Original Research

Knowledge-Building in Classroom: A Multimodal Semantic Wave Model

Qian Zhao[1] (iD) and Zezhi Zheng[1]

Abstract

The generation of meaning in discourse is regarded as the process of knowledge-building, which can be analyzed by semantic waves. Combining the deep learning systems and polynomial fitting method, we present a method to compute multimodal semantic waves. With this method, four representative examples are discussed for the knowledge-building of Chinese, English, Mathematics and Science subjects in Chinese primary schools. From the multimodal semantic waves of these subjects, we can clearly observe the differences between the processes of knowledge-building. Besides, it is advantageous for the proposed method that the value of multimodal semantic waves at any given future time can be predicted by an nth-order polynomial, which may help to study practical classroom discourse.

Plain Language Summary

This work aims to explore the process of knowledge-building in classroom discourse through multimodal semantic waves. Based on the deep learning systems and polynomial fitting method, a computational method is presented to study multimodal semantic waves, which may provide reliable data for multimodal corpora. This method is implemented by extracting multimodal values from classroom videos and plotting the multimodal semantic wave by an nth-order polynomial. The limitation of this work is that the authors only investigate the multimodal semantic waves in the classrooms of primary schools in China.

Introduction

Semantic waves are the key to cumulative development by activating recontextualization of knowledge through time and space (Maton, 2013b). The formation of meaning in discourse is inseparable from knowledge-building, and it can be effectively applied to guide teachers to improve teaching skills (Brooke, 2020) and benefit students' multimodal communicative competence (Dai & Wu, 2021). The practice of education and training has been shifting from traditional knowledge transmission approaches to constructivist approaches of knowledge-building (Verhoeven & Graesser, 2008). Therefore, making semantic waves in the knowledge expressed in classroom discourse (or other practices) is of great significance for the knowledge-building (Maton, 2015; Matruglio et al., 2013).

The knowledge-building plays an important role in classroom teaching. Through multimodal semantic waves, the knowledge-building processes in different subjects can be easily discovered. In the past 10 years, a lot of works on knowledge-building have been presented by researchers. With the help of semantic waves, Oteíza et al. (2018) found that language resources in history classrooms can affect the knowledge-building. Xie (2021) thought that English teachers' monologs would be beneficial for students to learn foreign language knowledge.

[1]Xiamen University, China

Corresponding Author:
Qian Zhao, Department of Chinese Language and Literature, Xiamen University, No.422, Siming South Road, Xiamen, Fujian 361005, China.
Email: hanmerzhao@outlook.com

Jiménez et al. (2016) identified movements of semantic waves between scientific and everyday discourse. Besides, Macnaught et al. (2013)'s work addressed how teachers can be trained to enable cumulative knowledge-building. We refer readers to Tang et al. (2021), Vrikki et al. (2019), and Hipkiss and Varga (2018) to learn more about knowledge-building.

Numerical methods have been widely applied in computational linguistics, corpus linguistics and other fields (Biber & Jones, 2009; Gries, 2009; Johnson, 2011). However, as far as we know, the computational analysis of the knowledge-building with numerical methods is still in its infancy. In this work, we will design a method based on the deep learning systems and polynomial fitting method to computationally study multimodal semantic waves for Chinese primary schools, which allows us to analyze the differences more accurately in knowledge-building processes between different subjects. Compared with qualitative methods (e.g., Jackson, 2016; Maton & Chen, 2015), multimodal semantic waves in the proposed method are plotted by the polynomial fitting method, whose values at any given future time can be predicted by an nth-order polynomial.

Theoretical Background

Semantic Waves

Different types of knowledge are presented by different types of discourse (Bernstein & Solomon, 1999). Bernstein (2000) proposed the model of "discourses" and "knowledge structures," and classified discourses into two types: "vertical discourse" and "horizontal discourse." It is believed that "vertical discourse" refers to educational knowledge (such as natural science discourse and humanities discourse), and "horizontal discourse" refers to daily knowledge (such as greeting discourse and other discourse used in daily life) (Maton, 2013b). According to Bernstein's model, Maton and Chen (2020) presented the legitimation code theory (LCT) including five dimensions, where the semantic wave theory was based on the semantic dimension.

Since semantic wave theory can provide scientific qualitative analysis and guidance for classroom discourse, it has received extensive attention in teaching research. Kilpert and Shay (2013) used the semantic wave theory to look for context-independent learning in students' assessments in a journalism curriculum. Subsequently, Georgiou et al. (2014) indicated that the LCT and semantic wave theory offered a way of analyzing the organizing principles of knowledge practices and their effects on science education. In recent years, Cranwell and Whiteside (2020) compared the complexity of spoken-language explanations of the same chemical process within UK secondary (high school) and university contexts, and
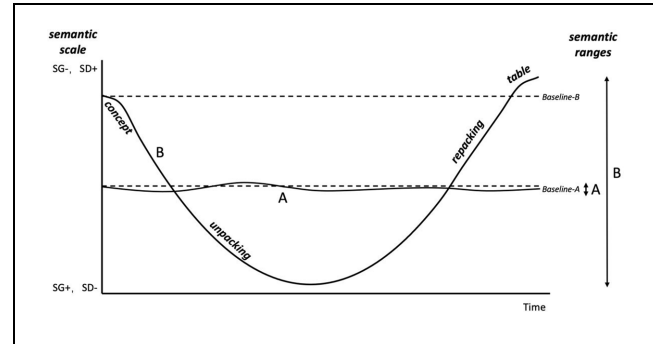

Figure 1. Semantic profiles and semantic ranges.

found that a larger-scale study of semantic waves could usefully inform specific-purposes language teaching. Besides, Barreto et al. (2021) draw on the LCT to understand the epistemic dimension of the higher education classroom discourse of a professor who is well evaluated by his students.

Semantic waves are the pulses of cumulative knowledge-building. The formation of semantic waves mainly depends on semantic gravity (SG) and semantic density (SD) (Wang, 2021). Duarte (2020) thought that teaching models based on semantic gravity and semantic density may become a pedagogical strategy to avoid the problem of segmental learning. The semantic gravity and semantic density of pedagogic discourse are investigated by Maton (2011), which may be relatively stronger (+) or weaker (−). Semantic gravity refers to "the degree to which meaning relates to its context." The stronger the semantic gravity (SG +), the more dependent the meaning is on its context; the weaker the semantic gravity (SG−), the less dependent the meaning is on its context. The semantic density refers to "the degree of condensation of meaning within socio-cultural practices." The stronger the semantic density (SD +), the more condensed the meaning is within practices; the weaker the semantic density (SD−), the less condensed the meaning is (Maton, 2013b).

It's much clear to demonstrate the process of knowledge-building by semantic gravity and semantic density. Figure 1 depicts two semantic waves according to Maton (2013b)'s work, where the X-axis represents the time, and the Y-axis represents the semantic scale of strengths of semantic gravity and semantic density. The baselines (dashed lines) in Figure 1 are determined by the semantic wave values at time zero. The semantic wave A and semantic wave B are drawn through different contexts, respectively. Compared with the baselines, the semantic wave A is relatively flat and the semantic wave B fluctuates greatly. The semantic wave B is obtained in an ideal teaching environment. It displays the specific process of the knowledge-building, including

four stages: "concept," "unpacking," "repacking" and "table." Specifically, classroom teaching starts with concepts. Then, with the strengthening of semantic gravity and the weakening of semantic density, knowledge points are continuously unpacked by teachers. Finally, the process of "repacking" is the stage in which teachers integrate and summarize knowledge points.

Semantic wave theory is a key to cumulative knowledge-building (Maton, 2013a, 2013b, 2014; Mouton, 2020). It has been used for teaching evaluations, designing teaching processes, discourse analysis, and so on (Wang, 2021). In view of semantic wave theory, teachers' expressions should be accurate and easy for students to understand in the daily teaching. Moreover, some elusive terms, symbols, theories, and abbreviations need to be explained patiently by teachers. When teaching new concepts, teachers should consciously reduce the semantic density and use multimodal discourse to illustrate important knowledge points. By controlling the strengths of semantic density and semantic gravity, students can complete learning tasks more effectively.

Multimodality Discourse

Multimodality refers to "communication in the widest sense, including gesture, oral performance, artistic, linguistic, digital, electronic, graphic and artefact-related" (Kress, 1997; Lytra, 2012; Pahl & Rowsell, 2006). Multimodal theory has been widely applied in textbook analysis and classroom teaching. O'Halloran (1998, 2015) researched the impact of different lexical-grammatical strategies in symbolic mathematics on the structural properties of oral discourse, as well as the multimodal phenomena in Mathematics classrooms. The interactive processes of multimodal teaching and learning were discussed in the work of Kress et al. (2006), where they pointed out that classroom teaching and learning should be completed by multiple modes and the language is no longer the only way of teaching. Therefore, in classroom teaching, new knowledge needs to be imparted through multimodal forms, such as teaching aids, multimedia equipment, colors, images, music, basic language symbols, teachers' body language, and so on (Ryan et al., 2010).

A multimodal corpus, defined as an annotated collection of coordinated content on communication channels including speech, gaze, hand gesture and body language (Knight, 2011), can be used in everyday language teaching (Frankenberg-Garcia, 2012). Botley et al. (2000), Sinclair (2004), and Kaltenböck and Mehlmauer-Larcher (2005) praised corpus-based

language teaching as the new revolution in language teaching. At present, a variety of multimodal discourse analysis methods have been proposed in the field of corpus research. For example, the linguistic annotation tool ELAN has been applied to analyze multimodal information (Brugman et al., 2004; Wittenburg et al., 2006), and Baldry (2006) presented that multimodal corpus authoring system allowed researchers to analyze multimodal film genres. In this paper, by videos from a corpus, we can plot multimodal semantic waves to guide teaching practice. The analytical expression of the multimodal semantic waves here is an nth-order polynomial, which may help to increase the utility and functionality of multimodal corpora.

Research Rationales and Questions

With the development of technology, primary school classrooms in China have become more diverse. Combined with multimedia devices such as slideshows and AR, teachers' behaviors, facial expressions, and intonations can better help students learn knowledge. According to the Curriculum Standards for Compulsory Education in China, teachers should be good at using multimodal means in classroom teaching (Ministry of Education of the People's Republic of China, 2022), which can influence knowledge-building directly. The semantic wave is regarded as a fusion of multimodal forms. It can assist teachers to scientifically design classroom discourse and promote the formation of knowledge structure and cumulative learning for students (Zhang & Qin, 2016). Therefore, it is very necessary to probe the semantic wave qualitatively and computationally to improve the teaching ability.

The semantic wave theory is crucial to study the information flow in classroom discourse (Maton, 2020). However, most of works on semantic waves still adopt qualitative analysis methods (e.g., Georgiou et al., 2014; Maton & Chen, 2015), which is a limitation for corpus-oriented data-oriented research. Moreover, few researchers have examined how semantic knowledge is used in discourse (Kintz & Wright, 2017). In classroom teaching, teachers thus cannot formulate appropriate teaching schemes based on the specific semantic wave values. In this study, we will try to combine machine learning methods to carry out computational research on multimodal semantic waves.

In view of the above discussions, this study aims to solve the following pressing questions:

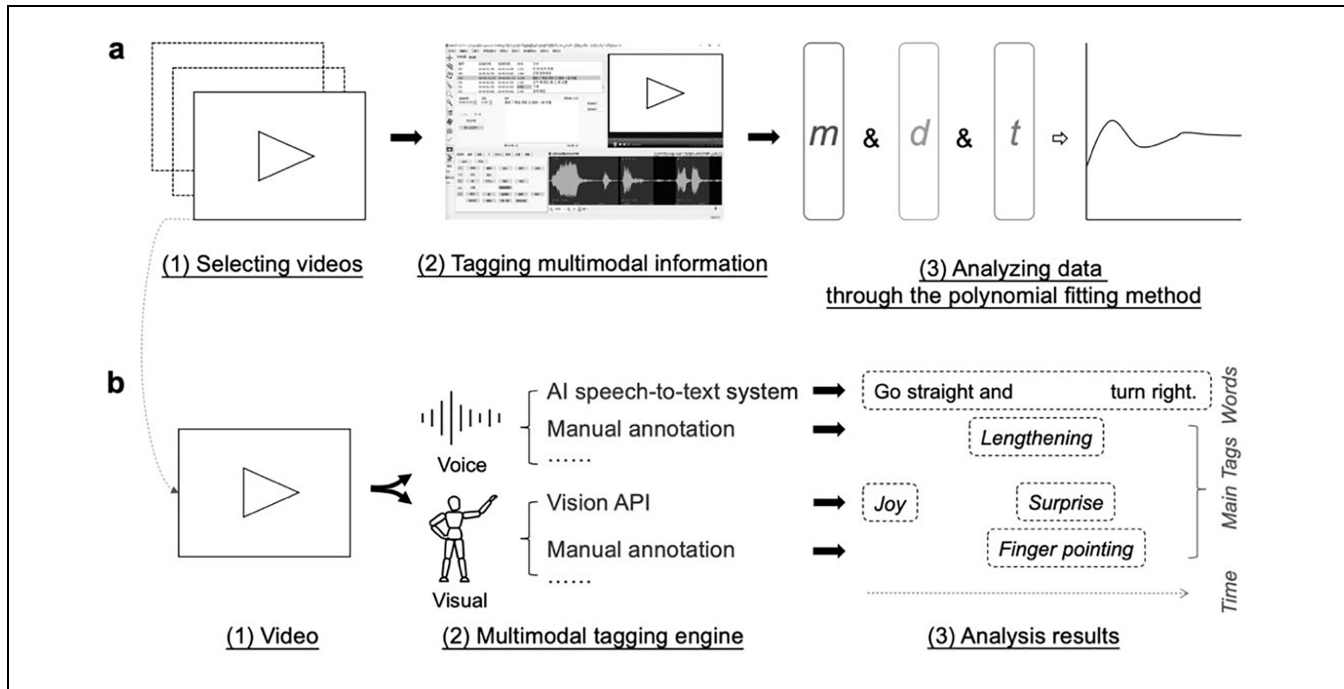1. How can we analyze multimodal semantic waves computationally?

Figure 2. Multimodal semantic wave analysis (a) main procedures of producing multimodal semantic waves, (b) flowchart of tagging multimodal information.

2. What are the processes of knowledge-building for different subjects in China by means of computational multimodal semantic waves?

Methods

Video Corpus and Ethics

This study analyzes multimodal semantic waves using a multimodal corpus that incorporates teaching videos of Chinese, English, Mathematics, Science, and other subjects in primary schools from various Chinese provinces and autonomous areas. These videos have been rated as excellent demonstrations by relevant education departments and published in China with the permission of teachers, students and schools. Therefore, they have reference value and can be used as examples for analysis. In the videos, the interactions between teachers and students in the classrooms outfitted with blackboard, electronic gadgets, teaching aids, etc. are captured on camera.

In 2021, data collection came to an end, and the video corpus totaled 3,000 min. All videos were analyzed and representative videos from four courses were selected as samples. These samples will be presented in this study. The phonetic transliteration process of the corpus is basically completed by the deep learning systems. The collection and use of the research data comply with the relevant laws and regulations of the Ministry of Education of the People s Republic of China and the Academic Committee

of Xiamen University. All the data processors pass the moral quality test and handle the potentially privacy-related information during the annotation process. The videos follow the principle of anonymity in the process of analysis and dissemination of research results, from which sensitive information have been removed. Facial data collected by computers are used for this study only and will not be shown in this paper. Moreover, in order to protect the minors, all computer image recognition resources involved in this study will not be used again.

Analyzing Data

The data in the videos are more complicated than text data. The processing and analysis of these data are customized according to the characteristics of the multimodal corpus within the research categories. Figure 2a shows the main procedures of producing multimodal semantic waves, including "selecting videos," "tagging multimodal information," and "analyzing data through the polynomial fitting method," where "m," "d," and "t" denote the "multimodal value," "dialog text," and "time," respectively.

Step 1: Selecting Videos. It is impossible to separate the study of theoretical language from that of actual classroom teaching. Complete classroom teaching videos can facilitate us to carry out a comprehensive analysis and to discover details that are difficult to observe or easily

overlooked in the teaching process (Kimura et al., 2018). As a result, referring to Yang et al.'s (2022) work, we decide to use representative classroom teaching videos (about 20–40 min) as research examples, where the content of the videos satisfies the following requirements:

1. Teachers speak with clarity and fluency.
2. Classes are interactive.
3. Teachers have the ability to describe relevant curriculum knowledge to students in the teaching process.

When conducting this research, we do not consider differences in students' familial environments, cognitive levels, or whether they have previously taken the same theme courses. Besides, we also do not investigate the gender, age, and experience of the teachers.

Step 2: Tagging Multimodal Information. Figure 2b clearly describes a flowchart of how to tag multimodal information, where the multimodal information from videos can be tagged by manual annotation and deep learning systems (e.g., AI speech-to-text system (Trivedi et al., 2018) and Vision API (Chen & Chen, 2017)). During the manual annotation stage, we apply the open-source software SubtitleEdit 3.5 (Olsson, 2022; Stauder & Ustaszewski, 2020) and XMUVT 1.0 (Zheng & Zhao, 2021) to manually tag specified behavioral modes and display voice-assisted waveforms. In addition, the multimodal discourse and physical behaviors of participants are transcribed according to works of Jefferson (2004), Mondada (2016), and Kääntä (2021).

In general, multimodal data in videos are massive and complex, and the scope of tags may vary by categories. Therefore, in order to save time and analyze data efficiently, we resort to the deep learning systems, where the speech transcription is completed by the AI speech-to-text system; the Vision API is used to improve the efficiency of tagging multimodal information. Finally, the main analytical results are obtained after implementing the manual annotation and deep learning systems.

As shown in Table 1, the multimodal tags are divided into five categories. The first category is "role," containing tags "teacher" and "student." Since teachers' props, tones, faces and body language are most perceived by students in classroom teaching, we classify the multimodal representations of teachers into four categories that are listed in Table 1 in detail. For example, when a teacher expresses the sentence "go straight and turn right," the main analysis results are as follows. Firstly, the AI speech-to-text system converts the speech into text. Then the Vision API tags the teacher's face as "joy" and "surprise." Lastly, two pieces of information (i.e., the teacher points his finger at the blackboard and

Table I. Multimodal Tags.

Category	Tag
Role	Teacher, student
Prop	Text, picture, animation, blackboard, AR, teaching aid, $\cdots$
Tone	Doubt, stress, lengthening, pause, $\cdots$
Face	Joy, sorrow, anger, surprise, $\cdots$
Body language	Nod, handclap, finger pointing, thumbs-up, jump, stamp, $\cdots$

lengthens the voice to say "and") are manually marked as "finger pointing" and "lengthening," respectively.

Step 3: Analyzing Data Through the Polynomial Fitting Method. In this step, a computational method based on the polynomial fitting method is designed to analyze multimodal information. After tagging the multimodal information in step 2, Equation 1 is applied to calculate the multimodal value

$$m = \sum_{i=1}^{\alpha} \frac{g_i f_i}{g} \qquad (1)$$

where g_i denotes the number of the ith multimodal information; f_i denotes the weight of the ith multimodal information; g denotes the total number of multimodal information corresponding to a word or phrase (Author, 2021).

By means of the polynomial fitting method, we are able to plot the multimodal semantic waves, which is finished by the function *polyfit* in MATLAB 2021b. The nth-order multimodal semantic wave based on polynomial curve fitting is analytically shown below:

$$w(t) = a_0 + a_1 t + \cdots + a_n t^n \qquad (2)$$

where $a_i (i = 0, 1, ..., n)$ is the coefficient of the nth-order multimodal semantic wave and w denotes the semantic wave value. From Equation 2, we can easily computationally study the trend of the semantic wave value w over the time t.

Results

Knowledge-building is a step-by-step process. Knowledge points need to be unpacked and repacked many times in the process of overall classroom teaching (Maton, 2013b), which is vividly illustrated by a multimodal semantic wave. In this section, we primarily test the performance of the nth-order multimodal semantic wave (or nth-order polynomial) that fits the multimodal values for dialog lines best in a least-squares sense.
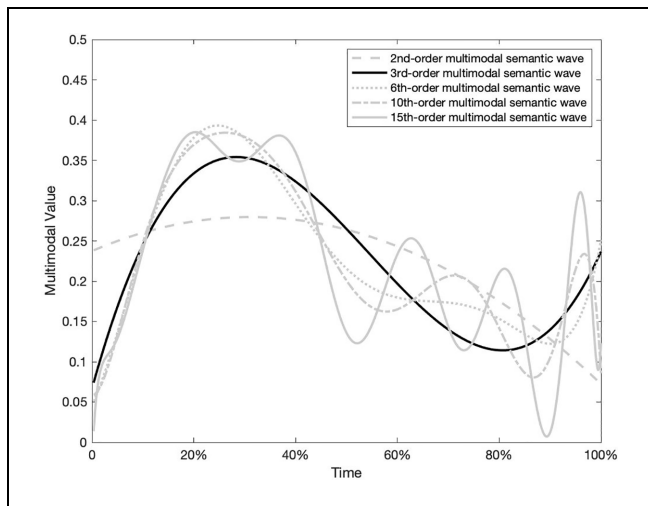
Figure 3. Multimodal semantic waves (the multimodal values are fitted by different multimodal semantic waves in a Mathematics classroom.).

Fitting Multimodal Values

Figure 3 shows the images of the 2nd-, 3rd-, 6th-, 10th-, and 15th-order multimodal semantic waves, where the X-axis represents the ratio of the elapsed time to the full video duration and the Y-axis represents the semantic wave values for the fourth grade Mathematics classroom. It can be clearly seen that the second-order multimodal semantic wave is relatively smooth, while the 15th-order multimodal semantic wave fluctuates severely. On the one hand, the more fluctuating the semantic wave curve is, the more helpful it is to discover the details of knowledge-building. On the other hand, the relatively smooth semantic wave curve makes it easier to identify the trend of knowledge-building.

Typically, the relative error between the semantic wave value and the real multimodal value decreases as the order of the fitted polynomial increases. However, in order to reduce the fluctuations of semantic waves and avoid the Runge phenomenon (Runge, 1901), the low-order polynomials are often used to fit multimodal values. Therefore, in the following discussion, we focus on the 3rd-order multimodal semantic waves.

Examples for a Computational Analysis of Multimodal Semantic Waves

For the purpose of conducting a computational analysis of multimodal semantic waves, four representative classroom videos from the fourth grade of primary schools are chosen as samples. Table 2 shows the relevant information for the sample videos, in which the English class is taught in English and the others in Chinese. Taking advantage of these videos of different subjects, we obtain the multimodal values and multimodal semantic waves shown in Figure 4.

Multimodal Semantic Wave in a Chinese Classroom. We first study a multimodal semantic wave in a Chinese classroom with the theme of "a beautiful new century."Figure 4a depicts the trend of the multimodal semantic wave over time. During the first 65% of the full video duration (see the time period A in Figure 4a), the semantic wave shows a downward trend corresponding to the teacher's introduction of the curriculum theme. Here, the teacher's tone and body language were varied. For example, the teacher applied gestures to guide students to "look here" and asked students to "think again" in a questioning tone. In the last 13 min (time period B) of the Chinese class, the teacher discussed with students and encouraged students to express their own thoughts correctly. During this time period, the semantic wave is essentially straight.

In primary education, the goal of Chinese classes is to develop students' understanding of beauty and improve their conversational skills. A variety of learning situations should be created in the Chinese classroom to stimulate students' curiosity, imagination and desire to learn. The knowledge-building in Chinese education is a cumulative process from less to more and from shallow to deep, and Chinese knowledge cannot be obtained by "understanding without teaching" (Su & Zhao, 2022). Therefore, teachers play an important role in classroom teaching. They should guide students in plain language and gently help students build knowledge in the classroom, which is reflected by the semantic wave curve in the time period A. There is a large amount of multimodal discourse used to guide students' learning. For example, in the video, the teacher resorted to slides with many pictures to show the changes of life in the new century, including food, clothing, housing, and transportation.

Multimodal Semantic Wave in an English Classroom. English teaching requires teachers not only to communicate with students in English, but also to help students complete the knowledge-building of the English subject. Aspects of language knowledge, cultural knowledge, language skills and learning strategies should be considered in the English classroom. Below we will study the multimodal semantic wave of an English class of about 30 min.

As shown in Figure 4b, the teacher was constantly "unpacking" language knowledge in the time period A. There were a lot of multimodal behaviors indicated by quantified semantic wave values. For example, the teacher asked students to "answer questions" through his gestures, raised his tone to emphasize grammar knowledge, and made the students understand the English

Table 2. Sample Video Information.

| Subject	Theme	Full video duration	Number of dialog lines	Total number of tags
Chinese	A beautiful new century (Read the text, connect with real life, and talk about the feeling of reading.)	00:35:05,800	668	1,183
English	Asking the way (Learn how to ask for directions in English.)	00:30:16,966	601	965
Mathematics	Mixed operations for addition and subtraction of decimals (Utilize mixed operations to solve simple practical problems.)	00:31:34,933	614	872
Science	Magical light bulbs (Learn about the internal structure and working of light bulbs.)	00:32:59,366	574	1,211
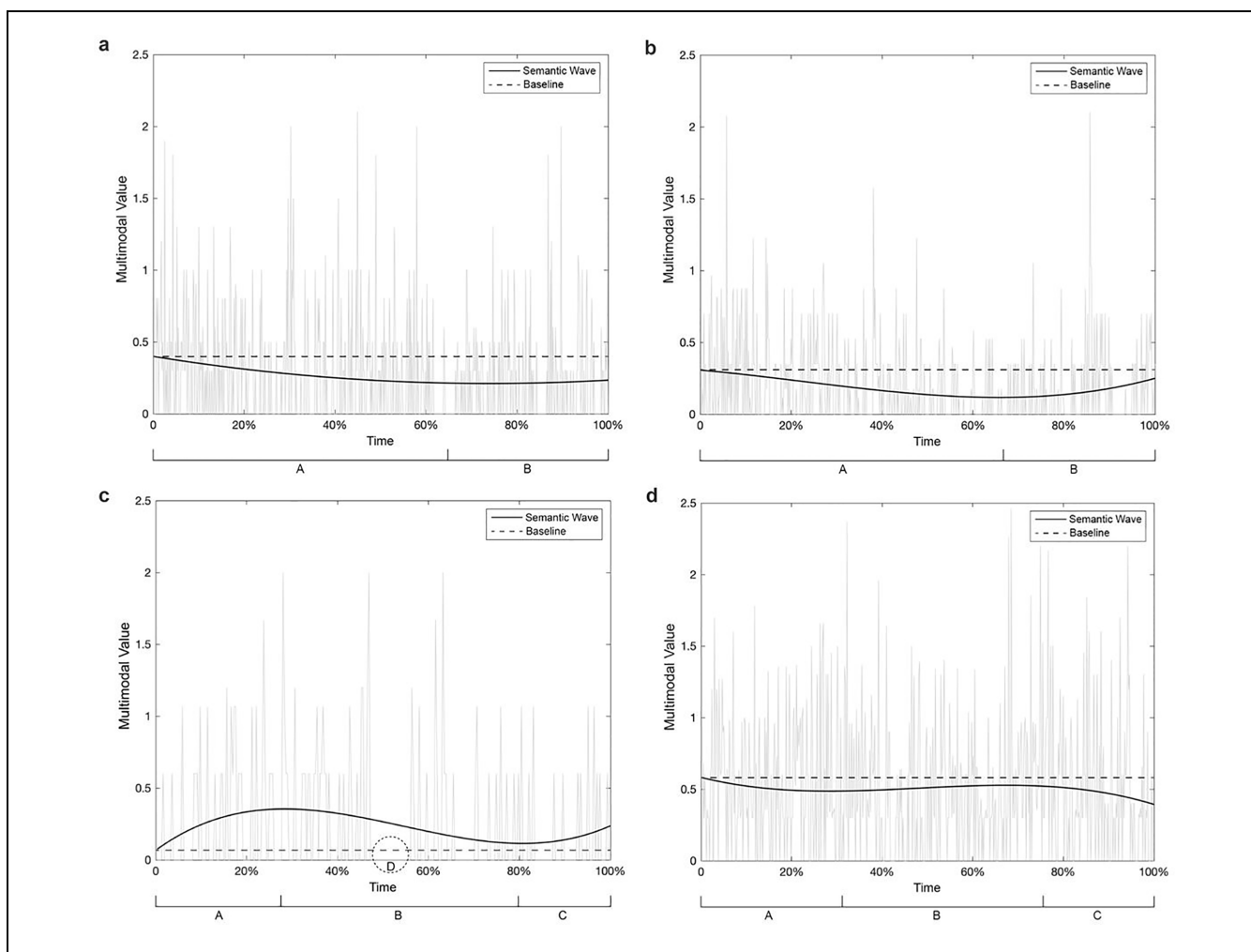
Figure 4. Examples of multimodal semantic waves (a) multimodal semantic wave obtained by a third-order polynomial for a Chinese classroom, (b) multimodal semantic wave in an English classroom, (c) multimodal semantic wave in a Mathematics classroom, where the multimodal values are zero in the time duration D, and (d) multimodal semantic wave in a Science classroom.

directional words through maps. Different from the semantic wave of the Chinese class, the semantic wave in Figure 4b demonstrates a upward trend in the time period B, which corresponds to the process of the teacher

"repacking" knowledge. In this process, the teacher first instructed the students to spell out the directional words with a puzzle, then played a children's song about way-finding on the electronic device, and finally requested the

students to draw a route map on the blackboard to judge if the students had mastered the language knowledge.

Multimodal Semantic Wave in a Mathematics Classroom. Because the mathematic knowledge is highly abstract, teaching it has traditionally been seen as a difficulty for teachers. Mathematics classroom teaching should focus on the hierarchy and diversity of mathematic knowledge and methods, and gradually expand and deepen the content of the curriculum. Besides, Mathematics teachers should be good at utilizing multimodal forms to improve their teaching level.

Compared with the multimodal semantic waves for the Chinese and English classrooms, the semantic wave for the Mathematics classroom is more complex. Figure 4c plots the multimodal semantic wave consisting of three parts. The first and third parts of the semantic wave are on the rise, and the corresponding durations account for about 50% of the full video duration. However, the second part of the semantic wave declines slowly in the time period B, where there is a time duration D with zero multimodal values.

Although less multimodal information was used in the Mathematics classroom for this example, the semantic wave in Figure 4c significantly shifts. During the time period A, the teacher presented "previously learned knowledge" through slides and rolled his fingers around his ears to inspire students to "think," while the corresponding semantic wave gradually rises. For the second part of the semantic wave, the teacher mainly explained some abstract concepts, where students discussed freely in the time period D with zero multimodal values. It can be easily found from Figure 4c that the open conversation among students is advantageous to unpack difficulties. Students can actively acquire knowledge by free discussion even though teachers are not directly involved in the knowledge-building. At the end of the course, the teacher helped students summarize the knowledge through handwritten cards, which causes the semantic wave to rise again (see the time period C).

Multimodal Semantic Wave in a Science Classroom. In Science classrooms, we advocate the implementation of diverse learning methods and encourage students to actively participate in the process of scientific inquiry. Teacher-student interactions should be strengthened in Science teaching. Meanwhile, students need to be guided to summarize, reflect, apply and transfer knowledge by multimodal forms.

In this example with the theme of "magical light bulbs," an experiment of interest to students was designed to spark students' curiosity in the Science classroom. In the time period A of Figure 4d, the teacher first showed the students a real light bulb and led the students to consider why the light bulb glowed. He then questioned students how to make a light bulb glow by connecting a battery with a wire in the time period B. During these two time periods, there are fewer knowledge points explained, so the semantic wave is relatively flat. A significant drop for the semantic wave can be clearly observed in the time period C, where the relevant principles were elaborated and applied to life practice.

Discussion

Through the analysis of the multimodal semantic waves, we can find that the knowledge-building in the classroom is closely related to the teaching goals and cultural background of the different subjects. Specifically, both the Chinese and English subjects focus on language learning, and their main teaching goal is to make students proficient in language knowledge. Hence the multimodal semantic wave curves of these two subjects are similar in Figure 4a and Figure 4b. However, unlike native Chinese subjects, English teachers also undertake the task of teaching common cultural knowledge to Chinese students unfamiliar with foreign cultures. In order to help students have a more comprehensive understanding of knowledge, teachers need to pay attention to the time period when the semantic wave rises in the English classroom.

One of the primary goals of Mathematics and Science classes is to develop students' ability to think abstractly. In classroom teaching, it is often difficult for teachers to explain abstract concepts in detail to primary students (Mulholland & Wallace, 1996). Thus, in Figure 4c and Figure 4d, the semantic waves fluctuate more frequently for these two subjects, where knowledge is built through a "repacking" process, an "unpacking" process, and an additional "repacking" or "unpacking" process. In view of the poor abstract thinking of primary students, it is a good choice to first carry out the "repackaging" process to introduce abstract knowledge in the Mathematics classroom, just like Maton (2013b)'s theory that concrete and simple meanings may offer a more engaging way into central focus of topic. The "repacking" process occurs during the time period when the semantic wave rises from a relatively low position. In this process, the construction of old and new knowledge complements each other. It is desirable that teachers start with simple old knowledge to teach abstract new knowledge. For example, in the time period A of Figure 4c, the Mathematics teacher reminded the students of the simple addition and subtraction operations through pictures, and then introduced the abstract concept of mixed operations. Instead, abstract knowledge can be explained at the beginning of the Science classroom ("unpacking" process). Students are able to

conduct hands-on experiments more effectively after mastering some necessary concepts and principles.

The process of knowledge-building is a process alternating between unpacking difficulties and repacking knowledge (Wang, 2021), in which the fluctuations of the multimodal semantic waves are induced by the superpositions of multimodal forms. In the computational analysis, we generally believe that the larger the multimodal values, the richer the multimodal forms. There are often relatively rich multimodal forms in the process of "unpacking" knowledge corresponding to a downward multimodal semantic wave curve. For example, in the Science classroom, the teacher used props, slides, gestures and other multimodal forms to show students the internal structure of a light bulb.

Poor multimodal forms are numerically manifested as relatively small multimodal values, which can be divided into following two cases. The first case is that in the process of imparting knowledge, teachers directly summarize knowledge and seldom use multimodal forms. This usually happens when knowledge is repackaged. For example, during the time period B in the English class, the teacher listed tables to directly summarize the English grammar. The second case is students' self-improvement in classroom activities. It can be found from the computational multimodal semantic wave in Figure 4c that the free discussion is the process of students "unpacking" knowledge on their own. Therefore, the self-learning without teachers explaining knowledge is very important for students' knowledge-building.

Multimodal semantic wave theory has been widely used to guide teaching practice (Blackie, 2014; Brooke, 2017; Mouton, 2020). With the help of semantic waves, teachers can easily analyze classroom discourse and rationally arrange teaching activities. For example, by referring to the semantic wave in the time period D of the Mathematics class (see Figure 4c), perhaps teachers can organize students to discuss freely in actual teaching to obtain a descending semantic wave curve. Moreover, given that the analytic formula of the computational semantic wave is an nth-order polynomial, the value of the semantic wave at any future time can be predicted by Equation 2, which is beneficial for teachers to design the teaching process. It is worth noting that the theoretical results for multimodal semantic waves cannot be replicated mechanically. Due to the flexibility of classroom teaching, teachers need to formulate appropriate teaching plans for different teaching environments.

Conclusion

Based on the deep learning systems and polynomial fitting method, the computational method has been presented to study multimodal semantic waves, which may provide reliable data for multimodal corpora. This method is implemented by extracting multimodal values from classroom videos and plotting the multimodal semantic wave by an nth-order polynomial. Multimodal semantic waves can be applied to describe the processes of the knowledge-building of different subjects and provide reference for practical classroom teaching, and the model is not limited by the actual language. In addition, according to the Equation 2, the trend of a semantic wave in the next time period can be predicted computationally. In this work, by analyzing the semantic waves of Chinese, English, Mathematics, and Science subjects in the fourth grade of primary schools in China, we easily find that the knowledge-building of different subjects is affected by various factors such as cultural background and teaching goals. Therefore, in the actual teaching, teachers should reflect and summarize their own multimodal behaviors by means of semantic waves and communicate with students frequently.

In the teaching process, multimodal semantic waves can provide teachers with a reasonable means to organize multimodal representations. Consequently, it is suggested that teachers should carry out teaching flexibly and complete the processes of unpacking difficulties and repacking knowledge according to the changing trend of multimodal semantic waves. Although the proposed method for computational analysis of multimodal semantic waves is applicable to different subjects, there are still some shortcomings here. First, according to the relevant privacy policy, we only investigate the multimodal semantic waves in the classrooms of primary schools in China, which is a limitation resulting from conducting research in a country with a specific and unique culture and language. To further test the performance of this method, we will next conduct some international collaborations to study classroom teaching in different regions. Second, the deep learning systems have been introduced to quickly acquire multimodal values, but it still takes a long time to implement this computational method. In the future, we intend to explore more efficient numerical methods to handle mass data.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Ethics Statement

This research follows the relevant policies of the Academic Ethics Committee of the College of Humanities of Xiamen University (https://rwxy.xmu.edu.cn).

ORCID iD

Qian Zhao 🆔 https://orcid.org/0000-0003-1760-249X

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Baldry, A. P. (2006). The role of multimodal concordancers in multimodal corpus linguistics. In T. D. Royce, & W. L. Bowcher (Eds.), *New directions in the analysis of multimodal discourse* (pp.173–193). Routledge.

Barreto, L. P., Rodrigues, A. A. D., de Oliveira, G. C. B., de Almeida, L. T. G., Felix, M. A. C., Silva, P. D. S., Quadros, A. L., Macedo, A. M., & Mortimer, E. F. (2021). The use of different translation devices to analyze knowledge-building in a university chemistry classroom. *Research in Science Education*, *51*, 135–152.

Bernstein, B. (2000). *Pedagogy, symbolic control, and identity: Theory, research, critique* (Vol. 5). Rowman & Littlefield.

Bernstein, B., & Solomon, J. (1999). Pedagogy, identity and the construction of a theory of symbolic control': Basil Bernstein questioned by Joseph Solomon. *British Journal of Sociology of Education*, *20*, 265–279.

Biber, D., & Jones, J. K. (2009). *Computational methods in corpus linguistics* (Vol. 2, pp. 1286–1304). De Gruyter Mouton.

Blackie, M. A. L. (2014). Creating semantic waves: Using legitimation code theory as a tool to aid the teaching of chemistry. *Chemistry Education Research and Practice*, *15*, 462–469. https://doi.org/10.1039/C4RP00147H

Botley, S., McEnery, A. M., McEnery, T., & Wilson, A. (2000). *Multilingual corpora in teaching and research* (Vol. 22). Rodopi.

Brooke, M. (2017). Using 'semantic waves' to guide students through the research process: From adopting a stance to sound cohesive academic writing. *Journal of the Scholarship of Teaching and Learning*, *7*, 37–66.

Brooke, M. (2020). Guiding teacher talk in the content and language integrated learning classroom using semantics from legitimation code theory. *Teaching in Higher Education*, *25*, 812–824.

Brugman, H., Russel, A., & Nijmegen, X. (2004). *Annotating multimedia/multi-modal resources with ELAN* [Conference session].

Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04) (pp. 2065–2068).

Chen, S. H., & Chen, Y. H. (2017). A content-based image retrieval method based on the Google Cloud Vision API and WordNet. In N. T. T. Nguyen, S. Tojo, L. M. M. Nguyen, & B. Trawiński (Eds.), *Intelligent information and database systems* (pp. 651–662). Springer International Publishing.

Cranwell, P. B., & Whiteside, K. L. (2020). Investigation into the semantic density and semantic gravity wave profile of teachers when discussing electrophilic aromatic substitution (SEAr). *Journal of Chemical Education*, *97*, 3540–3550.

Dai, Y., & Wu, Z. (2021). Developing multimodal communicative competence: Insights from "semantic gravity" and the Knowledge Process framework. *System*, *100*, 102561–102573. https://doi.org/10.1016/j.system.2021.102561

Duarte, E. G. (2020). Student-teachers' practicum as the scenario to build cumulative learning through semantic waving instructional models. *Open Journal of Social Sciences*, *08*, 77–88.

Frankenberg-Garcia, A.(2012). Integrating corpora with everyday language teaching. In J. Thomas, & A. Boulton (Eds.), *Input, process and product: Developments in teaching and language corpora* (pp. 36–53). Masarykova univerzita.

Georgiou, H., Maton, K., & Sharma, M. (2014). Recovering knowledge for science education research: Exploring the "Icarus Effect" in student work. *Canadian Journal of Science Mathematics and Technology Education*, *14*, 252–268. https://doi.org/10.1080/14926156.2014.935526

Gries, S. T. (2009). *What is corpus linguistics? Language and Linguistics Compass*, *3*, 1225–1241.

Hipkiss, A. M., & Varga, P. A. (2018). Spotlighting pedagogic metalanguage in reading to learn – How teachers build legitimate knowledge during tutorial sessions. *Linguistics and Education*, *47*, 93–104.

Jackson, F. (2016). Unraveling high school English literature pedagogic practices: A legitimation code theory analysis. *Language and Education*, *30*, 536–553.

Jefferson, G. (2004). Glossary of transcript symbols. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 24–31). John Benjamins Publishing Company.

Jiménez, J. P. C., Melo, G., Bacigalupo, F., & Manghi, D. (2016). Semantic waves in teacher-student interaction: Analysis of two natural science lessons in 6th grade. *Ciência and Educação*, *22*, 335–350.

Johnson, K. (2011). *Quantitative methods in linguistics*. John Wiley & Sons.

Kääntä, L. (2021). Multimodal perspective into teachers' definitional practices: Comparing subject-specific language in Physics and History lessons. In S. Kunitz, N. Markee, & O. Sert (Eds.), *Classroom-based conversation analytic research: Theoretical and applied perspectives on pedagogy* (pp. 197–223). Springer International Publishing.

Kaltenböck, G., & Mehlmauer-Larcher, B. (2005). Computer corpora and the language classroom: On the potential and limitations of computer corpora in language teaching. *ReCALL*, *17*, 65–84.

Kilpert, L., & Shay, S. (2013). Kindling fires: Examining the potential for cumulative learning in a journalism curriculum. *Teaching in Higher Education*, *18*, 40–52.

Kimura, D., Malabarba Kelly, T., & Hall, J. (2018). Data collection considerations for classroom interaction research: A

conversation analytic perspective. *Classroom Discourse*, *9*, 185–204. https://doi.org/10.1080/19463014.2018.1485589

Kintz, S., & Wright, H. H. (2017). Semantic knowledge use in discourse: Influence of age. *Discourse Processes*, *54*(8), 670–681.

Knight, D. (2011). The future of multimodal corpora. *Revista Brasileira de Linguística Aplicada*, *11*, 391–415.

Kress, G. (1997). *Before writing: Rethinking the paths to literacy*. Routledge.

Kress, G., Charalampos, T., Jewitt, C., & Ogborn, J. (2006). *Multimodal teaching and learning: The rhetorics of the science classroom*. Bloomsbury Publishing.

Lytra, V. (2012). *Multilingualism and multimodality, in: The Routledge handbook of multilingualism* (pp. 533–549). Routledge.

Macnaught, L., Maton, K., Martin, J. R., & Matruglio, E. (2013). Jointly constructing semantic waves: Implications for teacher training. *Linguistics and Education*, *24*, 50–63.

Maton, K. (2011). Theories and things: The semantics of disciplinarity. In F. Christie & K. Maton (Eds.), *Disciplinarity: Functional linguistic and sociological perspectives* (pp. 62–84). Continuum.

Maton, K. (2013a). *Knowledge and knowers: Towards a realist sociology of education*. Routledge.

Maton, K. (2013b). Making semantic waves: A key to cumulative knowledge-building. *Linguistics and Education*, *24*, 8–22. https://doi.org/10.1016/j.linged.2012.11.005

Maton, K. (2014). A tall order? Legitimation code theory for academic language and learning. *Journal of Academic Language and Learning*, *8*, A34–A48.

Maton, K. (2015). *Knowledge and knowers: Towards a realist sociology of education*. Routledge.

Maton, K. (2020). Semantic waves: Context, complexity and academic discourse. In J. R. Martin, K. Maton, & Y. J. Doran (Eds.), *Accessing academic discourse* (pp. 59–85). Routledge.

Maton, K., & Chen, R. T. H. (2015). LCT in qualitative research: Creating a translation device for studying constructivist pedagogy. In K. Maton, S. Hood, & S. Shay (Eds.), *Knowledge-Building: Educational studies in legitimation code theory* (pp. 27–48). Routledge.

Maton, K., & Chen, R. T. H. (2020). Specialization codes: Knowledge, knowers and student success. In J. R. Martin, K. Maton, & Y. J. Doran (Eds.), *Accessing academic discourse: Systemic functional linguistics and legitimation code theory* (pp.35–58). Routledge.

Matruglio, E., Maton, K., & Martin, J. R. (2013). Time travel: The role of temporality in enabling semantic waves in secondary school teaching. *Linguistics and Education*, *24*, 38–49. https://doi.org/10.1016/j.linged.2012.11.007

Ministry of Education of the People's Republic of China. (2022). *Compulsory education curriculum plan*. Beijing Normal University Press.

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, *20*, 336–366.

Mouton, M. (2020). A case for project based learning to enact semantic waves: Towards cumulative knowledge building. *Journal of Biological Education*, *54*, 363–380.

Mulholland, J., & Wallace, J. (1996). Breaking the cycle: Preparing elementary teachers to teach science. *Journal of Elementary Science Education*, *8*, 17–38.

Olsson, N. L. (2022). *Subtitle Edit*. http://www.nikse.dk/subtitleedit.

Oteíza, T., Henríquez, R., & Canelo, V. (2018). Language resources to negotiate historical thinking in history classroom interactions. *Linguistics and Education*, *47*, 1–15.

O'Halloran, K. L. (1998). Classroom discourse in Mathematics: A multisemiotic analysis. *Linguistics and Education*, *10*, 359–388. https://doi.org/10.1016/s0898-5898(99)00013-3

O'Halloran, K. L. (2015). The language of learning mathematics: A multimodal perspective. *Journal of Mathematical Behavior*, *40*, 63–74.

Pahl, K., & Rowsell, J. (2006). *Travel notes from the new literacy studies: Instances of practice* (Vol. 4). Multilingual Matters.

Runge, C. (1901). Über empirische funktionen und die interpolation zwischen äquidistanten ordinaten. *Zeitschrift für Mathematik und Physik*, *46*, 20.

Ryan, J., Scott, A., & Walsh, M. (2010). Pedagogy in the multimodal classroom: An analysis of the challenges and opportunities for teachers. *Teachers and Teaching*, *16*, 477–489.

Sinclair, J. M. (2004). *How to use corpora in language teaching* (pp. 1–317). John Benjamins.

Stauder, A., & Ustaszewski, M. (2020). Syntactic complexity as a stylistic feature of subtitles. In A. Bartoszewicz (Ed.), *Studia Translatorica* (pp. 177–188). ATUT.

Su, X., & Zhao, S. (2022). The necessity of teaching Chinese knowledge in the period of Chinese Literacy Education. *Applied Linguistics*. 1, 92101. https://doi.org/10.16499/j.cnki.1003-5397.2022.01.012

Tang, K. S., Tan, A. L., & Mortimer, E. F. (2021). The multi-timescale, multi-modal and multi-perspectival aspects of classroom discourse analysis in science education. *Research in Science Education*, *51*(1), 1–11.

Trivedi, A., Pant, N., Shah, P., Sonik, S., & Agrawal, S. (2018). Speech to text and text to speech recognition systems-a review. *IOSR Journal of Computer Engineering*, *20*, 36–43.

Verhoeven, L., & Graesser, A. (2008). Cognitive and linguistic factors in interactive knowledge construction. *Discourse Processes*, *45*(4-5), 289–297.

Vrikki, M., Wheatley, L., Howe, C., Hennessy, S., & Mercer, N. (2019). Dialogic practices in primary school classrooms. *Language and Education*, *33*, 85–100.

Wang, W. (2021). A review on semantic waves and its application in discourse analysis. *Contemporary Educational Researches Journal*, *5*, 183–188. https://doi.org/10.26689/jcer.v5i12.2868

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). *ELAN: A professional framework for multimodality research* [Conference session]. 5th International Conference on Language Resources and Evaluation (LREC 2006), pp.1556–1559.

Xie, M. (2021). Enabling students' knowledge building in English classrooms in China: The role of teacher monologue. *Text and Talk*, *41*, 261–285.

Yang, W., Luo, H., & Zeng, Y. (2022). A video-based approach to investigating intentional teaching of mathematics in Chinese kindergartens. *Teaching and Teacher Education*, *114*, 103716–103726.

Zhang, D., & Qin, J. (2016). The theory of semantic wave and its role in analyzing and constructing teachers' discourse. *Foreign Languages Education*, *37*, 4.

Zheng, Z., & Zhao, Q. (2021). A preliminary study of language measurement in elementary school mathematics classrooms from the perspective of multimodal representation. *Proceedings of the 20th Chinese National Conference on Computational Linguistics* (pp. 413–421). Huhhot, China. Chinese Information Processing Society of China.