

# **Vocabulary Explanations in Beginning-level Adult ESOL Classroom Interactions: A Conversation Analysis Perspective**

Kevin W. H. Tai<sup>a\*</sup> and Nahal Khabbazbashi<sup>b</sup>

*<sup>a</sup>UCL Centre for Applied Linguistics, UCL Institute of Education, University College London, London, United Kingdom; <sup>b</sup>Centre for Research in English Language Learning and Assessment, University of Bedfordshire, Luton, United Kingdom*

## **Correspondence details of the corresponding author:**

**Kevin W. H. Tai's** Email Address: [wht28@alumni.cam.ac.uk](mailto:wht28@alumni.cam.ac.uk)

Recent studies have examined the interactional organisation of vocabulary explanations (VEs) in second language (L2) classrooms. Nevertheless, more work is needed to better understand how VEs are provided in these classrooms, particularly in beginning-level English for Speakers of Other Languages (ESOL) classroom contexts where students have different first languages (L1s) and limited English proficiency and the shared linguistic resources between the teacher and learners are typically limited. Based on a corpus of beginning-level adult ESOL lessons, this conversation-analytic study offers insights into how VEs are interactionally managed in such classrooms. Our findings contribute to the current literature in shedding light on the nature of VEs in beginning-level ESOL classrooms.

**Keywords:** Conversation Analysis, ESOL classroom interaction, vocabulary explanations, sequential organisation

## 1. Introduction

Previous empirical research on second language (L2) vocabulary teaching and learning (e.g. Schmitt 2000) has offered insights on the nature of vocabulary knowledge, how vocabulary is learnt and factors that contribute to vocabulary learning. Numerous studies have also been carried out to investigate the overall approaches to vocabulary explanation (VE) such as inductive versus deductive (e.g. Nation 1990) and effective techniques for vocabulary instruction including strategies for guessing meaning (Nation 1990) and using first language (L1) in explaining vocabulary (Tian and Macaro 2012).

The focus of the majority of the above studies is on *what* needs to be achieved in vocabulary instruction; however, there is a lack of studies which illustrate *how* the process is achieved in real-life L2 classroom interaction (Mortensen 2011; Waring et al. 2013). Furthermore, most of the studies exploring VEs occurring in the L2 classrooms have predominantly analysed the verbal part of such VEs without taking into account their non-verbal accompaniments (e.g. Chaudron 1982; Mortensen 2011, Morton 2015). However, the works of Lazaraton (2004) and Smotrova and Lantolf (2013) have demonstrated the importance of L2 teachers' use of gestural resources to visually illustrate the meanings of L2 words and as 'fundamental means of communication' (Lazaraton 2004, 90).

Conversation Analysis (CA) is increasingly employed as a methodology to examine how VEs are interactionally managed, focusing both on the verbal and non-verbal resources employed by participants (e.g. Mortensen 2011; Waring et al. 2013; Morton 2015). Nevertheless, there is a dearth of studies that have attempted to identify the sequential patterns of VEs while taking gestures into account in beginning-level ESOL classrooms (Tai and Brandt 2018). Studying beginning-level ESOL vocabulary instruction is necessary since in such learning contexts, students have limited English proficiency and may not share a common L1 with the teacher and other classmates (Roberts et al. 2004). Several SLA studies have demonstrated the benefits of using L1 in L2 classrooms to facilitate the L2 learning processes and outcomes (e.g. Macaro 2009; Garcia et al. 2012). However, it can be assumed that not all ESOL teachers can employ students' L1s to offer translation equivalents for English VIs or provide examples from students' everyday life to explain abstract VIs (Roberts et al. 2004) unless the teachers

have acquired all their students' L1s or if there is technology, such as Google translate, available that can allow for translanguaging in the classroom. Similarly, students cannot merely draw on their L1s to overcome the language barrier when learning new VIs. Rather, they have to draw on their incomplete English repertoire to negotiate meanings or seek clarifications from their teachers and classmates in order to understand the meanings of particular VEs (Tai and Brandt 2018). Investigating the interactional resources that ESOL teachers employ in explaining English VIs to beginning-level ESOL students can potentially raise teachers' awareness of their interactional practices in doing VEs in beginning-level ESOL classrooms. Hence, this study aims to utilise CA to analyse the sequential patterns of VEs with gestures in beginning-level adult classrooms and examining how a teacher employs various linguistic and gestural resources to construct L2 VEs.

## **2. Explanations as Social Practice**

This study views explanation from an interactional and conversational analytic approach where explanations are seen as situated practices that are sequentially organised and interactionally produced. The main premise of the socio-interactional approach (Pekarek Doehler 2010) is that explanations are collaboratively achieved and produced. Hence, the participants' production of an explanation is a collaborative activity which requires participants to draw on what they know about the topic, employ language to illustrate their knowledge, and demonstrate interactional competence in participating in various social practices.

Baker (2009: 145) described explanation as an 'interactive contextual reconstruction, rather than an expression of problem-solving processes that occurred in an individual's mind'. In classroom settings, explanation sequences can be perceived as a 'question and answer language game', in which 'explanation is circumscribed by the (possibly implicit and hypothetical) question which precedes it' (Antaki and Leudar 1992, 183). However, Llinares and Morton (2010, 48) argued that explanation is not only initiated by a question, but it can be circumscribed by any preceding turns which 'do the work of identifying the explanandum' (i.e. indicating or projecting the language-based trouble in understanding) 'and the distributions of roles' (i.e. the explainer and the recipient of the explanation). Merke (2016, 2) further argued that in an explanation sequence,

participants need to ‘consider the delivered explanation (explanans) to be an appropriate and acceptable solution to the raised explainable (explanandum)’. Simply put, the connection between the explainable and explanation has to be established and accepted by both the explainer(s) and the recipient(s) in order to be seen as an explanation. This definition of explanation assumes that the participants agree on the matter which requires an explanation during the course of the interaction (Merke 2016).

Explanation is primarily offered by the explainer, but the recipient of the explanation can play a role in steering the explaining by initiating clarification questions or change of state tokens (Heritage 1984). In other words, the explanation can be marked off from the turn-by-turn interaction and in this sense, providing an explanation is inherently an interactive activity and in an important sense, it is a matter of interactional competence. Young (2008, 101) defined interactional competence as ‘a relationship between the participants’ employment of linguistic and interactional resources and the contexts in which they are employed’. This description of interactional competence can be related to what teachers and students in the classroom do when they construct explanations. They not only have to negotiate whether an explanation is needed, but also determine who serves the roles of explainer and questioner. They need to know how the explanation will be unfolded; for instance, establishing an explainable matter and moving on to the explanation turns. They also need to have the ability to change course when they receive feedback from other participants and know when and where explanations are needed and how explanations work in other social practices (Hall 1999).

It is important to note that scholars have attempted to classify the nature of vocabulary explanation into planned (PVE) and unplanned (UVE) vocabulary explanations. UVE is defined as ‘impromptu explanations of word meanings during a lesson in which an unfamiliar word or lexical concept appears’ (van Compernelle and Smotrova 2017, 194). A UVE typically involves moments when explanations are given contingently, usually due to some disrupting elements such as learners initiating uninvited responses (i.e. learner initiatives) or errors are initiated by the learners. Alternatively, PVE is referred to as prepared explanations of word meanings designed to teach the target vocabulary (Morton 2015). This usually refers to moments where the teacher’s explanations serve

as a pedagogic point of a lesson. For example, during a pre-listening activity, a teacher's introduction of a list of words and provision of its meanings indicate planned intent to explain target VIs. However, the distinction between UVE and PVE is not always clear, as it is possible for teachers to predict words that learners may not know and prepare PVE in advance so that they can address learner-initiated questions contingently and appropriately in the classrooms (See section 4 for further details). In this study, we define VE as explanations of word meanings that are constructed to clarify the meaning of VIs during lessons in order to make them clear and intelligible to the students. Generally, the explainer is assigned as the 'principal speaker' (Wald, 1978) and the explainer has the right to hold the conversational floor until the closing of the explanation. VE does not necessarily involve a detailed linguistic explanation of the VI. Rather, it can sometimes 'take a local answer, e.g. when only a synonym is provided' (Heller 2016, 255). It can also take a gestural action, e.g. when only using a gesture to represent the meaning of the VI, as demonstrated in this study. However, the explanations that are provided in the classrooms can potentially be constructed collaboratively by all participants including the teacher or the students or even the questioner (Donato 1994). Explanations can be co-constructed in which students elaborate on each other's explanations or provide alternatives in order to produce a coherent explanation.

### **3. The Nature of Vocabulary Explanations in Various Learning Settings**

In order to illustrate explanation as an interactive activity, this section will draw on a variety of studies that have explored the nature of L2 VE in different settings.

#### ***3.1 Vocabulary Explanations in Everyday Settings***

In previous CA-for-SLA studies that analysed L2 learning in everyday conversations outside the classroom, learning opportunities were mostly connected to vocabulary (e.g. Brouwer 2003; Lilja 2014). It is possible that everyday encounters provide a favourable learning environment for vocabulary learning (Lilja 2014). For example, Brouwer's (2003) analysis of L1 and L2 speaker talk illustrated how L1 speakers could often act as language experts and offer assistance to L2 speakers by helping them search for the correct word and answering L1 speakers' questions about the appropriateness of a word used. Brouwer also found that L2 speakers learnt the new words by integrating them

into their speech. Similarly, Kim's (2012) analysis of casual conversations between L1 and L2 English speakers illustrated that when an L2 speaker offered a description of a phenomenon, the L1 speaker would then provide a more specific VI in the talk to refer to the phenomenon which provided an opportunity for the L2 speaker to learn the more specialised and economical ways of referring to the phenomenon. A recent study by Svennevig (2018) shares similar findings with Kim's (2012); the study describes how a migrant worker, while working in construction site, positioned himself as a language learner who struggled to search for a word and oriented to his manager – who was an L1 speaker – as the language expert. It was found that the worker drew on multiple gestural resources to illustrate the referent of the word that was being searched for and the manager provided the word asked for. Lilja's (2014) study, in contrast, illustrated something different from the previous findings. Drawing on the analysis of a conversation between L1 and L2 speakers of Finnish, Lilja (2014) showed that when the L2 speaker encountered a problem in understanding abstract VIs (e.g. bearing), the L1 speaker employed not only linguistic but also gestural and material resources to facilitate his explanation.

### ***3.2 Vocabulary Explanations in Classroom Settings***

In classroom settings, Koole (2010) examined two different types of explanations between teachers and students in a L1 Dutch mathematics classroom. Koole suggested that discourse unit organisation is an information delivery format which involved the teacher's unilateral tellings and the student as the recipient. He found that discourse unit organisation often ends with the teacher explicitly inviting students to acknowledge understanding for example by uttering acknowledgement tokens or change-of-state tokens. In dialogue organisation sequences, the teacher typically engaged in question-answer sequences. However, he found that dialogue organisation often sets an expectation for the student not only to confirm but also to demonstrate understanding by 'doing some sort of analysis' and using 'that analysis in producing a next utterance' (Sacks 1992, 253). Koole distinguished two interactional features which illustrates a demonstration of students' understanding: demonstrations of 'having known' and demonstrations of having acquired access here-and-now. Demonstrations of 'having known' involves a claim of understanding (e.g. yes or no-answer) followed by an evidence which supports the claim. Alternatively, demonstrations of having acquired

access here-and-now involves students responding to a teacher's question (e.g. yes/no question, designedly incomplete utterance) by offering a correct answer. Heller (2016) also examined how mathematical terms are explained in L1 classrooms but focused specifically on how various semiotic resources were used in the explanation sequence. Findings suggested that participants' abilities to construct explanations largely depended on their access to and control of objects (e.g. pencil case), requiring them to draw on alternative semiotic resources in order to bring about coherent and intelligible explanations.

### ***3.3 Vocabulary Explanations in L2 Classroom Discourse***

Explanations as social actions have been examined in a range of settings including content classrooms (e.g. Koole 2010 in mathematics classrooms) as well as language classrooms. A series of studies employed CA to analyse the discursive constructions of VEs in L2 classrooms (e.g. Markee 1995; Mortensen 2011; Waring et al. 2013, Morton 2015). Markee (1995) studied the interactional organisation of teachers' responses to learners' clarification requests about the meaning of specific VIs. It was found that rather than giving answers to the learners directly, the teacher asked display questions for those VIs that the learners had difficulty with. Majlesi and Broth (2012) demonstrated how VIs were interactively established as relevant explainable matters and subsequently developed as an area of pedagogical focus in L2 Swedish classrooms. The findings illustrated that the classroom participants employed multiple visually available sources within the physical surrounding including material objects to explain concrete or abstract VIs. Investigating student-L2 tutor interaction, Belhiah (2013) demonstrated how the L2 tutor and students employed gestures to emphasise the meaning of verbal utterances, disambiguate the meaning of VIs and display alignment and understanding through gesture replication and gesture co-production. Merke (2016) explored the nature of student-initiated explanation sequences in Finnish-as-a-foreign-language classrooms and the analysis demonstrated that in order for the teacher to integrate students' identified linguistic problems into the classroom agenda, the students needed to raise the explainable matters through addressing the teacher with a question. During the explanation which was carried out collaboratively between students, a more competent student displayed her knowledge of the VI (the word 'blond' in Finnish) to assume the expert role and used the opportunity to share her knowledge with her peers.

Studying Content and Language Integrated Learning (CLIL) classrooms in a Spanish secondary school, Morton's (2015) CA analysis indicated that the CLIL teachers tended to repeat a word or display word(s) on the board, and then solicit learners' demonstration of understanding through synonyms and use of L1.

Although there are considerable number of CA studies analysing how VEs are done in L2 classrooms, only a few CA studies have systematically identified the sequential organisation of VEs in L2 classrooms. Mortensen (2011)'s CA study investigates how VEs are jointly constructed by teacher and learners in Danish L2 classrooms. He identified the following interactional organisation: (a) the teacher highlights a specific VI, (b) the learner repeats it (elicited by the teacher or learners self-repeated it spontaneously), (c) the teacher requests a word explanation, (d) the learner provides the word explanation (p.139). Mortensen (2011, 136) argued that these sequences 'evoke the institutional character of the language classroom and define the ongoing activity as "doing word explanation"'. Waring et al. (2013) identified two main types of VE in an intermediate level adult ESL classroom in the US: analytic and animated explanations. Analytic explanations entail heavy reliance on verbal and textual resources, while animated explanations involve using a range of multimodal resources. In line with Mortensen (2011), Waring et al. (2013, 254) offered an overall sequential description of the main elements of L2 VEs: '(1) set word in focus, (2) contextualise word, (3) initiate understanding-display sequence (UDS) (a two-part sequence which entails teacher checking learners' understanding and learners' display of understanding) or the teacher offers explanations by him/herself, (4) close the explanation with a repetition'. There are two terms that are particularly relevant here: contextualisation and UDS. Contextualising a word in the analytic explanation is done in a textual way, for example, by placing a word in a sentence, and therefore evoking a grammatical or semantic context. In animated explanations, contextualisation is done by employing gestures or acting out a scene to depict a situational context in order to engage the learners. Waring et al. argued that there are three types of animated explanations: talk+gesture, talk+environmentally coupled gesture and talk+scene enactment. Talk+gesture refers to gestures that elaborate on the talk, 'where each movement has its lexical affiliate' (p. 255). In addition, talk+environmentally coupled gesture means that the meanings of the gestures can be only understood with the material surround (e.g.



visual display). This means that the talk and gestures ‘cannot be stand alone as independent explanations without this material surround’ (p.258). Finally, talk+scene enactment refers to the teachers acting out a scene with gestures and verbal utterances to construct their VEs.

### ***3.3.1 The Role of Gesture in Explaining Vocabulary in L2 Classrooms***

Lazaraton (2004) was one of the earliest studies to address the lack of attention given to non-verbal aspects of classroom interactions by examining embodied resources as its main focus (2004). Lazaraton demonstrated the importance of gestures in illustrating the literal meanings of L2 VIs, particularly action-related meanings conveyed by verbs, nouns and prepositions. Similarly, Smotrova and Lantolf (2013) discovered that the Russian L2 teacher employed a variety of gestures (metaphoric and ironic) to visualise the contextual meaning of the problematic words. More recently, Sert (2017) demonstrated how an iconic gesture was used by an English-as-a-second-language teacher to aid her explanation of the word ‘each-other’. Kupetz’s (2011) CA analysis on a CLIL geography lesson has demonstrated how explanation was constructed through the linguistic resources, such as the L2, as well as gestural resources and physical objects, such as the overhead projector.

One of the very few CA studies which examined how an ESOL teacher constructed VEs in beginning-level ESOL classrooms was by van Compernelle and Smotrova (2017). The authors demonstrated how the timing of a teacher’s gesture and their synchronisation with verbal utterance were combined to make the meanings of unfamiliar VIs transparent and clear to beginning-level students. A recent study by Tai and Brandt (2018) adopted CA to examine how a teacher employed embodied enactments of hypothetical situations as a pedagogical resource to contingently explain VIs to learners in an adult beginning-level ESOL classroom.

As shown in the literature, the use of concepts such as ‘embodied enactment’ (Tai and Brandt, 2018), ‘animated explanations’ (Waring et al., 2013), and ‘embodied explanations’ (Sert 2017) illustrate the recent attention to the role of gestures in explaining in L2 classroom interaction research. Although these terms emphasise the role of gestures as part and parcel of the communicative process, these terms share some

similarities and differences. Arguably, the concepts of embodied explanations and embodied enactment are in some ways related to Waring et al.'s concepts of 'talk+gesture' format and 'talk+scene enactment' format. These two formats are connected to Waring et al.'s overarching notion of 'animated explanation'. As argued, the term 'embodied explanations' (Sert 2017) emphasises that visual behaviours like gestures are 'synchronised with the target words' (Sert 2017: 20). However, the examples provided by Sert (2017) – using iconic gestures to point to self and an imaginary other to explain 'you' and 'me' – only conceptualised gestures as a resource for elaborating on the talk where each gestural movement has its immediate lexical affiliate (i.e. VIs demonstrated by the gestures). Thus, we argue that while Sert's notion of embodied explanation aligns with Waring et al.'s definition of 'talk+gesture', a sub-category of 'animated explanation', it differs from embodied enactment as the former considers gestures as a supplement or aid to the teacher's VEs whereas the latter involves 'more than simply using [one's] body to emphasise pronunciation or add visual description of a concept' (Tai and Brandt 2018: 262). As shown in Tai and Brandt (2018: 262), the ESOL teacher physically creates a situational context for students to understand how the target language can be used in specific contexts'.

It is important to acknowledge that Waring et al.'s example of 'talk+ scene enactment', another type of animation, and Tai and Brandt's (2018) example of embodied enactment both show how participants enacted a scene with gestures and verbal utterances in the course of their VEs. Hence, there is a close alignment between the notions of 'embodied enactment' and 'talk+scene enactment' since both terms emphasise similar social actions (i.e. acting out a scene). The key difference between these two terms lies in the different sequential organisations in constructing the VE. In the examples given by Waring et al., (i.e. explaining the words 'trade' and 'passed away'), the situational contexts were established either through teacher's verbal utterances (e.g. using the subordinate conjunction 'if' and then building up an imaginary dialogue) or teacher's borrowing of the hypothetical contexts from the classroom materials (e.g. the conversational topics in the movies or textbooks). In contrast, the construction of embodied enactment in Tai and Brandt (2018) emphasises the need for participants to first physically and verbally construct a hypothetical context, which mirrors a real-life situation, before enacting a scene. Here, the hypothetical context in embodied

enactment is constructed through both physical movements and verbal utterances rather than solely relying on verbal descriptions of the scenario or classroom artefacts. Moreover, Waring et al. (2013) rightly argued that the contextualisation in animated explanations is ‘mostly done through visual demonstrations of specific activities or entities’ (p. 258). Thus, such practice aligns with the act of creating an imaginary context in ‘embodied enactment’ as Tai and Brandt (2018) argued that by physically creating a hypothetical context before the enactment, it can facilitate visual illustration and help signal the appropriate use of a target phrase or word in a specific context.

The above literature suggests that VE has received some attention in L2 classroom interaction research; however, most of the reviewed studies have been with students of intermediate and advanced levels of English (e.g. Lazaraton 2004; Waring et al. 2013; Smotrova and Lantolf 2013) and/or have focused on settings where there was a shared L1 between classroom participants (e.g. Morton, 2015). To date, there is limited research (with the exception of van Compernelle and Smotrova 2017 and Tai and Brandt 2018) that explore the construction of VEs in beginning-level ESOL classrooms where there is an absence of shared L1 between teachers and learners. In these learning contexts, there is an increased likelihood for teachers and students to rely more heavily on multimodal resources, such as gestures, to function effectively in the classroom. Our study therefore examines the issue of the interactional management of VEs in beginning-level ESOL contexts in order to allow teachers to better understand how VEs can be conducted in classroom contexts.

#### **4. Data and Method**

This study aims to address the following research questions (RQs):

- (1) How are L2 VEs in adult beginner-level ESOL classrooms sequentially organised?
- (2) What linguistic and multimodal resources are employed by the ESOL teacher in constructing L2 VEs?

The classroom video-data for this study were drawn from the Multimedia Adult English Learner Corpus (MAELC). The video-data were collected at Portland Community College. This corpus was compiled to allow researchers to conduct longitudinal studies

of beginning-level adult ESOL learners' SLA processes (Reder 2005). The full corpus includes over 4600 hours of beginning-level ESOL classroom interactions over 5 years.

The entire corpus included 900 lessons. The segments of data selected for this study were collected from several lessons at one beginning-level ESOL classroom from January to April 2002 (two lessons per week each lasting two hours). For this study, all lessons (a total of 30 lessons) from this particular beginning-level classroom from January to April 2002 were observed. The ESOL teacher was an experienced teacher who had studied German and Spanish at a US university. In the class, there were twenty-one adult learners of English, who came from various countries including Romania, Latin American countries, Russia, Africa, China and Korea.

This study utilised CA to illustrate how speakers understand and talk to one another (Psathas 1995). CA adopts a participant-relevant perspective on social action to study how social order is co-constructed by individuals through the detailed analysis of the interaction. CA as a methodology allows for the study of social interactions which includes semiotic resources such as gesture and body posture without pre-theorizing the relevance and importance of language-in-use.

#### ***4.1 Data Analysis Procedures***

The first stage of analysis involved taking a stance of 'unmotivated looking' (Psathas 1995) as the guiding principle when reviewing the video-recordings from MAELC. We watched multiple classroom-videos with an open mind (i.e. without any particular interest or research focus) to discover any interesting interactional phenomenon that is worthy of further exploratory analysis. In this process, it was noticeable that the teacher devoted a lot of her time in explaining unfamiliar VIs to her learners, not least because of the limited linguistic resources shared between the participants. The teacher occasionally demonstrated her understanding of the Spanish produced by her students but tended to rely on English to explain the vocabulary. The students, on the other hand, displayed some ability to understand the teacher's English although their English proficiency appeared to be rather limited. The teacher's VEs drew heavily on embodied resources to explain VIs which is likely due to the students' limited English abilities; this was considered worthy of further analyses. Thus, we reviewed the video-recordings

again and paid attention to teacher-student interactions in order to observe how teacher's VEs were done in the interactions. As explained earlier (see section 2), the distinction between UVE and PVE is not always clear in CA analysis, as establishing these distinctions would require analysts to conduct teacher interviews in order to understand their prepared lesson plans. In order for us to determine whether particular extracts involved VE sequences, we adopted a widely-accepted definition of VE as guidance for identifying such interactional sequences i.e. VE turns as any turns that are constructed to clarify the meaning of the VI in order to make it clear and intelligible to the students (Merke, 2016). This typically involves a recipient of the VE, and an explainer of the vocabulary, who can be either the teacher or a student.

The second stage of analysis entailed the transcription of all excerpts that involved VE sequences. We adopted Jefferson's (2004)<sup>1</sup> transcription conventions to transcribe the video-data. After transcribing the data, we carried out line-by-line analyses to closely investigate various sequences-of-talk which entailed teacher's VEs. The collections consisted of nineteen extracts on teacher's VEs as a result of going through the corpus systematically by selecting all teacher's VEs. We constructed the descriptions of the sequential organisations of VEs based on the full range of extracts. This led to the formation of three features of the VE sequences.

For reporting purposes, we can only present illustrative extracts; this can raise concerns regarding the extent to which selected extracts are an adequate representation of all analysed data. It was therefore important to examine all extracts for similar and/or deviant instances (ten have, 1990). In this study, Extracts 1-4 are typical VE sequences that represent the larger collection.

## **5. Findings**

Our findings revealed two types of VE sequences: (1) talk+gesture format and (2) embodied enactments as explanations. These will be described in the next sections and illustrated with extracts from the analyses.

### ***5.1 Talk+Gesture Format***

Talk+gesture format refers to explanations that include deployment of visual behaviours, such as the use of hand gestures. The notion of talk+gesture reinforces the idea that gestures can work effectively to enhance the verbal explanations of the VIs and animate the meaning of the VIs to create a vivid image (Waring et al., 2013). In this study, seven instances were found which illustrated the teacher's (T's) talk+gesture format. Extracts 1 and 2 are typical cases which demonstrate this interactional phenomenon. The talk+gesture format typically displays the following sequential organisation:

- (1) T emphasises the target word by repeating it or the learners self-initiate the target word;
- (2) T then provides explanations through using a combination of linguistic resources, including simplified language, and gestural resources;
- (3) Students then acknowledge the word explanation by uttering acknowledgement tokens and/or imitating T's gestures.

*Extract 1: Road and Street*

Extract 1 is an example of how the teacher utilises multimodal resources (e.g. gestures and drawings) to explain the meanings of 'road' and 'streets'. Prior to the initiation of T's explanation of the meaning of 'road', T was giving real examples of common road names in the city including 'beaverton' and 'hillsborough' to explain the meaning of 'road' (line 1). In this extract, S9 first initiates the target word ('road') for explanation in line 6. T then launches her explanation by differentiating the meanings of 'street' (lines 10-20) using iconic gestures, drawing cursive lines on the whiteboard and physically moving her arms. Students then claim their understanding of T's explanations in lines 17, 20 and 26.

01 T: +beaverton (0.5) +hillsborough (0.5) +um

+T holds her left-hand out and points to the direction facing the whiteboard

+T holds her left-hand out and points to the direction facing  
the whiteboard

+T moves her fingers on her left  
hand upwards and downwards,  
palms facing downwards

02 (0.9)

03 S9: °tv highway°=

04 T: =tv highway

05 +(0.3)

+S9 curves her arms in parallel towards her right-hand side #1



06 S9: the (road) (0.3) no?

07 (.)

08 T: no (.) a road (0.3) a road is (1.1) um

09 (0.6)

10 T: +streets are usually +(4.4) a road maybe

+T writing on the whiteboard and drawing straight lines

+T holds her arms in parallel and moves her arms up and down. #2 T then faced towards the students and enact the same gesture. #3



#2



#3

11 +(4.2)

+T drawing cursive lines on the whiteboard

12 T: +not so

+T curves her right arm to illustrate a curved road #4



#4



#5

13 (0.2)

14 SS: haha

15 (0.2)

16 T: +straight [is street]



+T extends her right arm, facing to the students #5

17 SS: [ya (.) ya]

18 (0.2)

19 T: +road (0.3) +[sh:]

20 SS: +[ya]

+T points to the ground

+T walks to her right to and fro and moves her finger to illustrate the curved road #6 #7 #8



#6



#7



#8

21 (0.5)

22 S3: road

23 (0.2)

24 T: okay? hahaha

25 (0.5)

26 S3: oh yeah

In line 3, student 9 (S9) provides an additional example of a road name ('tv highway') to display her understanding of the meaning of 'road'. This is accompanied with her use of iconic gesture (curving her arms in parallel, figure 1) to visually illustrate a road in line 5. S9 then produces a question ('the road (0.3) no?') in line 6 to seek T's confirmation regarding her response in line 3.

Nevertheless, T treats S9's prior turn as a question whether 'tv highway' counts as a 'road' and T rejects S9's example by giving a negative assessment 'no' in line 8. T then repeats 'road' twice to place the VI at center stage for public scrutiny. Notice that there is a long 1.1-second pause, followed by a hesitation marker 'um' uttered by T in line 8 which potentially indicates that T is searching for an appropriate response in order to repair S9's understanding (Gardener, 2001). After a 0.6-second pause, T simultaneously draws straight lines on the whiteboard and utters 'streets are usually'. During the 4.4-second pause, T holds her arms in parallel and move her arms from up to down (figures 2 and 3) in order to indicate that streets are usually straight and perpendicular to each other. After explaining the meaning of 'street', T explains the meaning of 'road'. T first utters 'a road maybe' in line 10 to establish its focal status. During a 4.2-second pause in line 11, T draws cursive lines on the whiteboard and T curves her right arm to visually illustrate a curved road. Accompanied with her utterance 'not so' in line 12, T suggests to her learners that roads are usually curved and not in straight lines.

Although T's explanation leads to students laughing, there is no clear evidence which demonstrates learners' understanding of T's explanation. In line 16, T takes another turn to summarise her explanation of 'street' by uttering 'straight is street' and extending her right arm in front of her students. This leads to students' claiming of understanding (Koole, 2010), as illustrated by the several acknowledgement tokens uttered by the students (line 17). In line 19, T summarises her explanation of 'road' by employing iconic gestures and body movements (pointing to the ground and moving her finger as she walks to and fro, figures 6,7 and 8) to indicate the curved road. T's summaries of the meanings of 'road' and 'street' lead to learners' acknowledgment of the word explanations by uttering acknowledgement tokens 'ya' in line 20 and 'oh yeah' in line 26.

As shown in this extract, T engages in a very visual and animated manner of explaining the difference between ‘street’ and ‘road’. T first sets the VIs in focus by repeating and then employ gestural resources and visual demonstrations (e.g. drawings on the whiteboard) to vividly depict the meaning of the VIs. T later also provides verbal explanation in line 16 to summarise the meaning of ‘street’ but she does not offer a verbal explanation for ‘road’.

*Extract 2: We and They*

Extract 2 is an example of how T employs both verbal explanations and iconic gestures to explain pronouns. Prior to the extract, T was teaching the pronouns: ‘I’, ‘she’, ‘he’, ‘we’, and ‘they’. Students were asked to observe T’s gestures to determine what pronoun she was referring to. In this extract, the target words (‘we’ and ‘they’) are first repeated by the students and teachers from lines 2-20. T then offers a verbal explanation of ‘we’, accompanied with iconic gestures in lines 24 and 27. T then provides another verbal explanation of ‘they’, accompanied with her arm movements, in lines 32 and 36. Eventually, S9 imitates T’s gestures in line 37 which potentially displays her understanding of T’s explanation.

01 +(2.0)

+T stretches out her right arm, points at the learners on her right-hand side and makes a circular motion repeatedly

02 S9: they

03 (0.5)

04 T: +they

+T stretches out her right arm and points at the learners on her right-hand side

05 (0.2)

06 S1: they=

07 SSs: = +they +(0.2) they (0.2) they

+S3 and S9 point to the learners on their left-hand side

+T points to the learners on her right-hand side and moves her right arm to make a circular motion

08 (0.3)

09 S3: +we

+S3 makes a circular motion at chest level

10 (1.6)

11 T: +we

+T touches S3's shoulder with her left hand, holds her right hand upwards chest level and moves her right hand forward ((pointing at S3)) and backward ((pointing at T)) repeatedly

12 +(1.6)

+T points at the learners on her right-hand side

13 S3: +they

+S3 points at the learners on her left-hand side #9



#9

14 (0.2)

15 SSs: they

16 (0.3)

17 T: they=

18 S3: =ya

19 (0.3)

20 S9: they

21 (0.2)

22 T: okay

23 (0.2)

24 T: +we is two=

+T holds up her index and middle fingers

25 S9: =two

26 (0.2)

27 T: +[me] +and

+T's right hand pointing to herself

+T's left hand touching S3's shoulder

28 S3: [yeah]

29 (0.7)

30 S3: ya

31 (0.5)

32 T: but +they

+T points at the learners on her right-hand side

33 (0.3)

34 S3s: they

35 (0.2)

36 T: +she and +she

+T stretches out her right arm and points at S8

+T stretches out her right arm and points at S5

37 +(0.4)

+S9 uses both of her hands to point at the learners on her left-hand side

38 S3: yeah

As Extract 2 begins, T and students are practising what T has earlier taught about 'we' and 'they' (from lines 1-23). Although the students correctly identify the pronouns: 'we' in line 9 and 'they' in lines 13 and 15, T does not take the students' demonstration of understanding as a cue to close the sequence. T launches her explanation of the difference between 'we' and 'they' in lines 24-36 as an opportunity to consolidate the meanings. T first provides a short definition of 'we' by stating 'we is two', accompanied by her gesture of holding her index and middle fingers upwards (line 24). She then points to herself while uttering 'me', and then touches S3's shoulder while uttering 'and' in order to demonstrate that 'we' refers to herself and S3. T further explains the meaning of 'they' in lines 32-36. T first utters 'but they' which suggests that the word 'they' does not share the same meaning as 'we'. T then stretches out her arm and points at students 8 and 5 (S8 and S5) while uttering 'she and she' in line 36 to

reinforce her physical distance between herself and the group of learners on the other side of the classroom. In line 37, S9 imitates T's gesture, which possibly illustrates her understanding of the meaning of 'they'.

In these two extracts, it is noticeable that T in Extract 1 mainly relies on embodied resources to explain the difference between 'road' and 'street', whereas in Extract 2, T synchronises her gestural resources with her verbal explanations of 'we' and 'they'. Despite the differences, both extracts illustrate that T employs a range of resources including body movement and simple linguistic utterances to explain the target words to beginning-level learners. However, such interactional practices are different from embodied enactments as explanations (see below).

### ***5.2 Embodied Enactments as Explanations***

In this study, ten instances were identified which illustrated the constructions of embodied enactments. Extracts 3 and 4 are typical cases which demonstrate this feature. It is crucial to note the difference in perspective taking (McNeill, 1992) between Extracts 1 and 2 and Extracts 3 and 4. In Extracts 1 and 2, T adopts the talk+gesture format to construct the VEs by taking an observer's viewpoint, which means that the speaker's use of gestures symbolise some entities from the narrative and this links to the third person singular in the verbal expression (McNeill 1992, 119). In other words, T represents the meanings of VIs by making iconic gestures of streets and roads (Extract 1) and iconic gestures to illustrate 'we' and 'they' (Extract 2) from a third person viewpoint. Nevertheless, in Extracts 3 and 4, T takes a character viewpoint by 'incorporating the speaker's body into the space' as if on stage (McNeill 1992, 119) to act out the situational contexts physically and verbally to explain VIs. As shown in Extracts 3 and 4, T allows learners to understand how the target VIs can be employed in specific real-life situations by physically and verbally creating a situational context. The following sequential organisation is identified for both T and learners in performing embodied enactments:

- (1) T invites students to follow T's instructions;
- (2) T verbally and physically establishes an imaginary context;
- (3) T brings students into the enactment;

- (4) During the embodied enactment, T switches footing by shifting back and forth between the hypothetical context to the instructional context;
- (5) Students claim their understanding of the embodied enactment through different ways, including producing change-of-state tokens.

*Extract 3: Excuse me*

Extract 3 is an example of how T aims to offer additional explanations regarding the alternative meanings of ‘excuse me’ through enacting a hypothetical context with gestures and speech. Some parts of Extract 3 were previously analysed in Tai and Khabbazzashi’s (2019) study. However, the full extract was not fully presented and analysed in Tai and Khabbazzashi (2019). In this paper, we will present and analyse the whole extract. Prior to this extract, T was teaching the meaning of ‘excuse me’ (i.e. asking someone to repeat their utterance) to her students. In this extract, T first asks S13 to stand up from her seat (line 13). Then, T walks towards S13 to create a hypothetical scenario (lines 15-17). After that, T brings the student into the enactment by requesting her to move aside (line 16). Then a student claims his/her understanding of the embodied enactment in line 17. After a long pause, T and S13 step out of the hypothetical contexts but T brings S13 back to the hypothetical contexts again in order to give her a chance to repeat the previous enactment (lines 23-26). S13 finally demonstrates her understanding of the meaning of ‘excuse me’ by self-initiating the enactment (line 31).



01 T: now (0.5) +there's excu↓se me? (0.3) +>excu↓se me?<

+T pointing at the phrase 'excuse me' with high intonation on the whiteboard on the left

+T pointing at the phrase 'excuse me' with high intonation on the whiteboard on the left

02 (0.5)

03 T: and (0.8) +excu↑se me↓

+T pointing at the phrase 'excuse me' with low intonation on the whiteboard on the right

04 (1.4)

05 T: +up (0.8) and (.) +down

+T moving her right hand from low to high position

+T moving her right hand from high to low position

06 (.)

07 S10: +this is no good? (0.2) +this is good?

+S10 pointing at the phrase 'excuse me' with low intonation on the whiteboard

+S10 pointing at the phrase excuse me with high intonation on the whiteboard

08 (0.2)

09 T: +different (0.7) different (0.3) different

+T spreads out her hands in opposite directions #10



#10

10 (0.4)

11 S10: ah

12 (0.2)

13 T: different (0.2) +example (0.4) °please stand up°

+T points at S13, making beckoning and inviting  
motion

14 +(0.8)

+S13 stands up

15 T: °please stand here° +(0.8) okay +(0.8) excu ↑ se me ↓

+T stretches out her arms and points to the ground

+S13 walks towards

+T moves towards S13

16 +(1.4)

+T touches S13's shoulder

+S13 moves to the right to offer space for T to walk through

17 S?: oh

18 (1.0)

19 S3: excuse me

20 (0.3)

21 SS: hahaha

22 (0.2)

23 T: +yeah (your turn)

+T turns around, facing S13 opposite direction #11



#11

24 +(0.3)

+S13 walks towards T

25 S13: +excu ↑ se me ↓ (0.4) +excu ↑ se me ↓ =

+T moves to the right to allow S13 to go through

+S13 spreads out her left arm

+S13 touches T's left arm

26 T: =okay (0.3) hahaha (0.7) +no no

+T patting S13's elbow

+T's eye gaze on other learners in the  
classroom

27 (0.2)

28 T: [hahaha]

29 SS: [hahaha]

30 (0.7)

31 S13: +ah (0.2) ah excu ↑ se me ↓

+S13 steps towards T

+S13 moves her left hand to the left-hand side indicating the pushing

32 (0.4)

33 T: yeah (0.2) yeah (0.3) uh ha

In lines 1-5, T explains to the students that there are two ways of pronouncing ‘excuse me’: ‘excu↓se me?’ and ‘excu↑se me↓’. This leads to a follow-up question asked by student 10 (S10) asking which pronunciation is considered as preferable (line 7). T explains that there are differences between the two pronunciations (line 9).

Although S10 initiates a change-of-state token ‘ah’ in line 11 which possibly displays her understanding of T’s explanation, T initiates a new turn in line 13 and asks student 13 (S13), a Chinese student, to stand up which projects that T will offer an additional explanation to S10’s question through the use of an example (line 13). T makes a request to S13 to ‘°please stand up°’ quietly, making a ‘standing up’ motion with her hands. S13 stands up and follows T’s request by moving closer to T (line 15), although at this stage S13 has not yet been informed of the reason for her to stand up. T establishes a hypothetical scenario by walking towards S13 in line 15 to indicate her walking direction, which signals to the class that T’s forthcoming action will be performative. T then enacts a hypothetical context by uttering ‘excu↑se me↓’ (line 15) and physically touching S13’s shoulder (line 16), to represent the embodied enactment itself. T signals a change of footing for the event (Goffman 1981) – shifting from describing the scenario to enacting a hypothetical scenario. A change of footing refers to how participants change their orientation of the frame for events that they are participating in (Goffman 1981). In line 16, S13 follows T’s request by moving to the right to offer space for T to walk through, which is a demonstration of S13’s understanding of T’s previous action. As shown, T leads the construction of the embodied enactment from lines 15-18 by acting as the pedestrian who initiates request

and S13 as the passive pedestrian who moves aside. After an unknown learner initiates a change-of-state token ‘oh’ to possibly indicate his/her understanding of the embodied enactment in line 17, there is a long 1.0-second pause which is not being taken up by T or S13.

At this point, both T and S13 have appeared to shift the footing by ‘stepping out’ of the hypothetical context (Goffman 1981). Nevertheless, T chooses to bring S13 back to the hypothetical context and invites S13 to repeat the previous enactment by saying ‘yeah your turn’ in line 23. S13 accepts T’s invitation and S13 first walks towards T in line 24, spreads out her left arm and utters ‘excuse me↓’ in line 25 which T responds by creating space for S13 to go through. S13 utters ‘excuse me↓’ again in line 25 which T offers a verbal acknowledgment ‘okay’ to acknowledge S13’s request in line 26. It is noticeable that unlike the previous embodied enactment where T initiates a request, in this embodied enactment the hypothetical roles are switched when S13 is given the chance to act as the pedestrian to initiate the request. Both T’s laughter and her action of patting S13’s elbow (line 26) is a demonstration of the non-serious nature of walking past one another in this way.

Furthermore, S13 attempts to demonstrate her understanding through self-initiating enactment. S13 utters ‘excuse me↓’ while stepping towards T and moving her left hand towards the left (line 31), which indicates the act of pushing. It is evidenced that T utters acknowledgement tokens: ‘yeah’ for twice and ‘uh ha’ in line 33, which confirms S13’s understanding of the meaning of ‘excuse me’.

#### *Extract 4: You first*

Extract 4 is another example of how a student and T co-construct the meaning of a VI through engaging in embodied enactment to visualise the contextual meaning of the VI to other students. Prior to this extract, T was explaining the use of ‘go ahead’ and ‘after you’ and emphasising that both VIs could be employed in situations where an individual invites a person to pass through ahead of him/her. In this extract, the VI ‘you first’ is introduced by S15 (line 7). Then T acknowledges S15’s response and invites S15 to come to the front of the classroom (line 11). T then launches the hypothetical context by verbally and physically illustrating the shape of a ‘door’ (line 13). In line 15, T brings

S15 into the enactment and acts out a scene (i.e. pretending that they are walking together). From lines 15-22, T switches back and forth between hypothetical and instructional contexts in order to offer verbal explanations of the enactment to the students. After the enactment, S13 acknowledges her understanding of the meaning of 'you first' in lines 24 and 36.

01 T: interesting

02 +(1.0)

+T glances at the whiteboard

03 T: please go ahead

((T reads aloud the phrase on the whiteboard))

04 (0.2)

05 S15: um um

06 (0.5)

07 S15: eh (0.4) you (0.2) first?

08 (1.0)

09 T: +you first +(0.2) yeah +(0.4) you can say that too

+T raises her index finger at chest level

+T picks up the whiteboard pen

+T writing 'you first' on the whiteboard

10 +(5.4)

+T writing 'you first' on the whiteboard

11 T: so example (0.2) +please come (3.3) here

+T holds hands out, facing S15, parallel to each other, palms

facing upwards, and bends fingers quickly upwards

12 +(0.2)

+S15 walking towards T

13 T: okay +we +do here (.) +so (.) +door (0.6) door

+T's both hands pointing to the ground

+S15 walks to the spot indicated by T

+T extends her arms and lowers her arms,  
both hands facing downwards, palms  
facing backwards

+T puts hands together in parallel #12



#12

14 (0.2)

15 T: okay? (0.3) +same +time

+T walks slightly forward

+T touches S15's right arm

16 (0.2)

17 S15: +eh oh

+S15 walks behind T

18 (0.5)

19 T: +and he says=

+T stops walking

20 S15: = +you +first

+S15 leans down

+S15 lowers his arms near his knee level, moves his arms to the right-hand side,  
right arm is slightly higher and further than his left arm #13 #14



+T squeezes her body to go through the pathway



#13



#14

21 (0.3)

22 T: you +first

+T extends her right arm at chest level

23 +(0.4)

+S15 lowers his right arm near his waist level, moves his right arm to the right-hand side

#15



#15

24 S13: ah okay

25 (0.2)

26 T: +so +I go first

+T points at herself with her left hand

+T extends her right arm at chest level

27 (0.6)  
28 S13: you (0.2) first  
29 (0.4)  
30 T: um hm  
31 (0.8)  
32 S13: you first (0.5) af- (0.7) af-  
33 (0.5)  
34 T: or after you  
35 (0.3)  
36 S13: ah okay [after you]  
37 S15: [after you]

Extract 4 begins with T uttering ‘interesting’ (line 1) and ‘please go ahead’ (line 3) which potentially refers to her reflections regarding her previous explanations of ‘go ahead’ and ‘after you’. As no one takes up the floor in line 4, student 15 (S15) self-selects to initiate a turn to suggest an alternative phrase (‘you first’, line 7) to T. S15’s self-initiation opens a new sequence to discuss a specific topic. T confirms S15’s suggestion in line 9 by repeating the phrase and provides positive feedback to S15 by stating ‘yeah (0.4) you can say that too’. T also writes down the phrase on the whiteboard to place it at centre stage for public scrutiny.

Similar to Extract 3, T projects that she will offer an explanation of ‘you first’ through employing an example (line 11). T invites S15 to come over to the centre of the classroom by saying ‘please come (3.3) here’ in line 11 and making an inviting motion. S15 follows T’s instruction without being informed of the purpose (line 13). T then establishes a new imaginary context by uttering ‘so (.) door (0.6) door’, lowering her arms and putting her arms together in parallel (figure 4), which encourages learners to imagine a door facing in front of T. T then produces a pre-closing ‘okay?’ (Beach 1995) which provides an opportunity for the learners to raise any questions before moving on. Since no one produces any questions (line 15), T continues with establishing the imaginary context by saying ‘same time’. After that, T physically enacts the scene by walking slightly forward and S15 is standing behind T. Here, T shifts the footing from

instructional frame to hypothetical frame in order to enact that imaginary context. T touches S15's right arm which requests S15 to walk behind T. S15 utters a change-of-state token 'oh' and follows T's instruction in line 17. By doing so, T and S15 are co-constructing a scenario where there are two people planning to go through the hypothetical 'door'. After the enactment, T shifts the footing from enacting the scene (hypothetical frame) to providing a verbal explanation to the learners (instructional frame). T utters 'and he says' in line 19 which indicates the next utterance that S15 needs to produce. S15 immediately utters 'you first' (line 20) to address T's explanation. While S15 is uttering 'you first', he leans down his body and makes a 'welcoming' gesture with his hands (figures 5 and 6) to invite T to walk through the imagined 'door'. This reveals that there was a change of footing from describing to enacting the scene again. T continues with the embodied enactment by squeezing her body to go through the 'door' (line 20). T then repeats S15's response, 'you first', and simultaneously moves her right arm to chest level in line 22 which emphasises her walking direction to the imagined 'door'. This signifies a shift from the completion of the embodied enactment to T's provision of verbal explanation as T takes this opportunity to confirm S15's enactment in line 20 and complete the turn that she first initiates ('and he says') in line 19.

In line 24, S13 utters 'ah okay' to claim her understanding of the meaning of 'you first'. T reinforces her explanation by stating 'so I go first' (line 28). This is accompanied by her use of deictic gesture (pointing to herself when she says 'so') and iconic gesture (moving her right-arm to chest level when utters 'go') to emphasise the idea that T is offered the chance to go after S15. Note that S13 utters an incomplete utterance 'af-' twice in line 32, potentially searching for the right word to express her ideas. Despite S13's unintelligible utterances, T illustrates her understanding of S13's utterance by uttering 'or after you' in line 34. The phrase 'after you' was previously explained in the lesson. By doing so, T is offering corrective feedback to S13, as well as linking the meaning of 'after you' with the target phrase 'you first'. In line 36, S13 utters a change-of-state token 'ah okay' and repeats 'after you' to claim her understanding (Koole, 2010).

In summary, both extracts illustrate how T enacts a hypothetical context through verbal and multimodal resources to facilitate learners' understandings of the VEs. Importantly, Extracts 1-4 demonstrate the same interactional phenomenon, which is using gestural and body movements to animate the meanings of the VIs to the students. The difference between Extracts 1-2 and Extracts 3-4 is that in Extracts 1-2, T adopts an observer viewpoint to represent the literal meanings of VIs, such as specific objects (e.g. roads, streets) or particular persons (e.g. 'we' and 'they'), whereas T adopts a participant viewpoint in Extracts 3-4 to represent the contextually relevant meanings of VIs which is tied to specific situation of use, such as using 'excuse me' when an individual wishes to ask another person to move aside (Extract 3) and using 'you first' when an individual invites a person to pass through ahead of him/her (Extract 4). Moreover, Extracts 1-2 and Extracts 3-4 have revealed different interactional practices. The actual interactional practices that are employed to accomplish the act of explaining are different. In Extracts 1-2, the teacher employs iconic gestures, arm movements and drawings on the whiteboard to explain the literal meanings of the VIs. In contrast, in Extracts 3-4, the teacher needs to do actions including inviting students to follow the teacher's instructions before enacting the imaginary context as well as bringing students into the enactment in order to mirror everyday life actions. These L2 VIs with context-specific meanings, such as 'excuse me', are tied to specific situation of use and they cannot be depicted by solely using hand gestures. Hence, depending on the meanings of VIs, this can potentially motivate T to choose between a talk+gesture format and an embodied enactment in order to present the most relevant and situated meaning of L2 VIs to students in a concrete and visible form.

## **6. Discussion and Conclusion**

In response to the first RQ about the sequential organisation of the VEs, we have shown that the VE sequences typically entail the following components:

- (1) T/students set the VIs in focus and/or establishes the context;
- (2) T/students provide explanations through employing different linguistic and semiotic resources;
- (3) Students display their understandings of the VIs and/or acknowledge the receipt of the explanations.

The sequential organisation of the VEs demonstrated some of the common components identified in CA studies by Waring et al. (2013) and Mortensen (2011). Similar to those studies, T typically first sets the VIs in focus by repeating them or writing them on the whiteboard for public scrutiny. The next stage (i.e. providing explanations) is similar to Waring et al. (2013)'s third stage (i.e. invite or offer explanations). Finally, similar to Waring et al., the present study acknowledges that the last stage of the VE entails the closing of the VE sequence. Although Waring et al. identified that T or the students typically close the explanation with a repetition, this study shows that the last stage typically involves a claim of the students' understanding, including through uttering acknowledgment tokens (e.g. Extract 1) and change-of-state token (e.g. Extract 4), as well as a demonstration of the students' understanding, such as self-initiating enactment in Extract 3.

In response to the second RQ regarding the teacher's use of linguistic and multimodal resources in constructing the VE sequences, we found that T employs a range of verbal and non-verbal resources, including providing definitions, gestures, and embodied enactments, to construct the L2 VE sequences. Physical resources including the whiteboard are also employed to make the VI more salient. Notably, T draws heavily on gestures in explaining VIs; a finding which closely aligns with Lazaraton (2004) and van Compernelle and Smotrova's (2017) where they showed that gestures allowed teachers to visualise the contextual meaning of the VIs thus making the VE more comprehensible and concrete for the learners. In addition, this study fills in the research gap identified by Tai and Brandt (2018) by exploring the sequential organisation of embodied enactment in multiple beginning-level ESOL classrooms.

Alternatively, this paper reinforces the importance for beginning-level ESOL teachers to select the appropriate gestural and linguistic resources to help convey aspects of the relevant and contextualised meanings of target VIs for their students. This, in turn, can potentially facilitate the students' comprehension of the VEs. In this paper, we have illustrated that talk+gesture format (Extracts 1-2) and embodied enactments (Extracts 3-4) share similar interactional phenomena in terms of using embodied resources to vividly represent the meanings of the VIs as a sequence of actions performed by the

teacher. In an ESOL context where the teacher and students do not share a common L1, which does not allow L1-L2 translation, the VIs would be difficult to communicate through verbal explanations alone, particularly for ESOL learners with relatively limited English proficiency. Therefore, the teacher's choice of adopting a talk+gesture format or an embodied enactment in constructing the VEs is potentially motivated by the aim of selecting the most effective way for explaining the VIs. We argue that the data extracts serve as an example of a teacher orienting to the needs of her students and selecting a multimodal method that is deemed appropriate for constructing the VEs.

The findings contribute to the current literature on L2 vocabulary teaching and learning in several ways. First, the three components including setting the VIs in focus, establishing contexts and providing explanations are similar to what has been identified in Waring et al.'s (2013) study as key elements of L2 VE sequences. Although Waring et al.'s findings were generated from intermediate adult ESL classrooms, this study demonstrates that Waring et al.'s CA findings can potentially be extended to beginning-level ESOL classroom contexts where the shared linguistic resources are somewhat limited between the teacher and the students as well as between students from different L1 backgrounds. As argued, this specific learning context is not well-explored in research on L2 classroom discourse (Tai and Brandt 2018). More research is needed to examine the nature of VEs in L2 classrooms where teachers and students share limited L2 repertoires (e.g. Tai and Brandt 2018). The present findings provide insights into the complexity of beginning-level ESOL classroom interaction and behaviour which enables ESOL teachers to reflect on their practices in explaining vocabulary to low-proficiency English learners. Understanding the interactional practices of beginning-level ESOL teachers as they do VEs has important implications for developing beginning level ESOL teachers' 'classroom interactional competence' (Walsh, 2012), which refers to the skill required by teachers to employ language and features of classroom interaction in ways that facilitate learning opportunity.

## **7. Limitations**

First, as we were analysing video-data that was previously collected, we were not able to obtain any information regarding the students nor could we interview the teacher regarding the characteristics of the students, which is a general limitation of corpus-

based studies. Second, it is not possible for CA analyses to distinguish between teacher's planned or unplanned VEs in the classrooms. Future research can conduct teacher interviews in order to better understand their pedagogical practices. Third, it could be argued that the findings are not generalisable to other learning contexts as some of the interactional features (e.g. remedying knowledge imbalances between the teacher and student in Extracts 5-6) may be idiosyncratic to this teacher or this classroom. It needs to be noted that CA findings cannot be generalised to other contexts due to the central role of the specific context under study. What a CA analysis provides is not empirical but analytical generalisation, where each interactional feature is evidence that 'the machinery for its production is culturally available, involves members' competencies, and is therefore possible (and probably) reproducible' (Psathas 1995, 50). In this regard, the findings are likely to be generalisable as descriptions of what other L2 teachers can do in other classroom contexts, given the similar array of interactional and linguistic competences as the beginning-level ESOL students in this study.

## Notes

1. In recent CA studies, it is common for CA researchers to conduct multimodal analysis in their studies by including descriptions of non-verbal conduct, and screen-shots of relevant actions captured in the video-recordings (e.g. Sert 2017). As this study looks at how teacher and learners draw on their multimodal resources in the classrooms, we employed a '+' sign to indicate the onset of non-verbal actions. '#' sign was employed for the screen-shots to indicate to the readers the exact locations of the figures in the transcripts.

## Acknowledgements

The authors would like to thank Professor John Hellermann from Portland State University for providing support with our use of The Multimedia Adult English Learner Corpus (MAELC). MAELC was compiled at The National Labsite for Adult ESOL (known locally as the 'Lab School') at Portland State University with support from grant [R309B6002] from the Institute for Education Science, US Department of Education. This was a partnership between Portland State University and Portland Community College. We would like to thank Professor Li Wei, Professor Ernesto Macaro, Dr.

Adam Brandt, Dr. Daniel Lam, Ms. Miaomiao Zuo, Mr. Samuel Tsang and Mr. Kason Cheung for giving us invaluable ideas and input for this paper. We are grateful for the feedback given by members of the Multimodal Analysis Research Group (MARG) at Newcastle University, UK. We gratefully acknowledge the research funds that we received for this study from the University of Oxford Department of Education and St. Cross College. Any remaining errors are our own.



## Appendix A: CA transcription conventions (adapted from Jefferson 2004)

### Sequential and Timing Elements of the Interaction

[	Beginning point of simultaneous speaking (of two or more people)
]	End point of simultaneous speaking
=	Talk by two speakers which is contiguous (i.e. not overlapping, but with no hearable pause in between)
OR	continuation of the same turn by the same speaker even though the turn is separated in the transcript
(0.2)	The time (in tenths of a second) between utterances
(.)	A micro-pause (one tenth of a second or less)

### Paralinguistic Elements of Interaction

wo:rd	Sound extension of a word (more colons: longer stretches)
word.	Fall in tone (not necessarily the end of a sentence)
word,	Continuing intonation (not necessarily between clauses)
wor-	An abrupt stop in articulation
word?	Rising inflection (not necessarily a question)
<u>word</u>	(underline) Emphasised word, part of word or sound
word↑	Rising intonation
word↓	Falling intonation
°word°	Talk that is quieter than surrounding talk
<b>hh</b>	Audible out-breaths
<b>.hh</b>	Audible in-breaths
w( <b>hh</b> )ord	Laughter within a word
>word<	Talk that is spoken faster than surrounding talk
<word>	Talk that is spoken slower than surrounding talk
\$word\$	Talk uttered in a 'smile voice'

### Other Conventions

(word)	Approximations of what is heard
((comment))	Analyst's notes
#	Indicating the exact locations of the figures in the transcripts
+	Marks the onset of a non-verbal action (e.g. shift of gaze, pointing)
<b>XX</b>	Inaudible utterances

## References

- Antaki, C, and I. Leudar. 1992. 'Explaining in conversation: Towards an argument model.' *European Journal of Social Psychology* 22 (2): 181–94.
- Baker, M.J. 2009. "Intersubjective and intrasubjective rationalities in pedagogical debates: Realizing what one thinks." In *Transformation of knowledge through classroom interaction*, edited by B. Schwarz, T. Dreyfus, and R. Hershkowitz, 145–58. London: Routledge.
- Beach, A. 1995. 'Conversation Analysis: "Okay" as a Clue for Understanding Consequentiality.' In *The Consequentiality of Communication*, edited by S. J. Sigman, 121-162. London, Routledge.
- Belhiah, H. 2013. 'Gesture as a Resource for Intersubjectivity in Second-Language Learning Situations.' *Classroom Discourse* 4 (2): 111–129.
- Brouwer, C.E., 2003. 'Word searches in NNSeNS interaction: opportunities for language learning?' *Modern Language Journal* 87: 534-545.
- Chaudron, C. 1982. 'Vocabulary elaboration in teachers' speech to L2 learners'. *Studies in Second Language Acquisition* 4 (2): 170–180.
- Donato, R. 1994. "Collective scaffolding in second language learning." In *Vygotskian approaches to second language research*, edited by J. P. Lantolf and G. Appel, 33–56. Norwood, New Jersey: Ablex Publishing Corporation.
- Jefferson, G. 2004. 'Glossary of transcript symbols with an introduction'. In *Conversation Analysis: Studies from the first generation*, ed. G. Lerner, 14-31. Philadelphia: John Benjamins.
- García, O., N. Flores, and H. Woodley. 2012. "Transgressing Monolingualism and Bilingual Dualities: Translanguaging Pedagogies." In *Harnessing Linguistic Variation to Improve Education*, edited by A. Yiakoumetti, 45–75. Bern: Peter Lang.

Gardener, R. 2001. 'A review of response tokens'. In *When listeners talk: Response tokens and listener stance*, edited by R. Gardener, 13-64. Amsterdam: John Benjamins Publishing Company.

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

Hall, J.K. 1999. "The prosaics of interaction: The development of interactional competence in another language." In *Culture in second language teaching and learning*, edited E. Hinkel, 137–51. Cambridge, UK: Cambridge University Press.

Heller, V. 2016. 'Meanings at hand: Coordinating semiotic resources in explaining mathematical terms in classroom discourse.' *Classroom Discourse* 7 (3): 253-275.

Heritage, J. 1984. *Garfinkel and ethnomethodology*. Cambridge, UK: Blackwell Publishers.

Kim, Y., 2012. 'Practices for initial recognitional reference and learning opportunities in conversation.' *Journal of Pragmatics* 44: 709-729.

Koole, T. 2010. 'Displays of epistemic access: Student responses to teacher explanations.' *Research on Language and Social Interaction* 43 (2): 183–209.

Kupetz, M. 2011. 'Multimodal resources in students' explanations in CLIL interaction.' *Novitas-ROYAL (Research on Youth and Language)* 5 (1): 121-141.

Lazaraton, A. 2004. 'Gestures and speech in the vocabulary explanation of one ESL teacher: A microanalytic inquiry.' *Language Learning* 54 (1): 79-117.

Lilja, N. 2014. 'Partial repetitions as other-initiations of repair in second language talk: re-establishing understanding and doing learning.' *Journal of Pragmatics* 71: 98-116.

Llinares, A. and T. Morton. 2010. 'Historical explanations as situated practice in

Content and Language Integrated Learning'. *Classroom Discourse* 1 (1), 65-84.

Macaro, E. 2009. 'Teacher use of codeswitching in the second language classroom: Exploring 'optimal' use.' In *First Language Use in Second and Foreign Language Learning*, edited by M. Turnbull and J. Dailey-O'Cain, 35-49. Bristol: Multilingual Matters.

Majlesi, A.R. and M. Broth. 2012. 'Emergent learnables in second language classroom interaction.' *Learning, Culture and Social Interaction* 1: 193-207.

Markee, N. 1995. 'Teachers' answers to students' questions: problematizing the issue of making meaning'. *Issues in Applied Linguistics* 6 (2): 63-92.

McNeill, D. 1992. *Hand and Mind*. Chicago, IL: University of Chicago Press.

Mortensen, K. 2011. 'Doing word explanation in interaction.' In *L2 Learning as Social Practice: Conversation-analytic Perspectives*, ed. G. Pallotti and J. Wagner, 135-63. Honolulu: National Foreign Language Resource Center.

Morton, T. 2015. 'Vocabulary explanations in CLIL classrooms: A conversation analysis perspective.' *Language Learning Journal* 43 (3): 256-270.

Merke, S. 2016. 'Establishing the explainable in Finnish-as-a-foreign-language classroom interaction: Student-initiated explanation sequences'. *Learning, Culture and Social Interaction* 9: 1-15.

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

Pekarek Doehler, S. 2010. "Conceptual changes and methodological challenges: on language, learning and documenting learning in conversation analytic SLA research." In *Conceptualising Learning in Applied Linguistics*, edited by P. Seedhouse, P., S. Walsh. and C. Jenks, 105-126. Basingstoke: Palgrave MacMillan.

- Psathas, G. 1995. *Conversation analysis: The study of talk-in-interaction*. Thousand Oaks, CA: Sage.
- Reder, S. 2005. "The 'Lab School'". *Focus on basics: Connecting research and practice* 8: 1-7.
- Roberts, C., M. Baynham., P. Shrubshall., D. Barton., P. Chopra., M. Cooke., R. Hodge., K. Pitt., P. Schellekens., C. Wallace. And S. Whitfield. 2004. *English for Speakers of Other Languages (ESOL): case studies of provision, learners' needs and resources*. NRDC: London.
- Sacks, H. 1992. *Lectures on conversation*. Oxford, England: Blackwell.
- Schmitt, N. 2000. *Vocabulary in language teaching*. UK: Cambridge University Press.
- Sert, O. 2017. 'Creating opportunities for L2 learning in a prediction activity.' *System* 70: 14-25.
- Smotrova, T., and J. P. Lantolf. 2013. 'The function of gesture in lexically focused L2 instructional conversations.' *Modern Language Journal* 97 (2): 397–416.
- Svennevig, J. 2018. 'What's it called in Norwegian? Acquiring L2 vocabulary items in the workplace.' *Journal of Pragmatics* 126: 68-77.
- Tai, K. W. H. and A. Brandt. 2018. 'Creating an Imaginary Context: Teacher's Use of Embodied Enactments in Addressing a Learner's Initiatives in a Beginner-Level Adult ESOL Classroom'. *Classroom Discourse* 9 (3): 244-266.
- Tai, K. W. H. and N. Khabbazzbashi. 2019. 'The Mediation and Organisation of Gestures in Vocabulary Instructions: A Microgenetic Analysis of Interactions in a Beginning-level Adult ESOL Classroom'. *Language and Education* 1-24. Epub ahead of print.

- ten Have, P. 1990. 'Methodological issues in conversation analysis'. *Bulletin de Méthodologie Sociologique* 27 (1): 23-51.
- Tian, L. and E. Macaro. 2012. 'Comparing the effect of teacher codeswitching with English-only explanations on the vocabulary acquisition of Chinese university students: A lexical focus-on-form study.' *Language Teaching Research* 16 (3): 361 – 385.
- van Compernelle, R.A., and T. Smotrova. 2017. 'Gesture, meaning, and thinking-for-teaching in unplanned vocabulary explanations'. *Classroom Discourse* 8 (3): 194-213.
- Wald, B. 1978. Zur Einheitlichkeit und Einleitung von Diskurseinheiten. In: Quasthoff U (ed.) *Sprachstruktur – Sozialstruktur*. Königstein: Scriptor, pp. 128–149.
- Walsh, S. 2012. 'Conceptualising classroom interactional competence.' *Novitas-ROYAL (Research on Youth and Language)* 6 (1): 1–14.
- Waring, H.Z., C.C. Creider and C.D. Box. 2013. 'Explaining vocabulary in the second language classroom: a conversation analytic account.' *Learning, Culture, and Social Interaction* 2 (4): 249–64.
- Young, R.F. 2009. *Discursive practice in language learning and teaching*. Oxford, UK: Wiley-Blackwell.

**Kevin W. H. Tai** is Economic and Social Research Council (ESRC) Scholar and PhD Candidate in Applied Linguistics in the UCL Centre for Applied Linguistics at the UCL Institute of Education, University College London (UCL). He holds an MSc in Applied Linguistics and Second Language Acquisition from the University of Oxford and a BA (Honours) degree in English Language and Literature with First Class Honours from Newcastle University, UK. Kevin is CELTA-qualified and his research interests include second language teaching and learning, sociocultural theories, classroom discourse, translanguaging practices in multilingual contexts, qualitative research methods and language teacher education. His research has appeared in international peer-reviewed

journals such as *Classroom Discourse* (2018), *Language and Education* (2019), and *Linguistics and Education* (2019).

**Nahal Khabbazzashi** is Senior Lecturer in Language Assessment at the Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire. She holds a DPhil in Education and an MSc in Applied Linguistics and Second Language Acquisition from the University of Oxford. Her work appears in journals such as *Language Testing* (2017) and *Language Learning Journal* (2014), as well as in the *Studies in Language Testing* series (Moeller, Creswell, and Saville, eds. 2016).