SAPIENZA
UNIVERSITÀ EDITRICE

Research article

First published online: May 30, 2023

Ersilia Incelli[*]

A MULTIMODAL CORPUS APPROACH TO UNDERSTANDING PARLIAMENTARY DEBATES: THE ROLE OF GESTURE AND INSTITUTIONAL DISCOURSE

Abstract

The aim of this study is to contribute to current research on the effects of multimodality on ESP education, with particular reference to the situational context of the university language classroom. Changes in the way society interacts and accesses information as a result of the present move from language-centred texts to multimodal texts, have had a profound impact on education. As a result, in order to adapt and respond to the continuously changing educational landscape, higher education teachers are increasingly obliged to develop new or alternative conceptions of meaning and communication. For this purpose, the primary objective and impetus of this work is to demonstrate how multimodality and the multi-semiotic dimensions of digital communication may be incorporated into an academic ESP curriculum. The study also intends to illustrate how corpus linguistic retrieval of phraseology can be used to enhance ESP language acquisition through research-informed analyses of a multimodal corpus of video clips.

Keywords: ESP, gesture analysis, institutional discourse, multimodal analysis, multimodal corpus.

---

[*] Department of Methods and Models for Territory, Economics and Finance, Sapienza University of Rome, Italy

1 Introduction

Due to advancements in the instant accessibility of video data and worldwide video networks, learning in the audiovisual context has radically changed conventional viewing in the last two decades leading to the need to conceptualize new theories in language acquisition in learning/teaching environments in order to help educators and teachers adapt to the ongoing evolution of the communication landscape and changing education scenarios. In view of this and on account of an unprecedented increase in the use of technology in education and language teaching, this paper explores how students can be guided to consciously become aware of how meaning-making works in multimodal texts, that is, how semiotic resources involving language, gesture, and images, often accompanied by visual and aural elements, interact with each other as verbal and non-verbal features to create textual coherence (Lim, 2011; 2021). For this purpose, the following work presents a case study of a research-informed teaching project implemented in an advanced (C1)[1] level English language module at the Faculty of Economics, Sapienza University of Rome. English is taught in an English for Specific Purposes (ESP) and English for Academic Purposes (EAP) context, with the aim of preparing undergraduate students of Economics for their professional working lives and future job markets.

Today, English language learners are increasingly exposed to a variety of channels of socialization and identity construction, accessing information and communicating in wide ranging ways for both academic and social purposes. All of this has obvious ramifications on education, which should try to anticipate the critical skills and competencies needed in the professions or society in general (Duff and Talmy, 2011). Therefore, as educators, it seems essential that we reflect on how to raise both cultural and critical awareness applied to texts in different genres. Scholars have recently referred to the 'multimodal literate' student as someone who should be able to critically read multimodal texts, i.e. texts which occur in multimedia and digital contexts (Jewitt, 2005; Machin and Myer, 2012). Hence, the case study reported here hopes to show through examples of practical learning tasks, how language students can be guided to consciously understand how the elements of multimodal sources work to make meaning, rather than passively or naively consume these texts (Machin and Myer, 2012).

The project is centred on a multimodal analysis of selected video fragments taken from a multimodal corpus specifically built and compiled by the English language staff for undergraduate students of Economics, (see Table 1 in section 4). For illustrative purposes, a multimodal analysis was carried out on video clips from two live videos of two government debates: one concerning the US House of Representatives regarding the Republican party's *Tax Reform Bill* (December, 2018), and the other consisting of a UK House of Commons debate regarding BREXIT called *Leaving the EU* (April, 2019)[2]. Admittedly these video clips are not taken from recent datasets, but as the aim

---

[1] See the Common European Framework of Reference for Languages level descriptors: https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions (accessed December 11, 2022).

[2] UK House debate on *Leaving the EU*: https://www.youtube.com/watch?v=X1uMe5FmqH8&feature=youtu.be (accessed 9 June 2019). US House debate on *GOP tax reform bill 2018*: https://www.youtube.com/watch?v=CDoJRFADpqs&t=13562s (accessed 6 May 2019).

is to show how multimodal analysis in the classroom can enhance language learning and multimodal literacy (Machin and Myer, 2012), the belief is that this sort of analysis is replicable on other similar case study video datasets and multimodal ESP projects. Another reason for a dataset from 2019 is due to the fact that the recent world Covid pandemic created restrictions in the classroom and limited access to developing more current studies. Nevertheless, we now plan to continue expanding the English language video collection. Furthermore, the work reported here is the continuation of previous research on understanding and interpreting identity and ideology in government debates using the same video debates (Incelli, 2021). However, the video clips for this case study are different. This second stage of the project has a special emphasis on the role of institutional discourse as a feature of turn-taking discourse in argumentative debates, as well as the role of gestures that accompany utterances in parliamentary speeches. The video clips in fact demonstrate how intensive multimodal analysis provides a wealth of data that can be exploited in many ways for different teaching purposes and learning objectives. More specifically, this case study in ESP teaching is concerned with engaging students in understanding the language of debates through both the verbal mode and semiotic modes comprising paralinguistic features (i.e. vocal production and intonation), and kinesic features (for example involving body movement, gaze, gesture), all of which play a role in the communicative and persuasive utterance and argumentative discourse.

1.1 Aims and Research Questions

The main aims and research questions which formed the basis of the research investigation prior to the teaching applications are summarized here. The impetus and ultimate aim are pedagogical, that is, to show how multimodal literacy can be incorporated into an ESP university curriculum and to illustrate how research-informed analyses of a multimodal corpus of video clips can be integrated into language classes to improve ESP language learning. The principle objective of the analysis was to show how linguistic features in the MPs speech were co-deployed with non-verbal elements, in order to raise students' awareness of the multimodal nature of parliamentary debates and how speakers arrange the various modes to persuasively and pragmatically put across a message to their audiences (Jewitt, 2009).

Specific research questions were addressed which later formed the basis of specific classroom activities, for example, how does the interplay of the two modes of verbal interaction and non-verbal features construct meaning in an argumentative context? The results were then then classified and related to a practical teaching and learning level. Specific questions for intended learning outcomes were also formulated, for example, do spatial, visual images conveyed concurrently with the verbal mode make meaning and understanding easier? Does it motivate students and facilitate learning? What cognitive knowledge do the students need to have of the situational context?

The paper is organized as follows: section 2 presents the theoretical background; section 3 introduces the teaching context and the multimodal dataset used for the teaching/learning activities; section 4 describes the learning activities; section 5 provides some conclusions on the pedagogical implications.

2 Theoretical background

The approach which is best suited for tackling the complex relationship between the various meaning or semiotic systems apparent in a multimodal text, is Kress and van Leeuwen's (2006) theoretical framework, namely social semiotic multimodality. This was extended from Halliday's (1994) systemic function theory, to be further expanded and defined by O'Halloran (2004) as systemic functional multimodal discourse analysis (SFMDA). In this way, the framework was appropriately used for developing critical viewing skills in the construction of meaning making and introduced pedagogic semiotics semiosis as an approach to multimodal literacy in education. Hence, this work is also inspired by various studies in teaching which highlight the importance of multimodal teaching materials in a multimodal genre (New London Group, 1996; Prior, 2013; O'Halloran et al., 2016; Lim and Yan, 2018).

Since this case study report focusses on the construction of argumentative speech in institutional settings, i.e. government, this work also enters the realm of critical discourse analysis (CDA), and critical multimodal discourse analysis (CMDA), (Machin and Mayr, 2012, among others). These works show how multimodal analysis can be affiliated to the principle aims of CDA, especially when uncovering the buried ideological content of the communicative purpose of a text or speech. This is done 'multi-modally' through the combination of various semiotic resources, as well as through language which can be used for ideological purposes.

Listening and video viewing as classroom activities have always been practiced in the language classroom as a basic skill. However, recent insights from the multimodal approach have added greater depth to language acquisition through the viewing process, for example, the teaching of meaning-making through the in-depth cognition of 'multimodal ensembles' (Royce, 2006), which integrate verbal modes with non-verbal modes such as gesture, gaze, postural stance, and facial expression. Such modes were initially studied separately (Kendon, 2004), and only more recently have they been applied to multimodal studies as 'multimodal ensembles'. For the purpose of this research, for gesture interpretation and function, it is useful to refer to the works of Kendon (2004), in order to reduce subjective evaluation, (e.g. gesture descriptions include beats, deictic, iconic, metaphoric, social), whereas Querol-Julian (2011) is useful for defining gesture coding parameters and taxonomies.

To address the need for methods and approaches which can be applied to non-linguistic phenomena in a multimodal classroom, more recently scholars have begun to create multimodal corpora (Bateman, 2008; Bateman and Knight, 2011), which place language use in richer multimodally constituted contexts. Such studies profess the benefits and advantages of the uses of multimodal corpora for the enhancement of oral communication on users who can analyze several aspects of audio-visual communication. However, due to the complexity of creating multimodal corpora, integrated ESP multimodal materials epitomizing technology-mediated spoken communication used by professional communities are still absent from language learning textbooks, especially with regards to the social sciences or ESP contexts like economics, law, and medicine. In this way the work here wishes to contribute to recent studies which attempt to integrate the multimodal dimension in the ESP university classroom (Crawford-Camiciottoli and Bonsignore, 2015; Crawford-Camiciottoli and Campoy-Cubillo, 2018).

3 The ESP teaching context

The last review of syllabi and curricula in the Faculty of Economics, Sapienza, acknowledged the importance of English as a primary source of communication and as an essential skill for the future job market. This has led to a marked increase in the teaching of the English language throughout the academic curriculum. Although students allegedly have basic competence of general English, before they graduate, they are required to develop language skills at a higher and specialized level, which will help them enter the professional workplace. The English language classes run parallel to other subjects in the Economics degree course, entailing English is studied in several domains, for example, with reference to banking, economic policy, marketing, and management.

3.1 Multimodal corpora for ESP

To address the learning needs of students, the multimodal corpora (which is constantly expanding) consists of selected audio-visual materials related to the student's academic course of studies and curriculum, mainly in the domains of banking, business, economics, finance, political and social science, reflecting current affairs and socio-economic policy issues, e.g. economic sustainability, environmental economics, immigration, universal credit (Table 1). The specialized corpus also represents a variety of genres ranging from academic lectures, institutional speeches to live talk shows and TV interviews with well-known economists, such as Mario Draghi, Christine Lagarde, as well as government and parliamentary discourse, and online Economics courses. So far, the compiled corpus consists of 18 videos of different registers, as well as a collection of films on finance and business.

Table 1. The multimodal corpus for the EAP/ESP English language course

Title	Genre	Tapescript – Total words	Number of video clips for analysis	Duration
Circular Economy in a nutshell - 'Rethinking Resources'	Podcast /Educational Video YouTube (16 Aug. 2022)	6,976	Whole video	42 min
CORE project The Economy (University of Chicago)	Education al online course for Economics	30,712	2	12 min
UK House of Commons *Amazon Deforestation* debate (7 Oct.2019)	Hansard archives – live debate	26,855	5	2 h
UK House of Commons – *Leaving the EU* debate,	Hansard archives – live debate	27,458	8	3 h 30 min

Westminster Hall (1 April 2019)	(accessible on YouTube)			
US House of Representatives – *Immigration Bill* debate (4 June 2019)	House of Representatives (HoR) video archives	3,437	6	1 h
US House of Representatives – *GOP Tax Reform Bill* (19 Dec. 2018)	HoR video archives (accessible on *YouTube*)	25,735	8	4 h 22 min
ECB, Mario Draghi Responds to Trump 'We don't target the exchange rate' (18 June 2019)	TV interview (YouTube)	435	Whole video	2 min
Mario Draghi ECB Thank you speech (19 Oct. 2019)	Speech	2,892	2	5 min
Mario Drahi (ECB President – *Whatever it takes*, London (27 July 2015)	Speech at the Global Investment conference (You Tube)	1,466	Whole video	4 min
Christine Lagarde incoming president of ECB (30 Oct. 2019)	Speech	1,255	Whole video	5 min
Films on finance, environmental economics, social network				
The Inside Job (2010) 120 mins; The Social Network (2010) 120 mins; The Company Men (2010) 104 mins; Margin Call (2011) 107 mins; Steve Jobs (2015) 122 mins; Money Monster (2016) 98 mins; Before the flood (2016) 96 mins.				

Source: Author's elaboration

Methodological note: two videos were selected from the multimodal corpus, from which video clips were then extracted to form a dataset, which I describe here.

3.2 Methodology

Prior to the teaching/learning activities, a multimodal research analysis was carried out on the videos and selected dataset, in order to pinpoint the main verbal and non-verbal resources that the students might run into during the listening and viewing activities. In this manner, I could also identify the best ways to involve the students in valid learning tasks. In this section I briefly explain how the dataset was built and the multimodal analysis of the video clips.

3.2.1 Multimodal dataset collection

Table 2 summarizes the dataset extracted from the larger multimodal corpus. The main topics of the two debates are related to economic policy issues, therefore the discourse consists of economic and financial lexical items, e.g. *capital gains, income, tax revenue, multinational companies*; nevertheless, the context is parliamentary and therefore reference is made to political discourse with features such as formulaic turn-taking utterances, and formal language used to address members of parliament (van Dijk, 2004; Partington and Taylor, 2017).

Table 2. Summary data of multimodal dataset

	Institutional context	Subject of Debate	Total Words -Token/Type Ratio	Duration	Number of Speakers	Number of clips
UK	House Of Commons[3]	Leaving the EU - Government e-petition (1 April, 2019)	27,458/3,269	3 hours 30 mins. (15 min. break)	18 (focus on 4 speakers)	8
US	House Of Representatives[4]	GOP Tax Reform Bill (19 December, 2017)	21,735/2,984	4 hours 22 mins. (1 hr. break to vote)	36 (focus on 4 speakers)	8

Source: Author's elaboration

Five speakers in the videos were chosen on the basis of their well-defined rhetorical devices and clear paralinguistic and kinesic features, which were identified using the software *ELAN 5.6* (Wittenburgh et al., 2006, see below) and then systematically classified to create appropriate learning tasks. Each video clip taken from the dataset after careful viewing of both full debates lasted from 30 seconds to 2 minutes depending on the speech act. Accompanying transcripts were also downloaded to carry out corpus linguistic analysis and identify formulaic phrases which could then be shown to the students[5]. Typical features also include prominent rhetorical devices such as parallelisms (repetition through bicolons and tricolons) and rhetorical questions (Partington and Taylor, 2010; 2017).

3.2.2 Multimodal analytical framework – *ELAN*

After the primary linguistic traits and prominent discourse strategies had been determined, the speech acts were carefully observed for accompanying non-verbal communication. I used the multimodal annotation software ELAN 5.6 (Wittenburgh et

---

[3] UK House debate on *Leaving the EU*
https://www.youtube.com/watch?v=X1uMe5FmqH8&feature=youtu.be (accessed 9 June, 2019)
[4] US House debate on *GOP tax reform bill 2018*
https://www.youtube.com/watch?v=CDoJRFADpqs&t=13562s (accessed 6 May, 2019)
[5] The US debate did not have a transcript, so an automatic online converter was used, i.e. https://www.veed.io/tools/audio-to-text (accessed 19 December, 2021).

al., 2006) to carry out a systematic classification of the interaction of resources. To facilitate a systematic classification of interactional semiotic resources. ELAN enables multimodal analysts and users to explore and annotate the ways in which verbal and nonverbal components interact to produce meaning. Information can be inserted onto the tool's layers or tiers. The video and audio files, which are uploaded separately, allow speech production and images to be time-aligned and synchronized. An example of an annotated analysis of a video clip from the US debate is shown in Figure 1. The tiers (circled) are explained in the left column. ELAN was created by psychiatrists; therefore the software does not necessarily require knowledge of linguistic concepts. In fact, students can be introduced to ELAN in a simple way and shown how to use the annotation and tier system which helps them understand the interaction between inter-semiotic modes more deeply. *ELAN* analysis added depth to the students' fine-tuned observations.

Figure 1. Elan annotation of an image fragment from a video clip of the US debate indicating the multimodal ensemble and tier analysis.

Time Clip 1.07.20:
<u>Multimodal ensemble:</u> right arm and hand raised up towards the house speaker (PFA) -metaphoric (ontological – good is up); Body – turned towards the House Speaker: stress on 'it', 'power' – slight pause after 'powerful' for emphasis.
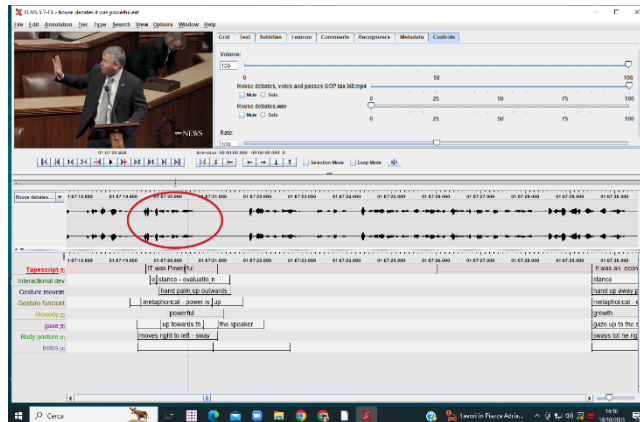
Source: Author's elaboration

In the fragment of the video in Figure 1, the speaker is facing and turned towards the House Speaker sitting higher up on a podium. The speaker's right arm is raised with his palm facing the House Speaker. We can interpret this gesture as emphasizing the authority of the House Speaker, to whom he directs his utterance, increasing the truth and reliability of his speech. Moreover, the gesture is ontological and metaphorical - 'good is up'. The phrase 'it was powerful' is reinforced by a slight pause after 'powerful'.

4 Language learning activities

The data for this study were collected from the ESP lessons held in the academic semesters of 2019 and 2020. In actual fact, some lessons were also held online during the Covid pandemic due to restrictions on frontal classes. The advanced (C1) English language course in the Faculty of Economics, consists of two, two-hour lessons per

week. A total of 10 hours was planned over a five-week period. The students were particularly motivated and committed to attending the course in light of their final assessment in the oral examination which involved power-point presentations. Feedback and observation notes were helpful in analyzing the learning outcomes and how language acquisition was shaped. Each lesson had a learning objective and intended learning outcome, with activities based on previewing, while-viewing and post-viewing stages. I give illustrative examples here of each lesson stage.

4.1. Identifying the context: reading and writing activities

The first 4 lessons were dedicated to building up the students' knowledge of the context, i.e. *Brexit* and the Republican party's 2018 *Tax Reform Bill*. We read articles from the *Economist*[6] and from the *Guardian*[7] acquiring familiarity with the language and specialized terminology. Activities involved gap fills for specialized lexis and common collocations, (e.g., *average tax rate, quarter, wage drop, to subsidize, to pay its debts, overstaffing*), and specific grammar practice for verb tenses, phrasal verbs, etc. e.g. (*bailed out)*. In this way students were able to gain cognitive knowledge of events and a cultural idea of the social and political events. Students were given reading comprehension questions which also encouraged cultural comparisons. Thus, the previewing lessons are traditionally language-centered with reading tasks and worksheets to assist comprehension of specialized vocabulary and complex lexical-grammatical constructions common to the genre, such as transitive and intransitive verbs (*decline, hamper, reduce*); the passive, e.g. *payers were granted.* This type of activity is useful for familiarizing students with difficult terminology which they would later meet in the video clips, e.g., *people can't afford to….*

4.2. Institutional discourse patterns

This lesson stage is dedicated to raising cultural awareness of the institutional context. To facilitate contextual and situational understanding students are exposed to the settings through activities that emphasize the significance of spatial semiotics, i.e. *proxemics*, such as the shape of the Hall, or parliamentary seating. Students can be asked to comment on where the debates take place. For example, in the House of Representatives (US) the House Speaker is seated higher than the member speakers. This high position represents authority and power. Whereas in the House of Commons (UK) the seating forms a horse-shoe shape in which members of different political parties sit facing each other representing opposing political ideas and opinions. The multimodal analysis also revealed another semiotic resource known as *chronemics*, or the influence of time on the speakers' speech and type of utterances. Parliaments tend to regulate the timing of speakers in order to maintain order in the room. Therefore, controlled timing will lead to planned speech, as the speaker needs to be concise and clear in the few minutes they have to speak. In this way, there are noticeable differences

---

[6]'Uncertainty about Brexit is hampering business investment in Britain Companies are still not sure what will happen', *The Economist,* Oct. 8th 2019.

[7]'Trump's tax cuts helped billionaires pay less than the working class for first time'*, The Guardian,* Oct. 9th 2019.

between the UK and US debates in terms of turn-taking strategies and argumentation features, which are both influenced by chronemics. Generally, in US House debates speakers are allocated from 1 to 3/5 minutes, strictly regulated by House regulations to prevent interruptions and provide order. This maintains courtesy, impeding direct questions or open comments. Whereas, the UK House of Commons debates tend to be more argumentative as spontaneous speech is allowed, along with controlled timing. This creates possible reactions from the audience while an MP is speaking, such as backchanneling, contemptuous remarks, and unstructured interventions (Bevitori, 2004). The House Speaker is responsible for upholding order, and one fundamental rule is that, in order to maintain politeness and civility in the house to reduce conflict, the third person singular must always be used (in both the UK and US Houses), mitigated through formal turn-taking devices (see Table 3). This was followed by an investigation into variations in the way the debates were conducted, identifying formulaic language for polite turn-taking strategies, e.g. *Will the honourable Lady give way*? (UK); *the gentleman from Kansas is recognized for one minute* (US House Speaker). Students were able to familiarize themselves with the cultural differences in the formal register of the UK and US institutionalized formulaic language.

A quick corpus retrieval analysis of the most frequent recurring sequences of words can be carried out using corpus software; for example, figure 2 shows sample concordance lines for *honourable* (175 instances). The same can be done for *give way,* or *gentleman/ lady/ yield / recognized* in the US corpus.

Figure 2. Sample concordance lines for hon./honourable in UK corpus

```
 1      Christine Jardine (Edinburgh West) (LD) Will the hon. Lady give way?  Marsha De Cordova (
 2    way?  Marsha De Cordova (Battersea) (Lab) Will my hon. Friend give way?  Catherine McKinnell
 3      give way?  Catherine McKinnell I give way to my hon. Friend the Member for Battersea
 4.  Brexit clock.Ian Murray (Edinburgh South) (Lab) My hon. Friend has been utterly fantastic on
 5    Ali (Bethnal Green and Bow) (Lab) To reinforce my hon. Friend's point, according to the Bank 6
 5      pm instead of 7.30 pm. We were listening to the hon. Member for Newcastle upon Tyne North
 7    action anything they say.  Catherine McKinnell My hon. Friend raises an important point. That
 8      revocation should be on the table. I support my hon. Friend's speech.  Catherine McKinnell
 9      my hon. Friend's speech.  Catherine McKinnell My hon. Friend speaks from the experience that
10  instances.Jo Swinson (East Dunbartonshire) (LD) The hon. Lady is absolutely right that those
11    the people has changed? Catherine McKinnell  The hon. Lady makes some excellent and important
12   lady have.  Anna McMorrin (Cardiff North) (Lab) My hon. Friend is making an excellent speech
13    it goes to a public vote?  Catherine McKinnell My hon. Friend succinctly says what I will say
14  Allen (Rutherglen and Hamilton West) (Lab/Co-op) My hon. Friend is making an excellent point
15  Andy Jenkyns (Morley and Outwood) (Con) I thank the hon. Lady for giving way. Hearing your
```

Source: Author's elaboration

A listening activity was able to check the students' understanding of formulaic language in parliamentary speech, the patterns they produce and their functions. We can identify clear institutionalized strategies, conventionally used to ask, for example, a current speaker to *'give way'*. The students can note down the procedural moves they hear, and list them (Table 3), to be then checked together. To raise students' awareness of appropriacy and register, they can be asked to reflect on the use of these formulas, e.g. Why is the turn-taking language so formal? Would you use this type of language in a classroom?

Table 3. Sample US /UK formulaic institutionalized language

Institutionalized (mitigation) strategies for turn-taking moves	
UK	**US**
I thank the hon. Lady for giving way.	The gentlelady from California is recognized for 3 minutes.
Will the hon. Lady give way?	
It is a pleasure to serve under your chairmanship.	I yield two minutes to the gentleman from California.
It is a big honour not only to serve under your chairmanship, ...but to speak in this debate.	Reclaiming my time, Mr. Speaker.
	The gentleman shall be allowed five minutes to speak in opposition to it.
I thank the hon. Lady for her intervention.	
The hon. Gentleman raises important point	I rise in opposition to the amendment.

Source: Authors' elaboration

To help students get through the potential issues of unfamiliar and problematic language features, questions helped focus the students, such as, what is the argument being put forward? In this way students are guided to understand how the speaker constructs his /her argument through lexical choices and prototypical features of institutional and political communication (Partington and Taylor, 2010). I followed this up with a "fill in the gap" task (based on the transcript) to assess listening comprehension and reinforce learning of vocabulary and grammatical structures, and specialized terminology as well as interactional and rhetorical devices.

4.3 Identifying non-verbal communication

The aim of this lesson stage is to explore the interplay and relationship between language and non-verbal cues, and how they contribute to the meaning-making process, in particular their communicative and persuasive roles. The focus is on identifying the non-verbal elements which co-occur with speakers' utterances. Table 4 reports the learning task carried out in the classroom.

While viewing the students filled in the various columns on the worksheet with printed images of the video fragments (Baldry and Thibault, 2006), the gesture description, the gesture function, the gaze and the body movements of the speaker. The interactional language device and the verbal transcript were provided in the first two columns to help guide the students while they listen carefully. They were asked to listen and underline any word stress or indicate a pause in the transcript. The students watched the video clips (from 30 seconds to 1 minute) at least three times, focusing on the non-verbal cues such as tone of voice, pauses, gestures, body movements and gaze. Questions were asked during the viewing to point students towards the interpretation, such as, where is the speaker looking? What do you think the gesture means? Why is the speaker hitting the paper in her hand? In this way, students become aware of the communicative message. Multimodal analysis is an intensive, time-consuming activity, for this reason, only a few video clips were chosen, nevertheless, the analysis is replicable on similar multimodal studies of this type. It was important to guide the students to be objective and reduce the subjectivity of their interpretations by referring to a table of coding taxonomies on gestures (Kendon, 2004).

Table 4 shows only a small number of fragments that are representative of prominent gestures in both debates. The columns in Table 4 have been filled in according to the results of the classroom activity, together with some additional

explanatory notes based on previous research analyses. Prosodic stress is underlined in the transcript alongside the image frame.

Table 4. Sample worksheet on the multimodal analysis of non-verbal semiotic resources: identifying non-verbal communication.

	Interactional Language device	Verbal transcript (word stress underlined)	Gesture description	Gesture function	Gaze / face/ body / voice
Time 1.04	Parallelisms for persuasive emphasis: bicolons, tricolons	*But it's the big corporations, the millionaires, the billionaires, they are riding the rules [...] this is where their values are – they are on display.*	Pointing index finger, beating the finger up and down on the bicolons and tricolons, e.g. *the millionaires, the ...* .	deictic, indexical, finger wagging is accusatory - pointing at who is to blame.	Gaze towards notes (planned speech); as the speech intensifies towards the opponents seated on the left; voice stress in rhythm with the beat of the gesture like 'rapping' (student observation)
Additional notes	The finger pointing up and down synchronized with the words creates illocutionary force, representing indignation that values are not in the right place, and ideological positioning				
Time 1.09,13	Parallel structures - rhetorical questions for persuasive rhetorical effect	*How do we get job creation up? How do we get those wages for working class families up?*	Arm raised, palm up facing away (PuFA) towards the House speaker	Metaphorical; 'good' is 'up', representing economic benefits, growth and power	Gaze and body turned towards the House speaker, and then towards opponents on the right
Time 3.40	Rhetorical strategy of appreciation / thanking	*I want to most of all commend and express my most profound admiration to the architect of this measure*	Hands clasped together, fore fingers touch in a (church) steeple	Iconic: thanking position, like a prayer of thanks, appealing to peers	Gaze towards co-party members on the left of the aisle; smiling, affirms peer approval
Time 3.40,12	Turn taking move, invoking the House Speaker	*Mr. Speaker the House is out of order*	Hands clasped, resting on the reader stand	Iconic; hands clasped - waiting position	Gaze towards the gallery and the disturbance, upright body position representing confidence, smiling
Additional notes	Calmly waiting for the disturbance to be escorted from the House before resuming his speech, expression shows he will wait as long as it is necessary, commands the situation and the House				

Time 3.41	Negative evaluation: refusing and denying	*...for years arrogant condescending paternalistic ideology ...,* (pause). *Look at where it got us; flat wages, ... stagnation is a breeding ground for a class-based society... .* *That is not the American idea*	Hand slicing motion, palm facing down, moving back and forth (PFDMSS), hand slice on each word (6 times)	Iconic: representing negation; the hand slicing motion denotes refusal and rejection of ideology which is not the *'American ideal'*	Gaze moves from the right (opponents), to the co-party members seated on the left, knowing he has their consensus and approval
Additional notes	The hand beating (slicing) back and forth synchronized on each word adds persuasive rhetorical effect: the gesture refuses anything which interferes with the 'American dream'; illocutionary force makes the speech act emphatic.				
UK House debate					
Time 1.13	Argumentation : counter – concession speech act	*You are absolutely right ..., but it's time to call for the withdrawal agreement to be put back to the British public ..., because it's the only democratic way out of this current impasse*	Hand and palm open beating down towards the paper	Deictic and iconic: the hand beats on 'back to the public' and 'the only democratic way'. Hitting the paper is emphatic, to convince the opponent	Gaze and body in confrontational stance facing the opponent, reacting to the previous comment
Time 2.45	Political rhetoric - presupposition	*Can she now be honest and say* this *has* nothing *to do with democracy or* rerun *of the vote? It's about revocation*	Hand flip in circular movement	Metaphorical and iconic: circular movement to represent 'rerun of the vote', action appears flippant and ironic	Gaze direct and confrontational , turned towards the opponent speaker (Catherine Mckinell)
Additional notes	The inference is mistrust in the 'democratic' intents of the speaker. MP Gray (conservative party) questions MP Mckinell's real motives insinuating she is dishonest				
Time 2.46	Denial	Here no verbal cues, only non-verbal gesture in response to MP Mckinell	Hands apart palms up facing each other in a question like manner	Iconic: dismissive meaning 'I don't know what you mean'. Denying the accusation that	Gaze first towards the speaker then away in dismissive manner, facial expression

		he was insinuating she was dishonest	complements the gesture
Additional notes	MP Gray's gesture denies and dismisses MP Mckinell's accusations, as if downplaying and ignoring the MPs complaint.		

Source: Author's elaboration

In fragment 1, the speaker is raising and pointing an index finger. Her verbal cue '*the big corporations, the millionaires'*, she sees as the culprits of misplaced '*values'*, so her non-verbal cue, finger-wagging, is an accusatory gesture. The speaker in fragment 2 faces the House Speaker to make his speech more persuasive by looking at authority, adding credibility to his utterance. The arm raised up symbols 'good is up', in fact, he refers to economic growth and job growth. Fragments 3-7 represent a series of iconic gestures. For example, in fragment 8 the speaker refers to a '*small'* oil excavation project in Alaska, the fingers touching together to form a small ring make the environmental impact look minimal.

Turning to the UK debate, we can see the speech becomes more argumentative, due to turn-taking between speakers. The speaker in fragment 9 is angry because she has been accused of being dishonest and misguiding the British people. Her hands beat down on each word. The response from the MP in fragment 10 is to ignore his accusation that she is being dishonest. As we can see a silent gesture with two hands flipping up dismissing her comment and confrontation. On the whole, we can see there are predominantly deictic gestures pointing to identity, and spatial and temporal contexts; indexical gestures with finger-pointing that frequently co-occur with accusatory speech acts supporting ideological stance. Sometimes indexical gestures also signal the desire to speak and interrupt someone, or on behalf of the House speaker to keep order in the House (UK debate). Some of the gestures become iconic when they stand for the meaning of a word or verb itself, e.g. in fragment 5 the hand slicing movement representing 'not' and rejection.

4.4 Consolidation activities

The objective of the post-viewing stage is to consolidate what students have assimilated and apply what they have learned to be later transferred to the real world. General questions are always useful to find out how much the students were motivated by the videos. For example, the teacher can ask how the semiotic resources work together to communicate a convincing speech, making one speaker more persuasive than another. Bearing in mind that the students had been given no political input, surprisingly, personal preferences were not based on political bias, but rather on how effective the speech was, depending on body behavior and voice.

In the post-viewing lessons, students can be given speaking activities to be carried out in a role-play debate where the students have the chance to reproduce what they have learnt and develop their critical thinking skills. Topics on social-economic issues are generally interesting for the students: for example, the environmental costs of fashion, the controversial exploitation of the Amazon forest, eco-friendly consumption. Students can be put into groups for role-play, for example, representing: the government, NGOs, an energy group, enterprises, and the local community. Each group prepares its arguments. During the debates, students were not able to incorporate gestures naturally, but they certainly appeared to be more aware of their own gestures

and how they accompany speech, and more attuned to the whole array of paralinguistic and kinesic features in the meaning-making process.

Some of our classes are composed of international students involved in Erasmus and Erasmus Mundi (e.g., from China, India, Europe). These were able to compare their own cultures in the context of government, institutions, etc. On the whole multimodal learning activities led to students feeling more self-assured and aware of their speech and body language, which they may eventually carry over to the professional world. Finally, student power point presentations also benefited from the multimodal analysis.

5 Conclusions

The purpose of this study was to show how a multimodal approach to language can be integrated into an ESP university language classroom, in so doing it has tried to demonstrate the benefits of encouraging learners to develop skills to understand and produce texts in the target language, and enable students to improve their ability to structure various types of discourse and communicate orally using language, intonation, and gestures. These types of learning activities based on multimodal literacy also highlight cultural differences in how people communicate non-verbally.

It seems reasonable to affirm that input other than the verbal mode is particularly relevant in situated communication context especially in ESP settings, where domain-specific linguistic discourse features might pose considerable challenges for language learners.

As regards methodology this classroom application has demonstrated the advantages of using multimodal corpora, confirming recent literature. We can see how such corpora allow language learners to carry out a fine-grained analysis of several aspects of oral communication, which goes beyond the 'monomodality' of written corpora (Bateman, 2008; 2013). Students have constant access now to spoken discourse from all types of multimodal resources and technological devices and they can learn to decipher not only the language but also how a message is conveyed through a variety of semiotic resources especially the role of gestures. Nevertheless, the complexity of building multimodal corpora still needs to be addressed more effectively since the natural complexity of multimodal data can overwhelm personal efforts. This is a clear limitation, the fact that the compilation and creation of multimodal corpora is labour-intensive and a time-consuming endeavor. The complexity of building multimodal corpora remains a challenge and still needs to be addressed more effectively since the high amount of multimodal data (in just one video) can overwhelm personal efforts. Another limitation is the amount of teaching and learning time needed for effective acquisition in the multimodal ESP classroom. Multimodal teaching applications require intensive listening in terms of paying attention to fine-tuned details. As far as language learning is concerned it would be more beneficial to the students for multimodal activities to occur over a series of lessons; however usually the students' curriculum is so tight with other disciplines in their curricular courses that not enough didactic space is allocated to language learning and consequently practice for intensive multimodal scrutiny is reduced.

In response to the research questions on whether the multimodal approach facilitates learning outcomes, this is difficult to verify empirically. More empirical data is required over a longer phase of teaching activities. Nevertheless, I believe my own observations and data retrieved from the students' feedback are replicable and valid on

a small scale. The world is constantly evolving with economic, social and technological change, more demanding of new types of work. This implies the need to prepare our students for a media-rich world of new demands. Hence, the findings of this case study can be applied to communication classes aimed at preparing future economists and others in related disciplines. Avenues for future research could involve creating multimodal corpora for academic and professional purposes, identifying specialized communities of practice which would benefit from tailored-made courses, thus enriching ESP course development among practitioners.

References

Baldry A. and Thibault P. J. (2006), *A multimodal transcription and text analysis: A multimedia toolkit and coursebook*. Equinox, London.

Bateman J. (2013), Multimodality and film. In: Chapelle C. A. (ed.), *The encyclopedia of applied linguistics*, 4030 - 4033. Blackwell, Oxford.

Bateman J. (2008), *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. Palgrave Macmillan, London.

Bevitori C. (2004), Negotiating conflict: Interruptions in British and Italian parliamentary debates. In: Bayley P. (ed), *Cross-cultural perspectives on parliamentary discourse*, 87-109. Benjamins, Amsterdam, Philadelphia.

Crawford Camiciottoli B. and Bonsignori V. (2015), The Pisa audiovisual corpus project: a multimodal approach to ESP research and teaching, *ESP Today*, 3 (2), 139-159.

Crawford Camiciottoli B. and Campoy-Cubillo M. (eds.) (2018), Special Issue: Multimodal Perspectives on English Language Teaching in Higher Education, *System*, 77, 1-114.

Duff P. and Talmy S. (2011), Second language socialization: Beyond language acquisition in SLA. In: Atkinson, D. (ed.) *Alternative Approaches to SLA,* 95 - 116. Routledge, London.

Halliday M. A. K. (1994), *An introduction to functional grammar* (2nd ed.). Edward Arnold, London.

Jewitt C. (2009), *The Routledge Handbook of Multimodal Analysis*. Routledge, London.

Incelli E. A. (2021), Engaging students in multimodal literacy practices in a university ESP context: towards understanding identity and ideology in government debates, *Multimodal Communication*, 11 (1), 49-61.

Kendon A. (2004), *Gesture: Visible action as utterance*. Cambridge University Press Cambridge.

Kress G. and Van Leeuwen T. (2006), *Reading Images: The Grammar of Visual Design* (2nd. Ed.). Routledge, London.

Lim F.V. and Tan K.Y.S. (2018), Developing Multimodal Literacy Through Teaching the Critical Viewing of Films in Singapore, *Journal of Adolescent & Adult Literacy*, 62 (3), 291-300.

Lim F. V. (2011), *A Systemic Functional Multi Modal Discourse Analysis Approach to Pedagogic Discourse*, Ph.D. thesis, National University of Singapore, Singapore.

Lim F. V. (2021), Investigating intersemiosis: A systemic functional multimodal discourse analysis of the relationship between language and gesture in classroom discourse, *Visual Communication*, 20, 34-58.

Machin D. and Mayr A. (2012), *How to do critical discourse analysis: A multimodal introduction.* Sage, Thousand Oaks, CA.

O'Halloran Kay L. (2004). *Multimodal Discourse Analysis: Systemic Functional Perspectives*. Continuum, London.

O'Halloran K.L., Tan S. and Smith B. A. (2016), Multimodal Approaches to English for Academic Purposes. In: Hyland, K. and Shaw, P. (eds.), *The Routledge Handbook of English for Academic Purposes,* 256-269. Routledge, London.

Partington A. and Taylor C. (2010), *Persuasion in Politics. A Textbook*. LED, Milano.

Partington A. and Taylor C. (2017), *The Language of Persuasion in Politics: An Introduction*. Routledge, Taylor & Francis Ltd, London, New York.

Prior P. (2013), Mulitmodality and ESP research. In: Paltridge B. and Starfield S. (eds.), *The handbook of English for specific purposes*. Wiley-Blackwell, Chichester (UK).

Querol-Julián M. (2011), *Evaluation in Discussion Sessions of Conference Paper Presentation: A Multimodal Approach*. LAP, Saarbrücken.

Royce T. (2006), Intersemiotic complementarity: A framework for multimodal discourse analysis. In: Royce T, and Bowcher WL. (eds.), *New Directions in the Analysis of Multimodal Discourse,* 63–109. Lawrence Erlbaum, Mahwah (NJ).

Van Dijk T. A. (2004), Text and Context of Parliamentary Debates. In: Bayley P. (ed.), *Cross-Cultural Perspectives on Parliamentary Discourse,* 339–372. John Benjamins, Amsterdam.

Wittenburg P., Brugman H., Russel A., Klassmann A. and Sloetjes Wittenbur H. (2006), *ELAN: a Professional Framework for Multimodality Research*, Max Planck Institute for Psycholinguistics, AH Nijmegen.