Recent Perspectives on Gesture and Multimodality

Recent Perspectives on Gesture and Multimodality

Edited by
Isabel Galhano-Rodrigues,
Elena Zagar Galvão
and Anabela Cruz-Santos

Cambridge
Scholars
Publishing

Recent Perspectives on Gesture and Multimodality

Edited by Isabel Galhano-Rodrigues, Elena Zagar Galvão
and Anabela Cruz-Santos

This book first published 2019

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-3536-3
ISBN (13): 978-1-5275-3536-7

TABLE OF CONTENTS

Acknowledgements vii

Introduction 1
Isabel Galhano, Elena Galvão, and Anabela Cruz-Santos

Chapter 1 8
Cohesive Gestures in Spontaneous Conversation
Manon Lelandais and Gaëlle Ferré

Chapter 2 19
An Exploration of Chilean EFL Teachers Interactional Practices
in Feedback Provision
Katherina Walper

Chapter 3 31
Time Reference in Weather Reports
Gaëlle Ferré

Chapter 4 41
Multimodal Clarification Requests in Human-Agent Interaction
Branislav Bédi

Chapter 5 58
A Review of Technologies for Gestural Interaction in Virtual Reality
Jorge C.S. Cardoso

Chapter 6 74
Musical and Expressive Gesture as Expression and Communication
Slavisa Lamounier and Paulo Ferreira-Lopes

Chapter 7 85
Multimodal Analysis of Translanguaging
Anna Ladilova

Chapter 8 95
Gestures in Hip-Hop Video-Clips: A Cross-Cultural Approach
Maria Clotilde Almeida and Bibiana de Sousa

Chapter 9 112
Harpo, de Funès: Le Combat contre les Mots
Isabelle Guaïtella

Chapter 10 121
Performance Mode under the Microscope: A Cognitive Semiotic Analysis of Eye Gaze and other Body Movements in a Contemporary Dance Improvisation
Carla Fernandes, Vito Evola and Joanna Skubisz

Chapter 11 136
Speech-Gesture in Early Language Acquisition in Portuguese Children
Teresa Maló Sequeira

Chapter 12 147
Gesture Production in Portuguese Toddlers with and without Hearing Impairment
Etelvina Lima and Anabela Cruz-Santos

Chapter 13 159
Some Recurrent Foot-Gestures of Rejection and Negation
Isabel Galhano Rodrigues

Contributors 171

ACKNOWLEDGEMENTS

This volume was made possible thanks to the collaboration of all the authors and their enthusiasm for sharing their interest in studying visible (and audible) actions.

Special thanks to Rita Almada for allowing us to use the drawing by her family member, José Almada Negreiros, for the cover of the book. Many thanks also to Ana Vasconcelos, of the Calouste Gulbenkian Foundation, for providing us with the image of the drawing, a masterpiece for gesture studies: it depicts a scene of spontaneous conversation between two fishwives in the North of Portugal.

INTRODUCTION

Gesture Studies is a highly interdisciplinary research field which was officially established in 2002 with the creation of the International Society for Gesture Studies (ISGS) in Austin, Texas. Within it, gesture has been approached from a wide range of theoretical perspectives and methodologies. The increasing number of research laboratories, multidisciplinary conferences, workshops, and academic programmes focusing on the study of gesture and multimodality in communication is ample testimony to the attractiveness of the field as well as to its notable expansion all over the globe. In addition, the emergence of new research paths is undoubtedly a sign of a very dynamic and productive scientific community. Gesture research has provided new insights into face-to-face communication, language acquisition, embodiment and conceptualisations of different dimensions, thus contributing to a deeper understanding of communication and language use in society.

In the past few decades there has been a proliferation of interdisciplinary research on gesture and multimodality in a variety of fields such as Anthropology, Ethnography, Psychology, Linguistics, Communication Studies, Language Acquisition, Neurosciences, Education Sciences and Deaf Education in particular, Performance Studies including Music, Dance, and Theatre, Human Kinestesics, Computer Science, Artificial Intelligence, and Information Technology.

This book offers a selection of the papers presented at the iGesto'17 – International Conference on Gesture and Multimodality, which took place at the Faculty of Arts and Humanities of the University of Porto (FLUP), Portugal, on 2-3 February 2017, as well as contributions from iGesto researchers, illustrating some of the dominant perspectives on gesture studies pursued in Portugal (Galhano-Rodrigues 2013).

In general, the chapters cover recent developments in the exploration of "gesture". This term is traditionally used both in a narrow sense, to designate upper-limb movements, and, in a broad sense, to refer to a wider range of kinesic modalities such as, for example, head and torso movements, gaze orientation, and facial expressions. For this reason, the term "gesture" may be often "vague and not useful", as criticised by Kendon (2017), above all in the case of comparative semiotic analyses of languages. Considering the name of the conference that motivated the

writing of these chapters, as well as the heterogenous approaches on visible body actions, the terms gesture and multimodality were maintained in the title.

This volume will be of interest to scholars, researchers, and students from various spheres of knowledge, whose common denominator is their multimodal outlook on communication. Besides contributing to the understanding of gesture and multimodality within different disciplinary perspectives, they also bring new insights into approaches to gesture in different languages and cultures, such as British and American English, European and Brazilian Portuguese, French, German, and Icelandic.

The chapters were organised according to the theoretical orientation followed by the authors and the specific objectives of their studies. Thus, as will be described in the following paragraphs, the first four chapters focus on the relation between gesture and speech in different kinds of interactional settings; the following two address technological issues concerning the development of avatars based on linguistic-oriented research on gesture; the next four chapters focus on communicative and emotive body movements in various performance contexts; the subsequent two discuss gesture acquisition in infants; finally, the last chapter deals with foot gesture as an alternative modality to manual gestures.

Spontaneous conversation as a product of social interaction has been investigated extensively from a linguistic perspective. Within this approach, gestures have been described according to their semiotic properties and their semantic relations to co-occurring speech. These studies seek to explain the emergence of gestures and the causes that have motivated the preference for some motion configurations. Speakers' physical experiences in the environment, their motion habits, and their consequent embodiment of mental representations of different abstract dimensions have been considered instrumental in the creation and stabilisation of recurrent gesture forms. To understand these issues, different modalities of visible actions have been analysed according to their forms, meanings, and functions as part of the utterance within a specific context (Müller, Ladewig, and Bressem 2013). Intertwined with speech and 'coming along' with words in utterance production, gestures (above all hand gestures) depict objects and ideas, locate items in the interactional space, focus on elements or parts of speech, and convey meanings beyond those expressed by words. According to their semiotic and more or less iconic properties, gestures may contribute to the referential meanings of the utterance, or they may be used like pragmatic elements of speech, fulfilling different types of discourse organisational and/or interactional functions.

One of the pragmatic functions attributed to gesture is the establishment of cohesion between elements of the utterance. This property has been described in terms of the reactivation of gesture forms co-occurring with the same verbal referent. McNeill (1992) identified this phenomenon in the context of narratives and called it *catchment*. Cohesion can also be considered from a multimodal perspective, as presented in **Chapter 1.** Based on the qualitative analysis of a spontaneous face-to-face interaction, the authors offer an accurate description of how cohesion is achieved through a combination of hand gestures, gaze orientation, head and eyebrow movements as well as prosody. They show how all these modalities contribute to delimiting appositive relative clauses in descriptive and argumentative discourse sequences.

Further contributions describe uses of gestures in more specific contexts. **Chapter 2** is dedicated to multimodal interaction practices in the classroom. The consideration of different communicative resources – such as teachers' gestures and prosody, which provide better and non-stressful conditions for learning – is an underexplored topic (Gullberg 2014). The inclusion of gesture studies in L2 acquisition research is significant, as it represents an extension of a field which traditionally emphasises verbal communication. The author focuses on the moments of feedback provision and describes how teachers' kinesics (body movements and gestures) convey effective supportive cues for eliciting students' responses.

Chapter 3 is concerned with the use of gesture in English and French TV weather reports, speech events that require multimodal resources and a conventionalised speech style. The information about the properties of different weather components (space, time and weather condition/temperature) is expressed by words or visually represented on screen through symbols and images. Pointing to locations on a map co-occurs with the verbal expression of specific information. The authors analyse how the references to the different weather components are achieved either by means of primary semiotic modes – the spoken words and the background screens – or by means of secondary semiotic modes (gesture and prosody). Through the comparison of English and French weather reports, the authors explore how the semiotic modes collaborate to convey information about these different dimensions.

Viewed in their broader sense of body actions, gestures, like language, are shaped by culture. Code switching or translanguaging provides an interesting field of study for cultural influence on bodily behaviour. This topic is explored in **Chapter 4,** which deals with intra- and intercultural spontaneous interactions between Brazilian and German speakers in the same setting. The analysis distinguishes some of the behavioural features

displayed by these speakers when switching from one language to the other.

Studies on gesture use in different interactional settings have contributed greatly to the development of **virtual agents** capable of interacting with humans. **Chapter 5** describes the multimodalities involved at a very specific moment of interaction, i.e., when a clarification is needed, and applies the results to the context of Human-Agent interaction in Icelandic. These clarification requests involve hearer signals, such as attention yielding, different kinds of positive or negative feedback signals, e.g. agreement, misunderstanding, misinterpretation, as well as different modes of expressing emotive reactions to a speaker's turn. These multimodal forms of expression share a common feature: they are generally performed to avoid a breakdown in communication. As these kinds of utterances are frequent in human interaction, they are essential to the development of Embodied Conversational Agents (ECAs) (Pelachaud, 2009)

In its turn, gestural (and bodily) interaction with virtual reality (VR) requires specific technologies. **Chapter 6** addresses the importance of considering a complete representation of the user's body for this interactional context. Recognising the limitations of commercial devices used by designers of VR experiences, the chapter offers a review of the technologies that can be used to provide different kinds of gestural detection. It presents a set of novel ways of exploring virtual worlds, which are not limited to the manual actions directed to an object, but are extended to real gestures.

Due to the richness and complexity of body motions as a means of displaying emotions, postures and other expressive body movements have attracted the attention of a number of researchers from artistic and performance studies, such as music, dance, theatre, and film. The material produced by actors in filmed theatre staging (Kipp and Martin 2009) and in other kinds of performances is well suited for the analysis of the expression of emotion, states of mind, and intentions. An application of gesture studies to the area of digital music is presented in **Chapter 7**. The authors hypothesise "that the human body and the musical instrument are configured as multisensory and interactional spaces". The emergence of expressive gestures and the construction of musical gestures are connected to the development of a Digital Sock, a prototype of a digital musical instrument created to record movements. Similar to gestural expressions in interaction, whose meanings and functions reveal the embodiment of sociocultural issues and space, musical gestures are determined by their multisensory complex environments. Here the instrument represents the multisensory space with which the subject interacts, making decisions

according to memory imprints, while manipulating and controlling the sounds. This tool is also designed for use in clinical settings, for instance as a device to establish a dialogue with autistic children, or as a pedagogical tool for other types of educational activities.

Chapter 8 describes and compares the postures, gestures, and facial expressions of Rap and Hip-Hop performers of two different cultural and language backgrounds: Portuguese and German. The motivations for specific bodily actions are presented as metonymic and metaphorical representations of central concepts within the Rapper/Hip-Hop movement, such as "RAPPERS ARE WEAPONS", "RAPPERS ARE FIGHTERS" and "RAPPERS ARE URBAN DOMINATORS". While a common cross-cultural 'gesture grammar' is identified, some of the gestures, such as 'hands-up', display cultural specificities.

In **Chapter 9**, the discussion is centred on the techniques developed by the "clown actors" Luis de Funès and Harpo to represent their characters. The author recognises two facets of the performing bodies: the actor's body itself, with its personal and cultural motion habits, and the semiotic body whose expressive movements convey specific meanings to the audience. Both actors had a strong preference for the use of relatively exaggerated body movements to express their art in a film industry era that was dominated primarily by words and puns. The author stresses Harpo's and Luis de Funès' remarkable acting skills, explaining how they perfected specific and transgressive nonverbal techniques to represent the interaction of their characters in different social and cultural contexts. Their behaviours function as effective means of social critique.

Chapter 10 describes an innovative experiment in a setting where the participants were asked to interact with each other without speech: the communicative event consists in an improvisation game conceived by a contemporary dance choreographer. The participants, both expert performers and non-performers, collaborated in improvising a dance piece, making "artistic" or "aesthetic" movements that should become part of it. The comparison of the participants' gaze shifts and communicative body movements yielded interesting results regarding their behaviours. The authors propose the existence of a "performance mode" that emerges from social practice and bodily interactional experience, and allows for an aligned and collective mindset that makes it easier to track other people's mental states and leaves more space for creativity.

After a series of chapters that examine a variety of gesture uses in adults, the next two chapters turn to the emergence and development of gesture and language in infants. Their focus is on how communicative gestures and speech develop during the early language acquisition process

and on how gestures become stable units that express meanings correlated with speech. **Chapter 11** presents a study about gesture acquisition in four children whose first language is European Portuguese and who were observed from the age of 7 to 24 months. The objective of this study is to detect the prevalence of each of the two communicative modalities (gesture-only acts, speech-only acts and acts composed of gesture-speech combinations) during the observation period. Although the validity of the results has to be verified in further research based on a more extensive corpus, the study provides interesting insights into gesture acquisition in European Portuguese – an underestimated and almost inexistent research topic in Portugal. **Chapter 12** explores the properties of the communicative gestures produced by Portuguese toddlers from 8 to 18 months with and without hearing impairment. It corroborates the findings presented in Chapter 11 as regards the sequence in which gesture types emerge in the process of gesture-language acquisition in Portuguese children. The authors point out the importance of identifying children's gestures in the pre-linguistic period, since it allows "for a better understanding of the process of gesture-language acquisition" and helps clinical interventions and educational assessment. For this reason, besides gesture type and its emergence in each of the groups analysed, the parameters for this comparison also include hand shape and stroke duration.

Finally, the last chapter raises the following question: "Can we gesticulate without our upper limbs?" The description of coverbal foot gestures performed by a speaker without upper limbs shows that this modality can function as an efficient alternative to hand gestures. Foot configuration and the semiotic properties of the stroke in foot gestures expressing rejection and negation were found to be similar to those of the hand gestures used for expressing these meanings. In other words, foot gestures appear to share some of the semiotic properties of, for example, open hand prone gestures with lateral movement (Kendon 2004). This study provides further evidence of the robustness of gesture while at the same time calling for the inclusion of a neurological perspective in studies of gesture use.

References

Galhano-Rodrigues, Isabel. 2014. "Gestures in Southwest Europe: Portugal." In Body - Language - Communication: An International Handbook on Multimodality in Human Interaction, edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva H. Ladewig, David McNeill, Sedinha Tessendorf. Berlin: Mouton de Gruyer, 1259-1266.

Gullberg, Marianne. 2014. Gestures and second language acquisition. In Body — language — communication: An international handbook on multimodality in human interaction, edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva H. Ladewig, David McNeill, Sedinha Tessendorf. Berlin: Mouton de Gruyer, 1868-1875.

Kendon, Adam. 2004. *Gesture. Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kendon, Adam. 2017. Languages as semiotically heterogenous systems. Commentary on the Susan Goldin-Meadow and Diane Brentari. *Behavioral and Brain Sciences* 40, e59. https://www.researchgate.net/publication/317262568_Languages_as_semiotically_heterogenous_systems [accessed Jan 13 2019].

Kipp, Michael and Jean-Claude Martin. 2009. "Gesture and Emotion: Can basic gestural form features discriminate emotions?" *Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII-09)*:1 – 8. DOI: 10.1109/ACII.2009.5349544 https://ieeexplore.ieee.org/document/5349544 [accessed Jan 13 2019].

McNeill, David. 1992. *Hands and Mind.* Cambridge: Cambridge University Press.

Pelachaud, Catherine. 2009. "Modelling multimodal expression of emotion in a virtual agent." Philosophical Transactions of the Royal Society B: Biological Sciences 12;364/1535: 3539–3548. doi: 10.1098/rstb.2009.0186

Isabel Galhano-Rodrigues
Elena Zagar Galvão
Anabela Cruz-Santos

January 2019

CHAPTER 1

COHESIVE GESTURES IN SPONTANEOUS CONVERSATION

MANON LELANDAIS AND GAËLLE FERRÉ

1. Introduction

This study discusses how gestures create different types of cohesion in spontaneous conversation. Cohesion in discourse refers to the set of resources for constructing relations beyond grammatical structure (Halliday and Hasan 1976).

Cohesion is relevant to examine online language production and comprehension, and presents implications for discourse modelling. Yet compared to the vast amount of research on cohesion either on speech (Ariel 1994; Detges and Weidhaas 2016) or on gesture and speech (Perniss and Özyürek 2015; Frederiksen 2016), the prosodic contribution to cohesion is often left out and its temporal organisation with the gestural devices is not addressed. Likewise, most of the studies on gestural cohesion focus on narrative sequences (e.g. Debreslioska et al. 2013; Perniss and Özyürek 2015), and little is known about other types of discourse sequences characterising spontaneous conversation, such as argumentation or description.

In face-to-face conversation, participants negotiate meaning through multimodal contributions, in which the linguistic resources of speech interface with gesture. These modes do not work independently of one another, although a particular mode may weigh more than the others at some point (Norris 2004). We propose a qualitative study on the temporal organisation of modality-specific features used to create cohesion in a discourse sequence.

2. Theoretical background

While few studies have analysed cohesion from a multimodal point of view, some gestural and prosodic features have been shown to participate in cohesion (Hoetjes et al. 2015; Perniss and Özyürek 2015).

2.1 Gesture

Gesture includes co-speech bodily movement considered part of an utterance (Kendon 2004). We focus on gaze, head and eyebrow movement, and hand gestures.

Representing referents through hand gestures is a cumulative process (Streeck 2009). Repetition throughout an interaction creates cohesion, as reference is maintained through cohesiveness of space, handedness, and style (McNeill and Levy 1993; McNeill 2005).

A gesture hold also explicitly represents two related ideas at the same time (Frederiksen 2016). Additionally, two speech segments can be related through their production in one gesture unit (Enfield 2009).

Gaze often moves away from the co-speaker for discourse elaboration when the speaking turn is secured (De Kok and Heylen 2009). A change in gaze direction towards the co-speaker announces a discourse boundary or an appeal (Holler et al. 2014).

Isolated head beats and eyebrow rises emphasise particular entities (Cavé et al. 1996) as they are linked to prosodic focalisation (Granström and House 2005). A cohesive stretch of speech is then likely to be produced without such isolated cues.

2.2 Prosody

Intonation can convey cohesive information that is not marked by verbal means (Bolinger 1984). Throughout a vocal paragraph, pitch height declines progressively (Wennerstrom 2001). Successive downsteps group different units in a single paragraph (ibid.).

To link a prosodic unit to an adjacent segment, pitch rises on the final syllable of the inserted segment, indexing it as prefacing further speech (Wells 2006). In this context, a final rising contour can be a mark of cohesion. While a rising-falling contour on the nuclear syllable codes emphasis (Ward and Hirschberg 1985), flat or falling-rising contours appeal to previous or entailed information (ibid.).

3. Corpus and methodology

3.1 Recording

The corpus used for this study, ENVID (described in Lelandais and Ferré 2016), is a two-hour collection of dialogues in British English. Each participant had a lavalier microphone (providing a separate audio track), and was filmed in a wide-angled shot.

3.2 Transcription and annotation

The corpus was transcribed verbatim in Praat (Boersma and Weenink 2013).

3.2.1 Prosodic coding

The corpus was segmented into tone-units (Crystal 1969; Wells 2006) based on dynamic pitch contours. An algorithm (Hirst 2007; Bigi 2012) notes pitch height (in Hz) on target syllables, allowing us to calculate mean pitch values for specific segments. Each measured pitch value is also compared to preceding ones, to note significant changes in the curve regarding the speaker's pitch range (Top, Bottom) or regarding the neighbouring tones (Upstep, Downstep, Same, Low, High).

3.2.2 Syntactic annotations

The syntactic annotations are oriented towards a multimodal study of subordination, which has thoroughly been linked with structural cohesion in Discourse Analysis (Lambrecht 1996). This study focuses on appositive relative clauses, which do not single out a nominal referent, but make an additional comment about a referent or a whole clause (ibid.). In "we went into a place called Tropicana, which was horrible" the appositive relative clause "which was horrible" evaluates "a place called Tropicana", which can however be identified independently as a referent.

Appositive clauses were allocated and coded on a separate track. We chose occurrences without an interruption, which were preceded and followed by another tone-unit by the same speaker. On a second track, the preceding tone-unit was labelled L (left co-text), and the subsequent one labelled R (right co-text).

3.2.3 Gesture annotations

Communicative gestures were manually coded in ELAN (Sloetjes and Wittenburg 2008), based on gesture phrases (Kendon 2004). Each phrase starts at the onset of the gesture and ends at the return to rest position. In the case of two consecutive gestures, the first phrase ends at a significant change in shape/trajectory. We annotated hand gestures following McNeill's typology (2005).

Head movements were labelled as nods (downward-upward movement on a vertical axis as in assent), shakes (side-to-side movement on a horizontal axis), tilts (inclination sideways on a diagonal axis), beats (downward chin movement on a vertical axis without any acquiescing value), or jerks (sudden backward chin movement). On separate tracks, gaze direction was annotated as either towards the co-participant or away, and eyebrow movement distinguished between rise and frown.

3.3 Working hypotheses

Our analysis focuses on the nature and temporal distribution of cohesive features. If cohesion is to be found in our extracts with the cues described in the literature, we test whether (1) appositive relative clauses are concerned; (2) both gestures and prosody are cohesive at the same time; (3) the gestural and prosodic cues differ across different types of discourse sequences.

4. Qualitative analysis

We focus on two occurrences of appositive relative clauses, positioned in two types of discourse sequences.

4.1 Appositive clause in a descriptive sequence

In extract (1), Joey describes a place in the Welsh countryside. Extract (1) is associated with Figure 1, where (a), (b), (c), and (d) correspond to different moments in its production.

(1)	Joey	my mum's school it's the school i used to go to have [(a) POINTING this place]
	L	in # [(b) HEAD BEAT Gwyn Fech]an #
	SC	which is [(c) POINTING # just a place it's not even a village] #
	R	hem # which [(d) POINTING and the nearest town is Crickhowell] #

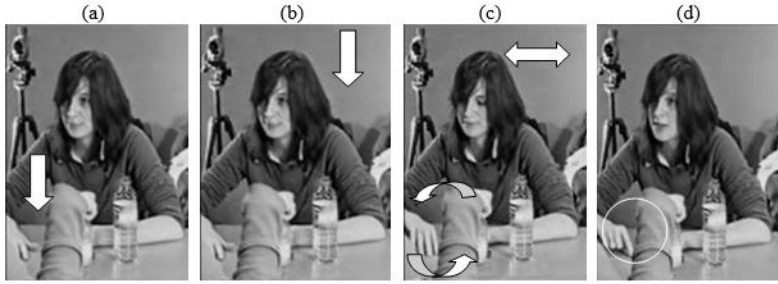

Fig.1-1: Several distinct pointing gestures and head movements during the production of (1).

This sequence shows distinct hand gestures in their configurations and trajectories. The appositive clause is produced in co-occurrence with a different hand gesture unit (c) from that in (a): clear articulatory boundaries are established in the sequence. However, these are all pointing gestures, also realised with the same hand. Joey positions her flat-palm right hand on the table in a deictic manner just before mentioning "this place" (a), and marks "Gwyn Fechan" as pragmatically important with a head beat in L (b). She then lifts her palm from the table and produces circling gestures around this anchoring point, in correlation with "just a place" in SC (c). The location of a given referent is modified, becoming less precise as Joey delimits a fuzzier space. This metadiscursive gesture circles around a specific entity whose lexical reference is also being verbally commented upon. It also acquires a modal value: Joey is not sure about the exactness of its location and uses a circle as a sign of approximation. The comment is accompanied by a head shake and a frown, carrying a negative modal value and reinforcing her subsequent negative verbal assertion ("it's not even a village"). The co-articulation of hand and head gestures allows Joey to take two different modal positions: one about the location of Gwyn Fechan as a referent, the other about its relevance in the discourse sequence. The tone-unit "it's not even a village"

is included in the circling gesture, although it is an independent clause from the point of view of grammar. After facing some planning difficulties, Joey produces a clear pointing gesture in R (d) with a single finger this time, along with "the nearest town".

Figure 2 below shows the vocal realisation of extract (1).

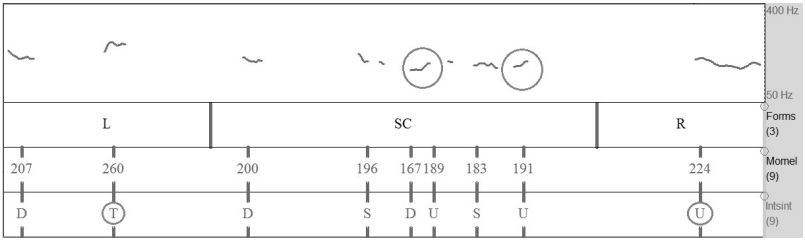

Fig.1-2: Intonation curve of extract (1) in Praat (first transcription track shows segments, second track gives Momel's corrected F0 values in Hz, and third track shows Intsint's coded values).

L ends with a very high pitch on "Gwyn Fechan", representing a pragmatic focus echoing the head beat and Joey's gaze towards the co-speaker in the gestural modality. Although SC stands out from L and R through silent pauses and differences in pitch height, the two tone-units "which is just a place" and "it's not even a village" show a similar rising contour, and are uttered at the same pitch height.

To sum up, in extract (1), the circling gesture in SC modifies an already-established referent, and ties two different speech segments in a same gesture unit. A similarity is also indicated vocally between these two tone-units. This gestural and prosodic cohesion, realised in two different manners, triggers the interpretation of "it's not even a village" as part of the appositive comment rather than part of the resumption of the description.

4.2 Appositive clause in an argumentation sequence

Extract (2), associated with Figure 3, is part of an argumentation. Rhianna lays out the reasons why she hates Ryanair.

(2) Rhianna L even [(a) METAPHORIC compared to (b) BEAT Easyjet]
SC which is [(c) BEAT another low cost company # (d) HOLD
Alex yeah # Easy- Easyjet's quite con-
i think it's considered to be better than Ryanair
Rhianna R yeah] (e) RETRACTION i hate Ryanair

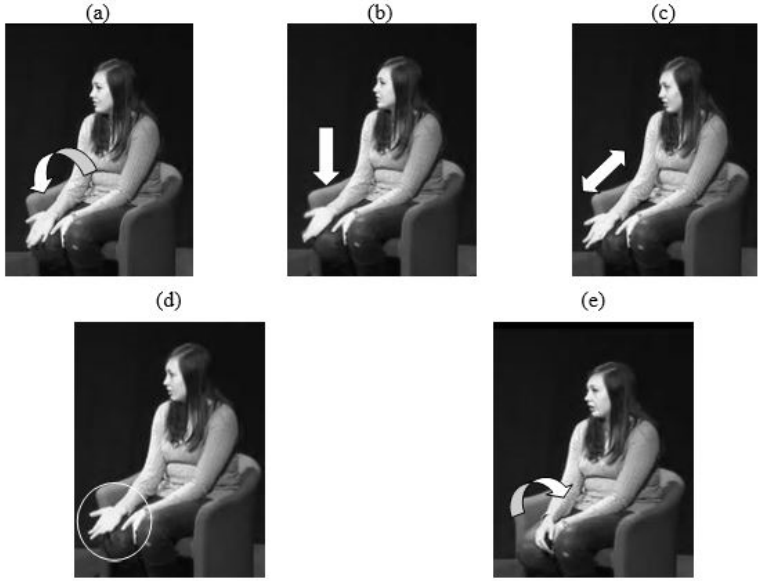

Fig.1-3: Repeated beat gesture (c) superimposed on an open palm-up hand gesture (a) during the production of extract (2). The palm-up configuration is held (d) until the production of R.

This sequence does not feature any representational gesture. Rhianna first states her opinion with a frown. She then produces a metaphoric (a) before she mentions Easyjet, opening a comparison. Her open-palm hand towards the co-speaker takes an interactional dimension as she alludes to a piece ofinformation that already belongs to the common ground (Easyjet), marked with a superimposed beat (b). Another beat in SC (c) reinforces both the comparison and the appeal to the common ground, as Rhianna's palm opens even more and is drawn a bit closer to Alex. Alex takes this opportunity to insert a backchannel comment, while Rhianna is holding her open-palm hand gesture (d). This hold has a turn-keeping value, as the speaker indicates that the space she has opened is not closed yet. She then restates her opinion in R and retracts her gesture (e). This extract is then characterised by metaphorical hand gestures acting on a pragmatic plane. The repeated beat gesture in SC appeals to the common ground between participants while the hold helps regulate the interaction.

Figure 4 shows the vocal realisation of example (2).

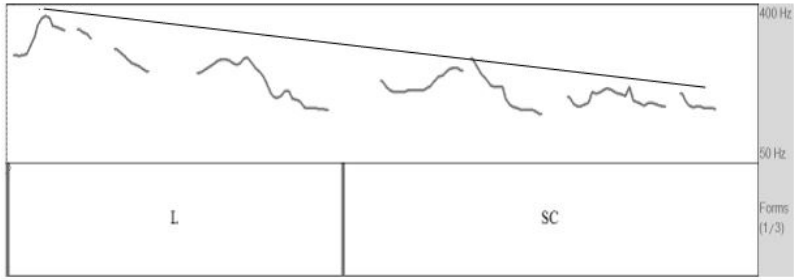

Fig.1-1: Intonation curve of L and SC in extract (2) in Praat (first transcription track shows segments; the oblique descending line is shown in black).

At first sight, this extract is characterised by distinct emphatic contours and a non-neutral interval between L and SC. SC's first syllable features a pitch upstep, indicating that L and SC make two distinct discursive moves. However, L's nuclear contour is replicated in a lower pitch key, which signals similarity.

To sum up, gestures show cohesion in terms of type, form, and held units in (2), while prosody indicates cohesion with a replication of pitch contour.

5. Discussion and conclusion

Cohesive cues are found in gesture and in prosody during the production of appositive relative clauses. However, different breaks are also identified in the sequences. Appositive clauses are realised under distinct tone-units instead of being integrated to the left co-text's tone-unit, and are also realised with specific gesture units, as seen with the circling gesture in the description, and with the beat gesture in the argumentation. These distinct gesture units suggest that speakers do not gesture less during such subordinate clauses.

Cohesion is achieved differently between gesture and prosody in the description. While the hand gesture integrates two tone-units in one gesture unit, prosody duplicates a first tone-unit on a second one. In the argumentation, the superimposed beats on the metaphorical gesture create a repetition pattern, which can be compared with the pitch contour duplication. However, the hold also relates different parts of speech together. In both sequences, prosody and gesture are not in correspondence at the level of the tone-unit, as they do not group speech in the same way during the production of appositive clauses. However, at discourse level,

both modalities are involved in the creation of cohesion over a speech segment including the subordinate construction.

In the description, hand gestures mainly show a representational function. However, the circling gesture used in the subordinate clause does not continue mapping referents, mainly giving information about the relevance of an established referent. The speaker is centred upon her own discourse during the production of the appositive clause. In the argumentation, hand gestures show a more abstract function at the level of discourse organisation. The hand beat produced with the subordinate clause has a modal function but also manages the common ground while the hold helps with turn-keeping at an interactional level. The speaker is more centred upon the co-speaker as she produces the appositive clause. Despite these differences between representational and abstract aspects, gesture creates cohesion in both sequences, involving techniques such as integration of two speech segments into the same gestural unit, repetition, and spatial modification.

As far as prosody is concerned, cohesion is achieved with the same prosodic device in both sequences, i.e. a duplication of pitch contour. However, the duplicated nuclear tones are rising in the first example, and rising-falling in the second. The latter also shows a lowering in key. No rhythmical cue has been found to contribute to cohesion in our examples.

A useful development to this study would be to enquire into other syntactic types of subordinate clauses across these same types of discourse sequences, and test whether the tendency between gesture and prosody is confirmed, reversed, or increased.

References

Ariel, Mira. 1994. "Interpreting anaphoric expressions: A cognitive versus a pragmatic approach." *Journal of Linguistics 30* (1): 3–42.

Bigi, Brigitte. 2012. "SPPAS: A Tool for the Phonetic Segmentations of Speech." In *Proceedings of LREC 2012*, 1748–55. Luxemburg: European Language Resources Association.

Boersma, Paul and David Weenink. 2013. *Praat: Doing Phonetics by Computer.* Accessed 30 January 2013 from http://www.fon.hum.uva.nl/praat/

Bolinger, Dwight. 1984. "Intonational Signals of Subordination." In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 401–413. Berkeley: eLanguage.

Cavé, Christian, Isabelle Guaïtella, Roxane Bertrand, Serge Santi, Françoise Harlay and Robert Espesser. 1996. "About the Relationship

between Eyebrow Movements and F0 Variations." In *ICSLP 96*, 2175–2178. Philadelphia: PA.

Crystal, David. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.

Debreslioska, Sandra, Asli Özyürek, Mariane Gullberg, and Pamela Perniss. 2013. "Gestural viewpoint signals referent accessibility." *Discourse Processes 50* (7): 431–456.

De Kok, Iwan and Dirk Heylen. 2009. "Multimodal end-of-turn prediction in multi-party meetings." In *Proceedings of the International Conference on Multimodal Interfaces*, 91-98. New York: ACM.

Detges, Ulrich, and Thomas Weidhaas. 2016. "Coherence and interaction. The French causal connectors comme and car."*Language Sciences 58*: 111–125.

Enfield, Nick. 2009. *The Anatomy of Meaning: Speech, Gesture and Composite Utterances*. Cambridge: Cambridge University Press.

Frederiksen, Anne Therese. 2016. "Hold + Stroke Gesture Sequences as Cohesion Devices: Examples from Danish Narratives." *San Diego Linguistics Papers 6*: 2 13.

Granström, Björn, and David House. 2005. "Audiovisual Representation of Prosody in Expressive Speech Communication." *Speech Communication 46*(3). 473–84.

Halliday, Michael A. K. and Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.

Hirst, Daniel. 2007. "A Praat Plugin for Momel and INTSINT." In *Proceedings of ICPHS*. Accessed 26 December 2012 from http://www.icphs2007.de/

Hoetjes, Marieke, Ruud Koolen, Martijn Goudbeek, Emiel Krahmer, and Marc Swerts. 2015. "Reduction in gesture during the production of repeated references." *Journal of Memory and Language 79*: 1-17.

Holler, Judith, Louise Schubotz, Spencer Kelly, Peter Hagoort, Manuela Schuetze and Asli Özyürek. 2014. "Social eye gaze modulates processing of speech and co-speech gesture." *Cognition 133* (3): 692–697.

Kendon, Adam. 2004. *Gesture: Visible Action as Utterance.* Cambridge: Cambridge University Press.

Lambrecht, Knud. 1996. *Information structure and sentence form.* New York: Cambridge University Press.

Lelandais, Manon and Gaëlle Ferré. 2016. "Prosodic boundaries in subordinate syntactic constructions." In *Speech Prosody*, 183–187. Boston: ISCA.

McNeill, David. 2005. *Gesture and Thought.* Chicago: University of Chicago Press.
McNeill, David and Elena T. Levy. 1993. "Cohesion and Gesture." Discourse Processes 16: 363–386.
Norris, Sigrid. 2004. *Analyzing Multimodal Interaction: A Methodological Framework*. London: Routledge.
Perniss, Pamela and Asli Özyürek. 2015. "Visible Cohesion: A Comparison of Reference Tracking in Sign, Speech, and Co-Speech Gesture." *Topics in Cognitive Science 7* (1): 36–60.
Sloetjes, Han and Peter Wittenburg. 2008. "Annotation by Category: ELAN and ISO DCR. In *Proceedings of LREC 2008.* Accessed 28 November 2013 from http://www.lat-mpi.eu/tools/elan/.
Streeck, Jürgen. 2009. *Gesturecraft. The Manu-Facture of Meaning.* Philadelphia: John Benjamins.
Ward, Gregory and Julia Hirschberg. 1985 "Implicating uncertainty: The pragmatics of fall-rise intonation." *Language 61* (4): 747–776.
Wells, John C. 2006. *English Intonation: An Introduction.* Cambridge: Cambridge University Press.
Wennerstrom, Ann K. 2001. *The Music of Everyday Speech. Prosody and Discourse Analysis*. Oxford: Oxford University Press.

Appendix. Transcription conventions

one line of transcription corresponds to one tone-unit

#	pause
[…]	gestural activity
-	interrupted construction
L	left co-text
SC	subordinate construction
R	right co-text

CHAPTER 2

AN EXPLORATION OF CHILEAN EFL TEACHERS' INTERACTIONAL PRACTICES IN FEEDBACK PROVISION

KATHERINA WALPER

Introduction

This paper reports on an exploratory study of English as a Foreign Language (EFL) teachers' interactional practices when providing feedback. It follows a multimodal conversation analytic (CA) approach (Schegloff 2007) with focus on teachers' embodied interactional practices to secure student responses. When approaching student groups, teachers often asked questions, or first-pair parts (FPP), to elicit the relevant grammatical or lexical items. However, these questions did not always result in student uptake, or second-pair parts (SPP), and repair trajectories ensued (Kasper 2006; McHoul 1990). The present study is especially concerned with teachers' interactional strategies in these situations. The two cases presented here illustrate the ways in which teachers dealt with lack of response (case 1) and a pedagogically-unfit answer (case 2).

With regards to formative feedback, Heritage and Heritage (2013) found that teachers use question-answer sequences as a diagnosis tool to identify the help needed. The present study seeks to expand on this by exploring the trouble sources that arise during this stage, most of which deal with vocabulary or grammar issues. Studies on vocabulary elicitation in foreign language classrooms which focus on teachers' interactional practices are scarce. Lazaraton (2004) highlighted that iconic gestures (McNeill 1992) are a relevant source of input, whereas Koshik (2002) found that teachers posed designedly incomplete utterances to elicit knowledge displays of vocabulary items. Olsher (2004) explored learners' embodied completions of turns that were verbally incomplete but with projectable trajectory and proved that these led to mutual elaborations

among students working in groups. Though not through a CA approach, van Compernolle and Smotrova (2017) explored the synchronisation of speech and gesture and highlighted how impromptu explanations were recipient-designed. The present study seeks to build on these and make teachers' embodied interactional practices visible as well as identify their role in securing student responses to move the pedagogical project forward.

Methodology

A jigsaw picture-story task was designed to elicit different kinds of interactions, such as peer work, teachers and whole class, teachers and student groups. The last type of interaction corresponds to the data subset reported in the present chapter and consists of approximately three hours of teacher-group interactions. Data was collected in six public secondary EFL classrooms in Southern Chile. Ethical approval was obtained at the University of York; permission to record was granted by the headmasters at each school, and all participants signed informed consent forms. Data was transcribed according to CA conventions and was annotated using ELAN (2018) with tiers for the verbal output and each of the multimodal means. The research question guiding the present study is: What are EFL teachers' interactional practices to secure student responses when providing them with feedback during group work?

The picture story task is based on the book 'The Great Escape' (Dupasquier 1996) and depicts a prisoner (Alf) that escapes from prison and goes into a cinema, a hospital, a circus, etc. Students work in groups organising the pictures and writing sentences about one of the events. They then share their sentences to make up the whole story. Teachers walk around the classroom giving them feedback and helping them complete the task.

Multimodal transcripts follow Mondada's (2014) transcription conventions in which each modality is transcribed below the verbal means, so as to reflect the temporality of embodied behaviour. Teacher gaze is signposted by %, and teachers' hand gestures by $. The symbols >> are used when the practice continues onto the next line. Screen shots accompany the main points and are marked with the symbol # on the transcriptions. Teaching materials are transcribed as TM, left hand as LH and right hand as RH. Due to space limitations, student gaze has not been added to the transcript, but further studies will look into how students display recipiency and use gaze shifts to mobilise help from their classmates.

Cases

Case 1: Dealing with lack of uptake

In this classroom, Teacher D provided students with questions about the pictures in the story. Students in this group did not understand the task, so they turned to the teacher for help. In this excerpt, the teacher is explaining to the students that they have to answer the questions given, and that, with these answers, they will be able to recount the story to their classmates later on.

One such question is 'Where is he?' (line 03), which makes it relevant for the students to answer: 'circus'. As there is no uptake (2.3 second gap, line 04), the teacher's interactional work focuses on showing students the meaning of the word 'where'. In other words, the teacher orients to students' lack of response as a language problem, meaning they do not understand the word 'where'.

```
01 Tea: these questions. if you answer the questions (0.4)
02      you can tell the story (.) for example (0.3)
03      where is he?
04      (2.3)
05 Tea:→where is he?
06      (0.8)
07 St2: ee::h
08 Tea:→>where<
09      (0.3)
10      →>where. (.) where is Nati. (.) over there<
11      (2.3)
12 Tea:→where is Ramón?
13      (0.7)
14 St2: dónde  es      [tá
        where  (he)  is
15 Tea:                [correct
16      (0.7)
17      so. where is alf?
18      (0.3)
19 St2: the circus
20 Tea: okay. <so alf is in the circus>
```

There are various instances (signposted with arrows on the transcript above) in which the teacher is further elaborating on the original elicitation turn (line 03), pursuing an answer from students. As students do not produce the second pair part, the teacher 'escalates' (Stivers and Rossano 2010) the embodied resources to mobilise a response: gaze shifts (Pomerantz 1984), deictic or pointing gestures, and iconic gestural practices are used to secure a response:

```
05 Tea: %where #is & he?
  Tgz: %at TM>>
  Thd:            &LH palm up traces TM>>
              #2-1
```

Fig.2-1: Gaze at TM, LH open palm up tracing TM

When producing the question ‘where is he’ (line 05), the teacher points and directs her gaze towards the relevant picture in the TM. Students shift their gaze towards the images, but they do not produce an answer.

The teacher orients to the absence of the SPP as a problem of understanding the word ‘where’ and initiates a sequence in which she exemplifies its meaning through pantomimic gestures, using the word in the real context with people in the room (Figs. 2-2 and 2-3):

```
10a Tea: >where.# (.)
  Tgz: >>at St2>>
  Thd: >>LHRH to forehead>>
               #2
```