A Multimodal Approach to Teaching Chinese as a Foreign Language (CFL) in the Digital World

Jinghe Han, Western Sydney University, Australia*

Qiaoyun Liu, Guangxi Art School, China

Ruiyan Sun, Northeast Forestry University, China

ABSTRACT

This research investigates a cohort of bilingual Chinese teachers' use of a multimodal approach in their Chinese as a foreign language (CFL) teaching. The data include the participants' CFL teaching practices and their reflections on multimodal teaching as recorded in their theses and a focus group discussion. The theoretical underpinning of this paper is based on Paivio's dual coding theory (DCT) and Kress's social-semiotic theory (SST). This research found that the teachers' multimodal use in CFL teaching demonstrated their research-informed committed endeavour in designing content specific activities to achieve pedagogical purposes, utilizing some digital technologies as a resource. The uniqueness of the written form of the Chinese language availed these teachers the opportunity to engage the multiple modes and advance their own understanding of multimodality as a concept. This research also found the teachers' meaning making through the multimodality did not always equate to that of their students' due to their social and cultural differences.

KEYWORDS

INTRODUCTION

The call to engage student language learning through a multimodal approach is not new (New London Group, 1996; Wess-Powell et al., 2016; Yi, 2014). The call to implement digital technologies to engage students in language learning is not new either (Rance-Roney, 2010; Smythe & Neufeld, 2010; Yi & Angay-Crowder, 2016). However, during the intervening years, it has been argued that the term "digital technology" and associated terms like "multimedia" have attached themselves interchangeably with multimodal and multimodality as Early et al. (2015, p. 454) alerted, "With

*Corresponding Author

emerging media and technologies, multimodality is often considered as digital, but 'multimodality is not synonymous with the digital.'"

This argument appears to have continued post-COVID 19, which saw the critical importance of digital technologies in the education of Australian children during lockdowns. In some cases, this tendency to blur the boundaries between digital technology, multimedia, and multimodality has created contention around "what is" multimodality and the forms of its practical realities in classrooms.

While second language education hinged on implementing digital technologies is receiving traction in the current literature (Jiang & Ren, 2021; Xu et al., 2021; Zhang et al., 2021), re-establishing the essence of multimodalities and multimodal pedagogy should be kept in sight. In defining what constitutes a mode, Lee et al. (2021) contend that:

Five main modes are identified as crucial for designing the meaning making outcome, ie. [sic] linguistic, visual, aural, spatial, and gestural, and any combination of the five elements is considered multimodal. (p. 66)

The points of reference are that multimodality refers to meaning accessed through these modes, whereas multimedia is the technology or digital platform (channel) that enables the multimodal "text" to be presented in an interesting, engaging, or interactive format. The multimodal phenomena exist (in terms of lesson planning and preparation) and have existed (across time) before digital technology and modern multimedia was created and enacted as the presentation method. The research reported in this article is situated within this premise. Multimodality is clearly distinguished from digital technologies, multimedia, and associated presentation formats, platforms, and terms.

In reviewing second language education literature containing the key terms "multimodal" or "multimodal approaches," the studies revealed two major trends – investigating teaching and learning strategies that involved a predominantly digital approach and a prevailing research methodology based on collecting data related to the participants' opinions. In Chinese as a second language (CSL) and English as a second language (ESL) research with a title or key focus on "multimodal," a literature search resulted in publications predominantly related to digital technology. Similarly, the term "multimodality" appears to be synonymous with digital technology and digital literacy. There appears to be a paucity of research focused on Chinese as a foreign language (CFL) practice and/or interventions implementing a multimodal approach in contrast to the trending of those with a focus on digital technologies.

Digital Technology-Assisted Language Teaching and Learning

Several studies were found on digital technology and multimodality; these considered multimodality as background knowledge rather than a research focus (Hafner, 2014; Smythe & Neufeld, 2010). For example, Hafner's (2014) research considered digital multimodal composing. Students were given the opportunity to collaborate with peers through mobile devices, apps, and online resources to compose digital videos, podcasts, and webpages (Hafner, 2014). However, it did not explore how the multiple modes were used in their compositions or what this meant for their multiliteracy achievements. Similar research was conducted on pre-service language teachers' perspectives of multimodal pedagogy (Li, 2020). In this research, students were provided with online digital tools and tasked with creating a multimodal project like digital books, flashcards, YouTube videos, and electronic posters using PowerPoint, Prezi, Storybird, and WeVideo. Rather than analyze the students' final products, the research gauged the participants' perceptions of the experience. It did not explore how each of the multiple modes was integrated into their project presentation. Another similar study explored the cognizance of college lecturers in respect of multimodal technologies and video games for language instruction (Dos Santos et al., 2020). This research reported the challenges created by the technology skills of individual teachers. These studies put technology as their research focus using "multimodal" and "multimodality" as descriptive words.

Other studies listed "multimodal approach" as the key component. Yet, they researched "multimedia." For example, Ganapathy and Seetharam (2016) explored the benefits of using a multimodal approach in meaning-making with secondary ESL students. Interestingly, this research gathered data from students concerning the amount of time they engaged with apps, mobile devices, and other devices in and out of school. It sourced the students' comfort levels when using online resources or mobile applications like iPads, laptops, or the internet. A study by Freyn and Fernandez (2017) researched students' perceptions and attitudes regarding their learning via a multimodal approach. The research involved teaching poetry through multimedia like video clips, television commercials, music videos, and film videos. It reported students' positive attitudes toward online video and audio resources in learning poetry. Actual poetry learning was not investigated or examined from a multimodal perspective. In addition, Jin (2018) investigated the function of WeChat as a communication platform for Mandarin Chinese learners. This research concluded that WeChat provided diverse affordances to Mandarin Chinese learners. Whether or how the multiple modes provided affordances to language learning was not in the scope of the study.

As reviewed, "multimodal" appears to be relegated to a descriptive term or by-product. It is often synonymous with DT rather than used as the key issue under investigation.

Visual- and Audio-Focused Multimodal Approach in Second Language Teaching

Studies researching a multimodal approach to language teaching and learning focused on the visual and audio modes rather than other modes (Grant et al., 2013; Jiang, 2016; Lan & Liao, 2018; Lin, 2009; Wilberschied & Berman, 2004). Creating digital learning environments and visual aids to support CFL learners' listening was explored. For example, Lan and Liao (2018) studied the impact of three-dimensional (3D) virtual and authentic contexts on the listening skills of language learners. They found that learners performed more ably when processing audio material if immersed in a virtual environment as compared to viewing pictures. It also found that visual aids facilitated by digital technologies lowered the learners' anxiety and motivated learning (Lan & Liao, 2018). Like this research, Grant et al. (2013) compared an online multiuser 3D virtual world simulation to a real-world classroom environment that was provided to students in their language learning. They concluded that a virtual context was advantageous in reducing students' stress. Wilberschied and Berman (2004) explored students' listening comprehension facilitated by a video clip only or video clips and screenshot pictures from the clips. This research found that screenshot pictures provide accurate clues to the video content and enhanced listening comprehension. All these studies demonstrate a consensus that incorporating visual and digital visual modes as contextual aids can improve learners' listening proficiency and/or reduces learning anxiety.

Due to the logographic features of Chinese characters[1] and its writing system, studies in teaching the written form of the language are abundant. Comparable to the research on CFL listening comprehension, most of these studies focused on the use of digital devices and visual aids (with only one paper incorporating visual and gestural modes). For example, McLaren and Bettinson (2016) studied the application of a software program in character-learning for beginning learners. They revealed that students held a positive attitude toward web-based digital devices to assist in character recognition and memorization. They also found that digital tools like an e-dictionary can promote students' engagement and stimulate motivation. Zhan and Cheng (2014) explored the use of the digital tools, animation, video, and PowerPoint in facilitating character learning. They concluded that a digital mode enhanced the quality and efficiency of character presentation. It helped students reinforce the connections between the form(s), sound(s), and meaning of characters. In a recent study, typewriting and handwriting of Chinese characters were examined in CFL students' learning (Zhang, 2021). Based on interviews with CFL learners and assessment tasks provided by CFL teachers, this research also confirmed that students performed to a higher standard when keyboarding compared to handwriting when creating characters. Wang and East (2020) argued that students should be encouraged to use digital devices to improve their typing/keyboarding skills for real-life communication. These studies prioritized investigations into the linguistic (written) mode by highlighting the facilitation role of digital technologies in achieving this aim.

Lu et al. (2013) investigated three modes of character teaching in relation to CFL learners' memorization. The first was the implementation of the traditional linguistic mode comprising pronunciation, semantic meaning, and the written form. The second was a linguistic-visual mode encompassing the three linguistic modes outlined. It also included an animated video depicting the characters' etymological form changes. The third mode was a linguistic-visual-gestural mode embracing the three traditional components, the animation video of character formation plus an extension to the animation video to include images of human body language, and gestures depicting the meaning and written form of the character. The results confirmed that the linguistic-visual-gestural mode enabled learners to outperform the other two modes in term of the students' total recall of the number of characters learned. As a second finding, the linguistic-visual mode proved to be more efficient for character memorization when compared to the traditional linguistic mode (Lu et al., 2013).

In summary, this literature revealed that digital technologies are a priority in current research into CFL and EFL in preference to a focus on a multimodal approach or pedagogy. It indicates that technology has been regarded as an innovation in and of itself, overshadowing a multimodal approach to foreign language teaching and learning. Critiques of this issue have been forthcoming. For example, Kubler (2018, p. 54) pointed out that "we must always keep our pedagogical goals foremost in mind and realize that technology is only the medium" and "it should always be the pedagogical goals that drive the technology and not the other way around." Similarly, Yu (2021, p. 4) concluded that "the essence of the multimodal language pedagogy may be how to appropriately design rather than what digital tools are involved."

Informed by the current literature, this article investigates multiple modes rather than a single mode or technology-focused teaching and learning in CFL education. Thus, the research question is: How has a multimodal approach been practiced by Chinese language teachers in Australian schools? Specifically, it explores how multiple modes have been or can be integrated into the teaching content to facilitate meaning-making and/or engage learners. Data from CFL teachers' multimodal practices were collected and examined to inform and contribute to the pedagogical development of multimodal teaching.

THEORETICAL FRAMEWORK: DUAL CODING THEORY AND A SOCIAL SEMIOTIC APPROACH

The theoretical underpinning of this article is based on Paivio's (1991) dual coding theory (DCT) and Kress's (2010) social-semiotic theory (SST). Multimodality is afforded consideration within both theoretical positions. DCT by Paivio (1978, 1986, 1990, 1991) provides a foundation for the development of multimodality in language learning. Based on this theory, memory and cognition are related to the sensory modalities and there is "an orthogonal relation between symbolic systems and specific sensorimotor systems" (Paivio, 1991, p. 257).

Dual coding in the theory refers to two independent but interconnected information processing systems – verbal and non-verbal. These two systems "symbolically represent the structural and functional properties of language and the non-linguistic world, respectively" (Paivio, 1991, p. 257). The verbal system comprises linguistic codes like the printed word, speech, and braille. The non-verbal system encapsulates the non-linguistic world of images, pictures, concrete objects, or events. The human mind creates separate verbal and non-verbal mental representations when processing incoming information; however, the two systems can and do interconnect with positive effects on recall. A representation in one system can activate an associated link or connection in the other. For example, if an image or the object a word represents is provided when learning written vocabulary, the learner's memory of the word can be supported (Williams, 2013).

While there is reference to a multi-faceted sensorimotor operation to this theory, the visual and audio modes take prominence. Reference to the other senses is less forthcoming. This tendency to emphasize visual imagery in preference to other modalities has led to the (mis)interpretation of DCT

as coding between the visual and verbal systems, between images and propositional representations (Paivio, 1991), or viewing visual and verbal as the two types of stimuli in learning (Murphy, 1990).

In contrast, SST extends DCT by considering meaning-making and sign-making in relation to social environments, agency, and cultural resources (Kress, 2010). This theory describes and analyzes all signs in all modes, as well as their interrelationships in meaning-making. SST perceives meaning-making as a multimodal enterprise. That is, the linguistic mode is "no longer [taken] as central and dominant, fully capable of expressing all meanings, but as one means among others for making meaning" (Kress, 2010, p. 79). Meaning-making can be enhanced through the linguistic (writing, speech), visual (image, picture, moving image), audial (music, soundtrack), spatial (layout), and/or gestural (gesture) modes.

SST promotes the concept of mode as being "socially shaped" (Kress, 2010, p. 79). The resources available to various modes within one society evolve according to the needs and demands of its social members. Thus, semiotic signs no longer represent "arbitrary relations of meaning and form" but rather are "newly made in social interaction" (Kress, 2010, p. 54). This theory emphasizes the agency and power of individuals who construct the meaning in contrast to the traditional semiotic theory that supports arbitrariness of the signifier-signified relationship (Kress, 2010).

SST also acknowledges that sign and meaning-making rely on "culturally given semiotic resource[s]" (Kress, 2010, p. 54). That is, one type of mode or example within a mode in one culture does not necessarily carry the same meaning in another. For example, in one society, a personal greeting or inappropriate behaviour may be managed through speech. It can be equally managed by hand or finger gesture in another society. This is equally true of the modes chosen to communicate understanding across cultural groups and within or across societies. It does not bode well to assume that "what is represented in speech in Culture A will also be represented by speech in Culture B" (Kress, 2010, p. 84). Similarly, a single mode, such as a gesture, that has an understood meaning in Culture A, may not be so in Culture B. Thus, according to SST, across all the available modes for communication, meaning-making has a social and cultural context that is informed by and informing the agency of its individual members.

In summary, aspects of both these theories are combined to inform the analysis of the data collected pertaining to the teachers' multimodal approaches implemented in their CFL teaching. DCT draws on cognition and learning theory, focusing on how multimodal information is processed or retrieved. It also explores how multiple modes can be combined to contribute to more efficient information recovery. SST is situated within a socio-constructivist perspective, focusing on meaning-making through multimodalities in social interactions (Kress, 2010). In this case information is not passively received or stored. Instead, multiple modes of meaning are made through agentic individuals who interact with socially constructed, culturally available resources (Kress, 2010, p. 54). The combination of the two theories enables an examination and identification of how specific modes are chosen for implementation. It also looks at their purposes and cultural embeddedness and how teachers use multiple modes as pedagogical tools to assist students' memory and cognition.

THE RESEARCH

This research studied three teacher-researchers who graduated with a Doctor of Philosophy or Master of Philosophy after completing an innovative post-graduate degree program, the ROSETE (Research-Oriented, School Engaged, Teacher Education), at an Australian university. The program implemented a teacher-as-researcher approach throughout the teacher education and Chinese language teaching components to their degrees. That is, as teacher-researchers, these higher degree research (HDR) students researched their own Chinese language teaching in Western Sydney schools as the basis for their thesis.

Each of the teacher-researchers planned and implemented Chinese language lessons with various year level students at local primary schools in the Western Sydney region. They had the agency to

nominate and research aspects of their Chinese language teaching, collect and analyze data within an action research framework, and produce a thesis. Some of these students chose to incorporate and study multimodal approaches in their Chinese language teaching (these are the participants in this research).

After searching the university's theses database, one criterion for participation was enacted – the thesis title needed to include "multimodal approach" or specific modes like "visual approach." Three theses satisfied this criterion. The authors/teacher-researchers were then invited to engage in a focus group interview as the second criterion to be eligible to participate. All three consented to participate. Thus, the research collected two sources of data. First, it collected the participants' CFL teaching practices. Second, it collected their reflections as recorded in the evidentiary chapters of their thesis and a focus group discussion (conducted for 90 minutes and audio taped).

FINDINGS

Data analysis revealed that multiple modes were rudimentary components within their teaching and learning activities. These were often achieved through the implementation of digital technologies like PowerPoint slides, smart boards, YouTube videos, and interactive apps. Four key findings were identified through the data analysis. First, these teachers demonstrated a research-informed understanding of multimodality and multimodal teaching. Second, they were committed to designing multimodal activities to achieve pedagogical purposes. In practice, modes were combined in a variety of ways, purposefully linked to the intended teaching and learning content, and with consideration given to their learners' cognitive, emotional, and behavioral engagement. Third, this research identified that the uniqueness of the written form of the Chinese language – the Hànzì – availed these teachers the opportunity to engage the visual mode in a way that advanced their practice of multimodality. Different from alphabetic languages, the written form of Hànzì characters has an inner connection and, therefore, allows an extended conceptual understanding of the visual mode. Last, and in relation to the analysis of data revealing the third finding, the teachers' meaning-making through implementing multiple modes did not always equate to that of their students'. Examples are provided throughout the findings in this article's sub-sections. It is argued that this phenomenon relates to the divergence in their social and cultural backgrounds.

Perceptions of "Multimodality" and Research-Informed Teaching Designs

Each of the participants in this research elected to explore Chinese language teaching through a multimodal sapproach as the focus for their higher degree research study. The teacher-researchers' understanding of multimodality and the design of multimodal teaching was gauged through their review of the relevant literature needed to inform their research studies. This was recorded during the focus group discussion:

- My initial interest in multimodal ways of teaching was from my research study. At the beginning stage, I read a lot around learning and cognition. One point I learned was that the human brain processes visual information many times faster compared to processing text. I then thought about designing my lessons through engaging more visual stuff and other modes. (Hu)
- Yeah, I agree. I wasn't clear about this topic either at the beginning. I was thinking if you make the presentation "colorful" in front of these young learners, you are more likely to attract their interest. Reading others' work definitely made me more confident in this approach. I think multimodal teaching or communication is always the case. It will be useful to find how different modes can effectively be combined to achieve good learning experiences for students. (Fen)
- Yeah. Young learners learn through involving all the senses. If you teach them the word "sweet" and you let them taste something sweet, that helps them understand what sweet means instead of only showing them how to say it or write it. (Dai)

- Exactly! A multimodal approach is not something new. I could imagine before language was created, early people would use body language to express themselves and combined this with the sounds they could make. I taught my students through the linguistic system. I also used visual and other modes to help the learning. (Hu)

This conversation excerpt reveals that the teacher-researchers developed a research-based understanding of multimodality teaching grounded in their reading of other scholars' work in this area. They expressed opinions that a multimodal approach to language teaching rather than purely a linguistic mode had the potential to engage students in successful language learning. The key issue was not whether multimodality should be used in CFL teaching, but rather how to combine the various modes to achieve the intended pedagogical purposes.

Content-Related Integration of Multiple Modes

Data revealed that digital technology was utilized across the three participants' teaching as a "default" channel to facilitate the integration of a multimodal approach. PowerPoint, YouTube videos, and smartboards were the most frequently implemented digital resources to facilitate the combination of modes within their CFL lessons.

However, data also confirmed that the written and oral linguistic modes provided the foundation for CFL teaching. This method included their strategies of modelling the verbal (pronunciation) and written (Hànzì) language forms with associated practice and assessment activities. Activities included in the lesson designs were "coloring in," matching characters and pinyin, and identifying the character through listening to its pronunciation. To achieve the planned linguistic outcome, the teacher-researchers supplemented the learning of the linguistic forms of the language (vocabulary, pronunciation, and Hànzì) with resource-rich lessons that combine visual, audial, gestural, and spatial modes. This is often combined as a digital delivery, engaging students with content to scaffold their meaning-making, understanding, and memorization. Table 1 lists and summarizes the range of modes and their combinations implemented by the teacher-researchers in their CFL content-related activities.

For the activities implementing a visual mode as the dominant form, the data revealed the use of video animation to set the lesson context. A physical object, image, and/or picture represents the Chinese characters. Then, characters can be visualized by connecting to their meaning. This mode was often integrated into the lesson content with the objectives to teach recognition and memorization of the written form and concrete words. Activities described in the data included matching the written words with pictures or images, deducing the meaning of words based on and carefully examining its written form, and practising Hànzì writing by tracing the written form as presented by an animation app.

Utilizing an audio-focused mode was identified in the data when the teacher-researchers planned to teach pronunciation and onomatopoeia words. The teacher-researchers incorporated rhythm, particularly with music, to facilitate pronunciation, make connections between a word meaning and the sound, and identifying the sound components in the written form of some Chinese words. For example, they designed "listen and fill-in the lyrics" and "chant the word" activities, introduced the onomatopoeia words that represent sounds produced by some animals and birds, and explained the sound radicals in characters that direct the pronunciation.

Using gestures and body movements was the mode implemented when the teacher-researchers' lessons aimed to teach tones, action verbs, and memorization of the stroke order and shape when writing Chinese characters. When teaching and practicing pronunciation tones, the teacher-researchers' data indicated they incorporated body movements, including limbs and head, to act as hooks for the young students to associate with the correct tone. Similarly, physically demonstrating action verbs and encouraging the students to do likewise allowed the gesture/word association to assist learning. An example included miming the action for "hit" while verbalizing "dǎ" ( ). Conversely, the teacher-researchers introduced a new Chinese word and approximated its meaning with a relevant gesture or action. They then invited the students to guess its meaning. Another example of incorporating gestures

Table 1.
CFL Teachers' multimodal practice

Mode	Examples of Mode Combinations	Content-Related Multimodal Activities
Visual Dominant (digital resources)	● Using video animation as context or facilitator ● Teaching the written form of word with image and picture as aids ● Visualizing the written form of words connecting to meaning	Teaching the written form Teaching concrete words ● Meaning-match written words to pictures or images ● Work out the meaning of words based on its written form ● Trace the written form through animation app
Audio-Focused (digital resources)	● Using rhythm to facilitate pronunciation ● Making the connection between word meaning and sound ● Using sound radicals to direct pronunciation	Teaching pronunciation Teaching onomatopoeia words ● Listen and fill-in the lyrics ● Chant the words (lyrics)
Gestural Featured (non-digital resources)	● Linking gesture to word meaning ● Acting action verbs ● Body movement to accompany pronunciation tone	Teaching action words Teaching written form of words ● Finger-dance strokes of words in the air ● Move limbs and head while practicing tone ● Guess word meaning by watching body gesture
Spatial-Related (digital resources)	● Deconstructing and reconstructing strokes and radicals (character parts) on a smartboard ● Using compound characters to explain the structure of the written form of the language	Teaching strokes and structure of characters ● Help students understand the orthography difference between English words and Chinese characters ● Use apps to understand spatial structure of written characters
Linguistic-Only Written and Oral (digital resources)	● Combining written and oral	Modelling written and verbal Testing written and verbal ● Color in ● Dictation (provide pronunciation for students to give written form) ● Match characters and pinyin ● Listen and recognize words

and/or actions as a mode was when the teacher-researchers encouraged the students to participate in "shū kòng" (translated: finger-writing in the air), a finger dance to practice Hànzì stroke formation.

The teaching of simple and compound Hànzì afforded the teacher-researchers' the opportunity to assist learning by incorporating aspects of spatial modality. Digital technology also assisted with these activities. The teacher-researchers used their digital expertise at the planning stage to navigate the internet to locate relevant information. It was also used to incorporate an app to assist students to understand the spatial structure of written characters. For example, each teacher-researcher demonstrated their searches for relevant teaching resources on the internet by quoting hyperlinks in their thesis. When teaching compound characters, the teacher-researchers deconstructed and reconstructed the strokes and radicals (character parts). Then, they drew the students' attention to spatial awareness through the spatial/directional cues of left to right, top to bottom, left to middle and right, top to middle, and bottom characters. Comparisons were made between the spatial structure of the Chinese language and the left-to-right letter by letter linear structure of English words.

Within four of the five modalities identified in this research, digital technology was utilized by the teacher-researchers to facilitate the delivery of various lesson content. CFL teaching was identified that used visual, audio, spatial, and linguistic-only modes as examples of lessons with digital technology. Those lessons drawing on physical body movements, gestures, and miming were action based. They did not involve digital technology.

To summarize the second finding in this research, the teacher-researchers demonstrated their commitment to implementing a multimodal approach in their teaching of CFL. They displayed their belief that the approach offered opportunities for content to engage students cognitively, emotionally, and behaviourally.

Multimodality as a Pedagogy to Assist Chinese Language Learning

The previous section provided a summary of the range of modes implemented by the teacher-researchers in their CFL teaching with young learners. This section presents the data informing the findings that a multimodal approach has the potential for successfully engaging students in CFL teaching and the teaching of Chinese characters, in particular. In addition, it found that several data vignettes attributed the meaning of "naturally" to some Hànzì by the teacher-researchers with Chinese background. However, it did not align with the meaning sensed by the local Australian students. Examples and explanations are provided throughout the following sub-sections.

Pictographic Features of Chinese and Visual Mode Teaching

This research reveals the uniqueness of the written form of Chinese language. It enabled these teacher-researchers to capitalize on the use of the visual mode. Acknowledging the pictographic and indicative features of the Chinese language, the connection between the written form of characters and their meanings was incorporated into their teaching.

In Fen's reflections on her CFL teaching, she accredited this positive connection as her "first" method when teaching the written form of the language. For example, when she taught , she stretched the square into a circle and changed the horizontal line in the middle into a dot point. At the same time, she provided an image of the sun and asked students to suggest a possible meaning. The students "sort of" agreed with her when she explained that the written form of many characters was initially based on the shape of what they represented. However, she noted, they needed to conform to the convention to fit each character into an imagined square. This visual method cognitively assisted students to look for cues for meaning-making from within the Hànzì form itself, thereby contributing to memorization.

However, the teacher-researchers also alerted that this visual method of explaining the Hànzì character and its meaning from the Chinese perspective did not always strike the same chord with the Australian students. Fen pointed out that "the clue does not always give them the right hint for the meaning making." In the focus group discussion, Hu acknowledged the problem too. She once taught the character (tián: field) and invited the students to suggest its meaning given that as a pictograph, its meaning is represented by its written form. She recalled that "one boy said it means 'hospital' as he sees the outside box as a fence and the cross inside indicates the Red Cross." She continued, noting that "another student said it means chocolate because it looks like a block of chocolate and another said it means window with four pieces of glass" (Hu). Once the students were familiar with Hu's explanation for (tián: field), she continued to extend the vocabulary by introducing a new (but related) character (nán: male). She recounted her explanation to the students based on its etymology: "The upper part ( ) means field or farm and the lower part (lì:) is a farming tool." She then asked the students to suggest a meaning for , providing the additional prompt: "Who does the farmwork in a family?" The students were non-responsive and looked confused.

Dai shared a similar experience when she attempted to teach the character (bǐ: pen). She described her explanation to the children and her reflection. "There are two components in this character," she said. "The upper part represents bamboo, and the lower part means animal fur. The students looked confused as they could not see any connection between pen and the plant bamboo and animal fur."

These data sound a warning. When CFL teaching with young children, the meaning-making through multiple modes, in this instance visual and oral modes, cannot be assumed as being synchronized between CFL teachers and learners. With character teaching, the content and meaning are context specific. Learning the characters based on the "stories" represented therein exposes a potential disconnect between teacher and learner based on personal histories or social and cultural differences.

Audio-Focused Mode in CFL Teaching

Data in this research also provided evidence of the successful use of an audio-focused mode with respect to teaching Chinese pronunciation and sound-meaning relationships. These teacher-researchers

used songs in association with the sound radicals in characters to assist the improvement of students' accuracy and fluency in pronunciation and tones.

The narrative by Dai recounted a vocabulary revision lesson based on a Chinese folk song (Kāngdìng Love Song), the lyrics of which included the previously learned vocabulary. Her reflections on that lesson were: "the rhythm of the song is catchy … after three times repeating the song, the whole class could sing along."

When teaching compound characters with sound radicals, Fen indicated she would initially deconstruct the word into parts. Then, she would identify the part that indicates the character's pronunciation. She would list other compound characters with the same "sound radicals," asking the students to identify these sound indicators and attempt the pronunciation of the newly formed compound. She commented that the students thought critically to make the connection. They became quite excited, saying "Aha! So many Chinese words share the same pronunciation!"

Hu's example of incorporating an audio mode in CFL teaching was with onomatopoeia words. Data pertaining to one of her teaching vignettes outlined how she compared the sounds made by the animals and birds in both English and Chinese (cat: meow and miāo; duck: quack quack and guā guā; bird: chirp chirp and jiū jiū). Her data reported that students were quite excited about learning this category of words where the pronunciation and meaning aligned. While acknowledging a slight difference, the students identified that these English and Chinese onomatopoeia words were both created based on mimicking the sounds made by animals and birds in nature.

While recounting successful episodes of implementing this mode in CFL teaching and learning, the teacher-researchers also acknowledged its limitations. For example, Dai and Fen observed that using music (lyrics and rhythm) was useful for students to practice pronunciation of learned vocabulary and gain a sense of the flow of language. However, it did not assist with students acquiring the correct tones. They also noticed that some songs with a slow tempo were less attractive to the young learners. Some songs contained lyrics that were unlearned vocabulary. Thus, "the effect was thin" (Fen).

Gesturing and Miming

Gesturing or miming the actions associated with the meaning of vocabulary incorporates a physicality that engages students, particularly younger CFL learners. An example from Fen's data related to the teaching of (xiào: smile). Her data noted that "I placed my hands on each side of my cheeks and put a dramatic smile on my face." This quick and easy gesturing captured the children's attention as they enjoy exaggeration and physical movement.

Dai said, "When I taught (d : hit), I would do the 'hit.' When I taught (chī: eat), I would do a biting and chewing action. I asked them to follow me. It is like 'learning by doing' or a 'totally physical response' method. They were engaged."

Hu introduced a novel action to assist students to memorize character writing and stroke formation. She demonstrated "shū kōng," whereby the index finger appeared to dance in the air while practicing the writing strokes. Although she admitted that shū kōng was neither creative nor interesting, it provided an alternative method for students to practice newly learned characters.

All three teacher-researchers indicated that they relied on body movement to facilitate the teaching of tones. Data from each participant confirmed their opinion that tones were difficult for Australian students to learn. They also noted that when practicing the four tones, engaging body movements (particularly the limbs, jaw, and head) in association with a tone was worthwhile. Fen reflected on this approach: "You can tell that students enjoy this activity."

Although the research data pointed to engagement and excitement when activities included gesture and mime, there was also reference to challenges when this mode was implemented. The three teachers admitted that, in general, younger learners (stages one and two) engaged with activities incorporating the gesture mode more willingly than older students. This was reflected in Dai's comment: "I could see my Year Five students sometimes felt embarrassed. My lower class simply followed without hesitation."

For Fen, the experience was slightly different. When teaching third tone to stage two students, she reflected that the activity was a success until she invited individual students to perform in front of the class. "Only a few students were willing to add body movement," she recalled. "Most of them just wanted to read with me instead of performing." Fen also reported "chaos" during a lesson when a finger/hand gesture had a different connotation for the students. In one lesson, she used gesture to teach numbers. "When they saw my gesture for the number eight – an open thumb (thumb up) and index finger pointing ahead – some boys apparently thought it meant 'shooting.' They started to shoot at others, which caused me some management problems."

The Use of Spatial Mode in Teaching Characters

The data revealed that a spatial modality was mainly employed in teaching or practicing the written form of the language/orthography. The teacher-researchers identified that when teaching students how to write characters, there was confusion around the writing/stroke sequence, particularly for compound words. They noted that students regarded characters as pictures rather than the written form of the language. Consequently, they often treated character writing as drawing. To counter this issue, all three teachers introduced a writing procedure that embraced the spatial mode.

Fen approached this concern by designing a lesson to showcase the differences in writing procedures between English words and Chinese characters. This involved highlighting the various spatial structures for writing simple and compound Chinese characters (strokes in order of left-to-right, top-down, and outside-in structures). Her explanation to students was: "For compound characters containing left and right parts, the writing starts from left and then moves to the right. For those containing top and bottom parts, the writing starts from the top and sequences down to the bottom. For those with one part within another, the strokes start from the outside and move to the center." She further recounted her explanation for writing simple characters: "The strokes must start at the top left and finish at the bottom right." According to Fen's data, this lesson was meaningful for students' CFL learning as she recorded observing the students' "Aha!" moment.

In Dai's CFL teaching, data revealed that students did not demonstrate a sense of space when practicing character writing. They were unaware of the need to have consistency in size when writing. Dai's found that the students wrote compound characters much larger than simple characters. Her data outlined a lesson devoted to foregrounding the differences between writing Chinese characters and English words. She explained and demonstrated to the students: "Each character needs to fit into one imagined square. You cannot use two or more squares/blocks for one character and one for another even though some of them look crowded in one square." The students were apprised that an English word with many letters can be written with as much space as needed; however, a Chinese character, whether containing many or a few strokes, must only use the same unit of space.

When experiencing similar issues with students' character writing, Hu incorporated digital technology to familiarize the students with the use of space and the proportion of the components in each character. Hu's lesson featured an interactive app on the smartboard. It allowed the strokes and radicals comprising simple and compound characters to be pulled apart and re-positioned. She recalled: "That was fun for the students. They always enjoyed jigsaw puzzle activity."

Each of the three teacher-researchers experienced a similar challenge with their students' writing of Chinese characters. They managed the challenge by drawing the students' attention to character writing through a spatial mode with verbal explanations, visual demonstrations, or digital technology.

An approach utilizing multimodalities has the prospect of aligning with the learning styles of more students to support their understanding and memorization, engage their interests and motivation for successful CFL learning.

DISCUSSION

As noted in the literature review, research into multimodal approaches to teaching have been dominated by research into digital technology (Li, 2020). This includes the use of online video and audio resources (Freyn & Fernandez, 2017), implementing games (Dos Santos et al., 2020), and using digital devices in teaching (McLaren & Bettinson, 2016; Wang & East, 2020; Zhan & Cheng, 2014). In other literature, multimodal teaching has been criticized for its singular focus on the digital tools incorporated into teaching and student satisfaction level with these. It often neglects to focus on teaching and learning designs (Yu, 2021).

This research explored how the participating teacher-researchers integrated a multimodal approach into their CFL teaching. Digital technology was used as a medium in many teaching episodes. However, the data collected provided evidence of multimodal teaching designs incorporating the visual, audial, gestural, spatial, and linguistic-only modes. Their data also indicated that the selection of the most appropriate modes to achieve educational purposes was influenced by their own research and teaching experience, the specific CFL content to be taught, and the needs of the young students.

A multimodal approach to teaching and learning, as indicated in other literature, has often foregrounded the visual and audio modes like visual resources to assist with character memorization (Lu, Hallman, & Black, 2013) and visual aids to assist CFL learners' listening skills (Grant et al., 2013; Jiang, 2016; Lan & Liao, 2018; Lin, 2009; Wilberschied & Berman, 2004). This situation can be traced back to the 1990s, when most multimodality teaching focused on the visual mode to the detriment of other multiple sensorimotor modes (Murphy, 1990; Paivio, 1991). In this research, the participant teacher-researchers provided evidence of CFL teaching practice that incorporated a balanced use of the multiple modes. The record of their teaching practices and reflections provided evidence of their understanding of the relationship between symbolic systems and specific sensorimotor systems. These findings do not support the argument made by earlier researchers.

This research found that there was a concerted effort and commitment by the teacher-researchers to engage students as fully as possible. The linguistic-only mode was one approach; however, the visual, audial, spatial, and gestural modes were all explored and implemented in their CFL teaching. This indicates they had an advanced understanding and use of multimodality in CFL teaching practice, possibly due to their research experience and the pedagogical issues covered in the ROSETE program.

Based on the traditional semiotic theory, learning to be literate is achieved through the linguistic mode by which meaning is arbitrarily attributed to a word. According to social semiotics, the learning is through engaging linguistic and non-linguistic processes like gauging the meaning from an image, sound, gesture, or space. One key finding from this research is that the participant teacher-researchers were able to explore multiple modes within one linguistic element – the characters. They utilized the "sound radicals" in compound Hànzì as sound indicators to scaffold students' word pronunciations and explore a spatial mode within characters to assist students to make sense of space and proportion between radicals and strokes when practicing character writing. In addition, they took advantage of the characters' visual (pictographic, indicative, or ideographic) features in its written form to guide students to deduce and memorize the meanings represented.

In this instance, these teaching practices are insightful. They demonstrate the teacher-researchers' innovative understanding and practice of multimodality. It extends beyond the assumed parallel relationship between the five modes to exemplify a new interweaving of modalities to achieve specific educational purposes and goals.

In the Chinese literacy system, meaning indicators in Hànzì enable an argument that signs, linguistic or non-linguistic, are less arbitrary than proposed by structuralist semiotic linguists. This enables a further argument that Chinese is a language providing more affordances than English to the learners as the signifier itself often is a reminder linking to the signified. This research produced clear evidence in disagreement with the proposed arbitrary relationship between language and the meaning and signifier and the signified that has been claimed by semiotic theorists such as De Saussure

(1974). These teachers demonstrated their capacity to implement social semiotic practices, during which individual teachers' and students' agency was empowered (Kress, 2010).

However, within a particular mode, the meaning-making afforded to the teachers may not be identical to that of their students. A sign may mean "eight" for the teacher and "gun" for the students. It may mean "field" for the teacher and "block of chocolate" or "hospital" for the students. This aligns with Kress's (2010) argument that multimodality and meaning-making are culturally resourced and socially shaped. When language teachers with a Chinese background are teaching student cohorts with other historical, social, and cultural backgrounds, some mismatch in meaning-making through different modes can occur.

CONCLUSION

This research studied three Chinese teachers' use of a multimodal approach in their CFL teaching. Data revealed that the uniqueness of the written form of the Chinese language and the digital classroom environment availed these teachers the opportunity to engage a multimodal approach comprehensively in their teaching. This research also found the teachers' meaning-making through the multiple modes did not always equate to that of their students' due to their social and cultural differences. The limitation of this study is the sample size with data drawn from the CFL teaching of three participants. While the findings may not be representative, the insights offered can contribute to further and larger scale research on multimodality and CFL teaching.

FUNDING

CONFLICT OF INTEREST

The authors of this publication declare there is no conflict of interest.

REFERENCES

De Saussure, F. (1974). Course in general lingustics. In M. Gottdiener, K. Boklund-Lagopoulou, & A. P. Lagopoulos (Eds.), *Semiotics*. Sage Publications.

Dos Santos, L., Von Gillern, S., Lockwood, J., & Geluso, J. (2020). Digital mindsets. *European Journal of Applied Linguistics and TEFL*, *9*(1), 131–152.

Early, M., Kendrick, M., & Potts, D. (2015). Multimodality: Out from the margins of English language teaching. *TESOL Quarterly*, *49*(3), 447–622. doi:10.1002/tesq.246

Freyn, A., & Fernandez, S. (2017). Effects of a multimodal approach on ESL/EFL university students' attitudes towards poetry. *Journal of Education and Practice*, *8*(8), 80–83.

Ganapathy, M., & Seetharam, S. (2016). The effects of using multimodal approaches in meaning-making of 21st century literacy texts among ESL students in a private school in Malaysia. *Advances in Language and Literary Studies*, *7*(2), 143–155.

Grant, S., Huang, S.-Y., & Pasfield-Neofitou, S. (2013). Language learning in virtual worlds. *Journal of Virtual Worlds Research*, *6*(1), 1–8. doi:10.4101/jvwr.v6i1.7027

Hafner, C. (2014). Embedding digital literacies in English language teaching. *TESOL Quarterly*, *48*(4), 655–685. doi:10.1002/tesq.138

Huang, S.-Y. (2015). The intersection of multimodality and critical perspective. *Language Learning & Technology*, *19*(3), 21–37.

Jiang, L., & Ren, W. (2021). Digital multimodal composing in L2 learning. *Journal of Language, Identity, and Education*, *20*(3), 167–182. doi:10.1080/15348458.2020.1753192

Jiang, Y. (2016). Validity of multimodality in autonomous learning of listening and speaking. *Journal of Language Teaching and Research*, *7*(2), 352–357. doi:10.17507/jltr.0702.14

Jin, L. (2018). Digital affordances on WeChat: Learning Chinese as a second language. *Computer Assisted Language Learning*, *31*(1–2), 27–52. doi:10.1080/09588221.2017.1376687

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

Kubler, C. (2018). Developing course materials for technology mediated Chinese language learning. *Innovation in Language Learning and Teaching*, *12*(1), 47–55. doi:10.1080/17501229.2018.1418626

Lan, Y.-J., & Liao, C.-Y. (2018). The effects of 3D immersion on CSL students' listening comprehension. *Innovation in Language Learning and Teaching*, *12*(1), 35–46. doi:10.1080/17501229.2018.1418242

Lee, S.-Y., Lo, Y.-H., & Chin, T.-C. (2021). Practicing multiliteracies to enhance EFL learners' meaning making process and language development. *Computer Assisted Language Learning*, *34*(1–2), 66–91. doi:10.1080/09588221.2019.1614959

Li, M. (2020). Multimodal pedagogy in TESOL teacher education: Students' perspectives. *System*, *94*, 102337. doi:10.1016/j.system.2020.102337

Lin, Y. C. (2009). *The effects of visual aids and text types on listening comprehension* [Unpublished master's thesis]. National Taiwan Normal University.

Lu, M. P., Hallman, G., & Black, J. (2013). Chinese character learning: Using embodied animations in initial stages. *Journal of Technology and Chinese Language Teaching*, *4*(2), 1–24.

McLaren, A., & Bettinson, M. (2016). Digital tools for Chinese character acquisition and their impact on student motivation. In R. Moloney & H. L. Xu (Eds.), *Exploring innovative pedagogy in the teaching and learning of Chinese as a foreign language* (pp. 235–251). Springer. doi:10.1007/978-981-287-772-7_13

Murphy, E. (1990). The effects of tactile and kinesthetic learning in improving spelling performance of a special education student. *Reading Improvement*, *27*(Fall), 207–211.

New London Group. (1996). A pedagogy of multiliteracies. *Harvard Educational Review*, *66*(1), 60–92. doi:10.17763/haer.66.1.17370n67v22j160u

Paivio, A. (1978). A dual coding approach to perception and cognition. In J. Pick, H. L. Pick, & E. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 39–51). Erlbaum.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press. doi:10.1093/acp rof:oso/9780195066661.001.0001

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255–287. doi:10.1037/h0084295

Rance-Roney, J. (2010). Jump-starting language and schema for English-language learners. *Journal of Adolescent & Adult Literacy*, *53*(5), 386–395. doi:10.1598/JAAL.53.5.4

Smythe, S., & Neufeld, P. (2010). Podcast time: Negotiating digital literacies and communities of learning in a middle years ELL classroom. *Journal of Adolescent & Adult Literacy*, *53*(6), 488–496. doi:10.1598/JAAL.53.6.5

Wang, D., & East, M. (2020). Constructing an emergency Chinese curriculum during the pandemic: A New Zealand experience. *International Journal of Chinese Language Teaching*, *1*(1), 1–19. doi:10.46451/ijclt.2020.06.01

Wess-Powell, C., Kargin, T., & Wohlwend, K. (2016). Enriching and assessing young children's multimodal storytelling. *The Reading Teacher*, *70*(2), 167–178. doi:10.1002/trtr.1491

Wilberschied, L., & Berman, P. M. (2004). Effect of using photos from authentic video as advance organizers on listening comprehension in an FLES Chinese class. *Foreign Language Annals*, *37*(4), 534–540. doi:10.1111/j.1944-9720.2004.tb02420.x

Williams, Z. (2013). *The use of multimedia material in teaching Chinese as a second language and pedagogical implications.* https://scholarworks.umass.edu/theses/1016

Xu, Y., Jin, L., Deifell, E., & Angus, K. (2021). Chinese character instruction online: A technology acceptance perspective in emergency remote teaching. *System*, *100*, 102542. Advance online publication. doi:10.1016/j.system.2021.102542

Yi, Y. (2014). Possibilities and challenges of multimodal literacy practices in teaching and learning English as an additional language. *Language and Linguistics Compass*, *8*(4), 158–169. doi:10.1111/lnc3.12076

Yi, Y., & Angay-Crowder, T. (2016). Multimodal pedagogies for teacher education in TESOL. *TESOL Quarterly*, *50*(4), 988–998. Advance online publication. doi:10.1002/tesq.326

Yu, Z. (2021). Exploring the effectiveness of multimodal language pedagogy during the COVID-19 time. Research Square. doi:10.21203/rs.3.rs-551047/v1

Zhan, H., & Cheng, H. (2014). The role of technology in teaching and learning Chinese characters. *International Journal of Technology in Teaching and Learning*, *10*(2). https://commons.erau.edu/publication/1097

Zhang, M., Akoto, M., & Li, M. (2021). Digital multimodal composing in post-secondary l2 settings. *Computer Assisted Language Learning*, 1–28. doi:10.1080/09588221.2021.1942068

Zhang, P. (2021). Typing to replace handwriting. *Journal of Technology and Chinese Language Teaching*, *12*(2), 1–28. http://www.tclt.us/journal/2021v12n2/zhangn.pdf

ENDNOTE

[1] Chinese characters and Chinese writing or written form is used throughout this paper as being synonymous with Hànzì.

Jinghe Han is a professor in language education in the Centre of Educational Research, School of Education, Western Sydney University. Her academic and research interests include translanguaging and multilingual teachers' professional development, English Medium Instruction, and HDR pedagogy. She has published four research books in recent years respectively on English Medium Instruction, culture theorising, post-lingual language pedagogy and internationalising Higher Degree Research education. Jinghe is one of the highest performing doctoral supervisors at the Western. She has successfully supervised 55 HDR students to completion on time in the past ten years.

Qiaoyun Liu is a lecturer in the Guangxi Art School. Her research interests include culture-integrated arts education, vocational education, and language education.

Ruiyun Sun is an Associate Professor in the College of Foreign Languages, Northeast Forestry University in China. Her specialization is Applied Linguistics.