Explaining different aspects of word knowledge: A multimodal analysis of lexical explanation sequences during online French tutoring sessions

Benjamin Holt

Lexis
Journal in English Lexicology

18 | 2021
Lexical learning and teaching

Explaining different aspects of word knowledge: A multimodal analysis of lexical explanation sequences during online French tutoring sessions

Benjamin Holt

Electronic version
URL: https://journals.openedition.org/lexis/5889
ISSN: 1951-6215

Publisher
Université Jean Moulin - Lyon 3

Electronic reference
Benjamin Holt, "Explaining different aspects of word knowledge: A multimodal analysis of lexical explanation sequences during online French tutoring sessions", *Lexis* [Online], 18 | 2021, Online since 18 December 2021, connection on 19 December 2021. URL: http://journals.openedition.org/lexis/5889

Explaining different aspects of word knowledge: A multimodal analysis of lexical explanation sequences during online French tutoring sessions

Benjamin Holt

We are grateful to the University of Lyon's ASLAN project (ANR-10-LABX-0081) for its financial support within the French government's 'Investissements d'Avenir' (ANR-11-IDEX-0007) program operated by the National Research Agency (ANR).

Introduction

1 Lexical acquisition is one of the core components of foreign language learning. The different ways of theorizing and conceptualizing lexical items and their acquisition, as well as the pedagogical recommendations that follow, contribute to an ongoing debate (Bogaards [1994]; Galisson [1991]; Grossmann [2011]; Nation [2013]; Tréville [2000]; Tréville & Duquette [1996]). Interactive, contextualized word learning is known to play a critical role in language acquisition, especially during the earlier stages and when contact with the language is limited (Hilton [2019]). While L2 lexical explanation sequences have been analyzed in fine detail in pedagogical (Fasel Lauzon [2014]) and other face-to-face settings (De Gaulmyn [1991]; Gülich [1991]; Lüdi [1991]; Miecznikowski [2005]), few studies have focused on lexical acquisition in online videoconferencing environments. This article aims to describe some of the ways in which desktop videoconferencing can be used by teachers to explain L2 lexical items.

2 Desktop videoconferencing is an increasingly popular tool for foreign language teaching and learning, and numerous studies have endeavored to better understand its peculiarities and potential as a tool for computer-assisted language learning.

Videoconferencing is useful for many things, including enriching intercultural competence (O'Dowd [2000], [2005]), facilitating spontaneous L2 production while lowering anxiety (Develotte, Guichon & Kern [2008]), and enhancing international mobility programs (O'Dowd [2016]). However, the pedagogical and technical skills necessary to teach effectively in this type of environment differ from those typically mobilized in face-to-face contexts and require specialized training and practice to acquire (Develotte & Mangenot [2007]; Guichon & Drissi [2008]). It is for these reasons that videoconferencing as a tool for teaching L2 oral skills has engendered such intense interest among researchers (Guichon & Tellier [2017]).

3 Desktop videoconferencing environments are semiotically rich in that they offer their users many affordances (Blin [2016]). In this article we take interest in the ways in which French language tutors use the audio, visual and textual affordances at their disposal to focus on different aspects of word knowledge and to involve their learners in the lexical explanation sequences. First, we will offer a brief overview of foreign language word learning and of multimodal interaction. Then, we will present our corpus and methodology. Finally, we will present our results with descriptive statistics and examples.

1. Theoretical framework: Multimodal lexical explanations

1.1. Teaching and learning lexical items in a foreign language

4 According to conservative estimates, adult speakers know upwards of twenty thousand word families[1] in their native language (Nation [2013]). Many of these words are learned incidentally, meaning that word learning was not the main purpose of the activity (Hulstijn *et al.* [1996]; Grossmann [2011]). Although this type of learning is advantageous (words learned in this manner are necessarily contextualized), this method is only feasible for advanced learners who already know lots of words (Folse [2004]). Explicit teaching and deliberate learning of the most frequently encountered words are therefore necessary during the earlier stages of language acquisition and remain so until the learner develops a rich enough vocabulary for incidental learning to take place.

5 Due to the complexity of lexical morphology, there are many ways of defining a word or lexical item (Mortureux [2008]). For example, a *lexical unit* (Bogaards [1994]) can be defined as the combination of a determined form and a stable meaning. However, this term only covers one definition of a word at a time, such that each of the many definitions of highly polysemous words such as *put*, *run*, and *set* would be counted as discreet *lexical units*. According to the polysemous approach that we adopt, knowing a word means knowing its different definitions and being able to select the appropriate one depending on the context. We therefore favor the term "*lexical item*" to designate words such as *coffee*, collocations such as *coffee break*, and idiomatic expressions composed of indivisible chunks of words such as *look a gift horse in the mouth*.

6 Opinions differ on what it means to *know* a lexical item. Some consider that knowing a lexical item means knowing two sides of a coin: the *form* (signifier) on one side and the *meaning* (signified) on the other (Van der Linden [2006]). Others envision word knowledge as a scale ranging from weak to strong (Waring [2002]) or from receptive to

productive (Tréville & Duquette [1996]). We adopt the classification proposed by Nation [2013] whereby word knowledge covers three main *aspects*: form, meaning and use. A word's *form* includes its spoken, written and word parts. A word's *meaning* includes its form-meaning connection, its concepts and referents, and its associations. Finally, a word's *use* includes its grammatical functions, collocations, and constraints on use such as register and frequency. Nation divides each of these subcategories into productive and receptive knowledge. We will adopt this classification to study the ways in which these different aspects of lexical items are explained by French tutors.

7 In addition to word knowledge, we will focus on the ways in which the tutors ensure the learners' active involvement in the explanation sequences. Fasel Lauzon [2014] notes that explanation sequences with the highest potential for problem solving and word learning are those in which the learners are encouraged to actively participate in the explanation sequence, and to reuse the lexical item that is being explained. Learners' active involvement is important because it favors deeper mental processing than simply recognizing a word as the correct response out of a list (Stahl & Vancil [1986]), leading to increased memorization. Learners' cognitive engagement in lexical explanation sequences was modelized by Laufer & Hulstijn's [2001] *involvement load hypothesis*, and elaborated on by Nation & Webb [2011]. According to this hypothesis, cognitive engagement is measured in terms of *need*, *research* and *evaluation* of lexical items. Socio-affective engagement must also be maintained (Celik [2008]; Mangenot [2011]), because it is a prerequisite of cognitive engagement (Jézégou [2010]). Therefore, in addition to analyzing the ways in which tutors focus on each of the three aspects of knowing a word, we will examine the ways in which they encourage learners' active involvement in the explanation sequences.

1.2. Explanation sequences and negotiation of meaning

8 One way of analyzing lexical explanation sequences is to consider them as negotiation sequences. It has been known for the past forty years that negotiation of meaning favors language acquisition, because it allows for the noticing of linguistic gaps, focus on form, focus on meaning, modified input, increased comprehension, and active participation (Ellis, Tanaka & Yamazaki [1994]; Hatch [1978]; Long [1981], [1983]; Pica [1994]). Negotiation sequences take place when the normal flow of conversation is halted or *pushed down* in order to address some language-related issue. These usually arise from a *trigger*, a linguistic element that is not understood by an interlocutor. Videoconferencing environments are particularly fertile grounds for negotiation sequences, sometimes even more so than synchronous or asynchronous text chat settings (Sarré [2011]). It is therefore interesting to study what multimodal resources online tutors use during negotiation sequences that are triggered by non-understanding (Holt, Tellier & Guichon [2015]). As it turns out, research has repeatedly shown that lexical items are the most common type of trigger for negotiation sequences, both in face-to-face (Fasel Lauzon [2014]) and in online interactions (Nicolaev [2010]; Renner [2017]).

9 When lexical explanation sequences are analyzed as negotiation sequences, it is common to analyze their unfolding into separate phases by defining the trigger, indicator, response, and reaction to the response (Varonis & Gass [1985]). This is the approach adopted by De Gaulmyn [1991], Gülich [1990] and Lüdi [1991] who proposed a

three-step model[2] for describing lexical explanation sequences, consisting of an opening phase, a core explanation phase, and a closing phase. This model was reimplemented by Fasel Lauzon [2008], [2014] to describe lexical explanation sequences that took place in a classroom setting, and by Wigham [2017] to analyze them in a videoconferencing environment.

10 Rather than breaking down the lexical explanation sequences into phases according to the model above and analyzing each one (as well as the transitions between them), we are interested in the ways in which tutors mobilize the multimodal resources at their disposal to focus on the different aspects of knowing a word. In addition, as mentioned above, we are interested in the ways in which tutors encourage learners' active involvement in the explanation sequences. In the next section, we will review the different multimodal resources that are available to tutors for carrying out various pedagogical interventions including lexical explanation.

1.3. The multimodality of videoconferencing interaction

11 The multimodal approach that we adopt for this study is inspired by mediated discourse analysis in which the unit of analysis is the *action* that is accomplished by means of images, gestures, sounds, and other *modes* (Jones [2014]). For our purposes, the *action* is focusing on a particular aspect of knowing a word, or involving the students in an explanation sequence. We are interested in the *multimodality* behind these actions, which has been defined by Kress & Van Leeuwen [2001: 20] as:

> the use of several semiotic modes in the design of a semiotic product or event, together with the particular way in which these modes are combined – they may for instance reinforce each other ('say the same thing in different ways'), fulfill complementary roles [...], or be hierarchically ordered [...].

12 Kress [2014: 60] defines a mode as:

> a socially shaped and culturally given resource for making meaning. *Image*, *writing*, *layout*, *music*, *gesture*, *speech*, *moving image*, *soundtrack* are examples of modes used in representation and communication.

13 The words *hierarchically ordered* are of utmost importance, because from a multimodal perspective, spoken language is part of a multimodal ensemble and is not automatically the most important mode at all times (Jewitt [2014]). Each mode plays a different role and has a different level of importance depending on the action that is being completed. For example, during a face-to-face conversation, the verbal mode could be the most important during the first half of a sentence when a speaker says *Today I saw a raccoon...* only to forfeit its status as most important mode to gestures when the speaker adds *who was going like this...* while gesturing a digging motion with their hands. Gesture and speech have no doubt coexisted since the beginnings of spoken language, but thinking of gesture as a discreet mode that is complementary to speech is a relatively recent development (Kendon [2004]; McNeill [1992]).

14 Although it can be argued that every interaction is multimodal in the sense that voice, gestures and facial expressions contribute to the meaning of a message (Goldin-Meadow [2003]; Norris [2004]), the multimodal nature of human interaction becomes conspicuous and unavoidable when it is mediated by a multi-channel (Lancien [1998]) tool such as videoconferencing. Most videoconferencing platforms allow the use of a microphone, a webcam, and a keyboard. In addition, many platforms – including the one used in our study – allow users to send multimedia files and documents. Learning

to juggle between the different modes that online platforms offer is part of what Cappellini & Combe [2017] call *techno-semio-pedagogical competence*. Different modes will be used depending on the immediate communicative context and the pedagogical regulation (Guichon & Drissi [2008]) being carried out, and the concurrent use of the available tools and affordances requires the user to focus on many things simultaneously (Develotte & Drissi [2013]). Following in the footsteps of Bange [1992] who mentions *bifocalization* – on the activity and on the language – here we are dealing with *trifocalization* (Holt [2020]) – on the activity, on the language, and on the manipulation of the videoconferencing tools.

15 Synchronous text chat, which was used alone for L2 tutoring before videoconferencing became widely available, is beneficial to negotiation of meaning (Smith [2003]). In videoconferencing environments, it allows tutors to give feedback to learners without interrupting them (Dejean-Thircuir, Guichon & Nicolaev [2010]; Guichon & Drissi [2008]), and the chat history is permanent, unlike ephemeral spoken language (Develotte et al. [2008]). When used in conjunction with speech, the keyboard can be used to focus on the written form of a word, thereby increasing the modes by which the learner is exposed to it (Mayer [2001]) and facilitating subsequent recall (Nation & Webb [2011]). The keyboard can also be used as a backup mode during technical breakdowns (Azaoui [2017]) or to solicit responses from learners.

16 The microphone allows participants to use verbal and vocal strategies to focus on different aspects of lexical items and to encourage active involvement in explanation sequences. For example, teachers can use pauses, emphasis, intonation, and logonymes such as *term* (Rey-Debove [1997]) to isolate and highlight lexical items. Other linguistic strategies for highlighting words include dislocation (Horlacher [2015]) and lemmatization (Mortureux [2008]). Words can also be introduced by formulations such as *we say* (Cicurel [2011: 70]), just as definitions can be introduced by *it's when* (Fasel Lauzon [2014: 144]). Definitions can of course be given orally, either in L1 or in L2 (Nation [2013]), formally or informally (Flowerdew [1992]).

17 The webcam has many uses. Not only can it increase the perception of online presence (Guichon & Cohen [2014]), but it can also increase language comprehension thanks to visible gestures (Sueyoshi & Hardison [2005]). Although most gestures are produced outside the webcam's field of view (Guichon & Wigham [2016]), webcam users seem to produce more gestures when they know that an interlocutor can see them (Mol, Krahmer, Maes & Swerts [2011]). During lexical explanation sequences, tutors can use the webcam to mark a word by using facial expressions such as raised eyebrows, and they can use exaggerated mouth movements to explain the proper pronunciation of a word (Holt [2020]). Teaching gestures (Tellier [2008]; Tellier & Cadet [2014]) are especially useful for explaining words in a foreign language, and there tends to be an inverse relationship between the size, duration and iconicity of a gesture and the learner's perceived level of proficiency (Lazaraton [2004]; Tellier & Stam [2012]). Previous studies on videoconferencing have shown that iconic and metaphoric gestures are often used in addition to verbal explanations to focus on the referential meanings of words (Holt [2020]; Holt et al. [2015]; Wigham [2017]).

18 In this article, we will analyze the ways in which tutors use the four main videoconferencing resources at their disposal – the microphone, the webcam, the keyboard, and the sending of documents and Internet links – in order to focus on the three main aspects of knowing a word: form, meaning and usage. In addition, we will

look at the ways in which tutors use these resources to ensure active learner involvement in the explanation sequences.

2. Methodology

19 Our data are from the ISMAEL corpus (Guichon & Tellier [2017]) which connected 12 French tutors with 18 learners during the fall semester of 2013. Each tutor was connected with one or two learners, and these two- or three-person groups remained mostly constant throughout the semester. The tutors were enrolled in a professional master's program at the University of Lyon and had to collaboratively prepare and execute six weeks of online interactions. The learners were undergraduate business students enrolled at Dublin City University who needed to work on their B1-level of French in order to complete an internship the following year in France. The six interactions, each lasting 30-40 minutes on the videoconferencing platform *Visu* (Guichon, Bétrancourt & Prié [2012]), were composed of several pedagogical activities based on business-related themes such as work culture, event planning, startup pitches, and job interviews. The main objective of the online courses was to improve oral skills, and vocabulary learning was not the main focus of the semester. The lexical explanation sequences therefore emerged organically from the interactions, except for a handful of sequences based on keywords that were so central to the activity that they were included in the lesson plans.

20 All of the participants were asked to sign consent forms in order to be included in the study. Those who declined to consent, as well as their partners, were excluded from the corpus. The final corpus therefore includes seven tutors and twelve learners. Fifteen and a half hours of interaction were recorded and transcribed using the ELAN multimodal annotation software (Wittenburg, Brugman, Russel, Klassmann & Sloetjes [2006]) and the ICOR transcription convention (Groupe ICOR [2007]).

21 Following transcription, we searched the corpus for lexical explanation sequences which we defined as sequences in which one or several aspects (Nation [2013]) of knowing a word were brought up, questioned, discussed, or explained by the participants. A lexical explanation sequence can result from a difficulty in comprehension, where a lexical item is not understood by the learner, or from a difficulty in production, where the learner cannot find the correct word (Grossmann [2011])[3]. After viewing our entire corpus several times, we identified 295 lexical explanation sequences that we divided into phases for preliminary analysis. Then, we annotated the multimodal resources used by the tutors to focus on the three aspects of knowing a word (form, meaning and use) and to involve the learners in the explanation sequences.

22 This article follows three other studies that have been carried out on lexical explanation sequences from the ISMAEL corpus. Holt & Tellier [2017] studied four tutors during the same session who, while following the same lesson plan, all explained the word *bobo*[4]. The multimodal analyses highlighted differences regarding the unfolding of the explanation sequences into phases, the different aspects of the word that were selected for explanation, and the timeliness and pedagogical effectiveness of each explanation. Wigham [2017] focused on fourteen lexical explanation sequences that arose from six sessions (three tutor-learner groups over two weeks), and her analyses shed light on the different multimodal behaviors that characterize each of the

three phases. Finally, Holt [2020] analyzed all 295 lexical explanation sequences from the ISMAEL corpus, while focusing exclusively on hand gestures and their various pedagogical functions. The results in the next section will also focus on all 295 lexical explanation sequences from the corpus, but will consider all of the multimodal resources (text chat, microphone, webcam, and document sharing) and not just hand gestures. Similar to the previous study and unlike the first two cited, we are more interested in resource use than in sequence structure.

3. Results

23 In this section, we will measure the use of each of the four main resources that *Visu* makes available to tutors: the microphone, the text chat, the webcam, and the sending of documents and links to online resources. We will measure them with respect to how they are used to focus on form[5], to focus on meaning, to focus on use, and to ensure active learner involvement.

3.1. Focus on form

24 Our analyses reveal that tutors focused on the form of lexical items to correct pronunciation and spelling, to enrich learners' productive vocabularies, to distinguish between phonologically similar words, and to increase the perception of words provided by tutors following a difficulty in production, thereby connecting the form of a lexical item to its meaning. Figure 1 lists the number of explanation sequences in which each strategy was used, out of 295. It should be noted that a sequence can exhibit more than one strategy.

Figure 1. Resources used to focus on the form of a lexical item

Channel	Method of focus on form	Number of sequences
Microphone	Logonyms	152
	Dislocations	57
	Vocal emphasis	27
Keyboard	Written repetition	98
Webcam	Facial expressions and head movements	47
	Looking directly into the camera	5
	Hand gestures	12

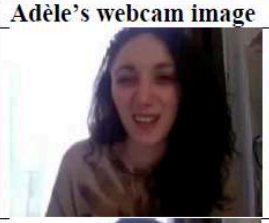
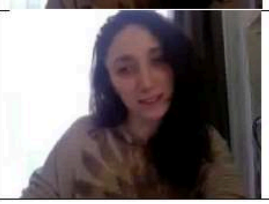
25 Tutors most often focused on the forms of lexical items by verbal means, using the microphone. Logonyms (Marra & Pallotti [2006]) such as *speak*, *word*, *say*, *French*, *called*, etc. were used by the tutors in just over half of the explanation sequences. Dislocations[6] were also used in one in ten sequences. Finally, vocal emphasis (prosodic markers such

as rising intonation, falling intonation, or increased volume) was used in one in nine sequences to introduce a lexical item.

26 The text chat was used in one in three explanation sequences to repeat and reinforce the forms of lexical items that were provided orally. What might be considered redundant – and therefore useless, or even harmful for memorization among native speakers (Sweller [2005]) – is not so for a learner who must memorize both the written and spoken forms of a new L2 word (Nation [2013]). These *intermodal repetitions* allow learners to gain access to both at the same time. The tutors' use of the text chat was often accompanied by meta-semiotic comments made orally (cf. Example 7), for example *I'm writing it in the chat window*, which reinforced the salience of the written form. The text chat was found to be particularly useful for resolving partial incomprehension of transparent words such as *equivalent*.

27 Visual resources were used by tutors to focus on lexical items' forms in one out of six explanation sequences. In eight sequences, exaggerated mouth movements to clarify the pronunciation of words were observed. Visible hand gestures, such as air quotes to highlight a word or gestures for marking different syllables of a word, were observed in 12 explanation sequences. In the following example, Adèle uses three multimodal resources at her disposal to draw attention to a lexical item whose difficulty she anticipates, and to gauge the necessity of an explanation. In the middle of an explanation sequence of the word *infusions*, she anticipates the incomprehension of the word *thym* (thyme) and halts the sequence to focus on its form.[7]

Example 1. Anticipated incomprehension of the word "thym"

Speech transcription		Adèle's webcam image	Observations
1 ADL	tu prends du **thym (0.5)** thym/ (0.3) euh: (1.3) thym/ (0.5) <((text chat)) thym>		Facial expression of uncertainty
1 ADL	you take **thyme (0.5)** thyme/ (0.3) uh: (1.3) thyme/ (0.5) <((text chat)) thyme>		
	(1.2)		Head tilted to the side while awaiting a response
2 ALN	**"thyme yeah"** (0.3)		
3 CTR	"thyme"		

28 By repeating the word *thyme* three times orally and by using an intermodal repetition via the text chat, Adèle directs the two learners' attention towards this particular lexical item. Her facial expression shows that she anticipates a comprehension difficulty caused by this lexical item, and uses three modes (voice, text chat and facial expressions) in a complementary manner in order to verify the comprehension. She fills the 1.2-second pause (second screenshot) with a new facial expression, this time accompanied by a head tilt, to communicate to the learners that she is checking their comprehension of the lexical item. This multimodal comprehension check is effective because the learners respond by producing the word in English (lines 2-3), thus affirming their comprehension.

3.2. Focus on meaning

29 There are many ways of focusing on a lexical item's meaning, such as performing demonstrative acts, showing objects or images, giving definitions in the learner's L1 or in the target language, or providing contextual information (Nation [2013]). Tutors can also use gestures (Holt [2020]) or send images (Wigham [2017]) to explain words. The table below lists the number of lexical explanation sequences in which each multimodal resource was used to focus on a lexical item's meaning.

Figure 2. Resources used to focus on lexical items' meanings

Channel	**Method of focus on meaning**	**Number of sequences**
Microphone	Oral translations	24
	Definitions, synonyms, examples, and reformulations	144
Keyboard	Written translations	7
Webcam	Hand gestures, facial expressions and head movements	48
Images	Sending images	7

30 Our annotations reveal that verbal explanations, provided via the microphone, were the most common. Translations, both oral and written, were used by tutors in roughly one out of ten explanation sequences. Tutors provided oral definitions in nearly one out of two sequences, and visual resources were used in one out of six sequences to give ostensive explanations through the use of gestures, facial expressions and images.

31 The following sequence emerged from a discussion about the attitudes, behaviors and rights of workers in France and in Ireland. This example showcases many strategies that the tutors employed to focus on the meanings of lexical items. Victor, the tutor, asks Liam if Irish workers go on strike in the same way that French workers do. Liam's incomprehension of the word *grève* triggers an explanation sequence.

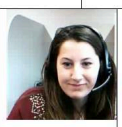
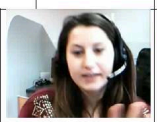
Example 2. A lexicographical definition accompanied by gestures

Speech transcription		Victor's webcam image	Observations
1 VIC	est-ce qu'ils font euh souvent la grève/ ou euh\ (0,2) pas trop\ pas autant qu'en France sans doute\ (1,2)		Victor touches his chin as he checks his learner's comprehension
2 LIM	euh: pardon/ (0,7)		
3 VIC	ils font pas souvent la grève\ (.) j'imagine\ (1,8) tu sais ce que **[c'est la grɛ]ve/**		
4 LIM	**[euh:] (1,3) euh non (0,4)**		
5 VIC	**ah d'accord\ (0,4)**		
1 VIC	do they go on strike often/ or uh\ (0.2) not that much\ certainly not as often as in France\ (1.2)		
2 LIM	uh: what/ (0.7)		
3 VIC	they don't go on strike much\ (.) I imagine\ (1.8) do you know what **[a strike is]/**		
4 LIM	**[uh:] (1.3) uh no 0.4)**		
5 VIC	**ah okay\ (0.4)**		
6 LIM	c'est q[uand les gɛ]ns/ (.)& [°euh°]		Iconic gesture to illustrate the act of chanting slogans
7 VIC	&arrêtent de de travailler/ (.) et ils sortent **dans la rue/ pour scander des (.) des** slogans:/ et\ (1,0) (et) vouloir de: de meilleurs conditions de travail\ (0,6) faire la grɛ[ve]\		
8 LIM	[o]k- (1,2)		
6 LIM	it's [when people]/ (.)& [°uh°]		
7 VIC	&stop working/ (.) and they go **out in the street/** to chant (.) slogans:/ and\ (1.0) ask for: for better working conditions\ (0.6) to go on [strike]		
8 LIM	[oka-] (1.2)		
9 VIC	en anglais/ **je ne sais pas ce que c'est donc euh**\ (1,7) mais tu: tu comprends\ (1,1)		Raised eyebrows and "thinking face" [Goodwin & Goodwin 1986]
10 LIM	oui e[uh je pɛ]nse/ que\		
11 VIC	[d'ac-] (2,6)		
12 LIM	hum (5,3)		
13 VIC	oh euh: d'accord\ donc t'as répondu/ ça va/		
9 VIC	in English/ **I don't know what it is so uh**\ (1,7) but you: you understand\ (1.1)		
10 LIM	yes [uh I think/] that\		
11 VIC	[oka-] (2.6)		
12 LIM	um (5.3)		
13 VIC	oh uh: okay\ you've responded/ so that's good/		

32 When the word *grève* is identified as the trigger (Varonis & Gass [1985]), Victor gives a lexicographical explanation accompanied by iconic gestures to illustrate the act of chanting slogans. What he says at line 9 reveals his pedagogical belief that a translation would be the best way of resolving incomprehension. Indeed, tutors resorted to translations in over 10% of the lexical explanation sequences that we analyzed. After his explanation, Victor checks his learner's understanding and accepts the latter's tepid response without any real verification. As Cicurel [2011] points out, it is incumbent upon the learner to provide proof that he or she has internalized the new information. Later on, during this same conversation, Liam asks for another explanation of the word *grève*, indicating that the initial explanation was ineffective. Victor gives a second verbal explanation[8] with an example and hand gestures, but it is not until the very end, when Victor finally remembers the English translation, that the incomprehension is resolved. This is a case in which a definition and an example, each one accompanied by gestures, are clearly less effective in quickly resolving incomprehension than a simple translation. It seems as though the videoconferencing environment does not change the fact that translations are the simplest (Nation [2013]) and most effective (Laufer & Shmueli [1997]) method of resolving incomprehension and learning new words in a foreign language.

33 Sending images is another visual resource that tutors can use to help focus on the meanings of lexical items. This resource was used seven times throughout our entire corpus, which represents only 2.4% of all of the lexical explanation sequences observed. In the following example, the participants are brainstorming a menu for a local food truck in Lyon when Samia, the tutor, sends an image of a *quenelle*, a dish for which the city is well known.

Example 3. High modal density and weak multimodal flexibility while sending an image

Speech transcription		Participants' webcam images and image sent	Observations
1 SAM 2 ANG 3 SAM 1 SAM 2 ANG 3 SAM	vous savez ce que c'est des quenelles/ euh: (0,2) oui (0,9) c'est avec (0,3) de la soupe/ (4,5) non/ (10,7) **<((image))>** do you know what quenelles are/ uh: (0.2) yes (0.9) it's with (0.3) soup/ (4.5) no/ (10.7) <((image))>		Samia sends a photo of a quenelle
4 SEN 5 SAM 6 SEN 4 SEN 5 SAM 6 SEN	(1,3) je vous ai envoyé une image/ pour que vous puissiez voir ce que c'est une quenelle\ **(6,4) ((rire))** (11,4) d'accord (0,3) les comme [les]& [d'accord/] &euh: (1,3) pommes de terre/ (0,6) (1.3) I've sent you an image/ so that you can see what a quenelle is\ **(6.4) ((laughter))** (11.4) okay (0.3) it's [like]& [okay/] &uh: (1.3) potatoes/ (0.6)	 	Samia smiles slightly after sending the photo; the learners look at it
7 SAM 7 SAM	oui (0,3) euh: ouais (0,4) sauf que c'est comme euh: (0,3) c'est pas vraiment des pommes de terre\ hein c'est pas des pommes de terre du tout même/ (0,4) c'est hum: (1,7) des gros pff c'est pas des pâtes/ (1,4) hu:m (0,3) -fin bon ce ça **ressemble** à ce qu'il y a sur l'image\ (0,8) ((rire)) yes (0.3) uh: yeah (0.4) except that it's like uh: (0.3) it's not really potatoes\ uh it's not even potatoes at all/ (0.4) it's um: (1.7) it's big uh it's not pasta/ (1.4) um (0.3) well anyway it **looks like** what's in the photo\ (0.8) ((laughter))		Barely visible pointing gesture

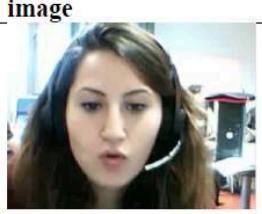
34 Samia uses the text chat to send two lexical items: *tarte aux pralines* (praline tart) and *quenelle*, but only checks the comprehension of the latter, perhaps because of the image that she had prepared to send. The tutor appears to be satisfied with this resource and unprepared to offer a verbal explanation of the meaning of the word *quenelle*, because she says *well anyway it looks like what's in the photo* (line 7). This meta-semiotic comment is accompanied by a pointing gesture that is barely visible on the screen. Here, the image-sharing channel attains a high modal density (Norris [2004]) because the sending of the image is followed by 18 seconds of silence and receives verbal commentary.

35 The modal density of the image-sharing channel is so high in this sequence that the explanation would be all but impossible without it, as the tutor seems unable to switch between modes. Building off of Cicurel's [2005] notion of *communicative flexibility* and Develotte et al.'s [2008] *pedagogical flexibility*, we propose the term *multimodal flexibility* to refer to the ability of online tutors to use and switch between different modes to complete a pedagogical action such as explaining a lexical item. This skill is useful during unplanned technical breakdowns (Azaoui [2017]) when a channel becomes unavailable, or during certain pedagogical situations when a given mode is ineffective due to lack of comprehension. In Example 3 above, the multimodal flexibility is relatively low because the tutor is unable to complement the sending of the photo with an explanation in another mode. Analysis of other lexical explanation sequences in the ISMAEL corpus has revealed that multimodal flexibility is not always low when sending images (see Wigham [2017: 101]).

36 Aside from sending images or using the webcam, it is also possible to give ostensive explanations by using the microphone. When the hardware and software are working properly, each sound made by the mouth is detected and amplified. This can be used

for pedagogical purposes, such as in the following example, where Samia uses the microphone to give a nonverbal explanation of the word *relaxation*.

Example 4. A sonorous ostensive explanation

Speech transcription		Samia's webcam image	Observations
1 SAM	pour moi c'est un moment de détente par exemple (3,3)		Samia blows into the microphone which amplifies the sound
2 ANG	un moment de: quoi/		
3 SAM	de DÉTENTE (0,5) où on se détend où on est euh (0,6) **((soufflement))** où on souffle (1,5)		
4 ANG	[d'accord]		
5 SAM	[où on a]rrête de travailler\ (1,6) où on se repose (2,5) un moment de détente		
1 SAM	for me it's a moment of relaxation for example (3.3)		
2 ANG	un moment of: what/		
3 SAM	of RELAXATION (0.5) where you relax or you uh (0.6) **((blowing))** where you take a breather (1.5)		
4 ANG	[okay]		
5 SAM	[or where you] stop working\ (1.6) and you rest (2.5) a moment of relaxation		

37 When Samia explains the act of relaxing, she blows into her microphone, which captures and amplifies the sound of the moving air. This sound is louder and more perceptible than it would have been in an open classroom with no microphone. This particular setup, with a webcam and a microphone headset, is well suited to this type of sonorous nonverbal explanation.

3.3. Focus on use

38 The tutors in our study focused on use in only 43 of the 295 lexical explanation sequences, which comes out to slightly less than one in six. The table below summarizes the number of sequences in which each channel was used to focus on use. The total number is higher than 43 because in many sequences, multiple channels were used.

Figure 3. Strategies and channels used to focus on use

Channel	Method of focus on use	Number of sequences
Microphone	Grammar	13
	Collocations	16
	Connotation and cultural charge	7
	Register and frequency	12
Keyboard	Grammar	3
	Collocations	6

	Register	1
Webcam	Grammar	1
	Collocations	2
	Connotation	1
	Frequency	1

39 Most of the focus on use was carried out using the microphone, which was used to give collocations, correct grammar, and provide information concerning a lexical item's register, frequency and cultural charge (Galisson [1991]). As mentioned earlier, some culturally charged words were part of the lesson plan and were therefore explained by the majority of the tutors. Holt & Tellier [2017: 78] showed that connotation was mentioned orally, as *bobo* is a word that *has become a bit derogatory* in France according to one tutor. Frequency can also be conveyed through the webcam, such as with an upward hand gesture to signal that a word *is very useful in France because people use it a lot* (Holt [2020: 20]).

40 The following excerpt exemplifies the use of the microphone, accompanied by a hand gesture, to focus on the grammatical restrictions of a word's use. This part of the explanation sequence arrived after the student incorrectly used a lexical item that the tutor had just provided.

Example 5. Focusing on a word's usage

Speech transcription		Adèle's webcam image	Observations
1 ADL	euh loyaux c'est le pluriel\ (.) si tu dis euh (0.4) cette personne est loyale/ (.) les français sont **loyaux**\		Pointing gesture towards the screen
1 ADL	uh loyal is the plural form\ (.) if you say uh (0.4) this person is loyal/ (.) French people are **loyal**\		

3.4. Actively involving the learner

41 We observed cognitive (Nation & Webb [2011]) or socio-affective (Jézégou [2010]) involvement in more than half of the 295 lexical explanation sequences that we observed. The table below counts the different strategies put in place by the tutors to ensure such involvement. Spatial constraints prevent us from listing the different modes that were used for each of these strategies, but the complete statistics are available in Holt [2018].

Figure 4. Strategies used by the tutors to involve the learners

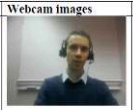
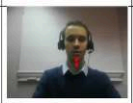
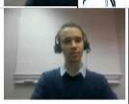
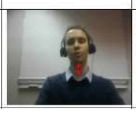
Engagement strategy	Number of sequences
Checking the comprehension of a lexical item	91

Asking the student to translate a lexical item into his/her native language	7
Triggering productive recall	4
Bringing a problematic element to the learner's attention	22
Asking the learner to provide a lexical explanation to the tutor	28
Asking the learner to evaluate the definition a word	7
Asking the learner to evaluate the appropriateness of a proposed word	55
Asking the learner to repeat the correct form of a word	15
Asking the learner to provide a lexical explanation to his or her partner	8
Making the learner carry out an interlinguistic reflection	1
Asking the learner to carry out an intercultural reflection	6
Personalizing a definition	8

42 Productive and receptive recall are known to facilitate acquisition of lexical items (Nation [2013]). Receptive recall takes place when the tutor provides the form of a lexical item and the learner must search for the meaning, and productive recall takes place when the meaning is provided without the form. In our corpus, triggering receptive recall was one of the most widely used strategies of cognitively involving the learners. Most often, it was initiated orally by the tutor and was accompanied by moving closer to the screen, raising eyebrows, or making a facial expression of uncertainty or a *thinking face* (Goodwin & Goodwin [1986]). Less often, the text chat was used to provide the learner with a visual representation of the written form in order to facilitate search for meaning. On two occasions, three channels were used together to trigger receptive recall (cf. Example 1).

43 In the following example, Victor triggers productive recall of the lexical item *chaise musicale*.

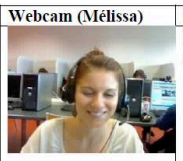
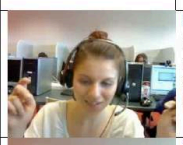
Example 6. Triggering productive recall

Speech transcription		Webcam images	Observations
1 VIC	**est ce que tu connais le jeu/ où hum tu sais il y a une musique (.) qui joue/ (0,6) et y a un nombre** limité de: de chaises/ (1,1)		Thinking face
1 VIC	do you know the game/ where um you know there's music (.) that plays/ (0,6) and there's a limited number of: of chairs/ (1,1)		
2 LIM	oui oui oui		Three head nods and a smile
	(0,3)		
3 VIC	**d'accord**		
4 LIM	**euh:** (0,2) c'est le: même chose que (0,8) hum: (1,6) euh j- j'ai pensé sur le: même chose		
2 LIM	yes yes yes		
	(0.3)		
3 VIC	okay\		
4 LIM	uhr (0.2) it's the: same thing as (0.8) um: (1.6) uh I I thought on the: same thing		
	(0,3)		Three head nods and a glance towards the keyboard
5 VIC	**d'accord** (1.1) alors ça en français/ (.) ça s'appelle/ (1,3) la CHAISE MUSICALE\ (0,3)		
	(0.3)		
5 VIC	okay\ (1.1) so in French/ (.) that's called/ (1.3) MUSICAL CHAIRS\ (0.3)		
	enfin chez moi ça s'appelle comme ça <((clavardage)) La chaise musicale.> (0,6)		Eyebrow raises
	at least where I'm from that's what it's called\ <((loud chat)) Musical chairs (0.6)>		
6 LIM	**((rire))**		Laughter
7 VIC	**((rire)) (1,6)**		
6 LIM	((laughter))		
7 VIC	((laughter)) (1.6)		
	c- et: comment vous dites en anglais/ **(1,6)**		Thinking face, then tilted head
8 LIM	**hum: c'est: (0,3) "musical chairs"\ (2,4)**		
	wh- and: what do you call that in English/ (1.6)		
8 LIM	um: it's: (0.3) "musical chairs"\ (2.4)		
9 VIC	**ah oui d'accord** (.) c'est la même chose\ (0,5)		Head nod and right shoulder raise
10 LIM	euh c'est la même chose oui		
9 VIC	ah yes okay\ (.) it's the same thing\ (0.3)		
10 LIM	uh it's the same thing yes		

44 Victor co-constructs this lexical explanation sequence with Liam and involves him in several different ways. By providing a definition of musical chairs and asking, with a thinking face, if Liam knows what it is (line 1), Victor provokes productive recall. Victor allows Liam to reveal the state of his understanding and nods to let him know that he's actively listening (lines 4 and 5). Liam says that he knows of the game, but is unable to recall its name. Victor then provides its oral and written forms (line 5). When Victor laughs along with Liam, we are witnessing *socio-affective alignment* because the emotions are shared. To finish the sequence, Victor asks Liam what the English translation is, thus putting him in the position of expert (Cappellini & Rivens Mompean [2015]). This constitutes a temporary rearrangement of the pedagogical contract (Brousseau [1980]) and is another way of involving the learner.

45 Tutors can also directly involve learners by asking them to provide explanations. These learner explanations can be addressed to the tutor, in the case of comprehension checks where the learner provides more than a yes/no answer, or to a peer, when there is an imbalance of lexical knowledge among the learners. Indeed, the lexical knowledge of a group is always greater than the lexical knowledge of an individual (Newton [2001]). When a word search is necessary, a tutor can involve a learner who had up to that point been a spectator, and ask for their help in finding the right word. This is what Melissa does in the following example when Ana searches for the word *collègue* but is unable to produce it.

Example 7. Recruiting a learner for productive recall

Speech transcription		Webcam (Mélissa)	Observations
1 ANA	amé- euh améliorer notre: compétence/ avec notre: (0,2) camarade de: travail/ (.) non c'est pas ca- (.) camar[ade c'est]: (0,5) [hum:]		Eyes squinted and big smile
2 MEL	[nan/] **(0,8) [Alejandra est-ce que t]u connais le mot/ (0,2)**		
1 ANA	imp- uh improve our: skill/ with our: (0,2) comrades of: work/ (.) no that's not it (.) comr[ade it's]: (0.5) [um:]		
2 MEL	[no/] **(0.8) [Alejandra do you] know the word/ (0.2)**		
	les (.) cama:**RADES (.) de travail/** comment on appelle ces personnes/ (0,3) les		Emblematic gesture to signify reported speech that is imprecise
	the (.) com:**RADES (.) of work/** what do we call these people/ (0,3) the		
	[gens qui **travaillent avec**]&		Head turned and tilted towards the screen
3 ALJ	**[euh:]**		
4 MEL	**&toi/** (5,9) je vais vous l'écrire c'est le mot/ (.) collègue\ (0,5)		
5 ANA	euh [les collègues]		
6 ALJ	[ah: collègue\]		
7 MEL	<((clavardage)) collègue> (0,5)		
8 ANA	euh		
9 MEL	un collègue/ et au pluriel (.) des/ collègues\		
	[people who **work with**]&		
3 ALJ	**[uh:]**		
4 MEL	**&you/** (5.9) I'll write the word for you it's/ (.) colleague\ (0,5)		
5 ANA	uh [colleagues]		
6 ALJ	[ah: colleague\]		
7 MEL	<((text chat)) colleague> (0.5)		
8 ANA	uh		
9 MEL	a colleague/ and plural (.) some/ colleagues\		

46 Melissa uses her voice, gestures and head to involve Alejandra, who had been a spectator watching on the sidelines before the beginning of this sequence. Ana searches for the word *colleague* but is unable to recall the correct form. She says *comrade of work*, which is a good example of what Lüdi [1994] calls *bricolage lexical*, and of what Dörnyei & Scott [1997] call *word coinage*. Melissa smiles as Ana tries to express herself, possibly indicating that she is aware of the pedagogical opportunity at hand. Then, Melissa cites Ana while gesturing quotation marks to represent the reported speech. By citing Ana, Melissa places value on her effort, which is one way of upholding the socio-affective dimension of the exchange (Celik [2008]). Melissa provides a definition of a colleague and leans towards the webcam in order to trigger productive recall (lines 2-4). When no response is given during the six-second pause, Melissa provides the oral (line 5) and written (line 7) forms of the correct word. Finally, she briefly touches on the lexical item's use by mentioning its grammatical functioning (line 9). Melissa orchestrates the interaction and plays the role of referee in order to make the learners collaborate with each other, even though it is she who ultimately provides the word.

4. Discussion and conclusion

47 This article has presented an overview of some of the ways in which different multimodal resources are used by online tutors to focus on the main aspects of knowing a lexical item – form, meaning, and use – and to involve the learners in the explanation sequences.

48 Our analysis showed that the microphone was the most widely used resource for focusing on the forms of lexical items. In addition to using linguistic strategies, tutors on many occasions multiplied and diversified their use of resources. The text chat was

used in a third of all lexical explanation sequences to provide written forms, and these intermodal repetitions were often accompanied by meta-semiotic comments to channel the learners' attention. Although redundancy can hinder information uptake and memorization in some learning contexts (Sweller [2005]), we consider intermodal repetitions to be beneficial here because of the learners' status as non-experts of the language, and because of the complexity of knowing a word. Gestures and facial expressions were also used in one fifth of the sequences to highlight spoken forms, which indicates that the webcam might play a role in guiding learners' attention.

49 The microphone was also the most widely used resource for focusing on the meaning of lexical items. All participants – tutors and learners alike – resorted to translations, which can be the fastest and most effective way of resolving incomprehension, as seen in Example 2. L2 definitions were also widely used, and our qualitative analysis corroborates Chaudron's [1982] finding that short and concise definitions lead to faster comprehension than complex and drawn-out ones. The text chat was not often used to focus on the meaning of lexical items, and this could be due to the fact that focusing on meaning requires more words than simply providing its written form, and since most tutors speak faster than they can type, the microphone was most likely favored for its rapidity and fluidity. The predominance of oral explanations is even less surprising considering the fact that the main purpose of the interactions was to improve L2 oral skills. In one sixth of the sequences, oral explanations were accompanied by gestures, even though a greater number of them were not visible because they were produced outside the webcam's field of view (Guichon & Wigham [2016]). Images were used less often, and were not always able to completely replace a verbal definition, as shown in Example 3. Online tutors should exercise caution and have a backup plan when using images during activities or lexical explanations. We proposed the concept of *multimodal flexibility* following Cicurel [2005] and Develotte et al. [2008] to designate the ability to use different modes to complete a given pedagogical task, and to change modes if necessary. Finally, nonverbal, ostensive explanations can also be produced using the microphone, as shown in Example 4.

50 Our analysis revealed that tutors did not focus much on the use aspect of lexical items, except for culturally charged words (Galisson [1991]) such as *bobo* (see Holt & Tellier [2017]). In fact, fewer than one in six lexical explanation sequences included any mention of this aspect at all. It could be that at this stage of language learning, tutors consider use to be less important than the other aspects of knowing a word.

51 Finally, we found that tutors made considerable efforts while using a variety of resources to involve their learners in the explanation sequences. This involvement can start as soon as a tutor asks a learner if there are any unknown words, or when a tutor opens a side sequence (Jefferson [1972]) in which a lexical difficulty can be resolved interactively and collaboratively. These collaborative spaces can be opened by moving closer to the webcam in order to manifest engagement, using meta-semiotic comments or gestures, looking straight into the webcam, smiling, or using a combination of resources to ask a learner to explain something. We proposed the notion of *socio-affective alignment* to designate moments of shared emotions between tutor and learner. These socio-affective alignments can reduce the discomfort caused by correcting or being corrected, as well as reduce the anxiety inherent in this type of interaction where the interlocutors' contexts are only partially mutually accessible (Guichon [2017]). We

therefore suggest that the involvement load hypothesis (Laufer & Hulstijn [2001]) be expanded to include this socio-affective dimension.

52 Lexical acquisition – and by extension lexical explanation sequences and word-learning techniques – was assigned varying levels of importance (Hilton [2019]) over the many decades before videoconferencing made inroads into language programs in higher education. Its received importance within the language learning community will most likely continue to fluctuate when videoconferencing as we know it evolves into something else. In order to examine videoconferencing in its current state, we have adopted a multimodal approach to take into account the different practices that are made possible thanks to the range of semiotic resources available. Our finding that the microphone was the most widely used resource for focusing on the different aspects of knowing a word and for ensuring learner involvement in the explanation sequences is perhaps not surprising, since videoconferencing is often thought of as a replacement for face-to-face spoken interaction and the purpose of the exchange between Lyon and Dublin was to improve oral skills.

53 This study contributes to a growing body of research that looks into which multimodal resources are used for which purposes. However, qualitative studies such as this one leave certain blind spots that deserve to be addressed by further research. One possible avenue of inquiry would be to implement an experimental protocol with control and experimental groups, predetermined word lists, and pre- and post-tests in order to determine which multimodal practices are the most effective in fostering lexical acquisition. Furthermore, since the present study was unable to determine with absolute certainty which modes and pedagogical regulations caught the learners' attention, eye-tracking technology would be useful for examining which strategies are noticed, and therefore cognitively processed,[9] by the learners.

BIBLIOGRAPHY

AZAOUI Brahim, 2017, « Faire face aux imprévus techniques », *in* GUICHON Nicolas and TELLIER Marion (Eds.), *Enseigner l'oral en ligne : Une approche multimodale*, Paris: Les Éditions Didier, 184-214.

BANGE Pierre, 1992, « À propos de la communication et de l'apprentissage de L2 (notamment dans ses formes institutionnelles) », *Acquisition et interaction en langue étrangère*, Vol. 1, 53-85.

BLIN Françoise, 2016, "The theory of affordances", *in* CAWS Catherine & HAMEL Marie-Josée (Eds.), *Language-Learner Computer Interactions: Theory, methodology and CALL applications*, Amsterdam/ Philadelphia: John Benjamins Publishing Company, 41-64.

BOGAARDS Paul, 1994, *Le vocabulaire dans l'apprentissage des langues étrangères*, Saint-Jean de Braye: Les Éditions Didier.

BROUSSEAU Guy, 1980, « L'échec et le contrat », *Recherches*, Vol. 41, 177-182.

CAPPELLINI Marco & COMBE Christelle, 2017, « Analyser des compétences techno-sémio-pédagogiques d'apprentis tuteurs dans différents environnements numériques : résultats d'une étude exploratoire », *Alsic*.

CAPPELLINI Marco & RIVENS MOMPEAN Annick, 2015, "Role taking for teletandem pairs involved in multimodal online conversation: Some proposals for counseling practice", *Language Learning in Higher Education*, Vol. 5, no. 1, 243-264.

CELIK Christelle, 2008, « Analyse de pratiques de tutorat dans un campus numérique de maîtrise de français langue étrangère à distance », *Alsic*, Vol. 11, no. 1, 93-119.

CHAUDRON Craig, 1982, "Vocabulary elaboration in teachers' speech to L2 learners", *Studies in Second Language Acquisition*, Vol. 4, no. 2, 170-180.

CICUREL Francine, 2005, « La flexibilité communicative : un atout pour la construction de l'agit enseignant », *Le français dans le monde. Recherches & applications*, 180-191.

CICUREL Francine, 2011, *Les interactions dans l'enseignement des langues*, Paris: Les Éditions Didier.

CONKLIN Kathy, PELLICER-SÁNCHEZ Ana & CARROL Gareth, 2018, *Eye-tracking: A Guide for Applied Linguistics Research*, Cambridge: Cambridge University Press.

DE GAULMYN Marie-Madeleine, 1991, « Expliquer des explications », *in* DAUSENDSCHÖN-GAY Ulrich, GÜLICH Elisabeth & KRAFFT Ulrich (Eds.), *Linguistische Interaktionsanalysen*, Tübingen: Max Niemeyer, 279-314.

DEJEAN-THIRCUIR Charlotte, GUICHON Nicolas & NICOLAEV Viorica, 2010, « Compétences interactionnelles des tuteurs dans des échanges vidéographiques synchrones », *Distance et Savoirs*, Vol. 8, 377-393.

DEVELOTTE Christine & DRISSI Samira, 2013, « Face à face distanciel et didactique des langues », *Le français dans le monde. Recherches et applications*, 54-63.

DEVELOTTE Christine, GUICHON Nicolas & KERN Richard, 2008, « "Allo Berkeley ? Ici Lyon... Vous nous voyez bien ?" Étude d'un dispositif de formation en ligne synchrone franco-américain à travers les discours de ses usagers », *Alsic*, Vol. 11, no. 2, 129-156.

DEVELOTTE Christine & MANGENOT François, 2007, « Discontinuités didactiques et langagières au sein d'un dispositif pédagogique en ligne », *Glottopol*, no. 10, 127-144.

DÖRNYEI Zoltán & SCOTT Mary Lee, 1997, "Communication Strategies in a Second Language: Definitions and Taxonomies", *Language Learning*, Vol. 47, no. 1, 173-210.

ELLIS Rod, 2016, "Focus on form: A critical review", *Language Teaching Research*, Vol. 20, no. 3, 405-428.

ELLIS Rod, TANAKA Yoshihiro & YAMAZAKI Asako, 1994, "Classroom interaction, comprehension, and the acquisition of L2 word meanings", *Language learning*, Vol. 44, no. 3, 449-491.

FASEL LAUZON Virginie, 2008, "Interactions et apprentissages dans des séquences d'explication de vocabulaire", Travaux neuchâtelois de linguistique, Vol. 48, 83-104

FASEL LAUZON Virginie, 2014, *Comprendre et apprendre dans l'interaction : Les séquences d'explication en classe de français langue seconde*, Bern: Peter Lang AG Internationaler Verlag der Wissenschaften.

FLOWERDEW John, 1992, "Definitions in science lectures", *Applied linguistics*, Vol. 13, no. 2, 202-221.

FOLSE Keith S., 2004, Vocabulary myths: applying second language research to classroom teaching, Ann Arbor: The University of Michigan Press.

GALISSON Robert, 1991, De la langue à la culture par les mots, Paris: CLE Internationale.

GOLDIN-MEADOW Susan, 2003, *How our hands help us think*, Cambridge, MA: Harvard University Press.

GOODWIN Marjorie Harness & GOODWIN Charles, 1986, "Gesture and coparticipation in the activity of searching for a word", *Semiotica*, Vol. 62, no. 1/2, 51-75.

GROSSMANN Francis, 2011, « Didactique du lexique : état des lieux et nouvelles orientations », *Pratiques. Linguistique, littérature, didactique*, Vol. 149-150, 163-183.

GROUPE ICOR 2007, *Convention ICOR*, http://icar.cnrs.fr/projets/corinte/documents/2013_Conv_ICOR_250313.pdf

GUICHON Nicolas, 2017, « Se construire une présence pédagogique en ligne », *in* GUICHON Nicolas & TELLIER Marion (Eds.), *Enseigner l'oral en ligne : Une approche multimodale*, Paris: Les Éditions Didier, 29-58.

GUICHON Nicolas, BÉTRANCOURT Mireille & PRIÉ Yannick, 2012, "Managing written and oral negative feedback in a synchronous online teaching situation", *Computer-Assisted Language Learning*, Vol. 25, no. 2, 181-197.

GUICHON Nicolas & COHEN Cathy, 2014, "The Impact of the Webcam on an Online L2 Interaction", *Canadian Modern Language Journal*, Vol. 70, no. 3, 331-354.

GUICHON Nicolas & DRISSI Samira, 2008, « Tutorat de langue par visioconférence : comment former aux régulations pédagogiques ? », *Les Cahiers de l'Acedle*, Vol. 5, 1-32.

GUICHON Nicolas & TELLIER Marion (Eds.), 2017, *Enseigner l'oral en ligne : une approche multimodale*, Paris: Les Éditions Didier.

GUICHON Nicolas & WIGHAM Ciara R., 2016, "A semiotic perspective on webconferencing-supported language teaching", *ReCALL*, Vol. 28, no. 1, 62-82.

GÜLICH Elisabeth, 1990, « L'approche ethnométhodologique dans l'analyse du français parlé », *in* CHAROLLES Michel, FISHER Sophie & JAYEZ Jacques (Eds.), *Le discours : représentations et interprétations*, Nancy: Presses Universitaires de Nancy, 71-109.

GÜLICH Elisabeth, 1991, « Pour une ethnométhodologie linguistique : Description de séquences conversationnelles explicatives », *in* DAUSENDSCHÖN-GAY Ulrich, GÜLICH Elisabeth & KRAFFT Ulrich (Eds.), *Linguistische Interaktionsanalysen*, Tübingen: Max Niemeyer, 325-364.

HATCH Evelyn Marcussen, 1978, "Discourse analysis and second language acquisition", *in* HATCH Evelyn Marcussen (Ed.), *Second language acquisition: A book of readings*, Rowley, MA: Newbury House, 401-435.

HILTON Heather, 2019, « Sciences cognitives et didactique des langues », *Rapport d'expertise pour le Conseil national de l'évaluation du système scolaire*, Paris: CNESCO.

HOLT Benjamin, 2018, « Séquences d'explication lexicale dans l'enseignement du français par visioconférence : une approche multimodale », Lyon: Université Lumière Lyon 2.

HOLT Benjamin, 2020, « Le rôle des gestes dans les explications lexicales par visioconférence », *TIPA Travaux interdisciplinaires sur la parole et le langage*, Vol. 36.

HOLT Benjamin & TELLIER Marion, 2017, « Conduire des explications lexicales », *in* GUICHON Nicolas & TELLIER Marion (Eds.), *Enseigner l'oral en ligne*, Paris: Didier, 59-90.

HOLT Benjamin, TELLIER Marion & GUICHON Nicolas, 2015, "The use of teaching gestures in an online multimodal environment: the case of incomprehension sequences", *Gesture and Speech in Interaction 4th Edition*, Nantes, 149-154.

HULSTIJN Jan, HOLLANDER Merel & GREIDANUS Tine, 1996, "Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words", *The Modern Language Journal*, 80, 327-339.

HORLACHER Anne-Sylvie, 2015, *La dislocation à droite revisitée : une approche interactionniste*, Louvain-la-Neuve: De Boeck Duculot.

JEFFERSON Gail, 1972, "Side Sequences", *in* SUDNOW David (Ed.), *Studies in Social Interaction*, New York: Free Press, 294-338.

JEWITT Carey, 2014, "An introduction to multimodality", *in* JEWITT Carey (Ed.), *The Routledge Handbook of Multimodal Analysis Second Edition*, London: Routledge, 15-30.

JÉZÉGOU Annie, 2010, « Créer de la présence à distance en e-learning. Cadre théorique, définition, et dimensions clés », *Distances et savoirs*, Vol. 8, no. 2, 257-274.

JONES Rodney, 2014, "Mediated discourse analysis", *in* NORRIS Sigrid & MAIER Carmen Daniela (Eds.), *Interactions, images and texts: a reader in multimodality*, Boston/Berlin: De Gruyter Mouton, 39-51.

KENDON Adam, 2004, *Gesture: Visible action as utterance*, Cambridge: Cambridge University Press.

KRESS Gunther, 2014, "What is mode?", *in* JEWITT Carey (Ed.), *The Routledge Handbook of Multimodal Analysis*, Second Edition, London: Routledge, 60-75.

KRESS Gunther & VAN LEEUWEN Theo, 2001, *Multimodal discourse: the modes and media of contemporary communication*, London: Hodder Education.

LANCIEN Thierry, 1998, « Images mobiles et multimédia », *Études de linguistique appliquée : revue de didactologie des langues-cultures, Hypermédia et apprentissage des langues*, 171-182.

LAUFER Batia & HULSTIJN Jan, 2001, "Incidental vocabulary acquisition in a second language: the construct of task-induced involvement", *Applied Linguistics*, Vol. 21, no. 1, 1-26.

LAUFER Batia & SHMUELI Karen, 1997, "Memorizing new words: Does teaching have anything to do with it?", *RELC journal*, Vol. 28, no. 1, 89-108.

LAZARATON Anne, 2004, "Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry", *Language learning*, Vol. 54, no. 1, 79-117.

LONG Michael, 1981, "Input, interaction, and second-language acquisition", *Native Language and Foreign Language Acquisition*, Vol. 379, 259-278.

LONG Michael, 1983, "Native speaker/non-native speaker conversation and the negotiation of comprehensible input", *Applied Linguistics*, Vol. 4, no. 2, 126-141.

LONG Michael H., 1991, "Focus on Form: A Design Feature in Language Teaching Methodology", *in* DE BOT Kees, GINSBERG Ralph B. & KRAMSCH Claire (Eds.), *Foreign Language Research in Cross-Cultural Perspective*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 39-52.

LÜDI Georges, 1991, « Construire ensemble les mots pour le dire. À propos de l'origine discursive des connaissances lexicales », *in* DAUSENDSCHÖN-GAY Ulrich, GÜLICH Elisabeth & KRAFFT Ulrich (Eds.), *Linguistische Interaktionsanalysen*, Tübingen: Max Niemeyer, 193-224.

LÜDI Georges, 1994, « Dénomination médiate et bricolage lexical en situation exolingue », *Acquisition et interaction en langue étrangère*, Vol. 3, 115-146.

MANGENOT François, 2011, « Spécificités du tutorat en langues », *in* DEPOVER Christian, de LIÈVRE Bruno, PERAYA Daniel, QUINTIN Jean-Jacques & JAILLET Alain (Eds.), *Le tutorat en formation à distance*, Bruxelles: De Boeck Supérieur, 213-225.

MARRA Antonietta & PALLOTTI Gabriele, 2006, « Les logonymes dans la classe de langue », *in* FARACO Martine (Ed.), *La classe de langue. Théories, méthodes et pratiques*, Aix-en-Provence: Presses de l'Université de Provence, 129-151.

MAYER Richard E., 2001, *Multimedia Learning*, Cambridge: Cambridge University Press.

MCNEILL David, 1992, *Hand and Mind*, Chicago: The University of Chicago Press.

MIECZNIKOWSKI Johanna, 2005, *Le traitement de problèmes lexicaux lors de discussions scientifiques en situation plurilingue : Procédés interactionnels et effets sur le développement du savoir*, Berne: Peter Lang.

MOL Lisette, KRAHMER Emiel, MAES Alfons & SWERTS Marc, 2011, "Seeing and Being Seen: The Effects on Gesture Production", *Journal of Computer-Mediated Communication*, Vol. 17, no. 1, 77-100.

MORTUREUX Marie-Françoise, 2008, *La lexicologie entre langue et discours*, 2ème édition revue et actualisée, Paris: Armand Colin. Collection Cursus Linguistique.

NATION I. S. P., 2013, *Learning Vocabulary in Another Language*, Second Edition, Cambridge: Cambridge University Press.

NATION I. S & WEBB Stuart, 2011, *Researching and analyzing vocabulary*, Boston: Heinle.

NEWTON Jonathan, 2001, "Options for vocabulary learning through communication tasks", ELT Journal, Vol. 55, no. 1, 30–37.

NICOLAEV Viorica, 2010, « Les négociations de sens dans un dispositif d'apprentissage des langues en ligne synchrone par visioconférence », *Les Cahiers de l'Acedle*, Vol. 7, no. 2, 169-198.

NORRIS Sigrid, 2004, *Analyzing Multimodal Interaction: A Methodological Framework*, New York: Routledge.

O'DOWD Robert, 2000, "Intercultural learning via videoconferencing: A pilot exchange project", *ReCALL*, Vol. 12, no. 1, 49-61.

O'DOWD Robert, 2005, "The Use of Videoconferencing and E-mail as Mediators of Intercultural Student Ethnography", *in* BELZ Julia A. & THORNE Steven L. (Eds.), *Internet-mediated Intercultural Foreign Language Education*, Boston: Thomson Higher Education, 2-30.

O'DOWD Robert, 2016, "Emerging Trends and New Directions in Telecollaborative Learning", *CALICO Journal*, Vol. 33.3, 291-310.

PICA Teresa, 1994, "Research on Negotiation: What Does It Reveal About Second-Language Learning Conditions, Processes, and Outcomes?", *Language learning*, Vol. 44, no. 3, 493-527.

RENNER Julia, 2017, "Negotiation of meaning and language-related episodes in synchronous, audio-based Chinese-German eTandem", *in* O'ROURKE Breffni & STICKLER Ursula (Eds.), *Language Learning in Higher Education*, Vol. 7, 137-159.

REY-DEBOVE Josette, 1997, *Le métalangage : étude linguistique du discours sur le langage*, Paris: Armand Colin.

SARRÉ Cédric, 2011, "Computer-mediated negotiated interactions: how is meaning negotiated in discussion boards, text chat and videoconferencing?", *in* THOUËSNY Sylvie & BRADLEY Linda (Eds.),

Second language teaching and learning with technology: views of emergent researchers, Dublin: Research-publishing.net, 189-210.

SMITH Bryan, 2003, "Computer-Mediated Negotiated Interaction: An Expanded Model", *The Modern Language Journal*, Vol. 87, 38-57.

STAHL Steven A. & VANCIL Sandra J., 1986, "Discussion is what makes semantic maps work in vocabulary instruction", *The Reading Teacher*, Vol. 40, no. 1, 62-67.

SUEYOSHI Ayano & HARDISON Debra M., 2005, "The role of gestures and facial cues in second language listening comprehension", *Language Learning*, Vol. 55, no. 4, 661-699.

SWELLER John, 2005, "The Redundancy Principle in Multimedia Learning", in MAYER Richard E. (Ed.), *The Cambridge Handbook of Multimedia Learning*, Cambridge: Cambridge University Press, 159-167.

TELLIER Marion, 2008, « Dire avec des gestes », *Le français dans le monde : recherche et application*, no. 44, 40-50.

TELLIER Marion & CADET Lucile, 2014, *Le corps et la voix de l'enseignant : théorie et pratique*, Paris: Éditions Maison des Langues.

TELLIER Marion & STAM Gale, 2012, « Stratégies verbales et gestuelles dans l'explication lexicale d'un verbe d'action », *in* RIVIÈRE Véronique (Ed.), *Spécificités et diversité des interactions didactiques*, Paris: Riveneuve éditions, 357-374.

TRÉVILLE Marie-Claude, 2000, *Vocabulaire et apprentissage d'une langue seconde : recherches et théories*, Québec: Les Éditions LOGIQUES.

TRÉVILLE Marie-Claude & DUQUETTE Lise, 1996, *Enseigner le vocabulaire en classe de langue*, Paris: Hachette F.L.E.

VAN DER LINDEN Elisabeth, 2006, « Lexique mental et apprentissage des mots », *Revue Française de Linguistique Appliquée*, Vol. 11, no. 1, 33-44.

VARONIS Evangeline Marlos & Gass Susan, 1985, "Non-native/Non-native Conversations: A Model for Negotiation of Meaning", *Applied Linguistics*, Vol. 6, no. 1, 71-90.

WARING Rob, 2002, "Scales of vocabulary knowledge in second language vocabulary assessment", *Kiyo* (Occasional papers of Notre Dame Seishin University, Okayama), Vol. 26, no. 1, 40-54.

WIGHAM Ciara R., 2017, "A multimodal analysis of lexical explanation sequences in webconferencing-supported language teaching", *Language Learning in Higher Education*, 24 January 2017, Vol. 7, no. 1, 81-108.

WITTENBURG Peter, BRUGMAN Hennie, RUSSEL Albert, KLASSMANN Alex and SLOETJES Han, 2006, "Elan: a professional framework for multimodality research", *Proceedings of the 5th International Conference on Language Resources and Evaluation* (LREC 2006), 1556-1559.

NOTES

1. Nation [2013: 11] explains that "a **word family** consists of a head word, its inflected forms and its closely related derived forms."

2. During the opening phase of a lexical explanation sequence, the unknown or non-understood lexical item is identified by the participants as the topic of the sequence. During the core phase,

the lexical item is explained or given. During the closing phase, the learner's understanding is confirmed and/or evaluated.

3. In our corpus, approximately 45% of the lexical explanation sequences were the result of comprehension problems, and approximately 55% were the result of production problems (See Holt [2018]).

4. This is a slightly derogatory term that means "bourgeois-bohemian" and designates the target market of a hypothetical food truck business that the learners had to invent.

5. Here, *focus on form* means focusing on a lexical item's form as opposed to its meaning or use (Nation [2013]). We are not referring to Long's [1991] broader definition of focus on form, which means isolating linguistic items that incidentally arise from interaction in order to check their comprehension or explain them. See Ellis [2016] for a review.

6. For example, *L'hygiène c'est la propreté* (Hygiene it's cleanliness) includes a dislocation because the subject is repeated. While this may seem odd or incorrect in English, dislocations are acceptable and common in French.

7. In the left-hand column, we transcribe the speech using the ICOR transcription scheme. The characters in boldface temporally correspond to the screenshots in the middle column. In the right-hand column, we describe the phenomena that are illustrated by the screenshots.

8. This second explanation is not transcribed here due to lack of space.

9. See Conklin, Pellicer-Sánchez & Carrol [2018] for an overview of the eye-mind hypothesis.

ABSTRACTS

This is a study on lexical explanation sequences that emerge naturally from online French tutoring sessions. From a 15.5-hour corpus of videoconferencing-based pedagogical interaction between seven tutors in France and twelve learners in Ireland, we identified nearly three hundred lexical explanation sequences. These were analyzed using a multimodal approach in order to determine which resources were used by the tutors to focus on the different aspects of knowing a word: form, meaning, and usage. In addition, we analyzed the ways in which tutors fostered the learners' cognitive and socio-affective involvement in the explanation sequences. Our descriptive statistics are accompanied by examples that illustrate typical and innovative uses of the affordances offered by the videoconferencing platform.

Cette étude porte sur des séquences d'explication lexicale qui émergent naturellement de l'interaction pédagogique médiatisée par visioconférence. À partir d'un corpus de quinze heures et demie d'interactions pédagogiques entre sept apprentis-enseignants en France et douze apprenants en Irlande, nous avons repéré près de trois cents séquences d'explication lexicale. Celles-ci ont été analysées par une approche multimodale afin de déterminer quelles ressources ont été utilisées par les apprentis-enseignants pour se focaliser sur les différents aspects des items lexicaux : la forme, le sens, et l'usage. Nous avons aussi analysé les procédés par lesquels les apprentis-enseignants ont assuré l'implication cognitive et socioaffective des apprenants lors des séquences d'explication. Nos statistiques descriptives sont accompagnées d'exemples qui illustrent des usages typiques et innovants des affordances de la plateforme de visioconférence.

INDEX

Mots-clés: explication lexicale, visioconférence, multimodalité, didactique du français langue étrangère
Keywords: lexical explanation, videoconferencing, multimodality, teaching French as a foreign language

AUTHOR

BENJAMIN HOLT

Université de Lille
benjamin.holt@univ-lille.fr