

## Article

# A Multimodal Approach to Catalan Pragmatic Markers: An Exploratory Study

Ana Llopis Cardona <sup>1,\*</sup> , Iris Hübscher <sup>2</sup> and Adrián Cabedo Nebot <sup>1</sup> 

<sup>1</sup> Val.Es.Co. Research Group, Faculty of Philology, Translation and Communication, University of Valencia, 46010 Valencia, Spain; adrian.cabedo@uv.es

<sup>2</sup> Institute of Language Competence, Zurich University of Applied Sciences, 8400 Winterthur, Switzerland

\* Correspondence: ana.b.llopis@uv.es

**Abstract:** This exploratory study aims to investigate the co-occurrence of Catalan pragmatic markers that fulfill an interactive function and multimodal cues, such as manual gestures, adaptors, head gestures, and eye gaze. To do so, we utilized spontaneous face-to-face conversations between ten different dyads, all of whom were undergraduate students discussing movies that they had recently seen. We processed the data with descriptive quantitative statistics, specifically employing exploratory multifactorial statistical visualization techniques, including multiple correspondence analysis (MCA) and cluster analysis. The results showed that pragmatic markers generally co-occur with multimodal cues. Although the multimodal system works independently, when pragmatic markers are uttered, both systems tend to harmonize in accordance with general communicative tasks, particularly turn management, alignment, and social affiliation.

**Keywords:** gestures; multimodality; pragmatic markers; interactive function

## 1. Introduction

Pragmatic markers, also known as discourse markers, are widely used in daily communication as devices to mark textual relationships, to show speakers' attitudes, manage the speaker turn, or intertwine interpersonal relationships (Fraser 1996; Briz 1998; Martín Zorraquino and Portolés 1999; Pons 2006; Cuenca 2013; among others). However, while there has been a great deal of research on pragmatic markers in different languages at the verbal level (see, e.g., Aijmer and Simon-Vandenberg 2006; Fischer 2006; Degand et al. 2013; among others), little is still known about the interplay between pragmatic markers and their multimodal manifestation (see, e.g., Ferre 2011; Esfandiari Baiat et al. 2013; Abuczki 2014; Hata 2016; Debras 2021 for exceptions) and, to our knowledge, no investigation of such nature has been undertaken in Catalan to date.

The omission of the investigation of multimodal cues might be explained by the fact that until recently, multimodal corpora were more scarcely available. This omission can also be understood within the context of the theoretical framework on pragmatic markers, where the focus predominantly lies on (co-textual) verbal components, often overlooking contextual elements, which are typically only explained at a theoretical level. Additionally, prosodic, gestural, and non-verbal components are commonly excluded from consideration. This inattention is incoherent in relation to one's own nature of the communicative process, which is inherently multimodal: a speaker performs kinesic movements at the same time that a listener perceives auditory but also visual cues to interpret messages.

In interactions, we normally use speech in conjunction with gestures that are orchestrated dynamically to create the utterance's meaning (McNeill 1992, 2013). Co-speech gestures are made up of their own system, which is integrated in the communicative process as an essential component helping comprehension and reinforcing the contact with the interlocutor (Lloberes and Payrató 2012; Maricchiolo et al. 2014).



**Citation:** Llopis Cardona, Ana, Iris Hübscher, and Adrián Cabedo Nebot. 2024. A Multimodal Approach to Catalan Pragmatic Markers: An Exploratory Study. *Languages* 9: 125. <https://doi.org/10.3390/languages9040125>

Received: 23 December 2023

Revised: 1 March 2024

Accepted: 21 March 2024

Published: 1 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Müller (2017) posits a continuum of increasing conventionalization, starting from singular gestures, progressing to recurrent gestures, and ultimately culminating in emblems (Kendon's continuum in McNeill 1992). This continuum correlates with more independent or less speech-linked usage of gestures. Singular gestures are spontaneous and created freely, yet within a shared cultural repertoire of techniques, so they are easy to interpret in context (Müller 1998, 2013). Recurrent gestures function as both conventional and idiosyncratic elements. They are often associated with gesture families with a common schematic meaning stemming from experiential frames, such as the 'Palm Up Open Hand' used to present a question or the 'Palm Away Open Hand' to reject a topic or idea (Kendon 2004, 2017; Müller 2004, 2013; Payrató and Teßendorf 2014; Teßendorf 2014). Their specific functions can be interpreted in context and generally serve pragmatic purposes (discourse, modal, performative, etc.; Ladewig 2010). Finally, emblems carry a conventional meaning and normally work in isolation.

In this exploratory study, we focus on gestures and other multimodal cues that accompany pragmatic markers that fulfill an interactive function, mainly singular gestures, and only occasionally recurrent gestures. From this basis, we aim to answer the following research questions:

- (1) How frequently do interactive pragmatic markers co-occur with multimodal cues?
- (2) Can discernible patterns be identified or is the co-occurrence completely irregular?

We examined specific co-speech gestures and other multimodal cues, namely manual gestures, adaptors, head gestures, and eye gaze. To do so, we utilized spontaneous face-to-face conversations between 10 different dyads (all undergraduate students). The data form part of a bigger multimodal corpus, which comprises various task-based interactions, all varying in their degree of experimental control (Section 3).

The data were processed with descriptive quantitative statistics to find out the co-occurrences between co-speech gestures and multimodal cues, and pragmatic markers more generally. Moreover, to enhance our analysis, we utilized exploratory multifactorial statistical visualization techniques, specifically multiple correspondence analysis (MCA) and cluster analysis. These techniques were implemented using the FactoMineR package developed by (Lê et al. 2008). By employing these methods, we were able to generate comprehensive profiles for the pragmatic markers under investigation.

After this introduction, Section 2 outlines the functional approach to pragmatic markers taken in this study, presents the theoretical background on multimodality, specifically regarding potential co-speech gestures, and lastly, reviews previous research on the multimodal approach to pragmatic markers. Section 3 describes in detail the corpus data, the set-up of the analysis, and the methodology applied. Section 4 addresses the research questions. Initially, it presents the findings concerning the co-occurrence of pragmatic markers and gestures (Section 4.1). The section continues by displaying the clusters resulting from a multiple correspondence analysis (Section 4.2), in which the most remarkable features of any cluster are described. Section 5 interprets the statistical results and delves into the most relevant implications. To sum up, we synthesize the main results achieved.

## 2. Theoretical Background

### 2.1. Pragmatic Markers with an Interactive Function

Although *discourse marker* tends to be the general term most frequently used, following Aijmer and Simon-Vandenberg (2006), we used the term *pragmatic markers* to refer to the items of interest in this study. We did this to highlight that they do not necessarily relate to discourse segments (Schiffrin [1987] 1997; Fraser 1996, 2009), but rather that they mainly fulfil a pragmatic and/or discursive function in the interaction (Aijmer and Simon-Vandenberg 2006, pp. 225–27).

As is known, the pragmatic marker set encompasses items belonging to disparate grammatical categories—adverbs, conjunctions, and interjections, but also nouns and verbs—whose lexical meaning is partly or completely blurred out and without morphological inflection as a result of a grammaticalization process. Thus, the most remarkable feature

is that they perform some function linked to discourse markedness and, consequently, they make up a functional class (Schiffrin [1987] 1997; Bazzanella 2006; Fischer 2006; Pons 2006; Cuenca 2013). Pragmatic markers generally provide clues to interpret the relationship among speech units, the speaker's attitude toward what (s)he is saying, as well as signaling aspects related to speech management.

Scholars distinguish a considerable amount of specific pragmatic and/or discourse functions and certain macro-functions that work as broad subcategories related to distinct dimensions of discourse, such as connection, argumentation, discourse structure, modalization, or interaction (Briz 1998; Martín Zorraquino and Portolés 1999; Portolés [1998] 2001; Fraser 2009; Cuenca 2013; Pons 2006; López Serena and Borreguero 2010). Some studies on pragmatic markers take macro-functions and/or specific functions as a starting point to later examine the forms that perform such functions. Our study is based on this functional approach.

We focused on pragmatic markers fulfilling an interactive function; hence, we analyzed pragmatic markers normally used in spoken Central Catalan and informal situations, such as *és que* ('that is'), *bé* ('ok', 'fine'), *no sé* ('I don't know'), *home, no?* ('man, huh?'), and so on. Unlike discourse markers typical of written language, these pragmatic markers are characterized by being quite polyfunctional, since they can display more than one function at the same time and can fulfil different functions depending on the context and position (Briz and Pons 2010).

Indeed, interactive functions are said to be closely linked to the exchange structure, specifically to the position with respect to a discourse unit, either turn or act (Briz and Pons 2010). These markers have more mobility (initial, medial, and final), and changes in position usually trigger a distinct function. For instance, in example (1), *no?* in the final position is asking for confirmation. In example (2), where *no?* is in the medial position, its function is that of attracting the hearer's attention:

- (1) Paula: Has vist la pel·li, *no?*  
[Paula: Have you seen the movie, right?]
- (2) Joan: Tracta d'un noi que té problemes *no?* i se'n va de casa  
[Joan: It's about a boy who has troubles, right? And goes home]  
Occasionally, they relate a turn to extralinguistic information (3):
- (3) Pere: Coneixes al profe d'Audiovisuals?  
[Pere: Do you know the professor of Audiovisual (communication)?]  
Marc: (*through a head nod*)  
Pere: *Però* si és molt famós a la Facultat  
[Pere: But if he is very well known at the faculty]

Specific interactive functions consist of the assorted communicative tasks performed by speakers by conversing. First, there are those related to speech management or turn exchange, such as taking, keeping, or yielding the floor. In addition, while taking the floor, pragmatic markers can be used for digressing from the current topic (4), introducing a new topic, or returning to the previous one (Fraser 2009):

- (4) Núria: *Per cert*, vols vindre el divendres al cinema?  
[Núria: By the way, would you like coming to cinema on Friday?]

Additionally, they can serve for showing agreement or disagreement with the interlocutor's opinion, for justifying themselves (5), or simply for expressing emotions regarding what has just been said (Briz 1998; Cuenca 2006; Pons 2006; López Serena and Borreguero 2010):

- (5) Elena: *és que* tinc partit de voleibol per la tarda  
[Elena: It's just that I have a volleyball game in the afternoon.]

Finally, as expected, they can play a role in the speaker-hearer relationship in different ways, such as drawing the hearer's attention (2), eliciting a response (1), expressing solidarity (6) (*tio*), and showing perception or understanding (6) (*ja*), among others:

- (6) Eduard: La vaig veure doblada *tio*  
[Eduard: I saw it dubbed dude]  
Pere: *ja*  
[Pere: yes]

## 2.2. Multimodal Cues in Face-to-Face Interaction

In face-to-face interactions, the communication process involves integrating speech and gestures as components of the utterance's meaning (McNeill 1992, 2013). Concerning hand gestures, McNeill (1992) differentiated propositional gestures (linked to the content) from non-propositional gestures (related to discourse activity). Similarly, Kendon (2004) distinguished between gestures carrying referential meaning and those with pragmatic meaning. From now on, we will concentrate on pragmatic gestures, also called "interactive" gestures (Bavelas et al. 1992) and "speech handling" gestures (Streeck 2009), among other terms.

Pragmatic gestures include "all actions of the hands (and a variety of other body parts, notably the face, head, and shoulders) by which aspects of the communicative interaction are displayed" (Streeck 2009, p. 10). They are very context-dependent, so they apparently do not have a conventional meaning. However, most of them are observed to be "used repeatedly in different contexts and its formational and semantic core remains stable across different contexts and speakers" (Ladewig 2010). These are named recurrent gestures (Ladewig 2010; Müller 2013, 2017) (Section 1).

As pragmatic markers, the pragmatic functions performed by non-propositional gestures have been categorized into different classes according to the communicative domain. Schiffrin ([1987] 1997) or Kendon (1995) stated that gestures can work autonomously as discourse markers. In Table 1, the main pragmatic functions identified by Müller (1998) and Kendon (2017) are displayed (see also Payrató and Teßendorf (2014) for an updated revision).

**Table 1.** Pragmatic functions of gestures according to different authors.

Authors	Pragmatic Functions
Müller (1998)	Performative Modal Discursive
Kendon (2017)	Operational <sup>1</sup> Modal <sup>2</sup> Performative <sup>3</sup> Parsing/discursive <sup>4</sup>

As can be seen, the complete repertory is made up of five pragmatic functions<sup>5</sup> that cover a great part of the communicative domains; thus, it is a good starting point to work with, though it might require further clarifications and an extension that takes into consideration other possible functions, such as face-threatening acts or face-flattering acts (politeness; cf. Hübscher et al. 2017), argumentative or meta-discursive movements, etc.

Regarding their interactive function, gestures serve to open/close conversations and manage turns, namely through actions such as "waving, greeting, (...), withdrawing, beckoning or halting, requesting or inviting turns at speaking, and so on" (Kendon 2017, p. 168). While Kendon (2017) recognizes this function, he does not elaborate on it in detail. It somehow corresponds to two functions distinguished by Bavelas et al. (1992, 1995): seeking a response (understanding, agreement, or help) and coordinating the turn (taking, giving, or opening it). Recent studies also emphasize that gestures can contribute to turn construction or the organization of turn-taking (Bohle 2014; Dressel and Satti 2021; Cabanes 2023). Among various turn-related activities (taking, keeping, yielding, appointing, etc.), turn-taking emerges as the one most prominently marked by distinct gestures (Schmitt 2005, in Mondada 2013; Dressel and Satti 2021; Cabanes 2023).

From the listener's perspective, "generic responses" (Bavelas et al. 1995) and "backchannels" (Knight and Adolphs 2008) are particularly noteworthy. The most typical signal produced during ongoing conversation corresponds to head nods, which may occur in isolation or in conjunction with markers, such as *mm*, *hm*, *uh huh*, *yeah*, and *right* (cf. Kidwell

2013). These kinds of verbal and non-verbal marks have been examined as “continuers” (Schegloff 1982) or “acknowledgments” (Jefferson 1972) in conversational analysis.

For Knight and Adolphs (2008, p. 178), backchannels are any device used “to provide feedback in communication”, such as verbalized signals, sentence completions, requests for clarification, head nods, and shakes. These researchers noticed that verbal backchannels usually functioned as continuers and convergence tokens. Continuers signal that “the addressee is listening, desiring the speaker floor holding narrative to continue”, while convergence tokens are used “to mark agreement” and “to help maintain good relations, by reinforcing commonality throughout the discourse” (Knight and Adolphs 2008, p. 181). They also identified verbal signals with nods in 41 instances (slightly more than 60%), with *yeah* being the most common, and in 27 instances without nods (Knight and Adolphs 2008, p. 184). According to Stivers (2008), these two functions can be fused in special contexts. She found that head nods in the mid-telling phase signify alignment with the discourse activity in progress, as well as express the listeners’ affiliation with the speaker’s stance. A recent study on turn-taking in Spanish conversations reached similar conclusions (Cabanes 2023).

### 2.3. Studies on the Interplay between Pragmatic Markers and Their Multimodal Manifestation

Research on pragmatic markers from a multimodal approach are scarce so far due to the fact that the availability of multimodal conversational corpora is still rather reduced. In fact, the first general study on this matter, conducted by Hata (2016), was not based on proper data but on reasonable assumptions on the interplay between co-speech gestures and discourse markers, which were illustrated by observations mainly coming from research carried out by experts on gestures (e.g., McNeill 1992; Bavelas et al. 1992; Kendon 1995). According to Hata (2016, p. 45), “there is a huge implication regarding the correlation between DMs and gestures-in-talk which marks the speaker’s interactive attempt to structure and manage ongoing discourse”. This postulation seems to be supported by the characteristic features of co-speech gestures, as well as discourse markers. On the one hand, co-speech gestures contribute to both semantic and pragmatic levels in discourse (cf. McNeill 1992). On the other hand, since context is key to interpret discourse markers in use, and context in interaction includes auditory and visual perception, multimodality should be taken into account. For instance, Hata (2016, pp. 45–47) underlines that discourse markers in conjunction with gestures can signal the structure of discourse, with discourse metaphorically taken as an object, or mark argumentative relations (e.g., *on the other hand*; cf. Kendon 1996 in Hata 2016). Co-speech gestures can also play a modal function role by emphasizing some messages, for instance, the contrast expressed by *but* would be intensified by means of a beat gesture (Hata 2016, pp. 47–48).

In relation to experimental research, only a few studies have been published to date, which focus on specific forms concerning French (Ferre 2011; Debras 2021), English (Esfandiari Baiat et al. 2013), or Hungarian markers (Abuczki 2014). Esfandiari Baiat et al. (2013) and Abuczki (2014) observed that the pragmatic markers under study (*well*; *mondjuk*, ‘let’s say’; *ugye*, ‘is that so?’; *amúgy*, ‘otherwise’) are not necessarily used with multimodal cues, whereas Ferre (2011) noticed that the frequency of co-speech gestures depends on every discourse marker (e.g., higher in *alors* than *in fact*). Debras (2021) also found that *je ne sais pas* is normally used with special multimodal cues. Regarding the distinction of functions, the presence/absence of co-speech gestures seems to offer insights into distinguishing functions in some cases (*mondjuk* and *ugye* in Abuczki 2014, or *je ne sais pas* in Debras 2021). Even the type of gesture could be relevant in that regard, as Ferre (2011, p. 674) observed in the inference and conclusion functions, where discourse markers are produced together with emblems and adaptors. Additionally, Debras (2021) detected a distinct, immediate kinesics environment in relation to lexical and pragmatic functions of *je ne sais pas*: it is accompanied by adaptors and hand referential gestures when used to deny, while it is uttered with a shrug component if it displays more pragmaticalized meanings (epistemic, subjective, and interpersonal).

Their specific results, however, are so disparate that it is not possible to draw definitive conclusions. Much work, and a high quantity of data, is needed to find out if there are any regular features in relation to forms or specific functions, and whether there are cultural differences or some more generalizable patterns.

In addition, an interesting issue concerns the status of the role of co-speech gestures that occur together with pragmatic markers. Generally speaking, discussion is set out as a dilemma: either co-speech gestures provide a contextual clue to interpret pragmatic markers, or they make up an autonomous system, harmonized in this case with pragmatic markers. [Maricchiolo et al. \(2014, pp. 1466–67\)](#) explained these two perspectives in terms of a facilitation or a conversational function: co-speech gestures accomplish a facilitation function, consisting of a verbal simplification and support to disambiguate or complete the message; however, the conversational function is about “improving the comprehension and the reciprocity between the interactants” ([Maricchiolo et al. 2014, p. 1467](#)).

### 3. Methodology

#### 3.1. Corpus Data

The data used in this study were collected at the Universitat Pompeu Fabra in Barcelona and form part of a bigger corpus previously published by [Brown et al. \(2023a\)](#) and [Brown et al. \(2023b\)](#). For the current paper, we used the data of the natural conversation task and focused solely on the interactions between friends. The corpus also contains data from the same task, where the main participant (not relevant for this study) interacted with a status superior, and three other tasks, namely a description of the “Tweety Bird” cartoon, a map task, and a roleplay. In the natural conversation task, in this paper referred to as the movie conversation task, the participants were instructed to talk about movies that they recently saw. The movie conversation task was selected because, despite being recorded in a laboratory and having a task-like format, it closely resembles a natural conversation, specifically focusing on discussing movies with the interlocutor.

The advantages of a corpus-based study methodology are manifold. Firstly, it provides a comprehensive and representative sample of language usage, allowing researchers to study language phenomena in a more natural and authentic context. Secondly, corpus-based studies enable quantitative analyses, providing reliable and statistically valid insights into language patterns and tendencies.

The corpus of this study consists of 15 conversations with contributions from 30 different speakers, all grade students (18 males and 12 females). They are all Catalan-dominant (71%) bilingual speakers of Catalan and Spanish and live in and around Barcelona. All interactions in the study occurred exclusively between speakers of the same sex, and each pair were friends. The participants consecutively took part in dyads in a sound-proof recording booth. Each dyad sat facing each other on chairs that were fixed to the ground. The interactions were video recorded from a side-on camera angle with a Panasonic HPX-171. The distribution by file can be examined in [Table 2](#).

**Table 2.** Distribution of times and pragmatic markers per conversation.

File	Duration (min: s)	Interactive Pragmatic Markers
1	4	41
2	2	19
3	4:31	37
4	3	17
5	2	26
6	2:43	50
7	3	35
8	3	33

Table 2. Cont.

File	Duration (min: s)	Interactive Pragmatic Markers
9	5	53
10	2:43	23
11	1:32	22
12	3:44	34
13	3	44
14	3:24	6
15	3:19	10
<b>Total: 15</b>	Total: 45:36 (min)	Total: 450

In total, each file had a duration ranging from 2 to 5 min. Therefore, the total annotated time was nearly 46 min. As shown in Table 1, while some conversations showed a high number of markers (50 in conversation 6), others had much lower occurrences (6 in conversation 14). Overall, 450 pragmatic markers, with varying forms and functions, were included in the analysis phase, in conjunction with the variables (Section 3.2).

### 3.2. Coding and Analysis

All multimodal cues and the functions and positions of the pragmatic markers were annotated in ELAN (verbal and visual).

In relation to multimodal cues, the starting point of the analysis was to identify when there was co-occurrence with pragmatic markers and when there was not. The second step consisted in selecting those multimodal cues more frequently used; thus, after a preliminary analysis, it was decided that multimodal cues with frequencies greater than 5 would be included in the final analysis (Table 3).

Table 3. Multimodal cues of the analysis.

Multimodal Cues	Values
Gaze	at interlocutor/averted/closed
Head_gesture	no/nod/tilt/shake/turn
Manual_gesture_type	no/iconic/beat
Adaptor	no touching/touching

Additionally, for previous studies on politeness and social distance, each manual gesture was coded in terms of its size (big vs. small on the vertical, horizontal, and front-back), handedness (both hands, yes or no, and in this case, right hand vs. left hand), and appearance of tenseness (hand shape, closed, and when three or more fingers were pressed together vs. open), which was also taken into account in the analysis as potentially interesting additional information on a greater dynamism in manual gestures (Table 4).

Regarding the functions of the pragmatic markers, the list of interactive functions shown below is based on previous studies (namely, Briz 1998; Portolés [1998] 2001; Pons 2006; López Serena and Borreguero 2010), as well as from the observation of our corpus data. They can be grouped into three categories in relation to the interlocutor (speaker/hearer) and type-tasks: management of the turn, speaker's communicative tasks, and verbal reactions (Table 5). Actually, the functions concerning the management of the turn often overlap with functions of speaker's communicative tasks or verbal reactions. For instance, when *no?* is used for seeking for confirmation, it is also giving the floor, or if *però* (*but*) starts a turn, it shows disagreement and at the same time, serves for taking the turn. Thus, the functions concerning the management of the turn are only reserved for those markers

that take up chiefly discourse functions due to their advanced stage in grammaticalization, such as *bé* (*well*). As mentioned, we opted for a more or less fine-grained list of functions.

**Table 4.** Specific marks of manual gesture.

Manual Gesture—Specific Marks	Values
Both hands	no/yes
Handedness	right hand/left hand
Hand shape	open/closed
Vertical	S (small)/B (big)
horizontal	S (small)/B (big)
Front–back	S (small)/B (big)

**Table 5.** The five labels are further annotations of a manual gesture.

<b>Management of the turn</b>	Turn-taking: <i>doncs</i> ( <i>so</i> ) Turn keeping (fillers): <i>eeh</i> ( <i>um, uh</i> ) Turn yielding: <i>no sé</i> ( <i>don't know</i> )
<b>Speaker's communication tasks</b>	Asking for reception or comprehension: <i>no? saps?</i> ( <i>huh? you know?</i> ) Seeking for confirmation: <i>veritat?</i> ( <i>right?</i> ) Drawing attention: <i>mira</i> ( <i>listen</i> ) Calling the listener/solidarity: <i>home, tio/a</i> ( <i>man, dude</i> )
<b>Verbal reactions</b>	Disagreement: <i>no</i> Partial disagreement: <i>bueno</i> ( <i>well</i> ) Justification: <i>és que</i> ( <i>it's just that</i> ) Contrast: <i>però</i> ( <i>but</i> ) Explanation: <i>doncs</i> ( <i>so</i> ) Showing agreement: <i>vale</i> ( <i>ok</i> ) Confirmation: <i>sí</i> ( <i>yes</i> ) Phatic: <i>sí, ja</i> ( <i>yes</i> ) Showing comprehension: <i>ah</i> ( <i>ah</i> ) Remembering something: <i>ah sí</i> ( <i>oh yes</i> ) Expressing emotions: <i>buah, ostres</i> ( <i>wow, crikey</i> )

As we took a functional approach, we included all the forms that perform an interactive function, regardless of whether they are genuine Catalan markers or are loan markers from Spanish, such as *vale* (*ok*) or *bueno* (*well*).

In relation to pragmatic marker positions, we needed a system of discourse units allowing us to specify the position occupied by the pragmatic markers under study. To do so, we based the system on the Val.Es.Co. Model (Briz and Val.Es.Co. Research Group 2003), particularly on those units necessary for our analysis, which are interventions and acts. On the one hand, interventions are the highest monological structural units generally characterized by the speaker shift, though they are not necessarily accepted by the listener/s (e.g., *ja ja*), in contrast to turns, which are accepted by the listener/s, since they cause a verbal (or non-verbal) reaction. On the other hand, acts are structural units that encompass propositional content (semantic), convey illocutive force (pragmatic), and often correspond to an intonative unit (prosodic). In Table 6, positions in relation to two discourse units used are shown.

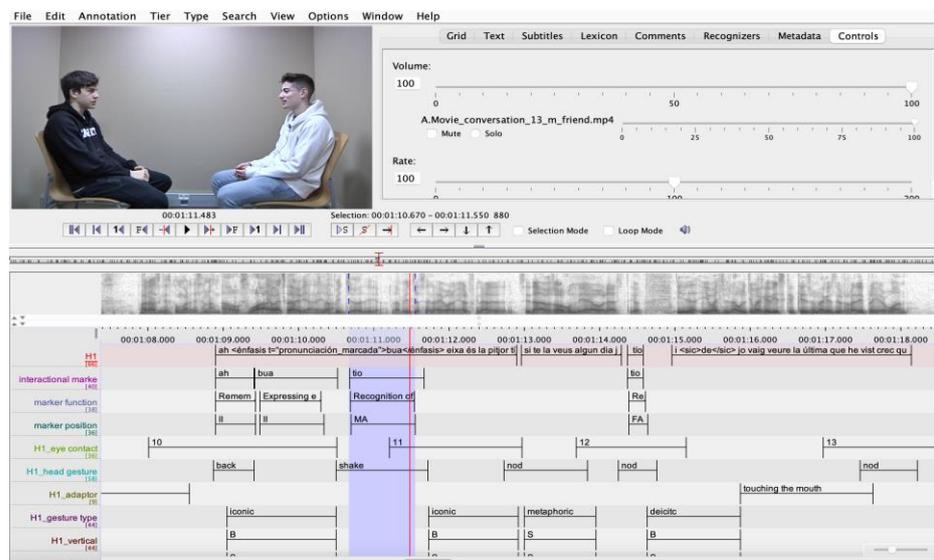
The medial position of interventions is in a white cell, since it is further detailed by means of the positions of acts. The “independent” position refers to autonomous uses of some pragmatic markers, such as *ah*, and *val/vale* (*ok*).

**Table 6.** Positions in relation to discourse units using the Val.Es.Co. Model.

Position/Discourse Unit	Intervention	Act
Initial	– <i>Bé</i> aniré més tard [Well I will go later]	– No recordo perquè no hi era / <i>Ah</i> ja sé estava malalt [I don’t remember why he wasn’t/oh now I remember he was sick]
Medial		– Aquest personatge era ehhh cómo diria [This character was um how to say]
Final	– Has vist Titanic <i>no?</i> [Have you seen Titanic, <i>right?</i> ]	– És una obra mestra <i>no?</i> / Jo l’he vista tres cops al menys. [It’s a masterpiece, <i>right?</i> I’ve seen it at least three times]
Independent	– <i>Val</i> ( <i>Ok</i> )	

### 3.3. Computational Treatment

In this particular investigation, the analysis variables involved the combination of pragmatic markers, their pragmatic function, and their position in conjunction with different multimodal annotation variables, such as facial gestures, manual gestures, head movements, eye gaze, and adaptors. A screenshot of the ELAN interface can be seen in Figure 1.



**Figure 1.** Screenshot from the ELAN interface.

Figure 1 exemplifies the annotation process by highlighting the pragmatic marker *tio* ‘dude’ within a mid-act position. This marker serves the function of acknowledging the interlocutor. Notably, this temporal occurrence of the pragmatic marker coincides with other multimodal variables, such as the absence of eye gaze and a head shake.

The ELAN framework facilitates the integration of these various annotation elements, enabling a comprehensive analysis of the video data. All the annotated information from the ELAN files was converted into a data frame in R<sup>6</sup>. Using an R library called SQLDF, the labels stored in the ELAN tiers were collapsed based on their temporal co-occurrence. Each tier was treated as a database, and SQLDF was used to combine the information from these tables that shared the same start and end times. This process resulted in obtaining co-occurrences of markers with multimodal variables and other annotation values, such as

the quantity of multimodal phenomena in a row or the speaker’s sex. Table 7 presents a brief sample of variables from the generated database.

**Table 7.** Sample displaying columns of the database.

File	Speaker	Sex Speaker	Pragmatic Marker			Multimodal Cues			
			Marker	Position	Functions	Eye Gaze	Head Gesture	Manual Gesture	Adaptor
1	H2	male	vale	initial inter.	agreement	at interl.	nod	no	no
1	H1	male	no	initial interv.	disagreement	averted	no	beat	touching
1	H2	male	tio-tia	final interv.	solidarity	at interl.	no	beat	no

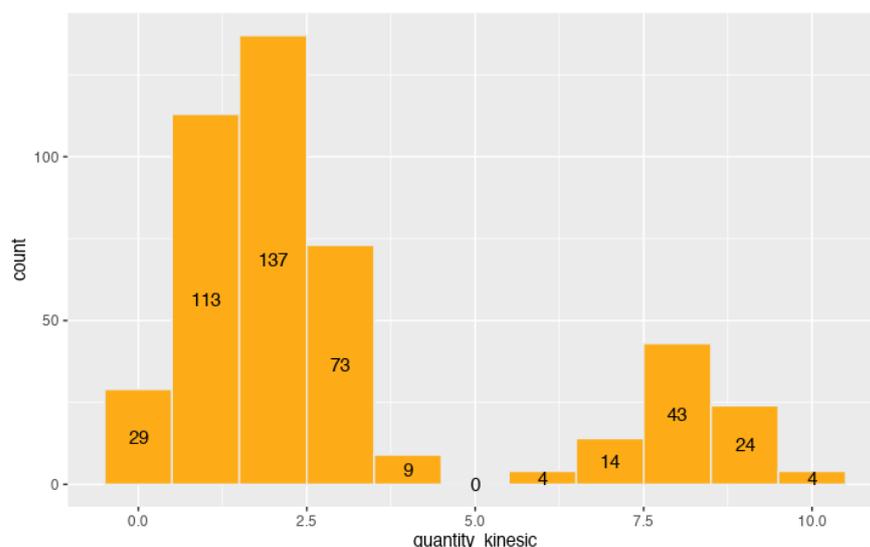
Next, we utilized boxplots and histograms, and conducted an analysis of variance (ANOVA) statistical test to gain insights into the relationships, distributions, and differences among multiple groups or variables (Levshina 2015).

#### 4. Results

In this section, we approach the research questions by presenting the data resulting from the statistical analysis<sup>7</sup>.

##### 4.1. Distribution of Multimodal Values Co-Occurring with Pragmatic Markers

Figure 2 shows the quantity of multimodal phenomena that co-occurred along with interactive pragmatic markers.



**Figure 2.** Histogram showing the quantity of multimodal elements co-occurring with interactive pragmatic markers.

The information that can be derived from the histogram is that pragmatic markers consistently exhibited a certain polarity in the quantity of multimodal elements utilized. They were frequently accompanied by multimodal cues, as evidenced by 421 out of 450 records. Additionally, there was a noticeable distinction between a small number of multimodal elements (ranging from 2 to 3 variables) and a larger number of co-occurring values (ranging from 8 to 9).

Pragmatic markers with no multimodal manifestation exhibited diversity, including cases such as *i* (5 occurrences), *sí* (4 occurrences), *bueno* (3 occurrences), and *vale* (3 occurrences), among others, totaling 29 out of 450 instances. These markers were typically very short and often appeared at the beginning of acts or interventions.

When there was a single multimodal cue accompanying a pragmatic marker, the most common expression was gaze at interlocutor, accounting for 71 out of 113 occurrences. Following that, head gesture—especially the value “nod”—was the next most common multimodal mark, appearing in 29 out of 113 instances. Similarly, when 2 multimodal values co-occurred with a pragmatic marker, the combination of gaze (mostly gaze at interlocutor) and head gesture was frequently observed, with 72 out of 137 occurrences; in addition, the combination of gaze and adaptor was another relatively common pairing, accounting for 26 out of 137 instances.

When analyzing instances in which there was a higher number of multimodal values (ranging from 6 to 10 resources), certain combinations of variables commonly appeared together. For instance, taking the highest frequency combination (8 co-occurring multimodal cues with 43 pragmatic markers on the histogram), the pragmatic markers that frequently co-occurred included *bueno* (6 occurrences), *saps?* (4 occurrences), *buah* (3 occurrences), *ja* (3 occurrences), *vale* (3 occurrences), and *no?* (3 occurrences). Among the multimodal variables, the combinations that co-occurred most frequently with pragmatic markers included gaze normally at interlocutor, head\_gesture, and manual\_gesture, as well as their specific marks—horizontal, vertical, front-back, right/left hand, both\_hands—appearing in 12 out of the 43 instances. These findings highlight certain patterns and relationships between pragmatic markers and their co-occurring multimodal variables, providing insights into the multimodal dynamics of the interactions being analyzed.

Beyond the data from the histogram, regarding the frequency of the multimodal cues, the most remarkable cues were related to eye gaze and head movements; namely, gaze at interlocutor accompanied pragmatic markers in 215 instances and nods were given with pragmatic markers in 179 instances. Adaptors were also noteworthy, so different kinds of self-manipulators (touching hands, face, nose, legs, etc.) were performed together with pragmatic markers in 109 instances. Concerning manual gestures, beat exhibited a lower quantity, with 94 instances of co-occurrence.

#### 4.2. Multiple Correspondence Analysis: Patterns of Co-Occurrence

Multiple correspondence analysis (MCA) is an exploratory statistical technique used to observe correlations among categories of several nominal variables. It is an extension of correspondence analysis (CA), which is used for analyzing contingency tables with two categorical variables. MCA allows for the analysis of contingency tables with more than two variables (Lê et al. 2008). MCA aims to uncover patterns and associations among categorical variables by transforming the data into a low-dimensional graphical representation. It helps in visualizing the relationships between categories and identifying similarities or dissimilarities among groups.

Figure 3 displays the plotting of the variables from the first two dimensions. Dimensions in MCA provide a way to represent and interpret the relationships and associations between categorical variables, allowing for a comprehensive understanding of the data structure.

We detected two great bunches of points distributed equally on the left and right of the chart. These dimensions explain about 30% of the inertia of the dataset. This is a mean value but, considering the number of variables and categories included (13<sup>8</sup> variables and 54 categories), the percentage of variation explained must be low.

MCA can be completed, for the definition of groups, with a hierarchical cluster analysis. Cluster analysis utilizes the output derived from dimensions in MCA. This is a commonly employed procedure explained thoroughly in the FactoMiner package in R. The aim is to group the records into meaningful conjuncts: each one of these groups could be explained by a combination of records (interactive pragmatic markers in this study) that frequently group together in the dataset. Figure 4 depicts the cluster analysis.

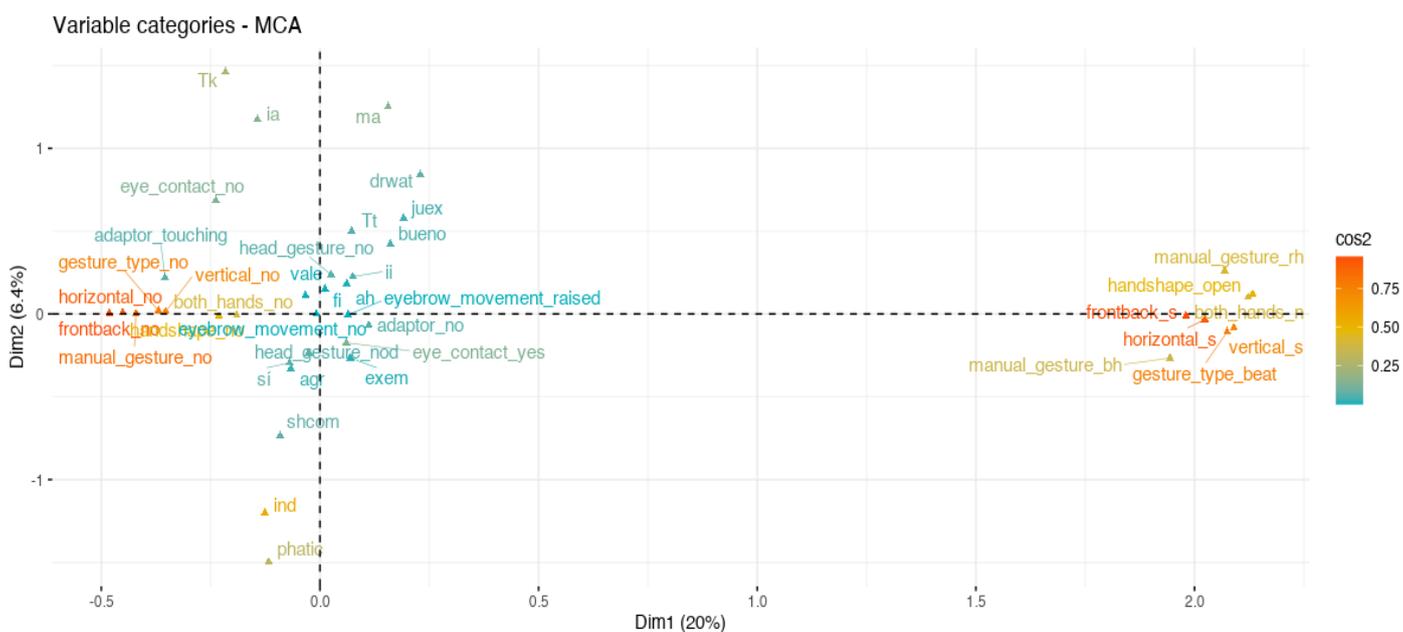


Figure 3. Plotting<sup>9</sup> of the variable set.

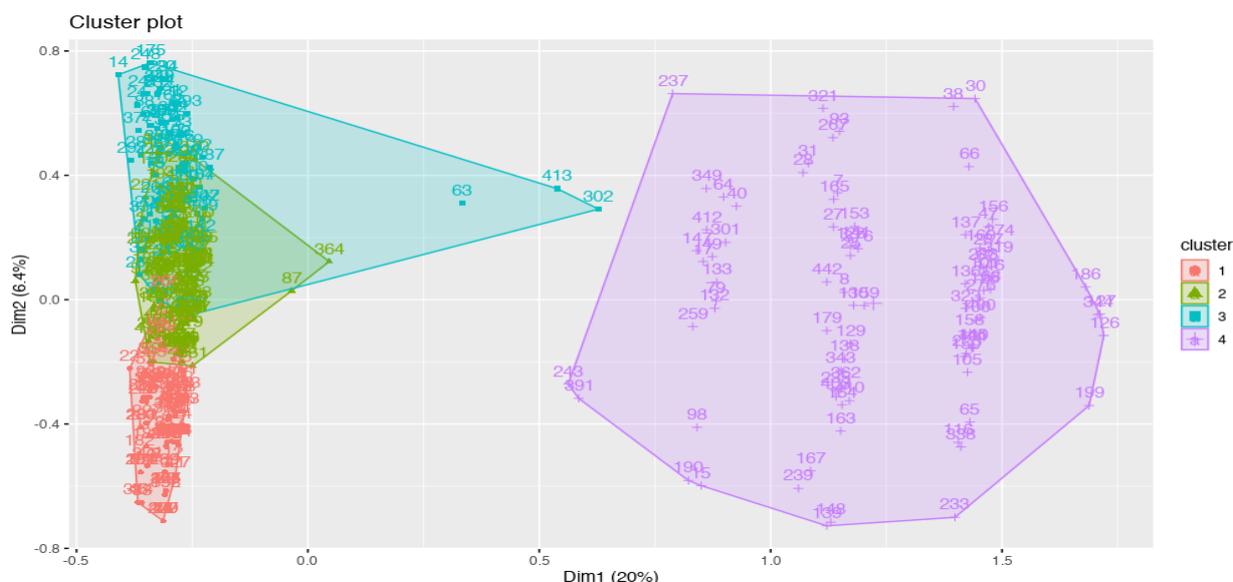


Figure 4. Cluster analysis.

Cluster 1 included practically half of the dataset records, with 185 pragmatic markers, whereas cluster 2 had 93 records, cluster 3 had 85, and finally, cluster 4 had 87 records. The explanation of each cluster is outlined below.

**Cluster 1** was formed by the markers *sí*, *ah*, *vale*, and *ja*, the position of independence or initial intervention, and the pragmatic functions of agreement, showing comprehension, expressing emotions, phatic, and confirmation. It chiefly included two multimodal values: gaze at interlocutor and nodding, as a head gesture. Figure 5 shows an example of both interlocutors uttering a *vale* ‘ok/so’ plus a co-occurring head nod, which was slightly produced to the side in both cases. They signal agreement here and produce it in unison, verbally, and gesturally.

Context: H1: la gent sempre barallant, H2: sí; H1 and H2: *vale*.  
 [H1: people always arguing; H2: yes; H1 and H2: ok].



**Figure 5.** An example of cluster 1 with co-occurrence between pragmatic markers, showing acceptance and certain multimodal cues.

**Cluster 2** was formed by the pragmatic markers *però*, *pues*, *bueno*, *doncs*, *no*, and *és que*. The position was at the beginning of the intervention. The pragmatic functions included turn-taking, contrast, disagreement, and explanation. The main multimodal value was touching some part of the body (hand, nose, head, etc.) as an adaptor, as well as averted eye gaze. In this example (Figure 6), the interlocutor on the left started the conversation by saying *doncs* ‘so’, after a pre-start with *vale*, and at the same time produced an adaptor. She touched her head and brushed her hair behind her ear.

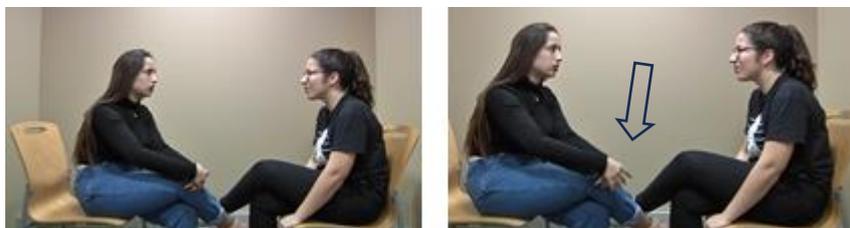
Context: beginning of the conversation, of the turn: H1: *vale* (1'') *doncs* [H1: ok so].



**Figure 6.** An example of cluster 2 with co-occurrence between a pragmatic marker introducing an explanation and certain multimodal cues.

**Cluster 3** was formed by the markers *no?*, *i*, *no sé*, *sí?*, *saps?*, *tio-tia*, and *vale?* All the other positions were related (medial act, ending act, and ending intervention). The pragmatic functions were turn keeping, drawing attention, seeking confirmation, and solidarity. The multimodal values tended to be gaze at interlocutor, head gesture turn, and overall, manual gestures, particularly small front-back, small horizontal, and small vertical in the right or both hands. In Figure 7, the interlocutor on the left produced a short beat gesture concomitantly while uttering *saps?* ‘you know?’ at the end of the intervention

Context: les fotos que <ininteligible/> com es veu com eee com les fan a la pel·lícula surten les que van ser de veritat *saps?* [pictures that (...) it is seen how um how they are taken in the movie/the ones that were real appear you know?]

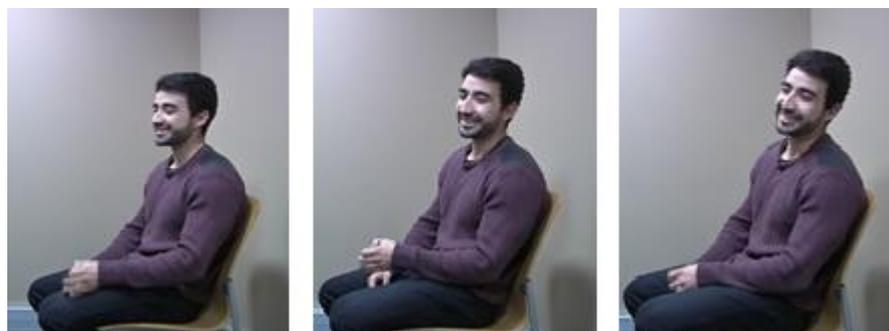


**Figure 7.** An example of cluster 3 with co-occurrence between a pragmatic marker seeking for confirmation and certain multimodal cues.

**Cluster 4** was really an open cluster that indeed covered a huge number of multimodal possibilities but not related to a position nor special pragmatic markers, except *buah*. Multimodal values included, especially, gaze at interlocutor, beats, and diverse manual gestures, quite often produced as big gestures, e.g., vertically or horizontally, right or both hands, and open hand shape. This final example (see Figure 8) shows the speaker producing a beat gesture together with a head tilt/turn and averted eye gaze from the interlocutor, while uttering *buah*. He was signaling his disbelief that his friend had not seen a certain movie.

Context: H1: no le no le vist però em sona alguna cosa. H2: no l'has vist? *Buaaah*.

[H1: I haven't seen it but something sounds familiar to me; H2: haven't you seen it? Crikey!]



**Figure 8.** An example of cluster 2 with co-occurrence between a pragmatic marker showing emotions and certain multimodal cues.

One limitation of MCA and cluster analysis is that they attempt to create groups or clusters without relying on predefined independent variables. This can pose a challenge when interpreting the results, as the clusters may not align with specific variables of interest.

Let us consider the pragmatic markers in the dataset. By exploring the previous chart, we may observe that clusters 1, 2, and 3 predominantly consisted of instances with certain pragmatic markers. However, when we view cluster 4, we realize that it is a more open and diverse cluster. It encompasses a bigger range of multimodal possibilities and various multimodal combinations, but it is not primarily associated with special pragmatic markers. In the next section, we will deal with how to interpret these results.

## 5. Discussion

### 5.1. Back to Research Questions: A Brief General View of the Results

The first research question addressed the frequency of co-occurrence between interactive pragmatic markers and multimodal cues. In the corpus under study, pragmatic markers were very often accompanied by multimodal cues, including gaze at interlocutor,

head gestures (especially nods), and sometimes adaptors and beats, among others. When these markers were accompanied by a larger number of multimodal cues, manual gestures with different specific marks (horizontal, vertical, front-back, right/left hand, and both hands) were observed. In addition to this high frequency, another relevant fact is that these multimodal cues were fully integrated in the interaction and in harmony with the interactive pragmatic markers. This integration revealed a high level of communicative dynamism in conversations between friends, where speakers behave more naturally, and among young people, who tend to gesture more energetically. Different results might have been obtained if the corpus had consisted of conversations between students and superiors (Hübscher et al. 2017).

The second research question explored potential patterns in the co-occurrence of multimodal cues and interactive pragmatic markers. Table 8 provides a summary<sup>10</sup> of significant combinations seen in the previous section.

**Table 8.** Combinations in the statistical analysis.

General Interactive Functions	Management of the Turn			Emotional Reactions	
	Turn-Taking		Turn Keeping/Yielding		
	alignment/affiliation	(slight) disaffiliation			
Multimodal Cues	gaze at interlocutor and nodding as a head gesture	adaptor, averted eye gaze	gaze at interlocutor, head gesture, manual gesture (small)	gaze at interlocutor, beat, manual gestures (generally big)	
Pragmatic markers	Forms	<i>sí, ja, vale, ah</i>	<i>però, bueno, doncs, és que</i>	<i>no?, i, no sé, sí?, saps?, tio-tia, vale?</i>	<i>buah, ostres, wide variety, no marker</i>
	Position	initial intervention independent	initial intervention	medial and final intervention	---
	Specific functions	agreement, comprehension, confirmation, phatic	turn-taking, contrast, disagreement, explanation	turn keeping, drawing attention, solidarity, seeking confirmation	----
Cluster (No. of records)	Cluster 1 (185)	Cluster 2 (93)	Cluster 3 (85)	Cluster 4 (87)	

The analysis suggested that these significant groups play important roles in four dimensions: (a) turn management, which often involves taking and sometimes keeping or yielding the turn, (b) alignment with conversational activity, (c) social (dis)affiliation with the interlocutor, and finally, (d) personal expressivity, demonstrated through reactions and displays of emotions.

In general, multimodal cues combined with interactive pragmatic markers chiefly serve an interactive function, specifically in managing the control of the turn (cf. Bavelas et al. 1995). When taking the turn, they often establish social (dis)affiliation by conveying this stance toward the interlocutor’s statement. Additionally, when speakers retain or yield the floor, they serve a structuring function by organizing discourse within the turn or coordinating the turn in the final position. Both co-speech gestures and pragmatic markers, albeit in different ways, play a role in fostering reciprocity between interlocutors and providing clues to facilitate the development and interpretation of the communication

process (whether the content is understood, is (dis)approved, draws attention, deserves further discussion, etc.).

### 5.2. Management of the Turn as an Essential Point

Speakers employ verbal and multimodal resources to shift or maintain their roles as speakers and listeners in conversation. Turn-taking involves clear verbal and multimodal cues. Speakers often seek eye contact with the listener and nod overtly. This behavior was observed in the pragmatic markers serving functions such as indicating agreement or showing comprehension (cluster 1). Furthermore, they may occasionally use hand movements, adjust their posture, make beats, or use adaptors, as observed in the pragmatic markers fulfilling functions related to justification, explanation, and contrast (cluster 2).

These results demonstrated that turn-taking in conversations is a communicative task that involves tension and negotiation, which explains the evident manifestations observed in turn-taking behavior (cf. Bavelas et al. 1995; Schmitt 2005, in Mondada 2013; Dressel and Satti 2021; Cabanes 2023). Unlike interviews or debates with a moderator, where turns may be more structured, speakers in everyday interactions must independently organize their turns. They do so by identifying opportune moments to participate or by contributing to the interaction when they have something to communicate. This exchange very often involves the use of multimodal cues integrated with pragmatic markers. Recent studies emphasize that gestures play a role in turn construction and, more specifically, in the organization of turn-taking (Bohle 2014; Dressel and Satti 2021; Cabanes 2023).

In addition, multimodal manifestations also occur in the final and medial positions (roughly corresponding to cluster 3), albeit to a lesser extent than in the initial position. The final position of pragmatic markers is considered a transition-relevant place, where multimodal cues, such as gaze at the interlocutor and manual gestures, tend to occur. When a speaker maintains the floor, multimodal manifestations notably decrease. The same pragmatic markers (*no? sí? vale? tio/tia*) can be used in both positions, although with different functions, such as turn keeping and drawing attention in the medial position or seeking confirmation in the final position, and varying quantities of gestures, with significantly fewer in the medial position. Compared to turn-taking, the functions of maintaining and yielding the turn pose less tension for the speaker. Indeed, yielding the turn, unlike taking it, is an invitation from the speaker to the listener to participate in the conversation, which would be a previously accepted exchange. Maintenance involves giving signals solely to keep attention without changes in communicative roles.

In summary, the frequency of spontaneous gestures combining with interactive pragmatic markers goes hand-in-hand with specific positions that interactive pragmatic markers occupy (a structural feature), as well as with the inherent tension in the exchange of communicative roles, particularly when the listener transitions to the role of speaker (a socio-communicative feature).

### 5.3. Alignment and Social (Dis)Affiliation in a Friendly Environment

Speakers often express their stance regarding what was just said through the use of both pragmatic markers and multimodal cues, implying (dis)affiliation toward the interlocutor. To further explore these implications (clusters 1 and 2), we provide a sample of the most frequently used markers, detailing their functions, positions, and co-occurring multimodal cues.

Pragmatic markers, such as *sí*, *vale*, and *ja*, contribute to social affiliation by indicating agreement. They typically occurred at the initial position of an intervention, often accompanied by nods and gaze at interlocutor (more than 70%), and occasionally other gestures. These pragmatic markers may also serve a similar purpose in listener responses, independently (in isolation). However, this happens less frequently. In instances where co-speech gestures are minimal (see Table 9), these markers show comprehension or serve as phatics, functioning as backchannels through which listeners align with speakers and promote ongoing interaction. This dual behavior is in line with the concepts of continuers

and convergence tokens, as discussed by Knight and Adolphs (2008). It also illustrates clear differences between alignment with ongoing discourse activity and social affiliation when pragmatic markers and multimodal cues co-occur (Stivers 2008; Olza 2022). This co-occurrence, particularly with nods, is more prevalent when pragmatic markers indicate agreement, thus reinforcing and integrating gestures become the most frequently observed co-speech gestures (Lin 2017).

In addition, pragmatic markers, such as *però* and *és que* (also sometimes *bueno*), express contrast, disagreement, or justification, resulting in certain or apparent disaffiliation. These markers typically occurred at the beginning of an intervention, accompanied by a diverse range of multimodal cues, including nods (often slight nods), eye gaze, and often adaptors and beats, some of which are related to expanding on topics, such as introducing an explanation.

**Table 9.** Sample of correspondence between selected pragmatic markers and the features under study.

Marker	Function	Position	eye_gaze	head_g	Manual Gesture	Adaptor
<i>ah</i>	emotion (11)	Ind (4) II (7)	yes (7)	nod (9)	beat (3)	yes (6)
	comprehension (14)	Ind (9) II (5)	yes (10)	nod (3)	-	-
<i>ja</i>	phatic (12)	Ind (12)	yes (3)	nod (4)	-	-
	agreement (7)	II (7)	yes (7)	nod (7)	beat (3)	yes (3)
<i>sí</i>	phatic (26)	Ind (26)	yes (2)	-	-	-
	agreement (44)	II (23) Ind (21)	yes (44)	nod (31)	beat (9)	yes (17)
<i>vale</i>	comprehension (9)	Ind (9)	yes (2)	nod (2)	-	yes (4)
	agreement (16)	II (11) Ind (5)	yes (16)	nod (16)	beat (5)	yes (2)
<i>bueno</i>	turn-taking (13)	II (13)	yes (12)	nod (4)	beat (4)	yes (1)
	partial agreement (9)	II (5) Ind (4)	yes (4)	nod (1)	beat (3)	yes (6)
<i>doncs</i>	explanation (10)	II (10)	yes (2)	nod (5)	beat (3)	yes (5)
<i>és que</i>	justification (12)	II (12)	yes (2)	nod (4)	beat (1)	yes (4)
<i>però</i>	contrast (22)	II (22)	yes (7)	nod (6)	iconic (3)	yes (7)

The findings suggested that interactive pragmatic markers and co-speech gestures “share common properties” (Bressemer and Ladewig 2011), yet they represent separate communication systems (Hadar et al. 1998; Krauss and Hadar 1999). To illustrate this independence, we considered one of the most frequent gestures, head nods. In markers indicating agreement or alignment, such as *val*, *sí*, *ja*, or *ah*, nod tends to be the most prominent, often accompanied by the multimodal cue of eye gaze at the interlocutor. However, with other pragmatic markers serving various interactive functions, such as turn-taking (*bé*, *doncs*), calling attention (*no?* *eh?*), or expressing emotions (*buah*, *ostres*), nods are accompanied by other multimodal cues. These cues may include adaptors or averted eye gaze, as well as specific manual gestures. Thus, head nods emerge as a very polyfunctional gesture associated with the function of agreement or confirmation, albeit not exclusively. Indeed, head nods can occur alone or in conjunction with many pragmatic markers. This comparison, including some specifications (e.g., nod, slight nod, and side nod), should be explored in future research.

Considering the total number of functions outlined in clusters 1 and 2, we observed that the co-occurrence between pragmatic markers and multimodal cues facilitated the progression of the interaction by showing alignment and especially contributed to social affiliation, in addition to turn coordination, as discussed in the previous section.

In our view, this behavior should be interpreted in light of certain interactional features: the genre (an informal interaction), the topic under discussion, the type of recording (speakers seated opposite each other), the interpersonal relationship between speakers (classmates and friends), and the cultural context (rooted in Spanish culture, characterized by solidarity and closeness). Different results are obtained when using distinct parameters, such as asymmetrical relationships between interlocutors or a different culture (Brown et al. 2023a, 2023b).

In these informal interactions, participants do not typically encounter conflicts in decision-making or engage in discussions about controversial topics. Instead, they conversed comfortably about movies, whether both parties had watched them or not. We must consider that movies are a topic associated with leisure time. Many people enjoy discussing potentially interesting films to watch in the future or films that should be discarded. Thus, participants strive to establish a shared ground and reciprocity between them. To achieve this, they make use of various devices, including pragmatic markers and co-speech gestures. Therefore, the predominance of combinations fostering alignment and social affiliation is influenced by the aforementioned factors, particularly the close relationships and culture rooted in solidarity. Additionally, they may employ other cues apparently associated with disaffiliation. When speakers introduced contrasting comments or justifications, they did not create social distance with the interlocutor; since they were friends, they could openly express their thoughts as a signal of confidence.

In short, both verbal and multimodal systems contribute to creating a distinct social atmosphere, typically fostering a friendly environment in these interactions.

## 6. Conclusions

Pragmatic markers are widely acknowledged for their varied roles in interactions. Gestures, including spontaneous and co-speech gestures, can also carry pragmatic meaning. Considering this, we enquired whether interactive pragmatic markers tend to co-occur alongside multimodal cues, whether discernible patterns can be identified, and which positions and functions are most noteworthy in cases of co-occurrence.

In line with possible expectations, this study on Catalan interactive pragmatic markers revealed a high percentage of co-occurrence between these markers and multimodal cues. While no crystal-clear patterns emerged, significant combinations of multimodal cues appeared to be associated with specific pragmatic marker functions and positions. Both systems harmonized and worked together in general communicative tasks, namely turn management, alignment, and social (dis)affiliation (Stivers 2008; Kidwell 2013; Bohle 2014; Maricchiolo et al. 2014, among others). Regarding turn management, turn-taking exhibited the most noticeable manifestations and correlations. Speakers also employed certain combinations to strengthen social affiliation with the interlocutor and to show alignment with the conversation's progression.

To conclude, the co-occurrence between pragmatic markers and multimodal cues constructs an intricate panorama with discernible patterns, revealing a complex and harmonized framework of communication.

**Author Contributions:** Conceptualization, I.H. and A.L.C.; methodology, A.C.N., I.H. and A.L.C.; software, A.C.N.; validation, A.C.N.; formal analysis, A.L.C. (pragmatic markers) and I.H. (multimodal cues); data curation, A.L.C., A.C.N. and I.H.; writing, A.L.C., A.C.N. and I.H.; supervision, A.L.C. and I.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** The collection of the data used in this study was supported by the URPP Language and Space fellowship awarded to the second author.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Notes

- 1 The term operational refers to gestures that confirm or deny what the speaker is saying at any level (sentence/discourse/interactive). For example, the Open Hand Prone (Kendon 2004) or the Away family (Müller 2013).
- 2 This function concerns the interpretation of an utterance, specifically related to epistemic, evidential (quotations), and other modal aspects (literal/ironic/fictional interpretation). A common example is using quotation mark gestures in reported speeches.
- 3 The performative function is chiefly centered on manifesting directive (questions, requests, and orders) and commissive (offers) speech acts, e.g., “finger bunch” (Kendon 2004).
- 4 This function involves marking discourse structure by highlighting the topical focus, contrasted/commented/digressive information, or indicating thematic shifts in discourse. The typical example is the beat gestures (see Loberes and Payrató 2012).
- 5 Regarding specific functions, although the primary function is easily identified in most gestures, many of them are indeed polyfunctional, since a slight modification of the gesture can give rise to a distinct pragmatic function. This is the case of the Brushing Aside gesture (Teßendorf 2014) and also the Palm Up Open Hand gestures (López-Ozieblo 2020), which are common gestures in Iberian Spanish conversations.
- 6 The script developed to process the data of this study is available at: <https://github.com/acabedo/multimodality>, accessed on 22 June 2022.
- 7 In this exploratory study, we included the specific marks of manual gestures as regular variables. This procedure is redundant, as they provide additional information of a general variable (e.g., both hands, right or left hand; the size of the gestures (big vs. small) on the horizontal, vertical, front–back plane, etc.). These cues indicate additional information as to how a gesture is being performed. In further studies, we will focus on specific interactive pragmatic markers and will organize the set of variables in a hierarchy.
- 8 The group of variables was made up of 10 multimodal signals and 3 variables of pragmatic markers (forms, functions, and positions), as indicated in Tables 3–6.
- 9 Abbreviations: Tk (turn keeping), Tt (turn-taking), ind. (independent turn), shcom (showing comprehension), juex (justification-explanation), fi (ending turn), ia (initial of act), ii (initial of turn), exem (expressing emotions), ma (medial act).
- 10 In Table 8, we did not include the specific marks of manual gestures that had also been analysed and processed (see Table 4).

## References

- Abuczki, Agnes. 2014. On the disambiguation of multifunctional discourse markers in multimodal interaction. *Journal Multimodal User Interfaces* 8: 121–34. [CrossRef]
- Aijmer, Karin, and Anne-Marie Simon-Vandenberg. 2006. *Pragmatic Markers in Contrast*. Amsterdam: Elsevier.
- Bavelas, Janet, Nicole Chovil, Douglas Lowrie, and Allan Wade. 1992. Interactive gestures. *Discourse Process* 15: 469–89. [CrossRef]
- Bavelas, Janet, Nicole Chovil, Linda Coates, and Lori Roe. 1995. Gestures specialized for dialogue. *Personality and Social Psychology Bulletin* 21: 394–405. [CrossRef]
- Bazzanella, Carla. 2006. Discourse markers in Italian: Towards a “compositional” meaning. In *Approaches to Discourse Particles*. Edited by Kerstin Fischer. Amsterdam: Elsevier, pp. 449–64.
- Bohle, Ulrike. 2014. Gesture and conversational units. In *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction*. Edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill and Sedinha Tessendorf. Berlin and Boston: Mouton de Gruyter, vol. 2, pp. 1360–67.
- Bressemer, Jana, and Silva Ladewig. 2011. Rethinking gesture phases. *Semiotica* 184: 53–91.
- Briz, Antonio. 1998. El español coloquial en la conversación. In *Esbozo de Pragmagramática*. Barcelona: Ariel.
- Briz, Antonio, and Salvador Pons. 2010. Unidades, marcadores discursivos y posición. In *Los Estudios Sobre Marcadores del Discurso en Español, Hoy*. Edited by Óscar Loureda and Esperanza Acín. Madrid: Arco Libros, pp. 327–58.
- Briz, Antonio, and Val.Es.Co. Research Group. 2003. Un sistema de unidades para el estudio del lenguaje coloquial. *Oralia* 6: 7–61. [CrossRef]
- Brown, Lucien, Iris Hübscher, Kim Hyunji, and Bodo Winter. 2023a. Indexing social distance through bodily visual practices in two languages. In *Multimodal Im/politeness*. Edited by Andreas Jucker, Iris Hübscher and Lucien Brown. Amsterdam and Philadelphia: John Benjamins, pp. 131–61.
- Brown, Lucien, Kim Hyunji, Iris Hübscher, and Bodo Winter. 2023b. Gestures are modulated by social context: A study of multimodal politeness across cultures. *Gesture* 21: 167–200. [CrossRef]
- Cabanes, Sandra. 2023. Análisis Multimodal en la Distinción Entre Intervención y Turno: Efectos en la Segmentación de la Conversación Desde el Modelo Val.Es.Co. Doctoral thesis, University of Valencia, Valencia, Spain.
- Cuenca, Maria Josep. 2006. *La connexió i els connectors: Perspectiva oracional i textual*. Eumo: Universitat de Vic.
- Cuenca, Maria Josep. 2013. The fuzzy boundaries between discourse marking and modal marking. In *Discourse Markers and Modal Particles: Categorization and Description*. Edited by Liesbeth Degand, Bert Cornillie and Paola Pietrandrea. Amsterdam: John Benjamins, pp. 191–216.

- Debras, Camille. 2021. Multimodal profiles of *je (ne) sais pas* in spoken French. *Journal of Pragmatics* 182: 42–62. [CrossRef]
- Degand, Liesbeth, Bert Cornillie, and Paola Pietrandrea, eds. 2013. *Discourse Markers and Modal Particles: Categorization and Description*. Amsterdam: John Benjamins.
- Dressel, Daniel, and Ignacio Satti. 2021. Embodied Coparticipation Practices in Collaborative Storytelling. *Gesprächsforschung* 22. Available online: <http://www.gespraechsforschung-online.de/fileadmin/dateien/heft2021/ga-dressel.pdf> (accessed on 30 October 2023).
- Esfandiari Baiat, Matt Coler, Marissa Pullen, Sinenaat Tienkouw, and László Hunyadi. 2013. Multimodal analysis of “well” as a discourse marker in conversation: A pilot study. Paper presented at 2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom), Budapest, Hungary, December 2–5; pp. 283–88. [CrossRef]
- Ferre, Gaëlle. 2011. Multimodal Analysis of Discourse Markers ‘donc’, ‘alors’ and ‘en fait’ in Conversational French. *ICPhS VXII*: 671–74.
- Fischer, Kerstin, ed. 2006. *Approaches to Discourse Particles*. Amsterdam: Elsevier.
- Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics* 6: 167–90. [CrossRef]
- Fraser, Bruce. 2009. Topic Orientation Markers. *Journal of Pragmatics* 41: 892–98. [CrossRef]
- Hadar, Uri, Dafna Wenkert-Olenik, Robert Krauss, and Nachum Soroker. 1998. Gesture and the processing of speech: Neuropsychological evidence. *Brain Lang* 62: 107–26. [CrossRef] [PubMed]
- Hata, Kazuki. 2016. On the Importance of the Multimodal Approach to Discourse Markers. A pragmatic view. *International Review of Pragmatics* 8: 36–54. [CrossRef]
- Hübscher, Iris, Núria Esteve-Gibert, Alfonso Iguualada, and Pilar Prieto. 2017. Intonation and gesture as bootstrapping devices in speaker uncertainty. *First Language* 37: 24–41. [CrossRef]
- Jefferson, Gail. 1972. Side sequences. In *Studies in Social Interaction*. Edited by David Sudnow. New York: Free Press, pp. 294–338.
- Kendon, Adam. 1995. Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics* 23: 247–79. [CrossRef]
- Kendon, Adam. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kendon, Adam. 2017. Pragmatic functions of gestures: Some observations on the history of their study and their nature. *Gesture* 16: 157–75. [CrossRef]
- Kidwell, Mardi. 2013. Framing, grounding, and coordinating conversational interaction: Posture, gaze, facial expression, and movement in space. In *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction*. Edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill and Sedinha Tessendorf. Berlin and Boston: Mouton de Gruyter, vol. 1, pp. 100–12.
- Knight, Dawn, and Svenja Adolphs. 2008. Multi-modal corpus pragmatics: The case of active listenership. In *Pragmatics and Corpus Linguistics: A Mutualistic Entente*. Edited by Jesús Romero-Trillo. Berlin and New York: De Gruyter Mouton, pp. 175–90.
- Krauss, Robert, and Uri Hadar. 1999. The role of speech-related arm/hand gestures in word retrieval. In *Gesture, Speech, and Sign*. Edited by Lynn Messing and Ruth Campbell. Oxford: Oxford University Press, pp. 93–116.
- Ladewig, Silva. 2010. Beschreiben, suchen und auffordern. Varianten einer rekurrenten Geste. *Sprache und Literatur* 41: 89–111.
- Levshina, Natalia. 2015. *How to Do Linguistics with R: Data Exploration and Statistical Analysis*. Amsterdam: John Benjamins Publishing Company.
- Lê, Sébastien, Julie Josse, and François Husson. 2008. FactoMineR: A Package for Multivariate Analysis. *Journal of Statistical Software* 25: 1–18. [CrossRef]
- Lin, Yen-Liang. 2017. Co-occurrence of speech and gestures: A multimodal corpus linguistic approach to intercultural interaction. *Journal of Pragmatics* 117: 155–67. [CrossRef]
- Lloberes, Marina, and Lluís Payrató. 2012. Pragmatic coherence as a multimodal feature: Illustrative cospeech gestures, events, and states. In *The Pragmatics of Catalan*. Edited by Lluís Payrató and Josep Maria Cots. Berlin and Boston: Mouton de Gruyter, pp. 215–45.
- López-Ozieblo, Renia. 2020. Proposing a revised functional classification of pragmatic gestures. *Lingua* 247: 102870. Available online: <https://www.sciencedirect.com/science/article/pii/S0024384120300784> (accessed on 10 July 2023).
- López Serena, Araceli, and Margarita Borreguero. 2010. La variación lengua hablada vs. lengua escrita. In *Los Estudios Sobre Marcadores del Discurso en Español Hoy*. Edited by Óscar Loureda and Esperanza Acín. Madrid: Arco Libros, pp. 425–95.
- Maricchiolo, Fridanna, Stefano Di Dominicis, Uberta Cancellieri Ganucci, Angiola Di Conza, Augusto Gnisci, and Marino Bonaiuto. 2014. Co-speech gestures: Structures and functions. In *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction*. Edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill and Sedinha Tessendorf. Berlin and Nueva York: Mouton de Gruyter, vols. 2 and 3, pp. 1461–73.
- Martín Zorraquino, María Antonia, and José Portolés. 1999. Los marcadores del discurso. In *Gramática Descriptiva de la Lengua Española*. Edited by Ignacio Bosque and Violeta Demonte. Madrid: Espasa-Calpe, pp. 4051–213.
- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- McNeill, David. 2013. Gesture as a window onto mind and brain, and the relationship to linguistic relativity and ontogenesis. In *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction*. Edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill and Sedinha Tessendorf. Berlin and Boston: Mouton de Gruyter, vol. 1, pp. 1461–73.

- Mondada, Lorenza. 2013. Conversation analysis: Talk and bodily resources for the organization of social interaction. In *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction*. Edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill and Sedinha Tessendorf. Berlin and Boston: Mouton de Gruyter, vol. 1, pp. 218–26.
- Müller, Cornelia. 1998. *Redebegleitende Gesten: Kulturgeschichte. Theorie. Sprachvergleich*. Berlin: Berlin Verlag.
- Müller, Cornelia. 2004. Forms and uses of the Palm Up Open Hand. A case of a gesture family? In *Semantics and Pragmatics of Everyday Gestures*. Edited by Cornelia Müller and Roland Posner. Berlin: Weidler, pp. 234–56.
- Müller, Cornelia. 2013. Gestures as a medium of expression: The linguistic potential of gestures. In *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction*. Edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill and Sedinha Tessendorf. Berlin and Boston: Mouton de Gruyter, vol. 1, pp. 202–17.
- Müller, Cornelia. 2017. How recurrent gestures mean. Conventionalized contexts-of-use and embodied motivation. *Gesture* 16: 276–303. [[CrossRef](#)]
- Olza, Inés. 2022. Patrones multimodales de (des)alineación conversacional. In *Perspectivas Integradas Para el Análisis de la Oralidad*. Edited by María Méndez Orense and Víctor Pérez Béjar. Sevilla: Universidad de Sevilla, pp. 131–56.
- Payrató, Lluís, and Sedinha Tefendorf. 2014. Pragmatic gestures. In *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction*. Edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill and Sedinha Tessendorf. Berlin and Boston: Mouton de Gruyter, vol. 2, pp. 1531–39.
- Pons, Salvador. 2006. A functional approach to the study of discourse markers. In *Approaches to Discourse Particles*. Edited by Kerstin Fischer. Amsterdam: Elsevier, pp. 77–99.
- Portolés, José. 2001. *Los Marcadores del Discurso*. Barcelona: Ariel. First published 1998.
- Schegloff, Emanuel. 1982. Discourse as an interactional achievement: Some uses of ‘uh huh’ and other things that come between sentences. In *Analyzing Discourse: Text and Talk*. Edited by Deborah Tannen. Washington, DC: Georgetown University Press, pp. 71–93.
- Schiffrin, Deborah. 1997. *Discourse Markers*. Cambridge: Cambridge University Press. First published 1987.
- Stivers, Tanya. 2008. Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction* 41: 31–57. [[CrossRef](#)]
- Streeck, Jürgen. 2009. *Gesturecraft. The Manu-facture of Meaning*. Amsterdam and Philadelphia: John Benjamins.
- Tefendorf, Sedinha. 2014. Pragmatic and metaphoric combining functional with cognitive approaches in the analysis of the “brushing aside gesture”. In *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction*. Edited by Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill and Sedinha Tessendorf. Berlin and Boston: Mouton de Gruyter, vol. 2, pp. 1540–58.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.