的魅力所在。当前 ASR 还未支持副语言与非语言成分的识别和理解。期待未来 ASR 存有足够的发展空间**图 3**，与其他 AI 技术相结合，可以构建出更加复杂的应用，将成为一个能支持二语教学的主流工具。因此我们不要偏激地、激烈地否定或反对相关技术，也不要盲目憧憬它，而是理性地进行取舍。

当然只借助输入等待 AI 输出的结果，中间没有任何认知负荷，也很少与它互动合作协助等交互过程，可否益于二语习得是个最大的疑问。我们知道语言与认知的交互作用，如社会能力、情感、感知、运动、记忆功能、思维与解决问题等能力都为语言发展的重要认知资源。同时, 注意、监控、调整、纠错和反馈等一系列语言习得过程又是认知功能深化的过程(今井等人，2014)。因论述范围太大，这里暂不讨论，另有机会再进行探索。

“亚洲学生远程讨论课”在华语地区合作单位的支持下，已实施 16 年之久，并已保存了 300 多小时的现场录像（砂冈，2016）。本文只能用手工转写语料，观察分析都靠目视和听觉。今后期望利用视觉传感技术（Visual image sensing）以及人脸表情识别等技术，深层挖掘多模态信息（陈华斌等人,2017;李淑婧等人,2015; 松居，2018;佐藤，2017），将建构多模态二语群体讨论知识库，为汉语自然语言理解与教学研究提供数据。

致谢：本论文在第 10 届国际汉语电脑教学研讨会（TCLT10-2018，於台湾师范大学）以及第 11 届中文现代化国际研讨会（AMCL11-2018，於澳门科技大学）上发表后，再在此基础上进行修改并补充完成。在会上得到同仁指导，至此致谢！

参考文献

Anzai, Y. (1992). *Towards a new paradigm of human-robot-computer interaction*. Proceedings IEEE international workshop on robot and human communication, 11-17, Tokyo, Japan: IEEE.

Anzai, Y. (2013, March). *Human-robot interaction by information sharing*. Plenary lecture given at the 8th ACM/IEEE International Conference on Human-Robot Interaction, Tokyo, Japan.

Anzai, Y. (2017). Theory of interaction by information sharing. *Cognitive science. 24*(2), 234-260. [安西祐一郎. (2017). 情報共有によるインタラクションの理論. *認知科学, 24*(2), 234-260.]

[16]王文斌於第三届全国高等学校外语教育改革与发展高端论坛: http://heep.unipus.cn/news/content.php?NewsID=4322 [2018-03-27]

Arita, Y., Ogata, J., Fujimoto, M., & Tsukada, K. (2003). Sports live speech recognition using acoustic and language model adaptation: Application to highlight scene detection. *The Journal of the Institute of Electronics, Information and Communication Engineers (IEICE) Technical Report*, *102*(618), 33-40. [有木康雄, 緒方淳, 藤本雅清, & 塚田清志. (2003).音響・言語適応処理を用いたスポーツ実況中継音声の認識: ハイライトシーン検出への応用. *電子情報通信学会技術研究報告*, *102*(618), 33-40.]

Asada, M. ((2010). *What is a robot: Challenge the mystery of human brain and wisdom*. Tokyo, Japan: NHK Publishing, Inc. [浅田稔. ((2010). ロボットという思想 : 脳と知能の謎に挑む.東京: NHK 出版協会.]

Chen, H., Kong, M., Lv, N. (2017). Status and development of vision sensors on intelligentized robotic welding technologies. *Electric Welding Machine*, 47(3) , 1-7. [陈华斌,孔萌, & 吕娜. (2017).视觉传感技术在机器人智能化焊接中的研究现状. *电焊机*. 47(3), 1-7.]

Chen, P. (2017). Three waves of development in China's language industry. *Chinese Journal of Language Policy and Planning*, *5*(11), 20-28 [陈鹏. (2017). 当代中国语言产业发展的三次浪潮. *语言战略研究*, *5*(11), 20-28.]

Chen, S., Wang, S., & Jin. Q. (2018). Multimodal emotion recognition in multi-cultural conditions. *Journal of Software*. *29*(4), 1060-1070. Retrieved from http://www.jos.org.cn/jos/ch/reader/create_pdf.aspx?file_no=5412&journal_id=jos [陈师哲, 王帅, &金琴. (2018). 多文化场景下的多模态情感识别. *软件学报, 29*(4), 1060-1070. 参见 http://www.jos.org.cn/jos/ch/reader/create_pdf.aspx?file_no=5412&journal_id=jos]

Chen, W. (2015). Activity design of the seventh grade English oral teaching based on human-machine dialogue. *English Teachers*, 17, 59-62. [陈卫兵. (2015). 基于"人机对话"背景下七年级英语口语教学的活动设计.*英语教师*, 17, 59-62.]

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC, US: American Psychological Association.

Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning? *@tic. revista d'innovació educative*, 10, 108-122. Retrieved from https://dialnet.unirioja.es/descarga/articulo/4334892.pdf

Da, J. (2015). The Application of speech recognition technology in Chinese language learning: What can be learned from a pinyin lab session. *Journal of Technology and Chinese Language Teaching*, *6*(1), 16-24. Retrieved from http://www.tclt.us/journal/2015v6n1/da.pdf [笪骏. (2015). 语音识别技术在中文教学中的应用：一堂汉语拼音练习课的启示. *科技与中文教学*, *6*(1), 16-24. 参见 http://www.tclt.us/journal/2015v6n1/da.pdf]

Da, J. (2018, June). *Speech technology in Chinese language learning and teaching*. Workshop given at the 10th International Conference and Workshops on Technology and Chinese Language Teaching (TCLT 10), National Taiwan Normal University, Taipei.

Dai, L., Zhang S., & Huang Z. (2017). Deep learning for speech recognition: Review of state-of-the-arts technologies and prospects. *Journal of Data Acquisition and Processing*, 2, 221-231. [戴礼荣, 张仕良, 黄智颖. (2017). 基于深度学习的语音识别技术现状与展望.数据采集与处理. 2, 221-231.]

Do, Q. T., Toda, T., Neubig G., Sakti, S., & Nakamura S. (2017). Preserving word-level emphasis in speech-to-speech translation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *25*(3), 544-556. doi: 10.1109/TASLP.2016.2643280

Du, J. (2018). Practical research on improving middle school students' English listening ability. *Popular Science*. 7, 29. [杜建萍. (2018). 提高中学生英语听力能力的实践研究. *科学大众(科学教育)*, 7, 29.]

Goffman, E. (1981). *Forms of talk*. Philadelphia, PA: University of Pennsylvania Press.

Graves, A., Wayne, G., Reynolds, M., Harley, T., Danihelka, I., Grabska-Barwiska, A., … Hassabis, D. (2016). Hybrid computing using a neural network with dynamic external memory. *Nature*, 538, 471–476. Retrieved from https://www.nature.com/articles/nature20101

Han, W., Zhang, X., Bai, S., Zhang, R., & Ma, M. (2016). Deep learning for speech recognition, *Journal of Military Communications Technology*, 3.91-97 . [韩 伟, 张雄伟, 白崧廷, 张瑞昕, &马鸣. (2016). 深度学习理论及其应用专题讲座(四)第 7 讲深度学习在语音识别中的应用.军事通信技术, 3, 91-97]

Hori, T., Araki, S., Yoshioka T., Fujimoto, M., Watanabe, S., Oba, T., … Yamato, J. (2011). Low-latency real-time meeting recognition and understanding using distant microphones and omni-directional camera. Transactions on Audio, Speech, and Language Processing (IEEE). *Audio Speech Language Process*, *20*(2), 499–513.

Imai, M., & Saji, N. (2014). *Language and physicality*. Tokyo: Iwanami Shoten. [今井むつみ, 佐治伸郎. (2014). *言語と身体性*. 東京: 岩波書店.]

Ishida, M., Inoue, K., Nakamura, S., Takahashi, K., & Kawahara, T. (2018). An attentive listening system generating empathetic responses and promoting user utterances. *Journal of the Japanese Society for Artificial Intelligence SIG-SLUD*, *B5*(3), 7-12. [石田真也, 井上昂治, 中村静, 高梨克也, & 河原達也. (2018). 共感表出と発話促進のための聞き手応答を生成する傾聴対話システム. *人工知能学会研究会 SIG-SLUD, B5*(3), 7-12.]

Ishiguro, H. (2015). Studies on interactive robots. *Journal of the Japanese Society for Artificial Intelligence*, *30*(3), 377-382. [石黒浩. (2015). 人と関わるロボットの研究. *人工知能*, *30*(3), 377-382.]

Kawahara, T., & Akita, Y. (2018). Captioning lectures using automatic speech recognition for hearing-impaired people. *Acoustical Science and Technology*, *74*(3),1-8. [河原達也, &秋田祐哉. (2018). 聴覚障害者のための講演・講義の音声認識による字幕付与. *日本音響学会誌*, *70*(3), 1-8]

Li, S., Ji, P., Deng, J., Sun, B. & Q. Liu. (2015). Component-based facial expression recognition. *Application Research of Computers*, 3, 917-921,941. [李淑婧, 嵇朋朋, 邓健康, 孙玉宝, &刘青山. (2015). 基于面部结构的表情识别. *计算机应用研究*, 3, 917-921,941.]

Li, X., & Wang, M. (2018). Construction and research of the teaching model of using automatic speech recognition APP in simultaneous interpreting training course—A case study of voice note as an auxiliary tool. *Media in Foreign Language Instruction (TEFLE)*, 179, 12-18. [李霄垅, &王梦婕. (2018). 基于语音识别 APP 的同声传译能力培养教学模式建构与研究—以科大讯飞语记 APP 为例. *外语电化教学*, 179, 12-18.]

Li, Y., Li, X., Xu, N., Zhong, W., Zhao, X., Cheng, X., … Yuan, J. (2017). Research on speech emotion recognition classification algorithm. *Journal of Nanyang Normal University*, 6, 28-33. [李银河, 李雪晖, 徐楠, 钟文雅, 赵新仕, 程晓燕, … 袁键. (2017). 语音情感识别分类算法研究综述. *南阳师范学院学报*, 6, 28-33.]

Liu, Q. (2002). The man-machine method used in teaching speaking in college English. *Media in Foreign Language Instruction*, 1, 15-17. [刘菁菁. (2002). 人机对话在大学英语口语教学中的应用. *外语电化教学*, 1, 15-17.]

Liu, Y. (2017). Recent advances in neural machine translation, *Journal of Computer Research and Development*, *54*(6), 1144-1149. [刘洋. (2017. 神经机器翻译前沿进展. *计算机研究与发展*, *54*(6),1144-1149.]

Liu, Z., Xu, J., Wu, M., Cao, W., Chen, L., Ding, X., … Xie, Q. (2017). Review of Emotional Feature Extraction and Dimension Reduction Method for Speech Emotion Recognition. *Chinese Journal of Computers*, 41(12), 2833-2851. Retrieved from http://cjc.ict.ac.cn/online/onlinepaper/lzt-20181213153256.pdf [刘振焘, 徐建平, 吴敏, 曹卫华, 陈略峰, 丁学文, … 谢桥. (2017). 语音情感特征提取及其降维方法综述. *计算机学报*, *41*(12), 2833-2851. 参见 http://cjc.ict.ac.cn/online/onlinepaper/lzt-20181213153256.pdf]

Luo, H., Pan, Z., & Yi, Y. (2017). Current development and analysis on the prospects of artificial intelligence translation. *Electronics World*, 21, 21-23. [罗华珍, 潘正芹, &易永忠. (2017). 人工智能翻译的发展现状与前景分析. *电子世界*, 21, 21-23.]

Matsui, T. (2018). Learning analytics 4: Multimodal learning analytics. *Information Processing*, 59(9),810-814. [松居辰則. (2018). ラーニングアナリティクス 4. マルチモーダルラーニングアナリティクス. *情報処理*, *59*(9),810-814.]

Nakamura, S., Sakti, S., Neubig, G., Toda, T. (2017). *Introduction of speech-to-speech translation: It aims to automatic translation by computer*. Tokyo, Japan: Koronasha Press. [中村哲, Sakti, S., Neubig, G., &戸田智基. (2017). *音声言語の自動翻訳:コンピュータによる自動翻訳を目指して*. 東京: コロナ社]

Pan, Y., Wan, Y., & Chen, G. (2018). Neural information system research: Current status and prospects. *Journal of Management Science*, *21*(5), 1-21. [潘煜, 万岩, & 陈国青. (2018). 神经信息系统研究:现状与展望. *管理科学学报*, *21*(5), 1-21.]

Pan, Z., Cai, W., Zhu, J. & Cui, D. (2015). Application of speech recognition in the diagnosis of affective disorders. *Chinese Journal of Forensic Sciences*, 6, 85-89. [潘忠德, 蔡伟雄, 朱 杰, &崔东红. (2015). 情感障碍的语音识别研究进展. *中国司法鉴定*, 6, 85-89.]

Sacks, H., Schegloff, E. A, & Jefferson, G. (1974). Turn-taking system lectures on conversation (Nishizaka, A., Trans.). Kyoto: Sekaishisosha-Kyogakusha Co., Ltd. [H. サックス, E. A. シェグロフ, G. ジェファソン. 会話分析基本論集(西阪仰译.). (1974): 順番交替と修復の組織.京都: 世界思想社.]

Satoh, Y. (2017). Collaborative visual sensing for understanding group attention and behaviors. *Journal of the Japanese Society for Artificial Intelligence.* 32(5), 714-720. [佐藤 洋一. (2017). 集合視によるグループの注視・行動のセンシングと理解 (特集; 人と調和して協働する知的情報処理). *人工知能, 32*(5), 714-720.]

Schegloff, E. A. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff, & M. Thompson (Eds.), *Interaction and grammar* (pp. 3-35). Cambridge: Cambridge University Press.

Shao, B., & Du, P. (2016). The method of speech emotion recognition based on convolutional neural network, *Science and Technology Innovation Herald*, 6, 87-90. [邵兵, & 杜鹏飞. (2016). 基于卷积神经网络的语音情感识别方法. *科技创新导报*, 6, 87-90.]

Shinoda, K. (2017). *Speech recognition*. Tokyo: Kodansha Press. [篠田浩一. (2017). *音声認識.* 東京: 講談社.]

Sunaoka, K. (2016). Cooperative peer learning activity of online cross-cultural communication by sharing Emoji. *Journal of Technology and Chinese Language Teaching*, *7*(1), 43-55. Retrieved from http://www.tclt.us/journal/2016v7n1/sunaoka.pdf [砂冈和子. (2016). 移動通信終端促成遠程討論參加者間形成合作學習作用分析. *科技与中文教学, 7*(1), 43-55. 参见 http://www.tclt.us/journal/2016v7n1/sunaoka.pdf]

Sunaoka, K. (2018a). The interactive modes of non-native speakers in international Chinese language distance class discussions: The analysis of smiling as a facial cue. *Technology-Mediated Chinese language teaching; Innovation in Language Learning and Teaching*, 12, 24-34. London: Taylor Francis Group UK.

Sunaoka, K. (2018b). The language management awareness of L1 students in the international distance Chinese discussion class: A comparative analysis of the roles of NS and NNS in language contact situation. *Applied Linguistics Research on Chinese*, 7, 40-48. Beijing, China: The Commercial Press. [砂冈和子. (2018b). 国际远程汉语讨论课中母语学生的语言管理意识: 语言接触场面 NS 与 NNS 角色对比分析. *汉语应用语言学研究*, 7, 40-48. 北京: 商务印书馆.]

Susan, M., & Alison, M. (2008). Input, interaction, and output in second language acquisition. In B.V. Patten, & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 283-302). New York: Routledge.

Tannen, D. (2007). *Talking voices: Repetition, dialogue, and imagery in conversational discourse.* Cambridge: Cambridge University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origin of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675–691.

Xu, Q. (2017). Research on the innovative professional interpreting training system against the backdrop of artificial intelligence. *Technology Enhanced Foreign Language Education*, 5, 87-92. [徐琦璐. (2017). 人工智能背景下的专业口译教学系统的创新研究. *外语电化教学*, 5, 87-92]

Yanagida, N. (2015). *Japanese native speakers' communication strategies in contact situations.* Tokyo: Koko Books, Ltd. [栁田直美. (2015). *接触场面における母语*

话者のコミュニケーション方略：情报やりとり方略の学习に着目して. 东京: ココ出版.]

Wang, P., & Li ,Y. (2015). Effects of negotiation of meaning and turn-taking structure: A case study of naturalistic interactions between a native speaker and a non-native speaker. *Chinese Teaching in The World*, *29*(3), 377-392 . [王萍丽 , & 李彦霖. (2015). 语义协商的效用与话步构成: 基于母语者和非母语者自然语言互动的个案研究. *世界汉语教学*, *29*(3), 377-392.]

Wei, W., & Zhang, J. (2018). An intelligent Chinese pronunciation teaching APP and the preliminary research of teaching experiment. *Journal of Technology and Chinese Language Teaching*, *9*(2), 83-97. [魏巍, &张劲松. (2018). 一款汉语智能语音教学 APP 及教学实验初步结果. *科技与中文教学*, *9*(2), 83-97.]

Zhang, X., & Wang, Y. (2012). On research on the effect of dynamic keyword captions on SLA. *Journal of Hebei University of Technology (Social Sciences Edition)*, *4*(1), 88-92. [张雪梅, &王怿旦. (2012). 动态关键词字幕对二语习得效果影响的研究. 河北工业大学学报(社会科学版), *4*(1), 88-92.]

Zhu, K. (2018). Research and implementation of multimodal teaching dialogue systems (Unpublished master's thesis). Beijing University of Posts and Telecommunications, Beijing, China. [朱坤鸿. (2018). 多模态教学对话系统研究与实现(未出版硕士论文). 北京邮电大学, 北京, 中国.]

一款汉语智能语音教学 App 及教学实验初步结果
(An Intelligent Chinese Pronunciation Teaching App and the Preliminary Result of a Teaching Experiment)

魏巍
(Wei, Wei)
北京语言大学
(Beijing Language and Culture University)
wwei906@163.com

张劲松
(Zhang, Jinsong)
北京语言大学
(Beijing Language and Culture University)
zjsblcu@qq.com

摘要：发音教学一直是汉语作为第二语言教学中的难点问题。近年来迅速发展的人工智能发音教学技术（CAPT）能够自动诊断学生发音中的偏误，并给出正音的引导信息，从而帮助学生不受时空限制的完成自主发音学习，被认为是克服二语发音教学问题的希望技术。但是由于出现时间尚短，技术也有待完善，其教学辅助效果到底如何是一个令语学界感兴趣的问题。北京语言大学智能语音习得技术实验室一直以来从事汉语作为第二语言的智能语音教学技术的研究，开发出以尔雅中文教材为模板的汉语 CAPT-APP 软件。为了探索 CAPT 实效问题，本文以该软件为工具，在北京语言大学的两个留学生班级进行了为期 6 周的教学实验。初步的实验结果显示，36 名学生在 6 周的时间内使用 APP 练习发音的总次数为 28101 次，平均每名学生练习次数为 780.6 次，其中，课后练习次数占到了 91.9%；发音偏误细节反馈信息能够帮助学生改善 83.2%的发音问题，并且改善一个发音问题平均只需要约 1.5 次的重复练习。这样的结果说明 CAPT 教学工具对于二语语音教学的确非常有效，不仅能够有效辅助学习者完成大量发音练习，而且有助于他们改善发音质量。

Abstract: Pronunciation has always been a challenge in the learning and teaching of Chinese as a second language. Two common best practices to improve learners' pronunciation include providing learners with sufficient meaningful practice opportunities and quick feedback from their instructors. However, due to limited classroom time and one-to-many relationship between teachers and students, it is difficult to follow the above two practices in classroom instruction. Driven by artificial intelligence technology, intelligent computer assisted pronunciation teaching (CAPT) system can help learners to practice pronunciation on their own and receive instant feedback. This paper reports the findings of a six-week teaching experiment in using *Erya* Chinese pronunciation APP at Beijing Language and Culture University. Our preliminary results showed that the total number of pronunciation exercises was 28,101 times, with an average of 780.6 times per student, of which the number of after-school exercises accounted for 91.9% in quantity; under the guidance of detailed feedback

information, students can remedy 83.2% of their pronunciation errors efficiently by only 1.5 exercises on average. This result showed that the intelligent CAPT system can help the learner to complete a large number of pronunciation exercises and improve the pronunciation quality.

关键词：二语教学，汉语发音，智能教学技术，教学实验

Key words: Second language teaching, Chinese pronunciation, intelligent teaching technology, teaching experiment

1. 引言

语音教学是二语教学的首要任务，是二语教学活动开始的第一个环节，也是一个重要环节。顺利习得正确的二语语音系统有助于激发学生对于后续学习的兴趣，构造起语言能力和社交能力发展的坚实基础。而发音的偏误则很容易被母语者感知到，不仅会带来强烈的外语口音感，还可能影响到言语的可懂度（张林军，2015）。因此，在针对母语非汉语的汉语教学活动中，师生都需要重视汉语基本语音系统的教授和学习。

语音教学也是个难点问题，其表现就是在二语学习者发音中普遍存在的非母语腔调问题。在对外汉语教学中，则被周知为“洋腔洋调”现象。不论是汉语的声母、韵母，还是声调等基本语音要素，对于母语非汉语的学习者来说都存在着各种不正确的发音现象，而且这种发音偏误在不同国家的学习者中又有着不同的表现（王韫佳等人，2004；吴门吉等人，2004；梅丽，2005；王茂林，2011）。

关于二语语音习得困难的原因，学界已经有很多的研究。其中代表性的理论——语言习得关键期假说（Lenneberg，1967）认为，人类的语言习得过程存在着最佳时期，在这段时期内，大脑认知机制最有利于语言习得和发展，一旦过了这段时间，认知机制就会发生变化，后续的语言学习就会变得困难。近期研究进一步从脑神经的工作过程揭示了不同年龄学习者存在脑部语言加工位置不同的事实（Kim 等人，1997）。由于人脑认知、加工机制带来了成人语音习得的困难，所以，大量的练习对于成年学习者来说具有必要性。

对于成年人来说，虽说语音学习比较困难，但是成人经过训练仍然可以在二语语音的知觉和产出上取得进步（梅丽，2014）。二语习得理论的输入假说理论便认为（Krashen，1985），学习者在接收“可理解性的语言输入”后，便可能触发二语能力的习得。针对大脑认知机制发展的研究也表明，即使过了儿童二语学习的敏感期，大脑仍然具有一定的可塑性，可以通过外在刺激的使用引发语言能力的发展（Li等人，2014）。相关的研究和实践表明，即使过了二语学习的最佳时期，学习者也可以通过大量的练习来改善学习效果、获得二语的语音能力（张林军，2009；孙悦等人，2013；Feng 等人，2014；李艳萍等人，2017）。

针对如何更加有效地获得二语语音能力的问题，输入假说理论提出“i+1”语言输入原则（Krashen，1985）：提供给学习者的语言材料应该是循序渐进的，是学习者容易理解和掌握的，即“i”部分（“i”代表 intelligible），并在已经习得内容上增添部分新的内容，即“1”部分，代表一部分新的内容。在语音能力训练过程中，关于输入语言的意义或者是偏误原因的反馈信息，都有助于学习者掌握所要学习的内容，有利于语习者从一个阶段进入另一个更高的阶段。

这样的理论表明，语音教学活动需要满足两个条件：一是学习者需要进行充足量的有意义的发音练习；二是在练习时学习者能够及时收到有针对性的正音反馈信息。例如在汉语发音教学中，关于 VOT 特征（Voice Onset Time）的及时反馈有助于学习者改善送气、不送气辅音的混淆问题（张林军，2009），关于音高调域高低的反馈有助于解决阳平、上声的混淆问题（孙悦等人，2013；Feng 等人，2014），关于第二共振峰变化的反馈有助于解决前后鼻音的混淆问题（李艳萍等人，2017），等等。这些运用音频编辑技术的语音训练方法，通过改变关键声学线索，帮助学习者有针对性地对比易混淆的两个音，并通过足量的练习改善发音问题。

在传统的课堂教学活动中，语音教学是师生间通过语音互动活动开展的。由于师生数量比例往往是一对多的关系，语音教学便只能通过顺序、分时进行。加上课堂时间总量有限，每位学生分摊到的训练时间有限，便难以得到充分的发音机会，也难以得到来自教师的及时且充分的正音反馈意见。因此，在课堂教学实践中，语音习得的两大必备条件都难以满足：学生难以在教师指导下进行大量的发音练习，也难以获得及时的正音反馈指导。这样一个由教师资源稀缺所造成的语音教学困境，进一步阻碍了二语语音教学的效果，加剧了二语语音习得的不正确腔调问题。

近年来迅速发展的计算机辅助发音教学（Computer Assisted Pronunciation Teaching，CAPT）日渐引起人们的关注（Godwin-Jones, 2011；Yang，2013；Liu 等人，2016）。首先，CAPT 系统能提供丰富的多媒体资源，如标准语语音的音频、发音动作讲解视频等信息，有助于增加学习者的二语语言资料输入量。其次，智能化的 CAPT 系统还能够评估学习者的发音质量，有助于学习者自主进行大量的发音练习。例如基于网络多媒体技术的“长城汉语”系统[1]，基于移动互联网的微软中文学习 AI 助手 Learn Chinese[2]等。一项问卷调查的结果显示（刘洋，2018），在被调查的本科及以上学历的外国留学生和孔子学院的学生中，有超过 90%的学生都下载过 APP 汉语辅助教学软件学习中文，学生对语音问题有较高的倾向性，多数学生认为汉语辅助教学软件可以帮助他们解决语音问题。

但是，由于智能 CAPT 技术出现时间尚短，其技术特征本身尚存在很大的待提升空间（Levis，2007）。目前绝大多数面世的智能 CAPT 技术所提供的反馈信息以发音质量评价信息为主，具体形式以整句评分或者表征发音质量“好/坏”的不同颜色形式为主。这样的反馈信息虽然能够帮助学习者了解发音是否存在问题，但是由于缺少关于具体发音问题的细节反馈，学习者并不知道如何去修正发音，从而影响

[1] http://www.greatwallchinese.com.cn/portal.do?method=user
[2] https://www.msra.cn/zh-cn/news/features/microsoft-learn-chinese-20171127

了发音学习效果。另外，由于缺少较为严格的教学实验验证，人们对于智能 CAPT 技术的教学效果也心存疑问，不知道作用到底如何。

本论文旨在通过教学实验来研究上述关于智能 CAPT 技术教学效果的问题。所采用的“尔雅中文”汉语智能语音教学 APP 由北京语言大学智能语音习得技术实验室（SAIT Lab）研制，在北京语言大学初级留学生课堂开展了为期 6 周的使用实践。该款 APP 利用大规模的汉语中介语语音语料库（曹文等，2009），运用深度神经网络模型（Deep Neural Network，DNN），能够实现较为精准的具体发音偏误检测，不仅能够就整句发音进行质量评分，还能够就具体的声母、韵母及声调偏误进行检测（Gao 等人，2015；张劲松等人，2016；Wei 等人，2017；屈乐园等人，2017；Lin等人，2018），并以颜色提醒、声音对比、动画展示、文字描述等多种方式给予发音教学反馈。

以下内容安排如下：第 2 节介绍“尔雅中文”APP 所实现的主要功能，第 3 节介绍发音教学实验设计，第 4 节介绍实验结果，第 5 节给予简单结论。

2. 智能语音教学 APP——尔雅中文 APP

从满足发音教学需要出发，尔雅中文 APP 设计了学生端[3]和教师端两大模块，各模块主要功能如图 1 所示。

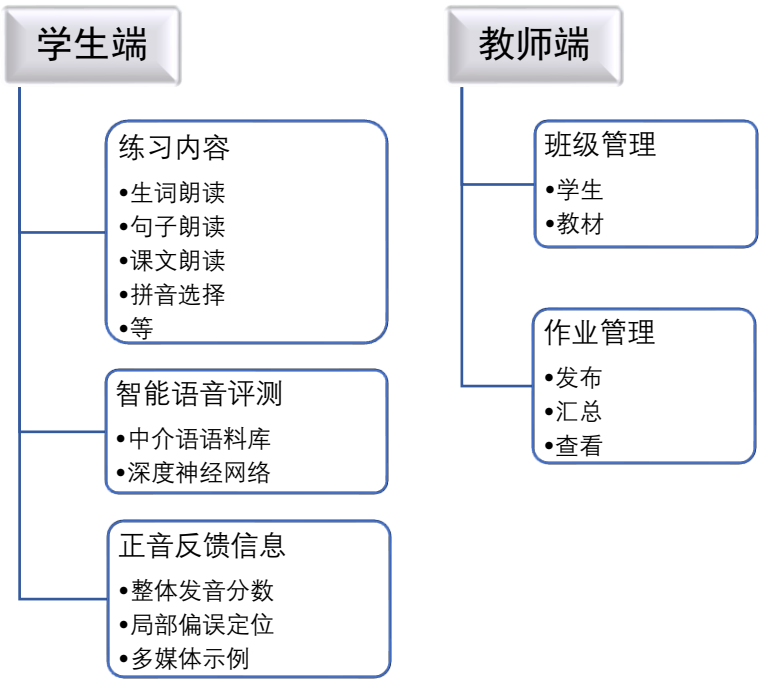
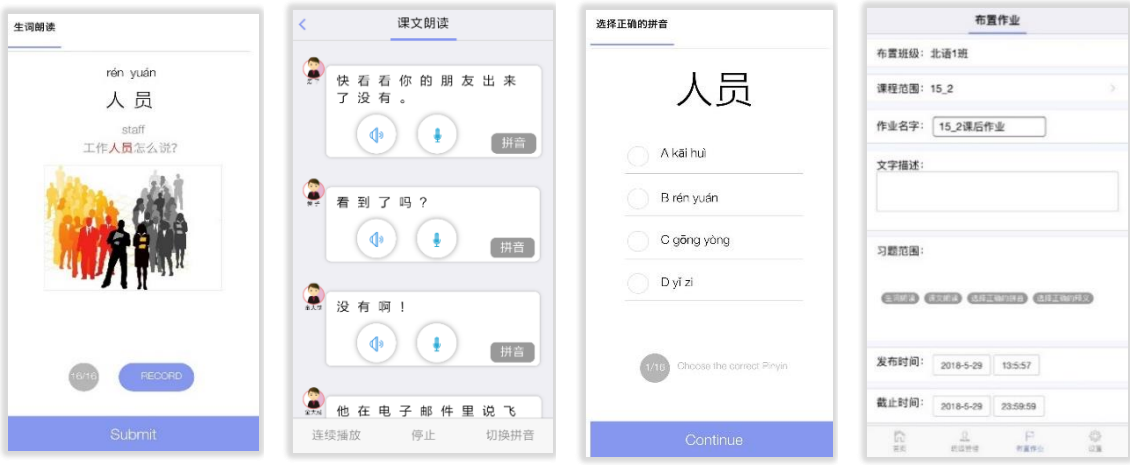


图 1. 尔雅中文 APP 功能模块结构图

[3] 尔雅中文 APP 学生端安装方式：1.安卓版可在 https://fir.im/tyzd 下载；2. 苹果版可在 App Store 搜索“SAIT 汉语”获取。

首先，学生端收录了汉语教材《尔雅中文》[4]中的练习材料，给学生提供生词朗读、句子朗读、课文朗读和拼音选择等练习，如图 2（a）（b）（c）所示。以发音练习模块为例，练习的初始页面呈现练习内容，并提供示例音频。学习者可以参考文字提示和音频示例，在手机上录制发音练习的音频，然后就会收到 APP 对发音进行评测后的反馈信息。

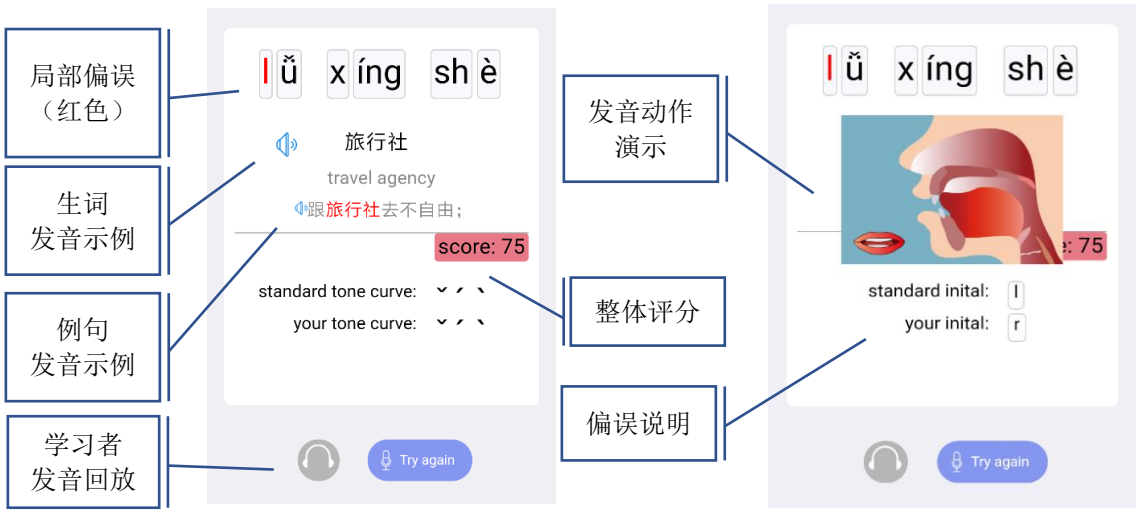


a. 生词朗读　　b. 课文朗读　　c. 拼音选择　　d. 发布作业

图 2. APP 模块示例图

其次，智能语音评测模块依托汉语中介语语料库，运用基于深度神经网络的语音评测技术，对学习者的发音进行整体打分和局部偏误的定位检测。发音评测的基本依据是学习者的发音和标准语音模型在声学上的相似度，如果相似度较低，则判断为存在偏误。APP 会识别学习者练习发音的音频内容中每个音节的声、韵、调，并对每个音节的声、韵、调进行相似度打分，依此对每个声、韵、调的正误进行判断，并得到整体发音的分数。

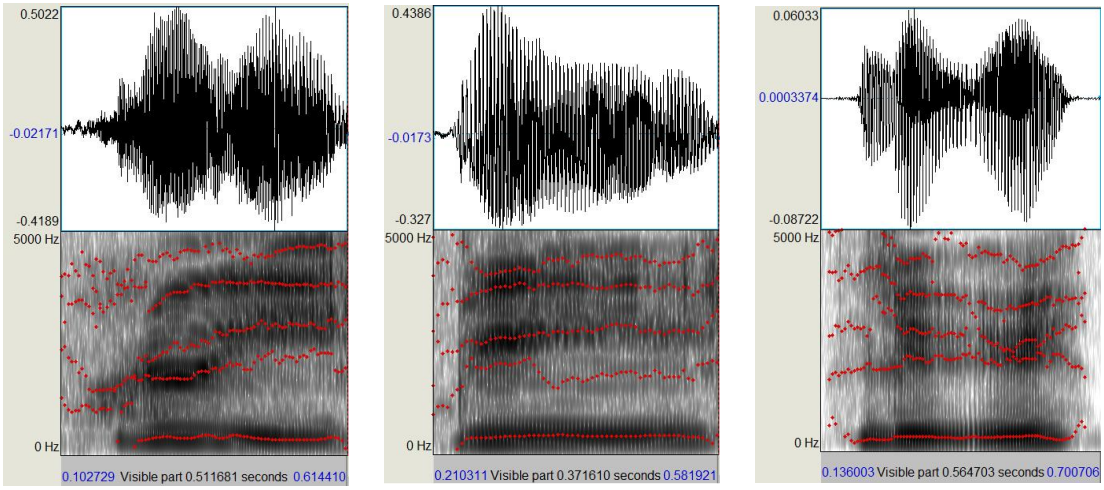
[4] 魏新红. (2013). *尔雅中文：初级汉语综合教程（上、下册）*. 北京：北京语言大学出版社.



a. 正音反馈信息示例　　b. 发音动作示例

图 3. 尔雅中文 APP 应用示例

再次，APP 采用“整体评分+局部偏误”的发音信息反馈方式，引导学习者纠正发音。以一名女性韩国学生的发音练习数据为例，如图 3（a）所示，学习者所发的“旅行社”的第一个音节的声母“l”存在发音偏误，APP 用红色区分有问题的发音，点击有偏误的声母，会收到图 3（b）所示的反馈信息，向学习者说明此处将“l”的发音混淆成了“r”。同时，APP 会向学习者反馈发音的整体分数。当学习者收到发音偏误的反馈信息后，可以有针对性地练习存在偏误的声、韵、调。如图 3（a）和图 3（b）所示，APP 提供了标准语音的音频，发音动作演示视频等多媒体教学资源，学习者可以通过反复听辨、视觉对比、发音模仿等方法来对比正误发音，通过自主式的反复摸索练习，进而掌握正确的发音方法，改善发音质量。如图 4 所示，图 4（a）是从图 3“旅行社”的语音中提取的“旅”字的语图，图 4（b）是学习者纠音后“旅”字的语图，图 4（c）是从中介语语料库中选取的女性中国发音人单音节“旅”字的语图。感知实验研究表明（陈静萍，2018），过渡段对中国被试感知声母 l、r 具有关键作用，普通话 l、r 的声学特征主要差异表现为过渡段的第三共振峰（F3）的模式，由于卷舌的作用，r 的过渡段的 F3 具有较大的变化范围。从图 4 中可以看出，学习者通过纠音练习，F3 的模式向母语者的模式靠近，这反映出了学习者发音质量的改善。



a. 偏误发音语图　　b. 纠音后发音语图　　c. 母语者语图

图 4. “旅”字声母纠音前后对比图

同时，APP 还有专门的教师端，以班级群组的形式管理课堂教学。教师在新建班级后，可以选择要使用的教材。目前，APP 收录了《尔雅中文：初级汉语综合教程》上、下两册，后期会根据教学的需求扩充教材资源。学生通过申请加入班级后，就可以使用指定教材的内容来练习发音。教师可以通过 APP 向班级中的学生发布作业，如图 2（d）所示。学生提交作业后，教师可以查看学生的作业完成进度，以及收听学生练习发音的音频。

3. 发音教学实验设计

3.1 实验目的

本研究采用行为数据分析的方法，考察二语学习者使用尔雅中文 APP 练习发音的情况，以探究智能语音教学工具在二语教学中对发音教学的作用。

3.2 实验原理及评价指标

SAIT Lab 研制的基于 DNN 的智能语音评测技术，在实验的环境下，检测准确率平均约为 90.4%（张劲松等人，2016；屈乐园等人，2017；Lin J.等人，2018）。APP 的智能语音评测模块就是在这些已有的技术积累下研制而成的，可以根据声、韵、调的相似度来检测发音偏误。发音练习是语音习得的一个过程，在这个过程中，发音质量是在一个范围内波动变化的。智能语音评测技术可以通过发音的分数将这个波动变化的过程进行量化。发音的分数反映了发音与标准语音的声学相似度，当相似度较高，即认为发音接近母语者的发音；当相似度较低，即认为发音存在明显偏误。学习者发音的评测结果从“有偏误”到“无偏误”，能够客观反映学习者的发音质量得到了较明显的改善。学习者根据 APP 的发音正误反馈进行的纠音练习，就是通过练习不断地改善偏误发音，最终接近标准语音的过程。

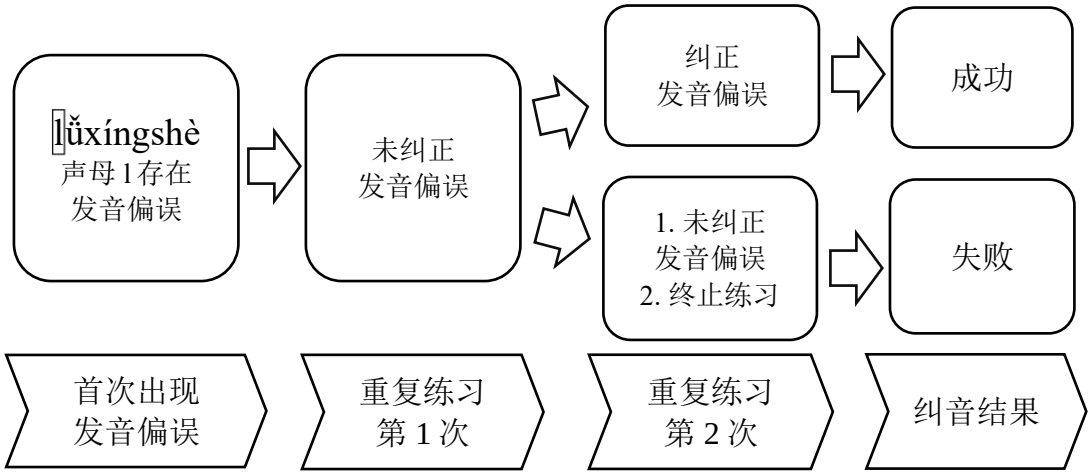


图 5. 纠音练习流程示意图

本实验考察学生使用 APP 的发音练习量和纠音效果。学生使用 APP 练习发音时，每练习一个条目，就会录制一条发音练习的音频文件，因此，音频文件的总数即是学生练习发音的总次数，音频文件的总时长可用于衡量发音的总时长。发音练习量以练习的次数和时长进行统计。纠音效果主要是为了反映“局部偏误”的反馈信息对学生纠正发音的有效性，以及学生的练习效果。

在练习发音的过程中，学生会根据 APP 的反馈信息对有偏误的发音反复地练习，以纠正发音偏误。实验中，我们先从发音练习的音频文件中提取学生重复练习的词条；再从这些重复练习的词条中，整理出每一个音节的声、韵、调的重复练习过程，并从中提取学生的纠音行为。如果通过反复练习将偏误的音改正，则视为纠音成功；如果最终没有改正，则视为纠音失败，如图 5 所示为“旅行社”这个词条中声母“l”的纠音过程。从首次出现发音偏误到结束纠音的过程称为一个纠音练习过程，记作 E，其中，纠音成功的记作 E_T，纠音失败的记作 E_F；在 E 中，重复练习的次数称为纠音练习次数，记作 N，图 5 所示纠音过程的练习次数 N=2。

目前，多数 CAPT 系统关注的是技术的创新，而不是教学的标准（Neri 等人，2002）。我们从改善发音质量的实际教学目标出发，针对纠音练习的具体过程，采用纠音成功率和纠音成功练习次数作为评价纠音效果的指标。由于纠音过程是在反馈信息的提示下进行的，因此，这两个指标能够直接反映反馈信息的作用。纠音成功率 P 从整体上反映所有 E 中 E_T所占的比例，定义如式（1）所示：

$$P = \frac{T}{T+F} \times 100\% \quad (1)$$

其中，T 为 E_T的总个数，F 为 E_F的总个数。纠音成功练习次数反映的是 E_T的具体细节，即纠正一个发音偏误所需要的重复练习次数，记作 N_T。实验将从 P 和 N_T两个指标分别评价声母、韵母和声调的纠音效果。

3.3 被试及实验语料

参加本次发音教学实验的是北京语言大学两个班级的留学生，共 36 名，分别

来自 19 个国家，各国家学生人数如表 1 所示，学生的汉语水平为初级。

表 1：各国学生人数统计表

国家 （同一格内按音序排列）	人数 （单位：人/国家）	小计 （单位：人）
韩国	7	7
巴基斯坦、俄罗斯	4	8
泰国、意大利	3	6
澳大利亚、乌兹别克斯坦	2	4
丹麦、哈萨克斯坦、墨西哥、斯里兰卡、特立尼达和多巴哥、委内瑞拉、乌克兰、西班牙、伊朗、印度尼西亚、英国	1	11
	总计	**36**

APP 自动记录了学生练习发音的相关数据，其中包括学生练习发音的音频文件，音频对应的发音练习文本（汉字及拼音），机器对音频语音识别的结果，机器对学习者发的每一个音节声、韵、调的打分及正误判断，机器对整体发音质量的打分，以及练习的时间等信息。从这些数据中可以得到学生使用 APP 练习发音的行为数据，如使用时间，纠音过程等。

通过对采集到的音频进行抽样调查，发现整体发音成绩小于 40 分的音频大多是由于录音操作不规范导致的音频未录完整。在分析发音质量时，将自动剔除得分小于 40 分的音频，这部分音频约占数据总量的 3.2%。

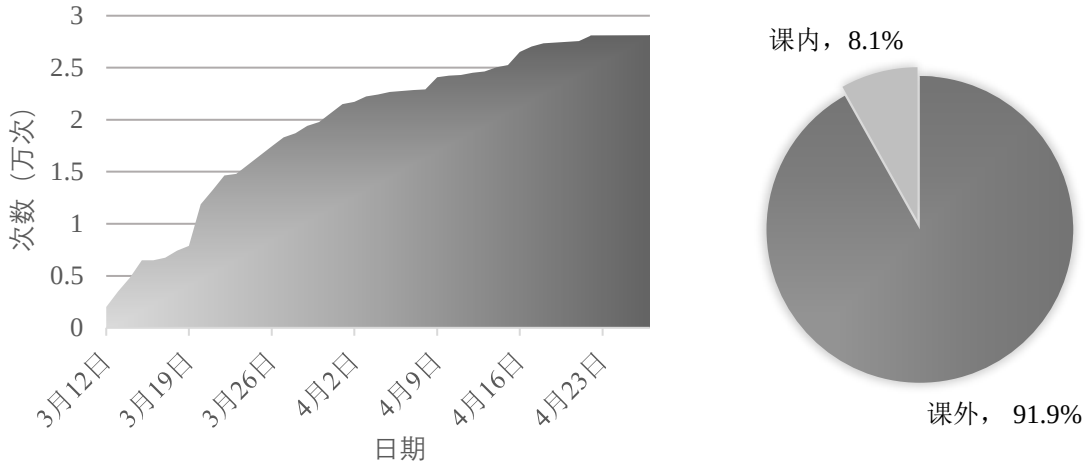
3.4 实验过程

实验在汉语课语音教学周期内同步进行，两个班的汉语课都安排在星期一至星期四，每天两课时，为期约为 6 周。在教学实验周期内，学生通过 APP 自主地完成发音练习。教师会根据授课进度和教学需求，通过教师端发布作业，并指定提交作业的截止时间，引导学生完成指定的发音练习任务。整个教学周期内，所有发音练习的行为数据均由 APP 自动地采集和记录，学生作业的成绩将计入最终的学业考核成绩中。

4. 教学实验初步结果和讨论

4.1 发音练习量统计结果

图 6 表示的是在整个教学实验周期内，所有学生使用 APP 练习发音的总次数，以及课堂内外使用 APP 所占的比例。图 6（a）反映了所有学生累计使用 APP 练习发音的总次数为 28101 次，人均练习约 780.6 次。所有发音练习的音频总时长为 21.0 小时，平均每个练习的音频约 2.7 秒。图 6（b）说明课堂外的练习次数占总练习次数的 91.9%，课堂内的练习次数占总练习次数的 8.1%。



a.累积次数统计图　　b.课内课外次数比例图

图 6. 发音练习次数统计图

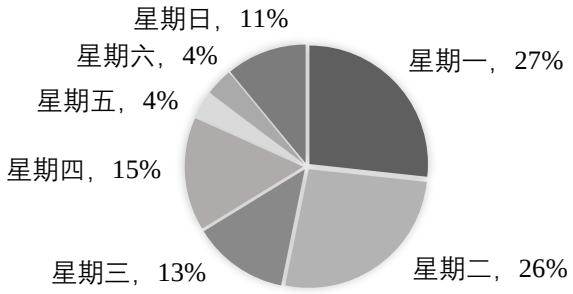


图 7. 发音练习次数按星期分布比例图

图 7 反映了发音练习次数按星期的分布情况，周一、周二、周三、周四和周日的练习量较多。教师端记录的数据显示，两个班的教师共发布了 31 次作业，发布时间为授课当天，其中有 20 次作业是要求学生当天完成，有 9 次作业学生需要在 1 至 2 天内完成，还有 2 次周四布置的作业要求在周日前（包括周日当天）提交。学生收到作业后，习题的完成量平均约为 88.7%。

4.2 纠音效果分析结果

行为数据中提取到的纠音练习过程 *E* 的总数约为 4.3 万个，其中，声母、韵母和声调的比例分别如图 8 所示。

声母、韵母和声调的纠音成功率 *P* 分别如图 9 所示。从图中可以看出，声母和韵母的 *P* 相近，均高于 84%；声调的 *P* 相对偏低，略低于 80%。

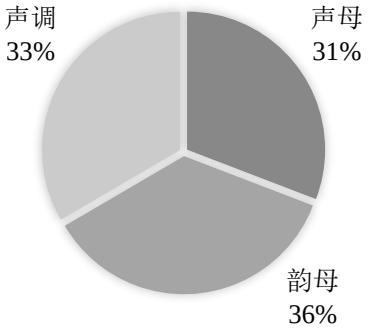


图 8. 纠音练习总个数比例图

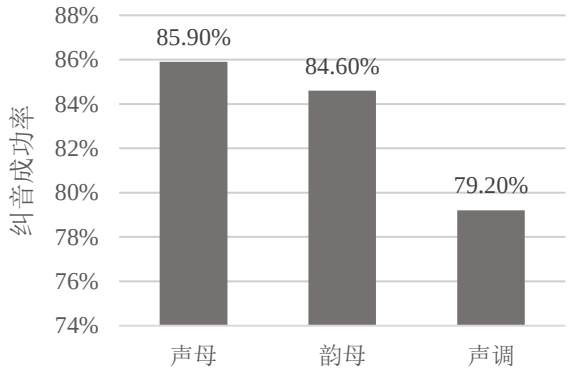


图 9. 纠音成功率统计图

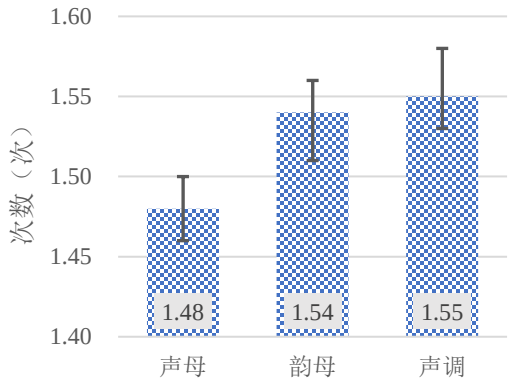


图 10. 声、韵、调纠音成功练习次数统计图

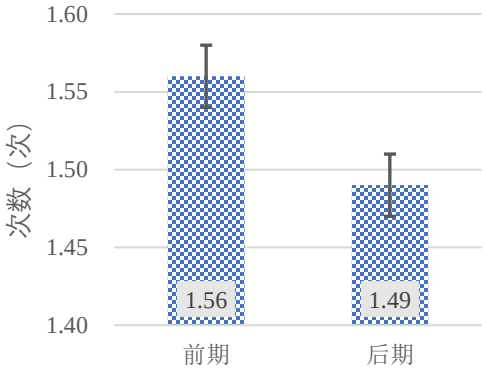


图 11. 前、后期纠音成功练习次数统计图

图 10 表示的是声母、韵母和声调的纠音成功练习次数 N_T 的平均值，分别为 1.48 次，1.54 次和 1.55 次。Levene 检验的结果显示，三组数据不满足方差齐性的条件（p=8.6×10^{-5}<0.01）。Kruskal-Wallis 秩和检验的结果显示，三组数据的均值具有显著性差异（p=2.5×10^{-4}<0.01）。两两比较的 Nemenyi 检验结果显示，声母组和韵母组具有显著性差异（p=5.3×10^{-3}<0.01），声母组和声调组具有显著性差异（p=4.9×10^{-4}<0.01），韵母组和声调组不具有显著性差异（p=0.85>0.1）。

将所有 E_T 按时间顺序平均分为前期和后期两部分，图 11 表示的是 N_T 的变化，前期为 1.56 次，后期为 1.49 次。置换检验结果显示，前期和后期的 N_T 具有显著性差异（p=7.1×10^{-7}<0.01）。

4.3 讨论

从发音练习量来看，练习次数和时长都反映出学生利用 APP 完成了比较多的发音练习，其中 90%以上的发音练习是在课后完成的。这说明 APP 能够帮助学生利用课后时间完成自主的发音练习，并且达到了一定的练习量。同时，课堂内的发音练习量反映出 APP 起到了辅助教师进行课堂教学的作用。

从总体上看，声、韵、调的平均纠音成功率约为 83.2%，说明在 APP 的辅助下，学生能够改善大部分音的发音质量。声、韵、调的平均纠音成功练习次数约为 1.5 次，即仅通过 1 至 2 次的练习，学生的发音就会有明显的改善。进一步的统计发现，纠音成功练习次数为 1 至 2 次的占到了纠音成功练习次数的 89.0%，整体上练习次数较少。这说明在大多数的情况下，发音偏误能够通过自主练习的方式比较容易地得到改善。

对于重复练习的词条，每个练习的平均时长约为 2.5 秒。纠音失败的练习次数平均为 3.2 次，即经过 3 至 4 次的重复练习，若偏误的发音没有得到明显的改善，学生就会结束该词条的练习。

从组间差异来看，纠音成功率从高到低依次为：声母>韵母>声调，纠音成功练习次数从低到高依次为：声母<韵母<声调。这样的结果说明声母的纠音效果较好，声调的纠音难度较大。

在所有的纠音练习中，后期的纠音成功练习次数显著小于前期的纠音成功练习次数。这种行为变化说明，学生通过纠音训练后，纠音的效率有明显提高。纠音的结果是心理认知和生理产出共同作用的结果，是语音能力的一种体现形式。因此，纠音效率的提高体现了学生发音能力的提高。

目前初步的实验结果仅从总体上分析了发音练习行为和纠音结果，在重复练习的纠音过程中，学生利用多媒体信息辅助练习发音的行为数据还有待进一步地分析，以详细考察影响学生正音效果的因素。学生语音习得的长效性还需要通过专门的测试来检验。

5. 结论

本文介绍了智能发音教学 APP——尔雅中文 APP 的功能模块，在此基础上，研究了智能语音教学工具在汉语作为第二语言教学中对提高学生发音质量的作用。初步的教学实验研究结果表明：在智能语音教学工具的辅助下，学生利用课后时间自主地完成了发音练习的任务，并且达到了一定的练习量；基于具体声、韵、调正误的发音反馈信息，有效地引导学生纠正了发音偏误，改善了发音质量。目前，正音反馈信息主要是从发音混淆的角度提示学习者存在发音偏误，今后的工作将结合发音偏误趋势，提供更加具体的发音反馈信息，以期获得更好的正音效果。

致谢: 本研究受北京语言大学语言资源高精尖创新中心经费（KYR17005）及校级重大基础研究专项（16ZDJ03）（中央高校基本科研业务专项资金）以及北京语言大学研究生创新基金项目（18YCX006）资助。作者感谢北京语言大学智能语音习得技术实验室解焱陆、李伟、郝丽霞、冯晓莉、及转转、杜丹、蔡睿、郑宝英等多位同仁的卓越贡献，也向大力支持教学实验的北京语言大学汉语学院曹文、张浩、刘敬华、唐翠菊老师表示感谢。

参考文献

Cao, W., & Zhang, J. (2009). The construction of a CAPL Chinese interlanguage corpus and its annotation. *Applied Linguistics*, 4, 122-131. [曹文, &张劲松. (2009). 面向计算机辅助正音的汉语中介语语音语料库的创制与标注. *语言文字应用*, 4, 122-131.]

Chen, J. (2018). *A comparative study of Chinese and Japanese phonetics of mandarin initial /r/ and /l/* (Unpublished master's thesis). Beijing Language and Culture University, Beijing, China. [陈静萍. (2018). *中日被试普通话声母/r/、/l/的语音对比研究*(硕士学位论文). 北京语言大学, 北京, 中国.]

Feng, X., Sun, Y., Zhang, J., & Xie, Y. (2014). A Study on the long-term retention effects of Japanese C2L learners to distinguish Mandarin Chinese Tone 2 and Tone 3 after perceptual training. *International Symposium on Chinese Spoken Language Processing* (pp.599-603). Singapore City, Singapore.

Gao, Y., Xie, Y., Cao, W., & Zhang, J. (2015). A Study on Robust Detection of Pronunciation Erroneous Tendency Based on Deep Neural Network. *Interspeech 2015: Conference of the International Speech Communication Association*. Dresden, Germany.

Godwin-Jones, R. (2011). Emerging technologies: mobile APPs for language learning. *Language Learning & Technology*, *15*(2), 2-11.

Kim, K. H. S., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, *388*(6638), 171-174.

Lenneberg, E., H. (1967). Biological foundations of language. *Hospital Practice*, *2*(12), 59-67.

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics*, 27, 184-202.

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex*, 58, 301-324.

Li, Y., Cao, C., Xie, Y., & Zhang, J. (2017). A study on training Japanese students to discriminate nasal finals in standard chinse. *The 14th National Conference on Man-Machine Speech Communication*, Lianyungang, China. [李艳萍, 曹冲, 解焱陆, 张劲松. (2017). 日本学生汉语普通话鼻音韵母知觉训练初探. *第十四届全国人机语音通讯学术会议*, 连云港, 中国.]

Lin, J., Li, W., Gao, Y., Xie, Y., Chen, N. F., & Siniscalchi, M., … Lee, C. H. (2018). Improving mandarin tone recognition based on DNN by combining acoustic and articulatory features using extended recognition networks. *Journal of Signal Processing Systems*, *90*(7), 1077-1087.

Liu, S., & Hung, P. (2016). Teaching pronunciation with computer assisted pronunciation instruction in a technological university. *Universal Journal of Educational Research*, *4*(9), 1939-1943.

Liu, Y. (2018). *A study of application of APP Chinese teaching software in extracurricular autonomous learning* (Unpublished master's thesis). Liaoning Normal University, Dalian, China. [刘洋. (2018). *APP 汉语辅助教学软件在课外自主学习中的应用研究*(硕士学位论文). 辽宁师范大学, 大连, 中国.]

Mei, L. (2005). Variability in pronouncing Chinese blade-palatal initials by Japanese SLA

learners. *Chinese Teaching in the World*, 1, 97-105. [梅丽. (2005). 日本学习者习得普通话卷舌声母的语音变异研究. *世界汉语教学*, 1, 97-105.]

Mei, L. (2014). Cross-language speech perception in adults. *Journal of South China Normal University (Social Science Edition)*, 4, 164-169. [梅丽. (2014). 成人的跨语言言语知觉. *华南师范大学学报(社会科学版)*, 4, 164-169.]

Neri, A., Cucchiarini, C., & Strik, H. (2002). Feedback in computer assisted pronunciation training: Technology push or demand pull? *International Conference on Spoken Language Processing 2002* (pp. 1209-1212). Denver, Colorado, US.

Qu, L., Xie, Y., & Zhang, J. (2017). A study of articulatory features based detection of pronunciation erroneous tendency. *Acta Scientiarum Naturalium Universitatis Pekinensis*, *53*(2), 239-246. [屈乐园, 解焱陆, &张劲松. (2017). 基于发音特征的发音偏误趋势检测研究. *北京大学学报(自然科学版)*, *53*(2), 239-246.]

Krashen, S (1985). *The input hypothesis: Issues and implications*. London: Longman.

Sun, Y., Zhang, J., Xie, Y., & Cao, W. (2013). Perceptual training of Japanese students to discriminate mandarin tones 2 and 3. *Journal of Tsinghua University*, *53*(6), 921-924. [孙悦, 张劲松, 解焱陆, 曹文. (2013). 日本学生汉语阳平和上声的知觉训练效果的初步分析. *清华大学学报(自然科学版)*, *53*(6), 921-924.]

Wang, M. (2011). An analysis on the production of Chinese monophthongs by the Indonesian and Korean students. *TCSOL Studies*, 4, 16-25. [王茂林. (2011). 印尼、韩国留学生汉语单元音韵母发音分析. *华文教学与研究*, 4, 16-25.]

Wang, Y., & Shangguan, X. (2004). How Japanese learners of Chinese process the aspirated and unaspirated consonants in standard Chinese. *Chinese Teaching in the World*, 3, 54-66. [王韫佳, & 上官雪娜. (2004). 日本学习者对汉语普通话不送气/送气辅音的加工. *世界汉语教学*, 3, 54-66.]

Wei, X., Chen, J., Wang, W., Xie, Y., & Zhang, J. (2017). A study of automatic annotation of PETs with articulatory features. *Asia-Pacific Signal and Information Processing Association Summit and Conference* (pp.1608-1612). Kuala Lumpur, Malaysia.

Wu, M., & Hu, M. (2004). An analysis of Vietnamese students' errors in Chinese tones. *Chinese Teaching in the World*, 2, 81-87. [吴门吉, 胡明光. (2004). 越南学习汉语声调偏误溯因. *世界汉语教学*, 2, 81-87.]

Yang, J. (2013). Mobile assisted language learning: review of the recent applications of emerging mobile technologies. *English Language Teaching*, 6(7), 19-24.

Zhang, J., Gao, Y., & Xie, Y. (2016). Mispronunciation tendency detection using deep neural networks. *Journal of Tsinghua University*, 56(11), 1220-1225. [张劲松, 高迎明, & 解焱陆. (2016). 基于 DNN 的发音偏误趋势检测. *清华大学学报(自然科学版)*, *56*(11), 1220-1225.]

Zhang, L. (2009). Perceptual training and the acquisition of Chinese aspirated/unaspirated consonants by Japanese students. *Language Teaching & Linguistic Studies*, 4, 85-90. [张林军. (2009). 知觉训练和日本留学生汉语辅音送气/不送气特征的习得. *语言教学与研究*, 4, 85-90.]

Zhang, L. (2015). The contributions of segmental and prosodic information to the perception of foreign accent and speech intelligibility. *Chinese Teaching in the*

World, 2, 242-249. [张林军. (2015). 音段和韵律信息对汉语外语口音感知及言语可懂度的影响. *世界汉语教学,* 2, 242-249.

A Survey of Mobile Apps for Learning Chinese Vocabulary (中文词汇的移动应用程序之调查)

Lin, Chin-Hsi
(林金锡)
The University of Hong Kong
(香港大学)
chinhsi@hku.hk

Zhou, Keyi
(周可依)
The University of Hong Kong
(香港大学)
u3551792@connect.hku.hk

Yang, Shiyun
(杨诗蕴)
The University of Hong Kong
(香港大学)
shiyun93@connect.hku.hk

Abstract: Mobile technology is continually developing. Research suggests that mobile apps are beneficial for language learning. Yet, because apps developed by researchers are not accessible to the public, the benefits of mobile apps on the market remain unclear. This study reviewed the 22 most popular Chinese-vocabulary learning apps on the market for native speakers. The results of content analysis show that only two types of these apps have been designed: dictionaries and games, neither of which provide their users with opportunities to practice new words. Although these apps facilitated the use of one or more learning strategies, they generally lack clear learning goals, explanations, opportunities to practice, and feedback.

摘要：移动技术在不断地进步发展，有研究发现，移动应用程序能为语言学习带来益处。然而因为这些由研究人员开发的应用程序无法公开，所以移动应用程序在市场上的益处仍然不清楚。本研究回顾了 22 个目前市场上最流行的汉语词汇学习移动应用程序，这些程序为汉语母语者所使用本研究的分析结果表明，这些应用程序的设计只有两种类型：词典和游戏，而且它们都没有给用户提供练习新词汇的机会。虽然这些应用程序促进了一个或多个学习策略的使用，但它通常缺乏明确的学习目标，解释说明，练习机会和反馈功能。

Keywords: Mobile learning, learning strategies, vocabulary learning, educational apps

关键词：移动学习，学习策略，词汇学习，教育应用

1. Introduction

Vocabulary is a central component of language learning (Grabe, 2004). A good vocabulary can make a person's communication smooth and precise, and can enhance all four language skills, i.e., listening, speaking, reading and writing (Schmitt, 2000). Vocabulary knowledge, which is comprised of breadth (i.e., the number of words one

knows) and depth (i.e., the extent to which one comprehends each word), correlates positively with both reading comprehension (Qian, 1999) and cognitive functions (Marchman & Fernald, 2008).

Learning Chinese vocabulary is an essential component of the curricula of Chinese primary schools, which include grades one through six. In Hong Kong, students need to learn 4,914 words by grade three, and another 4,792 words by grade six, or a total of 9,706 words by the end of the primary stage. Most of these words are commonly used in Mainland China and Taiwan (Bureau, 2008). However, most studies of vocabulary learning have focused on students of English as a second language (e.g., Gu, 2002; Lip, 2009; Yip & Kwan, 2006); little scholarly attention, however, has hitherto been paid to Chinese vocabulary learning by native Chinese speakers. The new Chinese-language curriculum in Hong Kong requires students to improve all four language skills at the primary school level, with the first three grades focusing on the development of listening and speaking, and primary four through six focusing on reading and writing. Chinese vocabulary involves phonological, orthographic, morphological and grammatical features (Yang, 2017), and is the basis for the cultivation of all four language skills. Therefore, the exploration of effective ways to teach it is especially important.

With the advancement of mobile technology, a growing body of research has examined the use of mobile apps for vocabulary learning, and has consistently reported that the impact of such use was positive (e.g., Chiu, 2013; Redd, 2011; 2015) with regard to both breadth and depth (Wu, 2015). For example, Wu found that students using the smartphone app Word Learning-CET6 significantly outperformed those who did not use apps to acquire new English words. Other studies indicated that mobile gaming apps could expand their users' vocabulary (Redd, 2011), and that using mobile apps for vocabulary learning increased students' motivation (Wang, 2017; Wang, Teng, & Chen, 2015).

However, while certain mobile apps have demonstrated some potential to facilitate individuals' development of vocabulary knowledge, most studies have focused on just one app developed especially for learning English vocabulary. Although the literature suggests that the use of vocabulary-learning strategies is positively related to the breadth and depth of vocabulary knowledge (X. Zhang & Lu, 2015; Y. Zhang, Lin, Zhang, & Choi, 2017), questions about which vocabulary-learning strategies are promoted in and by mobile apps remain largely unanswered. Without a clearer understanding of how these vocabulary-learning apps were designed, then, it would be difficult for Chinese-language teachers to recommend appropriate ones to their students. Accordingly, the present study examines the most popular Chinese-vocabulary learning apps to determine how they facilitate the acquisition of vocabulary knowledge, specifically, whether and how the design of these apps promotes the development of vocabulary-learning strategies.

2. Literature review

2.1 Vocabulary-learning strategies

Vocabulary-learning strategies are specific actions or techniques that learners use to develop their vocabulary knowledge. Several studies have examined the use of domain-general learning strategies in vocabulary acquisition, including cognitive and metacognitive strategies (Gu & Johnson, 1996; Y. Zhang et al., 2017). The former refers to the use of skills or techniques to process vocabulary information (e.g., association), and the latter, to individuals' ability to regulate their cognition (e.g., by setting vocabulary-learning goals). Similarly, though they did not focus on vocabulary learning, Oxford and Ehrman (1995) identified six language-learning strategies: cognitive, metacognitive, mnemonic, compensatory, affective, and social. Mnemonic strategies help learners to link one item or concept with another, or to learn and retrieve information in orderly strings (e.g., by creating acronyms). Compensatory strategies help them overcome their lack of knowledge (e.g., by guessing based on context). Affective strategies rely on learners' emotions (e.g., rewarding themselves for good performance). Finally, social strategies help them work with others, and to understand the target culture as well as the language (e.g., by verifying their knowledge through asking questions).

Using Schmitt's (1997) Vocabulary Learning Strategies Survey, Zhang and Lu (2015) conducted an exploratory factor analysis that yielded five vocabulary-learning strategies: learning the forms of words, "PIC/IMG," association, repetition, and word listing. The first strategy, learning the forms of words, refers to learning phonological and/or orthographic forms. Studying how to write a Chinese word is a typical example of this strategy. The second strategy, PIC/IMG, refers to linking words to images or situations. A typical example of using this strategy is to make an image of the words' meaning. The third learning strategy, association, refers to linking words to morphologically or semantically related ones. An example of using association would be studying words with the same radicals. The fourth learning strategy, repetition, is to repeat words or to continuously study the word frequently over a period of time. The fifth example, word listing, refers to the use of word lists to study and recite new words.

Zhang and Lu found that each of these strategies affected vocabulary knowledge in a different way, with forms and association being positively correlated with both vocabulary breadth and depth, and word lists negatively correlated to both of the same variables, while PIC/IMG was negatively correlated to vocabulary depth only.

2.2 Educational mobile apps

A recent meta-analysis of mobile-assisted language learning by Sung, Chang, and Yang (2015) found a moderate effect size of 0.55. However, most of the studies it covered focused on apps developed by researchers, many of which were unavailable to the public. Despite an increasing number of educational apps available from the iTunes App Store, Google Play, and similar platforms in Asia and elsewhere, little is known about how they have been designed for language learning, as systematic scholarly reviews of such apps are still in their infancy. The handful of reviews that have been

conducted suggest that the majority focus on drill-and-practice activities. Highfield and Goodwin (2013), for example, reviewed the iTunes App Store's math apps and lamented the limitations of their pedagogical designs. More recently, Papadakis, Kalogiannakis, and Zaranis (2018) examined Google Play's selection of educational apps for Greek preschoolers, which were again dominated by the drill-and-practice approach, and came to similar conclusions. Worse, Blair (2013) as well as Vaala and Levine (2015) highlighted how little feedback was provided in math and literacy apps respectively. Callaghan and Reich (2018) pointed out that such feedback (as was provided by mobile apps) encouraged learning through a trial-and-error approach, rather than helping users to understand the material better.

The present researchers were unable to identify any review studies that particularly focused on vocabulary apps available to the general public. This study therefore helps to fill in that gap by reviewing popular vocabulary apps for Chinese children.

3. Methods

Qualitative content analysis was used to examine the learning strategies supported by popular mobile apps for learning Chinese vocabulary. Qualitative content analysis is a qualitative approach to reveal information that is not available by studying the surface of the content. Rather, this approach involves systematic procedures to make valid inferences from text (Anderson, Rourke, Garrison, & Archer, 2001), including retrieving, evaluating, and coding a specific type or types of content. In this case, the content, or unit of analysis, consisted of mobile apps available on the iTunes App Store (China), the iTunes App Store (Taiwan), Google Play, and Xiaomi App Store as of October 2018. Content retrieval and evaluation will be described in section 3.1, and coding will be described in section 3.2.

3.1 Mobile-app selection

The research team reviewed the top ten free vocabulary-learning mobile apps for native Chinese speakers from each of the four app stores mentioned above. Two search terms, 词汇 [words] and 词语 [vocabulary], were used in both their traditional and simplified forms on all four platforms.

3.2 Coding framework

To capture the design aspects of mobile apps, both top-down and bottom-up coding approaches were employed. The top-down coding adopted Zhang and Lu's five vocabulary-learning strategy categories discussed above (learning the forms of words, PIC/IMG, association, repetition, and word listing). While this approach was capable of capturing such strategies in a general sense, it could not capture any special affordances of mobile devices. Thus, bottom-up coding was used to explore whether mobile apps used strategies other than those identified by Zhang and Lu.

3.3 Data analysis

Each author coded one of the four stores using the search terms mentioned above. After coding was completed, each app was demonstrated in a group meeting, at which time it was decided whether the codes accurately reflected its design features and underlying approach.

4. Results

4.1 Types of mobile apps

The research team identified 22 representative apps, comprising 10 dictionaries and 12 educational games (Table 1). All were listed on the first pages of their respective app stores' search results, and all had a large number of downloads and comments. It is worth noting that among many of the dictionary apps, "vocabulary" referred to Chinese idioms or fixed expressions, commonly in four-character forms (e.g., 亡羊补牢, [it's not too late to correct a problem]).

4.2 Vocabulary-learning strategies

In addition to the five strategies identified by Zhang and Lu (2015), a metacognitive strategy emerged from the coding process, which typically is promoted through educational games. Not all apps promoted the use of only one learning strategy. The strategy that was most prominent was forms (19 apps), followed by learning through games (12 apps), word listing (11 apps), learning through association (10 apps), PIC/IMG (seven apps), and repetition (three apps).

Table 1: A list of Chinese vocabulary-learning apps

Type	Apps	Forms	PIC/IMG	Association	Repetition	Word listing	Metacognitive strategy
Dictionary	成语辞典（Hisand Liu）	v		v			
	成語大詞典（商務印書館）	v		v	v	v	
	國語辭典-教育部重編國語辭典修訂本	v					
	漢語字詞典—最新版成語詞典&現代漢語詞典	v		v		v	
	現代漢語詞典（商務印書館）	v		v	v	v	
	中文字典—漢字拼音部首筆畫釋義查詢翻譯	v		v		v	
	漢語字典和漢語成語詞典—主持人配音	v				v	

	百度漢語—讓學習漢語更簡單	v				v	
	給兒童的注音及拼音詞彙工具書 (幼稚園)		v				
	成語詞典故事漢語造句大全					v	
Game	快樂中文 (Fun Chinese) 兒童漢語	v	v				v
	錯別字	v		v			v
	成語接龍-填填看	v					v
	中文連詞—漢語詞彙找茬			v	v	v	v
	我愛普通話—益智好玩的休閒娛樂文字遊戲單機	v	v				v
	成語玩命猜：1300 道看圖猜成語	v	v	v			v
	成語連連消：中文詞語消除謎題遊戲	v		v		v	v
	成語消消消—全民瘋狂玩消成語小遊戲	v					v
	看圖猜成語	v	v				v
	开心划词	v		v		v	v
	成語大挑戰（繁體版）	v	v				v
	搜詞之王（mobrix）	v	v			v	v
Total		19	7	10	3	11	12

The following sections provide examples of how each strategy was promoted by different mobile apps.

4.2.1 Learning Phonological or orthographic forms

Learning the forms of words consists of two parts: the orthographic (written form) and the phonological (spoken form). Most vocabulary-learning apps display the correct form, but in a variety of ways. For example, in the case of orthographic forms, the app 现代汉语词典 [Modern Chinese Word Dictionary] developed by The Commercial Press [商务印书馆] presents animations of stroke orders when the user clicks on a word (Fig. 1).

Figure 1. Stroke-writing animation in the app 现代汉语词典 [Modern Chinese Word Dictionary]

Some apps, such as 成語辭典 [Idiom Dictionary], developed by Hisand Liu, display words in a large size in an eight-segment grid to aid the user's clear understanding of their orthographic forms (Fig. 2).

Figure 2. Words displayed at a large size in grids in the app 成語辭典 [Idiom Dictionary]

To aid the learning of phonological forms, the apps provide several types of romanization. In the app 现代汉语词典 [Moden Chinese Word Dictionary], for example, when learners click the speaker button below a character, they will hear the sound of the word and also see it in the Roman alphabet. The vocabulary-learning apps targeted at Mainland China mostly provided Pinyin only.

In addition to Pinyin, vocabulary-learning apps targeting the Taiwanese market (e.g., 成語辭典 [Idiom Dictionary] developed by Hisand Liu) provide Zhuyin Fuhao (Bopomofo).

4.2.2 PIC/IMG

Most of the apps using PIC/IMG organized words semantically, and provided specific situations in which the learners would use them. For example, when learning words related to colors in the app 學中文 [Fun Chinese], users see a black-and-white room, then hear a word for a color, and have to click on the corresponding color of an object. When an answer is correct, the object will change to this color (Fig. 3).

Figure 3. Learning words through PIC/IMG in the app 學中文 [Fun Chinese]

While 學中文 [Fun Chinese] aims to help learners develop sound-to-meaning connections through PIC/IMG, mobile apps can also support meaning-to-form connections using the same learning strategy. The app 成語大挑戰 [The Idiom Challenge], for example, displays an image related to the meaning of a word; users need to guess that meaning based on the image, and then form the word by selecting four Chinese characters from among an unorganized group of characters given below (Fig. 4).

Figure 4. An image in the app 成語大挑戰 [The Idiom Challenge] that hints at the meaning of the Chinese idiom 食言而肥 [Break one's word]

Similarly, the app 給兒童的注音及拼音詞彙工具書 (幼稚園) [The Tool Book of Words with Pinyin and Bopomofo for Kindergarteners] presents images related to the meaning of a word above the position within that word that produce the same Chinese consonant or vowel as the initial sound of the depicted item (Fig. 5).

Figure 5. An image in the app 給兒童的注音及拼音詞彙工具書 (幼稚園) [The Tool Book of Words with Pinyin and Bopomofo for Kindergarteners] showing pictures of words that start with the same consonant

4.2.3 Association

The results of content analysis showed that three types of associations were used in Chinese vocabulary-learning apps: synonyms and antonyms, common morphemes, and topics. Several of these apps were intended primarily to help children learn the meanings of synonyms and antonyms. For example, 小学生同义词近义词反义词词典—辞海版新课标 [Thesaurus for Elementary-school Students Based on the New Curriculum] displays synonyms and antonyms for words that the user searches for (Fig. 6).

Figure 6. Synonyms and antonyms in the app 小学生同义词近义词反义词词典—辞海版新课标 [Thesaurus for Elementary-school Students Based on the New Curriculum]

Similarly, the app 詞語消消連—語文近義詞反義詞比拼 [Word Crush: The Thesaurus Challenge] helps users connect words with their synonyms and antonyms

using games. Correct pairings of synonyms and antonyms increase the user's score (Fig. 7).

Figure 7. Synonym-antonym pairing game in the app 詞語消消連—語文近義詞反義詞比拼 [Word Crush: The Thesaurus Challenge]

The second subtype of association strategy that the sampled Chinese vocabulary-learning apps used was based on morphemes, with the aim of creating lexical links between words. This was done in two ways: idiom solitaire and displaying words with similar morphemes. Idiom solitaire was used in the app 成語接龍闖關：玩成語謎題，看成語故事 [Chinese Idiom Solitaire], which asks users to select characters from a brief list (e.g., 莫, 斗, 禍) to complete a Chinese idiom (Fig. 8). At the time of writing, there are 500 levels in this app, with the difficulty gradually increasing as one moves up levels.

Figure 8. Idiom solitaire in the app 成語接龍闖關：玩成語謎題，看成語故事 [Chinese Idiom Solitaire]

The promotion of morpheme use for vocabulary learning was commonly used in the sampled dictionary apps. For example, 现代汉语词典 [Modern Chinese Word Dictionary] automatically presents several words (e.g., 醒脾, 脾氣) containing the same target morpheme (in this case, 脾; Fig. 9).

Figure 9. Words containing the same morpheme in the app 现代汉语词典 [Modern Chinese Word Dictionary]

Lastly, some of the vocabulary-learning apps organized words by topic to facilitate users' learning of similar words. Taking the app 成語詞典故事漢語造句大全 [Stories of Chinese Idioms: Sentence Bank] as an example, idioms were categorized by themes, such as seasons, mood, climate, and people (Fig. 10). Users can learn various idioms related to a theme by clicking the corresponding link.

Figure 10. Thematic classification of idioms in the app 成語詞典故事漢語造句大全 [Stories of Chinese Idioms: Sentence Bank]

4.2.4 Repetition

Repetition-based learning strategies are typically used in second-language learning. In traditional settings, learners generally make lists of target words they want to review, and do so often. Among the sampled vocabulary-learning apps, several featured word-list making (as discussed further in the next section), but only a few facilitated the use of repetition for vocabulary learning. For example, the app Quizlet [詞卡] allowed users to create a list for studying unfamiliar words; the app 中文 [Chinese] prompted users daily with new words, and; 汉语词汇学习机—汉语词汇轻松学 [10,000 Chinese Words and Phrases] tested users on a given word until they remembered it. However, these were rare exceptions to the rule that Chinese vocabulary-learning apps for native speakers do not promote repetition strategies.

Among the majority of the sampled apps that did not directly promote repetition, 现代汉语词典 [Modern Chinese Word Dictionary] allowed users to create a review plan for the words they added to their word lists based on their subjective levels of familiarity with each one (Figs. 11-12).

Figure 11. Desired words ranked by their familiarity in the app 现代汉语词典 [Modern Chinese Word Dictionary]

Figure 12. A word-review plan in the app 现代汉语词典 [Modern Chinese Word Dictionary]

4.2.5 Word listing

In traditional settings, learners can use the vocabulary sections of their textbooks or make their own word lists to review words they have learned. Chinese vocabulary-learning apps, despite being not directly comparable to textbooks, also facilitate word listing in various ways.

First, as mentioned in the section on association above, several apps' organize content semantically (e.g., 成語詞典故事漢語造句大全 [Stories of Chinese Idioms: Sentence Bank], Fig. 12). This allows learners to navigate words they want to learn, review, or practice. Another example is the app 开心划词 [Word Search], a game that asks learners to find a list of words based on scattered characters (see the words next to the lightbulb icon at the bottom of Fig. 13), using animals as prompts. The same app also allows users to create their own word lists with which to challenge their friends.

Figure 13. Word-search game in the app 开心划词 [Word Search]

In the dictionary app 中文字典-汉字拼音部首笔画释义查询翻译 [Chinese Dictionary: Search by Characters, Pinyin, Radicals with Meanings and Translations], after finding the meanings of a word, users can click a star icon in the top right corner to add it to their personal word list (Fig. 14) and later review that list in their dashboard (Fig. 15).

Figure 14. Adding to a word list in the app 中文字典-汉字拼音部首笔画释义查询翻译 [Chinese Dictionary: Search by Characters, Pinyin, Radicals with Meanings and Translations]

Figure 15. Word list in the app 中文字典-汉字拼音部首笔画释义查询翻译 [Chinese Dictionary: Search by Characters, Pinyin, Radicals with Meanings and Translations]

4.2.6 Metacognitive strategies

Metacognitive strategies emerged during our coding process, which were promoted through educational games in the apps. As Gee (2008) noted, games create a situated learning environment in which players can try different ways of solving a particular problem. During that process, players use multiple metacognitive strategies, such as setting goals and monitoring progress.

现代汉语词典 [Modern Chinese Word Dictionary] is a good example of an app that promotes both setting goals and monitoring progress. Setting goals was facilitated through selecting words that users wanted to study. Monitoring progress was promoted by asking users to rate their familiarity with words (see Fig. 11).

Metacognitive strategies were not promoted in isolation. Rather, in these games, they were promoted in conjunction with cognitive strategies identified by Zhang and Lu (2015). Learning a new word involves acquiring knowledge of its forms, sounds, and

meanings. The games that the sampled mobile apps feature typically provide one or more such components of a word and ask users to provide at least one additional component. Several examples have been mentioned earlier in this paper, including 學中文 [Fun Chinese], which asks users to click an object with the color corresponding to a color-word's sound (Fig. 3). Similarly, 成語大挑戰 [The Idiom Challenge] illustrates the meanings of Chinese idioms and asks users to pick up characters to form its words in an orderly manner (Fig. 4). Chinese Idiom Solitaire provides some orthographic parts of words and asks users to complete them (Fig. 8). One additional example, the app 开心划词 [Word Search], uses only one type of word information, i.e., it requires users to find words in a grid full of otherwise random characters (Fig. 13). Some of the vocabulary-learning apps also used lexical associations in their game designs: e.g., the pairing of synonyms and antonyms in the app 詞語消消連—語文近義詞反義詞比拼 [Word Crush: The Thesaurus Challenge] (Fig. 7).

5. Discussion

This paper's review of the 22 most popular Chinese-vocabulary learning apps on the market as of October 2018 yielded several interesting findings. These findings can be subdivided into two main categories: the apps' design aspects and their facilitation of vocabulary-learning strategy use.

5.1 Design

Only two types of apps were reviewed: dictionaries and games. While this finding is not especially surprising, it suggests that the affordances of mobile devices for vocabulary learning are not being fully exploited. While dictionary apps allow users to find information about words quickly, only a few allowed users to add their search results to their word list for later review. In other words, despite looking up the meaning(s) of a word as the start of learning it, not the end goal, none of the sampled apps provided their users with opportunities to proceed further.

The Chinese-vocabulary games examined exhibit a similar limitation, and in effect did not seem to have been designed for learning. While they created engaging environments in which users could use their word knowledge to complete various tasks, these games' designs all assumed that their users already knew the words embedded in them. For example, if users of the app 成語大挑戰 [The Idiom Challenge] (Fig. 6) did not know a given Chinese idiom, the game did not provide sufficient information for them to learn it. Conversely, players of idiom solitaire will be able to win the game if they have a certain baseline knowledge of the Chinese idioms in the app 成語接龍闖關：玩成語謎題，看成語故事 [Chinese Idiom Solitaire] (Fig. 10), even if this knowledge does not extend to the idiom they are using at any given moment. Word-search games are no different in this respect. Players can find the target word without knowing what it means (Fig. 15). In short, the sampled games have all been designed primarily to facilitate practice of users' existing knowledge, rather than their learning of new words.

5.2 Learning strategies

All five strategies identified by Zhang and Liu (2015) were used in these apps, though several incorporated only one or two. As mentioned earlier, given that all 22 apps were categorizable as either dictionaries or games, the amount of learning that occurs when using them is likely to be limited. Notably, none appear to take into account any learning goals, a key aspect of any learning strategy (Zeidner, Boekaerts, & Pintrich, 2000). Worse, none include any mechanisms whereby users can test their understanding of words. Other than game scores, wins, and losses, none of the apps provide feedback at any point. The games appeared to even have been designed based on an assumption that their users will already have sufficient knowledge to complete them, which may explain why hints and feedback are rare to nonexistent.

Several limitations of this study need to be noted. First, the researchers used only two search terms to identify Chinese vocabulary-learning apps for native speakers, and thus could have overlooked well-designed learning apps that did not happen to have "vocabulary" in their descriptions. Second, the coding framework was based on learning in traditional settings, and thus may not have been comprehensive enough to cover all the sampled apps' features. While bottom-up coding was used to allow new themes to emerge, a new theme, metacognitive strategy, was duly identified.

6. Conclusion

Families' access to mobile devices and the Internet has steadily increased (Plowman, McPake, & Stephen, 2010), and research has consistently demonstrated mobile devices' potential for language learning. However, while several prior studies have shown that vocabulary-learning apps may be beneficial, the present study's findings suggest that the most popular such apps may not have optimal designs for vocabulary learning, at least in the Chinese case. Lacking clear learning goals, explanations, opportunities to practice, and feedback, these apps are unlikely to provide sufficient scaffolding for users' learning of new words. App developers should give due consideration to how mobile apps can better support such features, rather than using drill-and-practice tasks or simply assuming users know words already.

References

Anderson, T., Rourke, L., Garrison, W., & Archer, W. (2001). Assessing teaching presence in a computer conference context. *Journal of Asynchronous Learning Networks*, *5*(2), 1-17.

Blair, K. P. (2013). Learning in Critter Corral: Evaluating three kinds of feedback in a preschool math app. In *Proceedings of the 12th International Conference on Interaction Design and Children* (pp. 372–375). New York, NY, USA: ACM. https://doi.org/10.1145/2485760.2485814

Bureau, E. (2008). Lexical items with English explanations for fundamental Chinese learning in Hong Kong schools [中英對照香港學校中文學習基礎字詞]. Chinese

Language Education Section, Curriculum Development Institute, Hong Kong Government.

Callaghan, M. N., & Reich, S. M. (2018). Are educational preschool apps designed to teach? An analysis of the app market. *Learning, Media and Technology*, *43*(3), 280–293. https://doi.org/10.1080/17439884.2018.1498355

Chiu, Y.-H. (2013). Computer-assisted second language vocabulary instruction: A meta-analysis. *British Journal of Educational Technology: Journal of the Council for Educational Technology*, *44*(2), E52–E56. https://doi.org/10.1111/j.1467-8535.2012.01342.x

Gee, J. P. (2008). Learning and games. *The Ecology of Games: Connecting Youth, Games, and Learning*, *3*, 21–40. Retrieved from http://cvonline.uaeh.edu.mx/Cursos/Maestria/MTE/DisenoyCreacion_Mat_Mult/Unidad%201/LearningGames.pdf

Grabe, W. (2004). 3. Research on teaching reading. *Annual Review of Applied Linguistics*, *24*, 44–69.

Gu, Y. (2002). Gender, academic major, and vocabulary learning strategies of Chinese EFL learners. *RELC Journal*, *33*(1), 35–54. https://doi.org/10.1177/003368820203300102

Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, *46*(4), 643–679. https://doi.org/10.1111/j.1467-1770.1996.tb01355.x

Highfield, K., & Goodwin, K. (2013). Apps for mathematics learning: a review of educational apps from the iTunes App Store. Retrieved from https://www.researchonline.mq.edu.au/vital/access/services/Download/mq:30079/DS01

Lip, P. C. H. (2009). Investigating the most frequently-used and most-useful vocabulary language learning strategies among Chinese EFL postsecondary students in Hong Kong. *Electronic Journal of Foreign Language Teaching*, *6*(1), 77-87. Retrieved from http://e-flt.nus.edu.sg/v6n12009/lip.pdf.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9–F16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

Oxford, R., & Ehrman, M. E. (1995). Adults' language learning strategies in an intensive foreign language program in the United States. *System*, *23*(3), 359–386. https://doi.org/10.1016/0346-251X(95)00023-D

Papadakis, S., Kalogiannakis, M., & Zaranis, N. (2018). Educational apps from the Android Google Play for Greek preschoolers: A systematic review. *Computers & Education*, *116*, 139-160. https://doi.org/10.1016/j.compedu.2017.09.007

Plowman, L., McPake, J., & Stephen, C. (2010). The technologisation of childhood? Young children and technology in the Home. *Children & Society*, *24*(1), 63–74. https://doi.org/10.1111/j.1099-0860.2008.00180.x

Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, *56*(2), 282–308. https://doi.org/10.3138/cmlr.56.2.282

Redd, J. (2011). Supporting vocabulary growth of high school students: An analysis of

the potential of a mobile learning device and gaming app. Retrieved from https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=3090&context=etd

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 199–228). Cambridge, UK: Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, UK: Cambridge University Press.

Sung, Y.-T., Chang, K.-E., & Yang, J.-M. (2015). How effective are mobile devices for language learning? A meta-analysis. *Educational Research Review, 16*, 68–84. https://doi.org/10.1016/j.edurev.2015.09.001

Vaala, S., Ly, A., & Levine, M. H. (2015). Getting a Read on the App Stores: A Market Scan and Analysis of Children's Literacy Apps. Full Report. *Joan Ganz Cooney Center at Sesame Workshop*. Retrieved from http://files.eric.ed.gov/fulltext/ED574396.pdf

Wang, B. T. (2017). Designing mobile apps for English vocabulary learning. *International Journal of Information and Education, 7*(4), 279-283. Retrieved from http://www.ijiet.org/vol7/881-C09.pdf

Wang, B. T., Teng, C. W., & Chen, H. T. (2015). Using iPad to facilitate English vocabulary learning. *International Journal of Information and Educational Technology, 5*(2), 100-104. Retrieved from http://ijiet.org/papers/484-S00003.pdf

Wu, Q. (2015). Designing a smartphone app to teach English (L2) vocabulary. *Computers & Education*, *85*, 170–179. https://doi.org/10.1016/j.compedu.2015.02.013

Yang, Y. (2017). *Vocabulary learning strategies used by L1 English speakers learning Chinese as a foreign language*. King's College London. Retrieved from http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.718557

Yip, F. W. M., & Kwan, A. C. M. (2006). Online vocabulary games as a tool for teaching and learning English vocabulary. *Educational Media International*, *43*(3), 233–249. https://doi.org/10.1080/09523980600641445

Zeidner, M., Boekaerts, M., & Pintrich, P. R. (2000). Self-regulation: Directions and challenges for future research. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 749–768). New York: Academic Press.

Zhang, X., & Lu, X. (2015). The relationship between vocabulary learning strategies and breadth and depth of vocabulary knowledge. *The Modern Language Journal*, *99*(4), 740–753. https://doi.org/10.1111/modl.12277

Zhang, Y., Lin, C.-H., Zhang, D., & Choi, Y. (2017). Motivation, strategy, and English as a foreign language vocabulary learning: A structural equation modelling study. *The British Journal of Educational Psychology*, *87*(1), 57–74. https://doi.org/10.1111/bjep.12135

Augmented Reality in Foreign Language Education: A Review of Empirical Studies
(增强现实技术在外语教学中的应用：文献综述性研究)

Zhang, Shenglan
(张胜兰)
Iowa State University
(爱荷华州立大学)
shenglan@iastate.edu

Abstract: This literature review examines how Augmented Reality (AR) has been used in foreign language learning. AR is a live view of reality that is augmented by computer-generated sound, image, or videos. It allows the user to interact with the real physical environment in an enhanced way. This study provides an overview of what AR is, its history, different definitions, and how it has been used in education in general. It summarizes how AR has been used in all aspects of foreign language education, including skill development (listening, speaking, reading and writing), vocabulary, grammar, culture, the aspect of affect in language learning, what AR tools were used, and a discussion of the advantages and disadvantages of AR in language learning. At the end of the article, the author suggests further research needed to support the widespread adoption of AR in foreign language education in general and in Chinese-as-a-Foreign-Language, specifically.

摘要: 增强现实技术是将计算机生成的虚拟图像、视频、3D模型等材料叠加到现实世界，使之与现实世界融合从而增强对现实的表现。该技术近年来在很多领域得到了普遍应用，特别是在军事、旅游、医疗、商业等领域。在教育领域，其应用也越来越普遍。在外语教学方面，增强现实技术的应用则正处于起始阶段。通过对相关英文学术期刊进行穷尽式搜索与综述，本文试图回答以下几个问题：过去五年内，外语教学中增强现实技术是不是得到了应用，是如何应用的，这些应用有没有益处，如果有，有何益处？穷尽式搜索得到十篇实证性研究，多数侧重在词汇教学。研究结果表明，增强现实技术能够更好，而且多方面地提高学习效率。文章最后探讨了如何更好地将这项技术运用到外语教学，包括中文作为外语的教学中。

Key words: Augmented Reality, AR, Foreign Language Teaching, Second Language Teaching, Multimedia, CALL

关键词： 增强现实技术，外语教学，混合现实，多媒体，中文教学，计算机辅助教学

1. Introduction

Augmented Reality (AR) is "a real-time direct or indirect view of a physical real-world environment that has been enhanced/augmented by adding virtual computer-generated information to it." (Carmigniani & Furht, 2011, p.1). It is one of the latest technologies adopted in Computer Assisted Language Learning (CALL). The word "augmented" comes from "augment" which means to add or enhance. Augmented reality intends to create an enhanced user experience by adding graphics, sounds, and touch feedback to the natural world. Azuma (1997) defines AR as systems with three important features: a combination of real and virtual dimensions, interactions in real time, and adoption of 3-D. There are also less restrictive ways in defining AR. For example, Eric Klopfer, professor of urban studies and planning and an expert on augmented reality games at Massachusetts Institute of Technology, included "location-based games on handhelds and mobile phones which provide additional virtual data or information at given locations" in the realm of augmented reality (Jenkins, 2008).

1.1 AR and VR

Sometimes AR is confused with virtual reality (VR) and mixed reality. Milgram, Takemura, Utsumi, & Kishino (1994) used a reality-virtuality continuum to describe how these different types of reality relate to each other. See Figure 1. In this continuum, AR lies in the region called mediated reality, between the physical environment and an entirely virtual environment. Mediated reality includes both augmented reality, which lies closer to physical reality, and augmented virtuality, which lies closer to virtual reality.

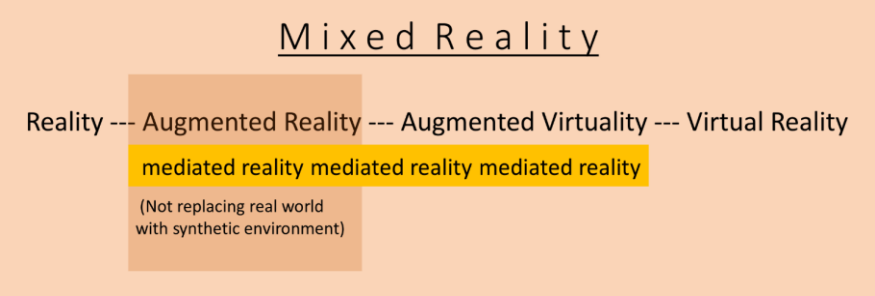


Figure 1. The location of AR in the continuum. Adapted from Milgram, Takemura, Utsumi, & Kishino (1994, p. 283).

Unlike virtual reality, which gives users an interactive computer-generated experience within an entirely virtual environment, AR takes advantage of the existing natural environment and superimposes virtual information on top of it. The co-existence of both virtual and real elements in AR provides users with a new and improved natural world with virtual information as assistance.

1.2 A short history of AR

The origins of AR studies date back to the 1950s. Morton Heilig (1926-1997), a cinematographer, thought that cinema should have the ability to draw viewers into the onscreen activity by utilizing all their senses in an effective manner (Carmigniani et al.,

2010). The term Augmented Reality was coined by Tom Caudell in 1990. He and his colleague, David Mizell, designed a head-mounted digital visualization system to guide technicians when they located cables on airplanes (Siltanen, 2012). Milgram, Takemura, Utsumi, & Kishino (1994) proposed a mixed AR technology model which represented the combination of reality and virtuality. This model had the main tenets of AR technology during the process of development (Chen & Tsai, 2012). In 1997, Azuma conducted the first survey study on the use of AR. Recently, AR has developed rapidly, with the mobile applications such as the Participatory Augmented Reality Simulations (PARS) and Mobile Augmented Reality (MAR).

1.3 AR technologies

Tracking technology is the essential part of AR. Two basic categories of tracking are distinguished: marker-based tracking and location-based tracking. Marker-based tracking refers to tracking by a fiducial marker, i.e., an optical image (or an interest point), which is an object placed in a scene as a fixed point of reference of position (see Figure 2). These markers, which sometimes use Quick Responsible (QR) code, provide an interface between the physical world and the augmented reality content such as 3D models or videos. Location-based AR tracking ties the augmented reality content to a specific, real-time location. Digital geo-spatial data are gathered and placed over the actual physical surroundings, allowing users to access the data. Accuracy is a critical factor in this case (Juegostudio, 2018; Kudan, 2017). The popular game Pokémon GO is an example of location-based tracking.

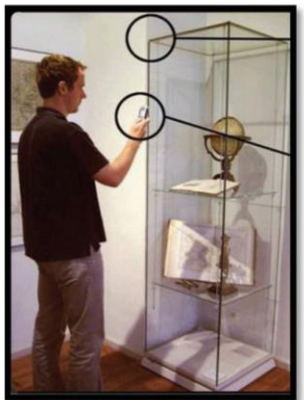


Figure 2. Using AR in a museum. (Carmigniani & Furht, 2011, p. 31)

Technologies used in AR are not limited to any particular type of display technology such as head-mounted display (HMD), and neither are they limited to the sense of sight. AR can potentially be used to augment smell, touch, and hearing (Carmigniani, 2011). “Technology has caught up with the idea of augmented reality as devices have become cheaper, smaller, and sufficiently powerful to run applications” (Salmon & Nyhan, 2013). In addition to HMD, smartphones, webcams, and digital projectors have also been used to augment reality. Smart phones play an important role in AR because they are equipped with cameras, apps and the internet connection. They make it possible to overlay virtual graphics and media over a physical object, such as a picture or an object. Using the smart phone camera to point at the picture or object, the user is exposed to the augmented virtual layers, such as video, 3D animation, or text explanation. This makes it feasible to “transform a classroom setting to a virtual learning environment: for example, real cultural artifacts triggering images or videos of their history, wall maps displaying geographical locations, portraits bringing to life real interviews” (Carmigniani & Furht, 2011, p. 20). See Figure 3 for an example of using AR in a classroom setting: The mobile scans a marker on the textbook and videos about that part of the grammar/culture are triggered.

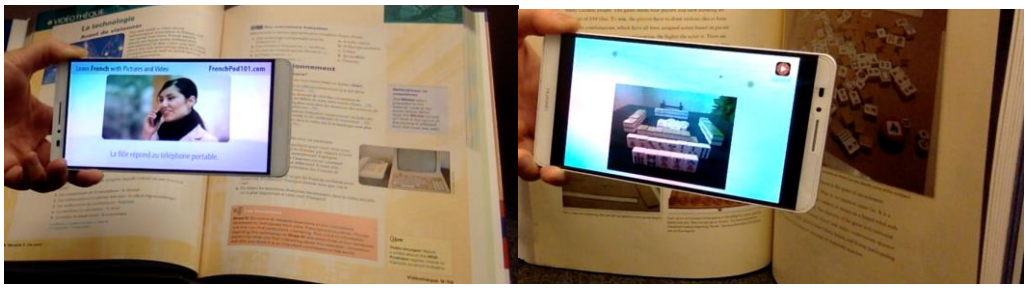


Figure 3. Using AR with a mobile device in a classroom setting

1.4 The application of AR

AR has been widely adopted in many areas of modern life, such as in advertising and marketing (Hampp, 2009; Yuen, Yaoyuneyong, & Johnson, 2011), in architecture and construction (Rankohi & Waugh, 2013), in entertainment (Von Itzstein, Billinghurst, Smith, & Thomas, 2017), medicine (Samset et al., 2008), and the military (Sisodiaa et al., 2007). It has also been used in education (Bacca, Baldiris, Fabregat, Graf, & Kinshuk, 2014; Yuen, Yaoyuneyong, & Johnson, 2011), for example, in science and math as well as in literacy education (Billinghurts & Duenser, 2012; Bower, Howe, McCredie, Robinson, & Grover, 2014). Research has demonstrated many favorable outcomes for AR, such as enhancing problem solving and collaboration (Dunleavy, Dede, & Mitchell, 2009; Wasko, 2013), creating contextual awareness (Ivanova & Ivanov, 2011), helping better understand abstract concepts by concretizing them (Dori & Belcher, 2005; Klopfer & Squire, 2008), offering authentic learning experiences (Klopfer & Squire, 2008; Yuen, Yaoyuneyong, & Johnson, 2011), and motivating learners and enhancing learners' enjoyment and attitude (Billinghurts & Duenser, 2012; Dunleavy, Dede, & Mitchell, 2009; Jerry & Aaron 2010). AR has been employed for learners at all levels, from elementary school students to students at university (Klopfer & Squire, 2008).

2. Literature Review

While AR has been adopted in educational fields, it is at the beginning stages in language education. Only four studies have discussed the use of AR in language learning. Ramya & Madhumathi (2017) reviewed five publications concerning AR in literature study to assess the overall effect of the incorporation of AR on learning outcomes. Of the five studies, only two were published in peer-reviewed journals. Their overall conclusion was that the studies "favored the efficiency of Augmented Reality" (p. 358). However, no methodology in searching relevant literature was specified, which resulted in missing quite a few relevant studies that should be included.

Godwin-Jones (2016) introduced AR as an emerging technology in language learning and explored different ways in which it is being used in this area from a practical point of view. He first discussed the potential that such games as Pokémon GO could provide for language learning, such as the social interactions that accompany gameplay. Then he explained what marker-based AR was and how students could benefit by creating markers and helping to create the augmentation (Bower, Howe, McCredie, Robinson, &

Grover, 2014; Slussareff & Boháccková, 2016). In addition, Godwin-Jones (2016) introduced place-based AR. He took *Mentira*, one of the better-known place-based mobile game for Spanish language learning, as an example and illustrated how it worked. He cited the result of the *Mentira* developers and researchers, Holden and Sykes's (2011; 2013) research to show that playing the game increased the students' awareness of pragmatic issues in Spanish. At the end of the article, Godwin-Jones explained how to use some AR tools such as ARIS and TaleBlazer and how AR could be used in promoting the learning of target historical, cultural or literary topics. Godwin-Jones's article provided a very comprehensive introduction to how AR could be used in the field of language learning.

Salmon & Nyhan (2013) proposed an evaluative framework that could be used as a baseline for making decisions on the use of AR applications for language teaching in a classroom setting. They used five exemplary studies that examine effective language teaching and learning and the role of technology in order to build the framework. They first summarized effective language teaching and learning principles with the use of technology. Then they established a framework based on the principles, and from the framework developed criteria for evaluating AR technology for language teaching and learning. For example, the software should have clearly stated pedagogical objectives; it should focus on specific language skills; it should enable the integration of different language skills; it should allow for task-based, exploratory language learning; and it should enable learners to build on existing knowledge and to develop this knowledge further. The last part of the article examined this framework using four AR applications: the Specialist Schools and Academies Trust (SSAT) AR Quizzes, *ZooBurst*, *Mentira*, and Context-Sensitive Microlearning of Foreign Language Vocabulary.

Davis & Berland (2013) focused their research on participatory AR simulations (PARS). They evaluated the possible merits and difficulties of utilizing PARS with English language learners in K-12 science classrooms. They identified elements of PARS that met the requirements for effective learning of English as a second language and these elements included modality, engagement, collaboration, and language use.

All of the above studies have reviewed literature of AR in the field of language education. Each of them approached AR from different perspectives, either to provide a big picture of how AR is used in a practical sense, or to establish a framework for evaluating the use of AR. However, a more comprehensive review of the recent empirical studies in the recent years is needed in order to understand the use of AR in language education and what research has said about the advantages and disadvantages of using AR in language education.

3. Research Questions

The current study explored the use of AR in teaching any language as a foreign or second language. It intended to answer the following research questions.

1) What does the literature say about using AR in foreign/second language education in the past five years?

2) Is there evidence that shows the advantage of adopting AR in foreign/second language education? If such evidence exists, what advantage(s) does the adoption of AR bring to foreign/second language education?

4. Methodology

Five databases were used in this study: Communication Abstracts (EBSCO), Education Resources Information Center (ERIC), J-STOR, Google Scholar, Linguistics and Language Behavior Abstracts (ProQuest). The year of publication was set from 2013 to the present in the searches. That is to say, as long as a study that was published in the previous five years met the following criteria, it was included in this study.

There were five criteria that a study had to meet to be included. First, it must have been published in a peer-reviewed journal. Second, it had to be an empirical study. Third, it must have been related to second-language or foreign-language learning in any language. Fourth, the study must have been written in or have been translated into English. Fifth, the term "Augmented Reality" had to be in the title. Because the definition of AR is so broad, if the term "Augmented Reality" was not in the title, it would have been extremely difficult to justify whether the core technology in the articles was defined as AR. In order to ensure the search was exhaustive and relevant, in addition to the query word "Augmented Reality" in the title, one of the following query words was also used in the database search: Language, English, Spanish, German, Russian, French, Chinese, Japanese, Korean, Indonesian, Malay, Arabic, TEFL, TESL, ELL, Vocabulary, Grammar, Writing, Reading, Listening, Speaking.

A few types of studies were, however, excluded because they concentrated on one specific application or on culture instead of language learning, such as studies that focused on the usability of the application developed with AR and/or the gamification idea with foreign language learning. If the study used language cards and used AR solely to determine if the cards could promote culture learning, the study was excluded. If the study developed a web portal and/or a container application with image recognition and AR technologies, but the main purpose of the study was to test if the web portal or the application worked, this study was also excluded.

5. Findings

The search resulted in ten relevant studies. Two obvious trends can be observed from these studies. First, the number of published studies about AR in language learning has progressively increased year by year (Figure 4). Second, the journals in which these studies were published are very scattered. Only two of the studies were published in the same journal. The journals where these studies were published include: *Procedia: Social and Behavioral Sciences* (in which two studies were published)*; British Journal of Educational Technology; The Journal of Educators Online; International Journal of Game-Based Learning; Research and Practice in Technology Enhanced Learning; Human-Computer Interaction; Computer and Education; Educational Technology &*

Society; Eurasia Journal of Mathematics Science and Technology Education.

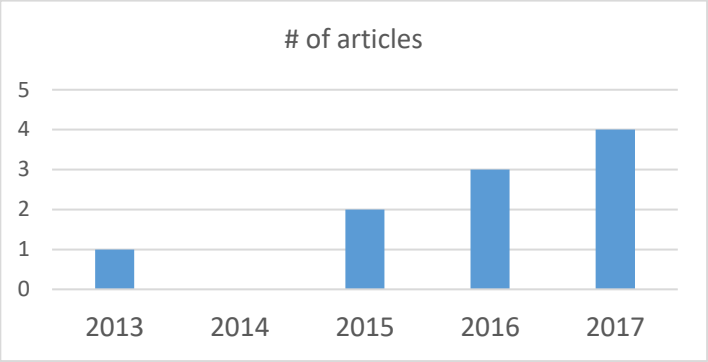


Figure 4. Publication year of the studies.

5.1 A general description of the studies

Of the ten studies, there were five experimental studies, three case studies, one descriptive study, and one exploratory study (Figure 5). While some studies addressed more than one language, others focused on only one. Nine studies focused on the application of AR in teaching English as a second language. Four studies examined the use of AR in teaching Basque, French, German, and Filipino, respectively (Figure 6).

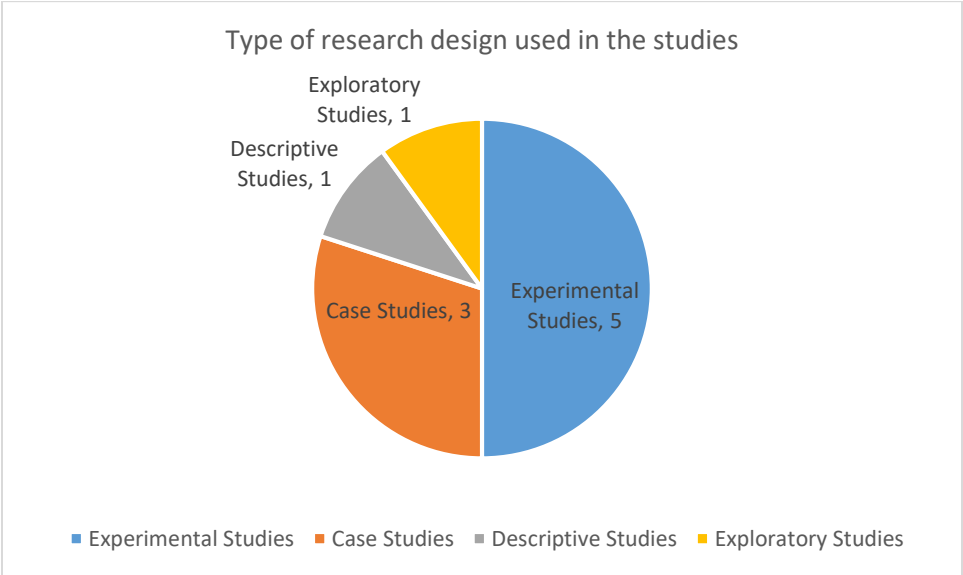


Figure 5. Research methods adopted in the studies

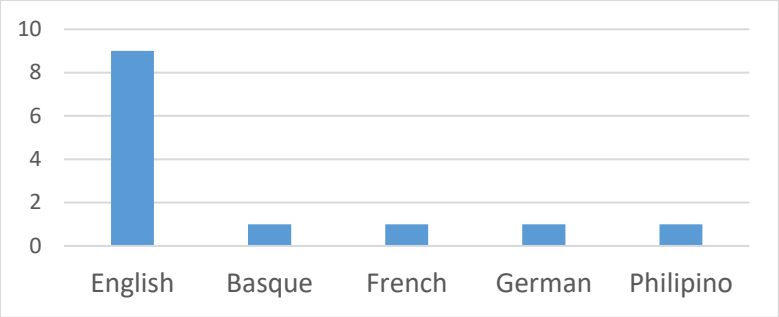


Figure 6. Focus languages represented in the studies

While some studies examined multiple aspects of learning and the learning experience, some had only one focus. Six of the ten studies focused on vocabulary learning and three on the learners' learning experience with AR, which included the flow experience, cognitive load, and learning anxiety. One of the ten studies focused on writing, one on alphabet learning, and one on culture understanding (Figure 7).

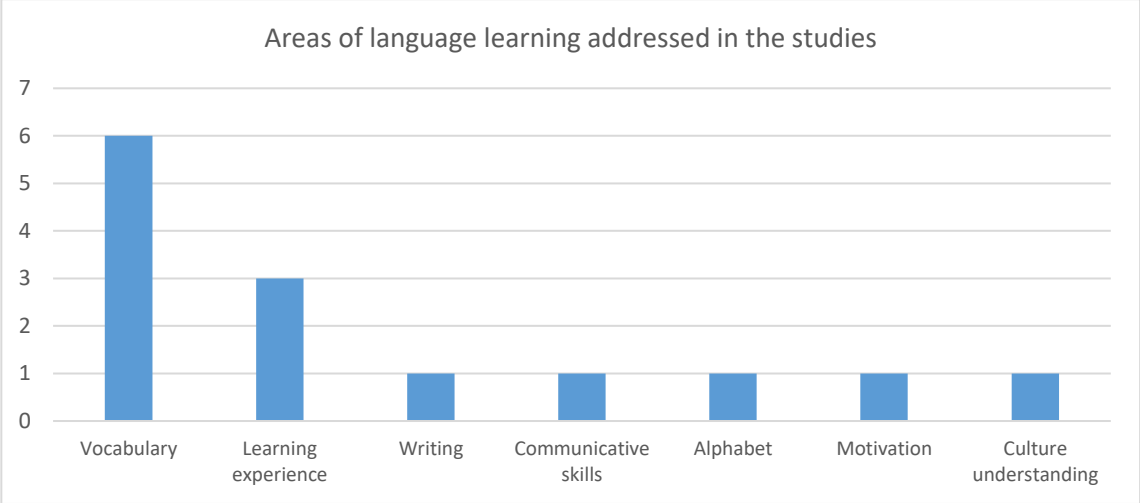


Figure 7. Areas of language learning that were addressed in the studies

Most the studies were conducted with students at the university level. Only one had graduate students as participants, one had 3rd graders as participants, and one had kindergarteners as participants. Participants in one study included both undergraduate students and employees from different professions in the local community (Figure 8).

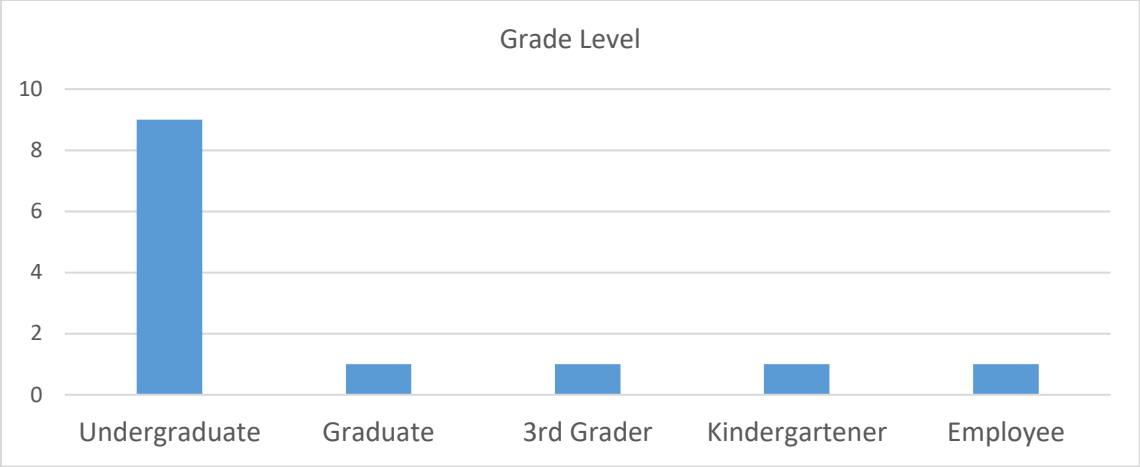


Figure 8. Participants' grade level

Most of the studies (6) examined how beginning level language learners performed with the implementation of AR technology. One study worked with participants who were intermediate level learners, and one worked with advanced level learners. Two studies did not specify the language level of their participants (Figure 9). Table 1 presents a summary of the ten studies.

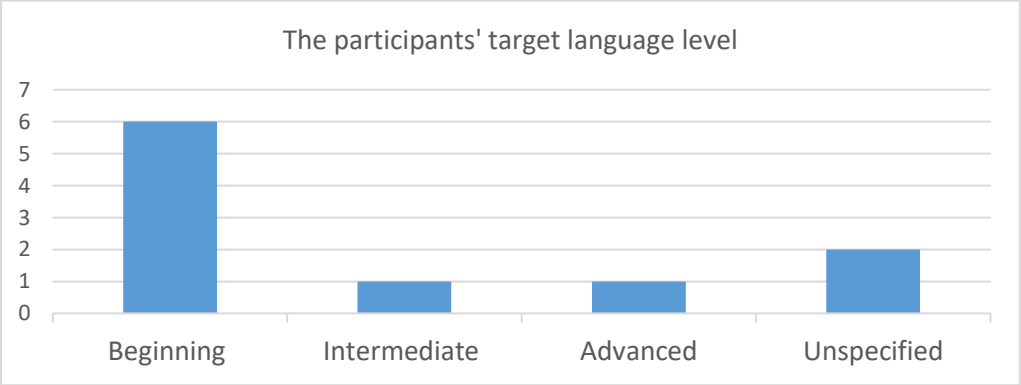


Figure 9. Participants' target language level

5.2 AR tools used in the studies

Five of the ten studies applied global positioning technology and AR in foreign language learning. Three of the five gamified it to motivate students. Two used open-source platforms for creating and playing AR games: ARIS (Augmented Reality and Interactive Storytelling platform) or Aurasma. Perry (2015) used ARIS to create *Explorez*. *Explorez* is similar to *Mentira*, the AR mobile game for learning Spanish. *Explorez* was said to have transformed the university campus "into a virtual francophone world, where students interact with characters, items, and media as they improve their French language skills and discover their campus" (p. 2308). It was designed to bridge the "gap between gaming and education through quest-based learning and augmented reality" (p. 2309).

Table 1. Empirical studies using AR in teaching a second or foreign language published in peer-reviewed journals

Study and year	Research Method	Target Language	Aspect of language	Participants' Language Level	Grade Level	Tools Used
Liu & Tsai (2013)	Case Study	English	Writing	Not specified	College	AR-based mobile learning materials
Perry (2015)	Case Study	French	Learning experience	Beginning	College	Mobile learning tool: *Explorez*
Solak & Cakir (2015)	Descriptive	English	Vocabulary; Motivation towards vocabulary learning	Beginning	College	A material designed with support of AR technology
Liu, Holden, & Zheng (2016)	Exploratory	English	Vocabulary learning; Culture understanding; Communicative skills	Intermediate	College	AR mobile game: *Guardians of the Mo'o* designed with ARIS
Richardson (2016)	Multiple case studies	English	Learning experience	Advanced	College	Location-based AR game: Mission Not Really Impossible using *Aurasma* mobile app
Santos, et al. (2016)	Experimental	Filipino; German	Vocabulary	Beginning	Graduate	Handheld AR system
Safar, Al-Jafar, Al-Yousefi (2017)	Experimental	English	Alphabet	Beginning	Kindergarten	AR apps: AR Flashcards Animal Alphabet and AR Alphabet
Ho, Hsieh, Sun & Chen (2017)	Experimental	English	Learning performance (Vocabulary)	Not specified	18-30 years old: college students and employees	Ubiquitous Learning Instruction System with AR features (UL-IAR)
Hsu (2017)	Experimental	English	Flow experience; Cognitive load; Foreign language learning anxiety, Learning effectiveness (vocabulary)	Beginning	3rd graders	Two AR educational games system
Ibrahim et al. (2017)	Experimental	Basque	Vocabulary (nouns)	Beginning	College	*ARbis Pictus* – A novel system for immersive language learning through dynamic labeling of real world objects in AR.

Liu, Holden, and Zheng (2016) created a mobile game called *Guardians of the Mo'o* using ARIS. It was a non-linear, "place-based quests using GPS tracking functions on mobile devices without an on-site program" (p. 370). The game theme was based on Hawaiian mythology. In the game, the students, i.e., the game players, were the helpers or guardians of the Mo'o, a figure of Hawaiian mythology, who was sick and needed help.

Richardson (2016) used AURASMA to develop a location-based AR game, *Mission Not Really Impossible*. This game provided a series of challenging language tasks for the learners when they moved around the city of Karlsruhe in Germany. The game had a video introduction to the game. After the introduction, the learners were given a map which they needed to find a "trigger" image in the gameplay area. Once the player found and scanned this image, the mysterious Danlod, the game's main character, appeared and showed a short video clip telling the players what still needed to be done.

Two studies used global positioning technology but were not designed as games. Ho, Hsieh, Sun, and Chen (2017) developed a Ubiquitous Learning Instruction System with Augmented Realty Features (UL-IAR). The design "incorporated AR technology and extensive computing power within a context-aware u-learning system to provide adaptive learning strategies to assist English learning" (p. 179). The site for which the UL-IAR was developed was Kaohsiung West Bay in Taiwan. The reason for choosing that site was that the West Bay has many popular scenic spots and food from street vendors, on which the learning materials were based.

In Liu & Tsai (2013), the EFL learners executed the AR-based mobile learning material on campus. When they used the mobile phone to point in a specific direction, their location was identified, the embedded camera automatically captured the peripheral images, and related information such as names and descriptions of the buildings were shown on the screen. If the learner wanted to know more about a certain building, the learner could click on that piece of information and details would appear.

Five of the ten studies did not use global-positioning technology when they applied AR. Four of the five studies used AR to connect multimedia (images, animations, and text) with real objects in real environment. Although these four studies integrated multimedia materials in the designed AR system, they were based on different concepts and underlying theories in the design. For example, Santos et al. (2016) treated AR "as a type of multimedia that is situated in authentic environments" (p. 1), and they applied multimedia learning theory as a framework for developing the application. They created a handheld AR system which displayed different combinations of multimedia in sound, image, animation, and text in real environments.

Hsu (2017), however, designed two AR systems based on the way learners approached the system: self-directed or task-based linear approach. One of the findings was that the AR educational system based on a self-directed learning approach did not restrict the learning sequence and provided more support in learning. Ibrahim et al. (2017), on the other hand, designed *ARbis Pictus*, a novel system for immersive language learning through dynamic labeling of real-world objects in AR. Learning in this study occurred in a controlled learning environment while learners used Microsoft HoloLens with an AR head-

mounted display. The application was set up in a room where all the objects were placed. The learner could walk around the room with the head-mounted display to locate the object, see labels, and interact with the labels to get multimedia information about the object.

One study chose ready-made AR applications. Safar et al. (2017) "reviewed a wide range of AR software available at the online App Store designed to be viewed using iPad tablets" and chose two applications. One was "AR Flashcards Animals-Alphabet" and the other was "AR Alphabet Flashcards." Both apps presented English alphabet initials of different pets and their predators. When the pets were clicked, a three-dimensional letter would appear accompanied by the sound and animated movements of each animal.

In short, of the ten studies, half used global positioning technology with AR and the other half utilized AR as multimedia. While nine were designed by a research team, one study utilized ready-made apps.

5.3 Advantages that AR has brought to language learning

The published studies reviewed here suggest that using AR in language learning brings many benefits. It engages not only beginning-level and intermediate level foreign language learners in learning (Ho et al., 2017; Liu & Tsai, 2013; Perry 2015) in both the virtual and the physical world (Liu, Holden, & Zheng, 2016), but it also engages and challenges advanced-level learners (Richardson, 2016). It improves learner satisfaction (Ibrahim et al., 2017; Santos et al., 2016; Solak & Cakir; 2015), gives a satisfying feeling of accomplishment (Solak & Cakir, 2015), enlivens learning experience (Solak & Cakir, 2015), increases learners' flow experience, and lowers learners' cognitive load (Hsu, 2017). In addition, it motivates learners (Liu, Holden, & Zheng, 2015; Perry, 2015; Solak & Cakir, 2015) and increases the effectiveness of language learning (Ibrahim et al., 2017; Liu & Tsai, 2013; Safar et al., 2017), regardless of whether the AR was designed using a self-directed or a task-based approach (Hsu, 2017). Specifically, the use of AR leads to better word retention (Ho et al., 2017; Hsu, 2017; Ibrahim et al., 2017; Santos et al., 2016; Solak & Cakir, 2017), helps students construct linguistic and content knowledge (Liu & Tsai, 2013), and supports learners in gaining content knowledge (Liu & Tsai, 2013). Furthermore, it is beneficial to adopt AR at the primary stage of education such as pre-school and kindergarten (Safar et al., 2017), with college students (Ho et al., 2017; Perry, 2015; Liu et al., 2016; Richardson, 2016; Solak & Cakir, 2015; Hsu, 2017; Liu & Tsai, 2013), and with graduate students (Santos, et al., 2016).

There are a variety of reasons why the use of AR offers these positive outcomes. First, AR increases learners' active interaction with the learning materials (Hsu, 2017; Safar et al., 2017; Santos et al., 2106; Solak & Cakir, 2015). Well-designed AR enhances the presentation of information and helps attract learners' attention to the learning materials by using "eye-catching" formats of presentation, such as stories, pictures, audios, animations stimulate curiosity (Solak & Cakir, 2015).

Second, AR technology combined with the GPS-positioning allows learning to take place outside of the classroom and provides a contextual and immersive learning experience for learners (Liu & Tsai, 2013; Perry, 2015; Solk & Cakir, 2015). This learning experience makes learning more relevant to the learners' life and therefore enhances the meaningfulness of learning. Learning is naturally tied to authentic activities and contexts (Brown, Collins, & Duguid, 1989). AR provides an opportunity to participate in authentic activities in semi-realistic contexts. Furthermore, learning that takes place out of the classroom is a change from the day-to-day setting in the classroom repetition and the change itself makes learning exciting (Perry, 2015).

Third, AR-enhanced games provide opportunities for dynamic negotiation and collaboration (Liu, Holden, & Zheng, 2016). In an AR-enhanced game, collaborative negotiation in the target language is frequently required in order to proceed in game playing. Through negotiation with peers and with strangers in the social context, learners are able to apply the language they are learning in real contexts. Furthermore, collaboration helps learners build a sense of community and relationships in a course, which can also maximize learning (Dörnyei & Ushioda, 2011; Richardson, 2016).

Fourth, the use of AR provides opportunities for the learner to use the language in a spontaneous and unplanned way (Richardson, 2016). In most classroom settings, students lack the opportunity to step out of their comfort zone and speak on the spot. Intermediate and advanced learners should be exposed to situations like these to improve their language skills and to prepare them to use the target language in real life (Richardson, 2016; White, 1971).

5.4 Other findings

In addition to the findings about the benefits of AR in language learning, some studies focused on other aspects of learning. For example, Solak and Cakir (2015) found that the motivation to use AR technology in language learning was not affected by gender or the students' majors. In other words, the use of AR could help increase college students' motivation in learning the target language without privileging specific majors or gender. Furthermore, Hsu (2017) found that the self-directed learning approach implemented in the AR educational game system could cause some learning anxiety and that this increased anxiety promoted effective learning more than the task-based learning approach, which limited the learning sequence.

5.5 Problems of using AR in language education

The studies also found some problems in using AR in language education. The biggest problem was that learners need "sufficient familiarity with the AR-based mobile learning material" (Liu & Tsai, 2013, p. 4) and they must know how to use the device before any learning can happen (Liu, Holden, & Zheng, 2016; Perry, 2015; Richardson, 2016). Therefore, in implementing AR, sufficient time should be scheduled to familiarize learners with operation of the devices as well as the application's rules. Furthermore, learners needed to be provided with adequate game-play scaffolding in the gaming system (Perry, 2015). As Gee (2003) pointed out, well-developed video games provide sufficient

guidelines for new players at the beginning. Perry (2015) stated that when creating an AR game system, developers and instructors should keep in mind to not only "provide sufficient language-learning scaffolding, but also include adequate game-play scaffolding" (p. 2313).

One limitation of using AR is that instructors cannot readily gauge to what extent students use the target language and cannot monitor usage so that the students do not revert to their first language (Perry, 2015). Instructors can participate in the field experience with the learners to monitor their use of language. However, it is not practical that all the learners can be monitored. A solution to this problem is needed.

In designing an AR system, two additional elements should be kept in mind. The first is that more information should be provided so that students could access a wide variety of information, selected from the same topic (Liu & Tsai, 2013). For example, to explain about the cultural phenomenon "playing Mahjong" in China, not only the rules of Mahjong should be introduced, but also who plays Mahjong, its significance for the players, and the diverse cultural images of Mahjong need to be provided. The second is that sufficient and effective feedback should be available to students. Feedback is important, especially in the game-based AR system. It should be instantaneous, and it should let the players know whether they are engaging the game in the appropriate manner (Gee, 2003).

6. Conclusion

AR is being increasingly explored in the field of language education. The advantage of using it in supporting language learning have been shown in the studies reviewed here. These studies have adopted various research methods to answer different research questions. English as a foreign language is the most studied language among the ten studies, and vocabulary learning has gained much attention. Although the ten studies were published in non-language learning journals, language-learning journals are likely to attract more articles on AR in language education in the near future.

Future studies that focus on adopting AR in foreign language education must, however, overcome one significant hurdle. As Safar et al. (2017) pointed out, there is a limited availability of specialists and experts in the field of AR technology. To have more research conducted in this area, language teachers need to learn how to develop AR. Of the ten studies, only three specified what tools they used to develop AR, and two tools were mentioned: ARIS and Aurasma. Both ARIS and Aurasma are open-source AR platform for creating and playing video games. ARIS consists of an authoring tool to create augmented reality games or interactive stories. It also has an application for IOS devices to play the games. Aurasma is an AR platform which uses a smartphone's camera to identify real-world images and overlay the images with an "aura". The aura can be an image, video, audio, animation, 3D model or a webpage. Each aura can be triggered by using the phone's camera to point at the corresponding image of the pre-selected object. Both ARIS and Aurasma do not require the developer to have programming knowledge. However, there is a steep learning curve to mastering the platform (Perry, 2015).

An additional suggestion for research on AR is that more design-based research is needed. Design-based research usually focuses both on developing the solution to a problem and on describing the result. By describing the details of the design process, both researchers and practitioners can learn how to apply AR in teaching the target language. However, only one of the ten studies reviewed here adopted a design-based research approach.

Furthermore, research on how AR can change language classroom dynamics is needed. Only after we have learned how the classroom dynamics have or have not been changed by the adoption of AR can we make the corresponding decision on how to change the other aspects of our teaching which do not involve the use of AR.

Unfortunately, none of the studies focused on learning Chinese as a foreign language. As a tonal and logographic language, Chinese presents students with additional challenges in the language learning process. There is, therefore, an urgent need to determine how the adoption of AR technology and techniques could make the learning of Chinese both more fun and more effective. A series of measured and deliberative steps needs to be taken. For example, a small-scale study on the adoption of AR could investigate how to improve learners' tones, character learning, and speaking. Gradually, gamification can be added with location-based techniques. Technology support is a critical key to success. However, its absence should not be a barrier. Teachers and researchers can try to start with using Mayer's (2005; 2009) multimedia design principles to design AR and produce multimedia, e.g. as the AR as Santos et al. (2016) did.

References

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments 6*(4), 355-385.

Bacca, J., Baldiris, S., Fabregat, R., Graf S., & Kinshuk. (2014). Augmented reality trends in education: A systematic review of research and applications. *Journal of Educational Technology & Society, 17*(4), 133-149.

Billinghurst, M. & Duenser, A. (2012). Augmented reality in the classroom. *Computer, 45*, 56-63. doi: 10.1109/MC.2012.111

Bower, M., Howe, C., McCredie, N., Robinson, A., & Grover, D. (2014). Augmented reality in education – Cases, places and potentials. *Educational Media International, 51*(1), 1-15. doi: 10.1080/09523987.2014.889400

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

Carmigniani, J. (2011). Augmented reality methods and algorithms for hearing augmentation. (Unpublished master's thesis). Florida Atlantic University, Boca Raton, FL.

Carmigniani, J. & Furht, B. (2011) Augmented reality: An overview. In B. Furht (Ed.), *Handbook of augmented reality (pp. 3-46).* New York, NY: Springer. doi: 10.1007/978-1-4614-0064-6_1

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., & Ivkovic, M. (2010). Augmented reality technologies, systems, and applications. *Multimedia Tools and Applications, 51*(1), 341-377. doi: 10.100711042-010-0660-6

Chen, C. M. & Tsai, Y. N. (2012). Interactive augmented reality system for enhancing library instruction in elementary schools. *Computers & Education, 59*(2), 638-652.

Davis, D., & Berland, M. (2013). Supporting English learners with participatory augmented reality simulations. *On the Horizon, 21*(4), 294-303.

Dori, Y. J. & Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *Journal of the Learning Sciences, 14*(2), 243-279.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation*. Harlow, UK: Pearson.

Dunleavy, M., Dede, C. & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology, 18*(1), 7-22.

Gee, J. (2003). *What video games have to teach us about learning and literacy*. New York, NY: Palgrave Macmillan.

Godwin-Jones, R. (2016). Augmented reality and language learning: From annotated vocabulary to place-based mobile games. *Language Learning & Technology, 20*(3), 9–19. Retrieved from http://llt.msu.edu/issues/october2016/emerging.pdf.

Hampp, A. (2009). Marketers hop on augmented reality bandwagon to promote 'Avatar'. Retrieved from http://adage.com/madisonandvine/article?article_id=140661.

Ho, S., Hsieh, S., Sun, P., & Chen, C. (2017). To activate English learning: Listen and speak in real life context with an AR featured U-learning system. *Educational Technology & Society, 20*(2), 176-187.

Holden, C., & Sykes, J. (2011). Leveraging mobile games for place-based language learning. *International Journal of Game-Based Learning, 1*(2), 1–18.

Holden, C., & Sykes, J. (2013). Place-based mobile games for pragmatics Learning. In N. Taguchi & J. Sykes (Eds.), *Technology in interlanguage pragmatics research and teaching* (pp. 1–15). Philadelphia, PA: John Benjamins.

Hsu, T.C. (2017). Learning English with augmented reality: Do learning styles matter? *Computers & Education, 106*, 137-149.

Ibrahim, A., Huynh, B., Downey, J., Höllerer, T., Chun, D., & O'Donovan, J. (2017). ARbis Pictus: A study of language learning with augmented reality. *Human-Computer Interaction, 1*. Retrieved from https://arxiv.org/pdf/1711.11243.pdf.

Ivanov, S., & Ivanov, M. (2011). Triple bottom line analysis of potential sport tourism impacts on local communities – A review. *Proceedings of the Black Sea Tourism Forum "Sport tourism – Possibilities to extend the tourist season*" (pp. 168-177). Varna, Bulgaria.

Jenkins, H. (2008). Augmented learning: An interview with Eric Klopfer. Retrieved from http://henryjenkins.org/2008/07/an_interview_with_eric_klopfer.html

Jerry, T. F. L., & Aaron, C. C. E. (2010). The impact of augmented reality software with inquiry-based learning on students' learning of kinematics graph. In V. Mahadevan & G. S. Tomar (Eds.), *The 2nd International Conference on Education Technology and Computer* (pp. V2-1–V2-5). Piscataway, NJ: IEEE.

Juegostudio. (2018). *Geo location based augmented reality apps*. Retrieved from https://www.juegostudio.com/blog/geo-location-based-augmented-reality-apps.

Kaufmann, H., & Schmalstieg, D. (2003). Mathematics and geometry education with collaborative augmented reality. *Computers & Graphics, 27*(3), 339-345. doi: 10.1016/S0097-8493(03)00028-1

Klopfer, E. & Squire, K. (2008). Environmental detectives --The development of an augmented reality platform for environmental simulations. *Educational Technology Research and Development, 56*(2), 203-228.

Kudan. (2018). *Augmented reality fundamentals: Markers*. Retrieved December 11, 2018 from https://www.kudan.eu/kudan-news/augmented-reality-fundamentals-markers/.

Liu, P.-H. E., & Tsai, M.-K. (2013). Using augmented-reality-based mobile learning material in EFL English composition: An exploratory case study. *British Journal of Educational Technology, 44*(1), E1-E4. Doi:10.1111/j.1467-8535.2012.01302.x

Liu, Y., Holden, D., & Zheng, D. (2016). Analyzing students' language learning experience in an augmented reality mobile game: An exploration of an emergent learning environment. *Procedia: Social and Behavioral Sciences, 228*, 369-374. doi: 10.1016/j.sbspro.2016.07.055

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp.31-48). Cambridge, UK: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning*. Cambridge, UK: Cambridge University press.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1994). Augmented reality: A class of displays on the reality-virtuality continuum. *Proceedings of Telemanipulator and Telepresence Technologies, SPIE 2351* (pp. 282-292). Retrieved from http://etclab.mie.utoronto.ca/publication/1994/Milgram_Takemura_SPIE1994.pdf

Perry, B. (2015). Gamifying French language learning: A case study examining a quest-based, augmented reality mobile learning-tool. *Procedia - Social and Behavioral Sciences, 174,* 2308–2315.

Ramya, G. & Madhumathi, P. (2017). Adopting Augmented Reality for English language teaching and learning. *Language in India, 17*(7), 352-360.

Rankohi, S., & Waugh, L. (2013). Review and analysis of augmented reality literature for construction industry. *Visualization in Engineering, 1*, 1-18. doi: 10.1186/2213-7459-1-9

Richardson, D. (2016). Exploring the potential of a location based augmented reality game for language learning. *International Journal of Game-Based Learning, 6*(3), 34-49. doi:10.4018/IJGBL.2016070103

Safar, A. H., Al-Jafar, A. A., & Al-Yousefi, Z. H. (2017). The effectiveness of using augmented reality apps in teaching the English alphabet to kindergarten children: A case study in the State of Kuwait. *Eurasia Journal of Mathematics, Science & Technology Education, 13*(2), 417-440.

Salmon, J., & Nyhan, J. (2013). Augmented Reality potential and hype: Towards an evaluative framework in foreign language teaching. *The Journal of Language Teaching and Learning, 1*, 54-68.

Samset, E., Schmalstieg, D., Vander Sloten, J., Freudenthal, A., Declerck, J., Casciaro, S., … & Gersak, B. (2008). Augmented Reality in surgical procedures. In *Proceedings of SPIE 6806, Human Vision and Electronic Imaging XIII*. Retrieved

from https://www.researchgate.net/publication/44179341_Augmented_reality_in_surgical_procedures

Santos, M., Lübke, A., Taketomi, T., Yamamoto, G., Rodrigo, M., Sandor, C., & Kato, H. (2016). Augmented reality as multimedia: The case for situated vocabulary learning. *Research and Practice in Technology Enhanced Learning, 11*(4). Retrieved from https://link.springer.com/article/10.1186/s41039-016-0028-2

Siltanen, S. (2012). *Theory and applications of marker-based augmented reality.* VTT Technical Research Centre of Finland, Finland.

Sisodiaa, A., Bayerb, M., Townley-Smith, P., Nash, B., Little, J., Casarly, W., & Gupta, A. (2007). Advanced helmet mounted display (AHMD). *Proceedings of SPIE 6557: Head- and Helmet-Mounted Displays XII: Design and Applications*. Orlando, FL. doi:10.1117/12.723765

Slussareff, M., & Boháčková, P. (2016). Students as game designers vs. 'just' players: Comparison of two different approaches to location-based games implementation into school curricula. *Digital Education Review, 29*, 284–297.

Solak, E., & Cakir, R. (2015). Exploring the effect of materials designed with augmented reality on language learners' vocabulary learning. *The Journal of Educators Online, 13*(2), 50-72.

Von Itzstein G., Billinghurst M., Smith R., & Thomas B. (2017) Augmented Reality Entertainment: Taking Gaming Out of the Box. In N. Lee (Ed.), *Encyclopedia of Computer Graphics and Games*. Cham, Switzerland: Springer.

Yuen, S., Yaoyuneyong, G., & Johnson, E. (2011). Augmented reality: An overview and five directions for AR in education. J*ournal of Educational Technology Development and Exchange, 4*(1), 119-140.

Wasko, C. (2013). What teachers need to know about augmented reality enhanced learning environments. *TechTrends, 57*(4), 17-21.

White, R. V. (1971). Activating advanced ESL students: A problem and a solution. *TESOL Quarterly, 5*(1), 231--238. doi: 10.2307/3585695

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