



HAL
open science

Exploring the status of filled pauses as pragmatic markers: the role of gaze and gesture

Loulou Kosmala

► **To cite this version:**

Loulou Kosmala. Exploring the status of filled pauses as pragmatic markers: the role of gaze and gesture. *Pragmatics and Cognition*, 2023, 29 (2), pp.272 - 296. 10.1075/pc.21020.kos . hal-04071933

HAL Id: hal-04071933

<https://hal.science/hal-04071933>

Submitted on 18 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pragmatics & Cognition

Exploring the status of filled pauses as pragmatic markers: the role of gaze and gesture --Manuscript Draft--

Manuscript Number:	PC-21020R4
Full Title:	Exploring the status of filled pauses as pragmatic markers: the role of gaze and gesture
Short Title:	
Article Type:	Special Issue Article
First Author:	Loulou Kosmala, Ph.D.
Corresponding Author:	Loulou Kosmala, Ph.D. Asnières-sur-seine, FRANCE
Funding Information:	
Section/Category:	Special Issue - "Discourse-Pragmatic Markers, Fillers and Filled Pauses"
Keywords:	filled pause; gesture; multimodality; face-to-face interactions; register variation
Manuscript Classifications:	Language and cognition; Pragmatics
Abstract:	<p>The present study aims to explore the status of filled pauses as pragmatic markers, by taking into account their accompanying visual and gestural behavior. Filled pauses (FPs), serve several pragmatic functions in speech, mainly planning but also turn-holding and emphasis, and their use is also highly determined by register and setting. This research explores the different pragmatic functions of FPs by analyzing their distribution in two different communication settings (conversation vs presentation setting), combining quantitative and qualitative approaches. Particular attention was paid to the co-occurring gestural activity of uh/ums and gaze behavior. Analyses show that the pragmatic functions of FPs are also embodied in kinetic activities which differ according to the setting: more pragmatic and referential ones were found during FPS in conversation than in the presentation setting, as well as more eye-contact, which reflects their potential communicative role during interactional sequences.</p>
Author Comments:	
Order of Authors Secondary Information:	

This is the author accepted manuscript.
The published version can be found here:

<https://doi.org/10.1075/pc.21020.kos>

Exploring the status of filled pauses as pragmatic markers: the role of gaze and gesture

Abstract:

The present study aims to explore the status of filled pauses as pragmatic markers, by taking into account their accompanying visual and gestural behavior. This aspect has not yet been widely explored, and the current study breaks new ground by demonstrating that the analysis of gaze and gesture can shed substantial light on the pragmatic functions of filled pauses and other pausing phenomena. Filled pauses (FPs), serve several pragmatic functions in speech, mainly planning but also turn-holding and emphasis, and their use is also highly determined by register and setting. This research explores the different pragmatic functions of FPs by analyzing their distribution in two different communication settings (conversation vs presentation setting), combining quantitative and qualitative approaches, following Kosmala & Crible (2019)'s study on the same data. Particular attention was paid to the co-occurring gestural activity of *uh/ums* and gaze behavior. Analyses show that the pragmatic functions of FPs are also embodied in kinetic activities which differ according to the setting: more pragmatic and referential ones were found during FPs in conversation than in the presentation setting, as well as more eye-contact, which reflects their potential communicative role during interactional sequences.

Keywords: filled pause, gesture, multimodality, face-to-face interactions, register variation

1. Introduction

Filled pauses, such as *uh* and *um* in English, or *euh* and *eum* in French, have been analyzed extensively in the past fifty years in a variety of research fields such as psycholinguistics (e.g., Clark 1996), phonetics (e.g. Shriberg 1994), conversation analysis (Schegloff, 2010) or second language teaching (e.g. Rose 2008), and have received particular attention in disfluency research (Bortfeld, Leon Bloom Schober & Brennan 2001; Schnadt & Corley 2006; Smith & Clark 1993; Shriberg 1994). Filled pauses, among other disfluency markers, have generally been viewed as a window into our cognitive systems (Shriberg, 1994), leaving

overt traces of speech planning or processing (Hieke 1981). As speakers spontaneously plan their utterances, they are often prompted to interrupt themselves temporarily to monitor their own production, which often requires additional time. The large number of filled pauses in spontaneous speech is thus often viewed as the result of processing load (Holmes, 1988), as filled pauses tend to occur more often before new information (Arnold et al. 2004), between major boundaries (Rose, 1998; Swerts, 1998) or when speakers experience uncertainty (Brennan & Williams 1995).

More recently however, a number of authors have challenged the view of filled pauses as markers of “disfluency”, and have opted for more neutral or positive terms, such as “communication management” (Allwood, Nivre & Ahlsen, 2005), “collateral signals” (Clark & Fox Tree 2002) or “planner” (Jucker 2015, Tottie 2014). This body of research tends to be less production-oriented and aims to target pragmatic aspects of filled pauses’ usage. For instance, Tottie (2011, 2014, 2015, 2016, 2019) conducted several corpus-based studies dedicated to the pragmatic uses of filled pauses in British and American English. She further put forward the idea that “uh” and “um” should be considered as a class of pragmatic markers, strongly determined by setting and register. Her studies revealed a number of effects in age, gender, socio-economic class, context, and register (Tottie, 2011, 2014). In line with Clark & Fox Tree’s (2002) hypothesis that filled pauses signal a delay in speech to keep or cede the floor, or to attract attention (e.g. Kjellmer 2003), Tottie (2011) stressed the fact that “uh” and “um” functioned as a planning device, and thus suggested the term “planner”, followed by Jucker (2015). She further argued that filled pauses could serve several overlapping functions, such as structuring upcoming discourse (also found in Swerts, 1998), but they can also be used intentionally with a stylistic purpose, notably in writing (Tottie, 2019). In sum, a number of researchers have documented the several roles filled pauses play in discourse, and have put forward the idea that they should not be merely viewed as “a filler

of pauses, a sign of hesitation or disfluency” (Tottie, 2014: 21) but rather as a pragmatic stance marker. More recently, Kosmala & Crible (2021) offered a more nuanced perspective to filled pauses, following previous functionally ambivalent approaches to (dis)fluency (Crible, Dumont Grosman & Notarrigo 2019). Based on the analysis of two different corpora in French, they triangulated evidence from multiple variables (register variation, language proficiency, and degree of familiarity) to argue in favor of a dual view of filled pauses, exhibiting several functions and patterns of co-occurrence across different contexts of use. This dynamic view of filled pauses is also reflected in the work of Cienki (2012, 2015) who considers non-lexical sounds (“mm” “uh” among others, such as filled pauses) as fluid categories with different degrees of lexicalization and conventionalization. Despite being nonlinguistic per se, filled pauses have the potential to gain a symbolic relation to certain degrees of fluency or lexicality, and therefore acquire a symbolic status.

Following this body of work, the present study is grounded in a corpus-based and discourse-functional approach to filled pauses, and aims to integrate a *multimodal* approach (Mondada 2016; Stivers & Sidnell 2005; Morgenstern 2014; Kosmala 2021a), based on the analysis of visual-gestural features of language (hand gestures, gaze direction, body movement etc.). As we shall see, very little is known about the relation between filled pauses and gesture, as a majority of studies tend to focus on vocal and phonetic aspects of filled pauses (e.g. Duez 2001; Shriberg 1994). While the existing literature on filled pauses is quite large and mainly covers prosodic, phonetic or functional aspects of discourse (which is not the focus of this paper), the present analysis sheds light on a far less documented aspect of filled pauses, which is gesture. Similarly, studies in gesture analysis and multimodality do not tend to analyze filled pauses in detail or explore their co-occurrence with manual gestures (except for a few, see section 2). This paper thus aims to bridge this gap by exploring how the

same pragmatic functions served by filled pauses in the speech channel may further be manifested in manual actions and gaze behavior.

2. Filled pauses, disfluency, and multimodality¹

The term *multimodality* or *multimodal communication* is an overarching term mainly used in the field of interaction studies and gestures studies to refer to the plurality of communication channels and modalities deployed in interaction, such as facial expressions, hand gestures, body postures, grammatical structures, prosody, and the like (Stivers & Sidnell, 2005). While the field of gesture studies is gaining more and more prominence in current linguistic research, the relationship between visible bodily behavior and filled pauses remains quite underexplored, since the latter tend not to occur extensively during gestures. For instance, Christenfeld, Schatcher & Bilous (1991) carried out two experiments in which they examined the co-occurrence of filled pauses with gestures. Overall, they found that while speakers produced fewer filled pauses when they were gesturing than when they were not, they also found a tendency for filled pauses *not* to occur during gestures, which made it difficult to draw definite conclusions regarding their relationship. More recently, a number of studies have examined the temporal relationship between speech and gesture, by looking at instances of speech suspension synchronized with gesture suspension. For instance, Seyfeddinipur (2006) investigated the coordination of disfluencies and gestures in a corpus of German semi-spontaneous speech. She focused more specifically on the production of overt and covert repairs in relation to the different phases of gestures (Seyfeddinipur, 2006: 107-09, based on Kendon, 2004), listed below:

1. Preparation: a movement of the hands to a location where the stroke is deployed.
2. Hold: When hands are in a static position, other than the rest position

¹ The present paper, as stated in the introduction, is based on Kosmala & Crible's (2021) study on the exact same data. Their paper includes a more complete literature review on the different functions and positions of filled pauses, which are not included here to avoid repetition and respect the Journal's word limit. The aim of the present paper is to shed light on gesture, a far less documented aspect of filled pauses.

3. Stroke: a phase which displays the core meaning of the gesture.
4. Retraction: when hands move back into rest position (on the lap, arm rests, or arm folded in front of the chest).

Her results showed that several gestures tended to be suspended prior to the production of disfluencies: out of 432 speech suspensions, 306 were accompanied by gestures.

Seyfeddinipur further illustrated cases of gestural suspension (i.e. hands dropping back into rest position, or held in a static position) temporally coordinated with a vocal speech suspension (i.e. a disfluency). Similar results regarding the relationship between gestural and speech suspension more specifically were also reported in other studies. Esposito & Marinaro (2007) conducted a study on pauses and holds among adult and child speakers during an elicitation experiment, and showed a high number of overlaps between holds and pauses in both groups. This was also found in Yasinnik, Shattuck-Hufnagel, & Veilleux (2005) who observed a higher number of gestures which were temporarily held during disfluent speech in recordings of academic lectures. In a similar vein, Graziano & Gullberg (2013, 2018) conducted two studies on the temporal coordination of disfluencies and gesture, based on a corpus of retellings from different groups of speakers (competent L1 speakers, adult and child L2 learners). They looked at the distribution of disfluencies in relation to the different gesture phases (preparation, hold, stroke etc.) but they also annotated the functions of gestures, i.e. *referential* and *pragmatic*, following Kendon (2004). Referential gestures refer to gestures which have referential content and are used to depict a person, object, or a concept, while pragmatic gestures deal with the interactive task at hand and are used to mark emphasis, yield a turn, or comment on the speakers' utterances. Graziano and Gullberg's findings yielded similar results as the ones reported in previous studies, mainly that gestures occurred significantly more during fluent stretches of speech, than disfluent ones, and that certain gestures tended to be held during disfluent speech. However, their results also indicated that

speakers (of all groups) produced a majority of pragmatic gestures during disfluencies, as a comment on the communicative breakdown. More recently, Kosmala (2021b) conducted a similar corpus study on the distribution of gestures during fluent and disfluent cycles of speech in native and non-native productions of French and English, and corroborated Graziano & Gullberg's findings. Kosmala's results also showed that speakers tended not to gesture when they produced disfluencies, but when they did, the disfluencies were accompanied by pragmatic gestures.

So far, the literature suggests a possible systemic relation between disfluency and gesture, which calls for further research in the field of pragmatics, interactional linguistics, and gesture studies. However, as pointed out earlier, while some studies have looked into the relationship between gesture and speech suspension (e.g. Beattie & Aboudan 1994; Chawla & Krauss 1994), very few of them have targeted the relationship between gesture and *filled pauses* specifically, except for Jehoul, Brône & Feyaerts (2016), who have investigated the relationship between filled pauses and gaze, which has received even less attention in the literature. They focused on several patterns of gaze behavior, such as mutual gaze and gaze aversion (i.e., when a speaker averts one's gaze, either looking back or looking away) in relation to filled pauses. Overall, they found a correlation between gaze patterns and filled pause use, based on their form ("uh" versus "um"). In particular, um-type filled pauses were found to be more often accompanied by gaze withdrawals than uh-type filled pauses. This nasal variant, the authors suggested, is more often associated with planning, reflecting a cognitive thinking process. This observation was also supported in Kosmala & Morgenstern's (2019) study on filled pauses in a question-answering task, where 88% of filled pauses were found to be accompanied by gaze withdrawals when participants answered to questions. This further echoes Goodwin & Goodwin's (1986) qualitative study on word searches and co-participation, where speakers were found to withdraw their gaze as they began to search for a

word. They explained that gaze withdrawals often occurred during “perturbations in the talk” (Goodwin & Goodwin, 1986: 67), which can be characterized by filled pauses.

3. Theoretical grounding and research assumptions

The present paper offers an integrated approach to filled pauses, combining different theoretical backgrounds and methodologies. The methodology used for this study is further described in the following section. As pointed out before, the aim of this paper is to further shed light on the pragmatic functions of filled pauses across different contexts of use, in line with usage-based approaches to filled pauses and (dis)fluency (see Crible, Dumont, Grosman, & Notarrigo, 2019; Kosmala & Crible, 2021), further situated within the field of gesture studies and multimodal interaction (e.g. Cienki 2015; McNeill 1985; Müller 2014; Kendon 2004; Streeck 2009; Morgenstern 2014). In addition, the present study is grounded in a conversation-analytic framework (Mondada 2007, Sacks, Jefferson & Schegloff 1974) where filled pauses are viewed as a potential token of interactional achievement, further shaped by the sequential development of the talk.

Therefore, it will be argued throughout this paper that filled pauses should not solely be considered as results of internal cognitive demands, but as dynamic and fluid categories with different degrees of conventionalization and lexicality, the interpretation of which is further shaped by a number of situational and contextual factors (Tottie, 2011, 2014; Cienki, 2012, 2015), as well as their co-occurring visual-gestural behavior (Jehoul, Brône & Feyaerts 2016; Graziano & Gullberg, 2018; Seyfeddinipur, 2006, Kosmala 2021c). This leads us to the following research questions:

- RQ1: How may gesture further our understanding of filled pauses and their pragmatic status in multimodal interaction?

- RQ2: Are filled pauses accompanied by different types of hand gestures or gaze behavior depending on the communicative situation?
- RQ3: At a qualitative level, can we find recurrent multimodal patterns associated with filled pauses' functions?

4. Data and Methodology

4.1. Corpus under study

The present study is based on Kosmala & Crible's (2021) previous work on the same data which is taken from the *DisReg Corpus* (Kosmala 2020). This corpus comprises video recordings of 12 French students from Sorbonne Nouvelle University (aged 18-21) in two different communication settings. They were first recorded during individual oral presentations, which were carried out in front of their teacher and classmates in a classroom in real time. The presentations were prepared at home, and usually consisted in analyzing an excerpt from a novel, play, or poem, as required by the instructor. The same students were then recorded in pairs in a more casual conversational setting, based on a semi-elicitation task. They were asked to spontaneously talk about a number of topics (last film seen on TV, funny episode at university etc.) but they were also free to talk about anything they wanted. In short, the present contribution aims to investigate the impact of the conversational context on the multimodal deployment of filled pauses. The same students were recorded in the same physical environment, mainly a classroom at university, but they performed entirely different tasks. The two situations differ significantly on a number of levels. The first one takes place in an institutional setting where the presenter is alone and audience co-participation is not expected; in the second one, however, the setting is more "ordinary" (Heritage 1997), casual, and familiar, so the co-participants are invited to demonstrate forms of involvement and

engagement throughout the exchange². This paper will examine the possible effect of these differences on filled pauses. The participants were assigned code labels (A1, A2, B1, B2 etc.) as well as pseudonyms to preserve their anonymity.

4.2. Methods

The present integrated theoretical framework is also reflected in my mixed-methods methodology (e.g. Peltonen 2019; Stivers 2015) which combines quantitative treatments and qualitative multimodal analyses. On the one hand, I conducted quantitative annotations of filled pauses in the dataset based on a number of variables (form, position, co-occurrence, see Kosmala & Crible, 2021 for a more detailed description³), as well as annotations of gestures based on earlier functional classification systems (e.g. Kendon 2004; Cienki 2004; Müller 1998). Gestures were annotated on the basis of their functions⁴, which comprise the following:

- a. Referential (Kendon 2004): gestures that relate to referential content, by ways of representing or pointing.
- b. Discursive (Cienki, 2004, Müller 1998): gestures used for discourse and segmentation purposes (i.e. emphasis, discourse structuring etc.)
- c. Interactive (Kendon 2004, Bavelas, Chovil Lawrie & Wade 1995): gestures used to regulate interaction (turn-taking, holding, yielding etc.), indicate a stance, or a speech act.
- d. Thinking gestures (Gullberg 2011): gestures used to indicate a deep search, oftentimes a wordsearch, used as a metapragmatic comment on the ongoing search.

² It should be noted that this conversational setting involved a certain degree of researcher control over the data (with semi-elicitation techniques) but still offers ecological validity in the sense that it includes semi-realistic real-life situations of students within a shared institutional and social environment, the university.

³ Read Kosmala & Crible (2021) for a detailed description of the method for the annotation of filled pauses. They were identified depending on their phonological variant, annotated for duration in milliseconds and four positions were distinguished.

⁴ The author is aware that annotating the functions of gestures is a highly subjective task. While inter-coder reliability has been conducted in previous work on the same data (e.g. Kosmala, 2021) this was not the case for this specific study.

Following Jehoul, Brône & Feyaerts (2016) changes in gaze direction were also annotated during filled pauses, as either “away”, “towards interlocutor”, “towards paper” and “in different directions”. Finally, the different phases of gesture were also annotated, following Seyfeddinipur (2006) and Kendon (2004), mainly “rest position” “preparation phrase” “hold” “stroke” and “retraction”. Gestural actions were only annotated during the production of filled pauses for the purpose of this study, but a more detailed account of their distribution can be found in Kosmala (2021c).

To analyze this data, several inferential statistical tests were used: log-likelihood tests to measure frequency differences across corpora, z-scores to assess the significance of differences between proportions, t-tests to compare means of numerical variables, and chi-square of independence to measure the association of two categorical variables.

These quantitative analyses were complemented with micro-analyses of emerging filled pauses in specific interactional sequences, using tools from Conversation Analysis and interactional linguistics (Mondada 2007; Goodwin 1981; Morita & Takagi 2018). Specific attention was paid to turn-taking analysis and the sequential development of the exchange, Participation Framework (Goodwin, 1981), intersubjectivity, and preference structure (Yule 1996; Sterponi & Fasulo 2010).

All the analyses were carried out with the software ELAN (Sloetjes & Wittenburg 2008) which allows for multi-level analyses of the data by aligning the production of filled pauses with other types of visible behavior in specific multimodal interactive practices. This type of method can be used for both quantitative and qualitative analyses, therefore shedding light on different, but complementary, aspects of the present investigation. The annotations conducted on ELAN were then converted to an EXCEL file to conduct statistical treatments (e.g. average, means etc.) and perform the statistical tests.

5. Quantitative results

5.1. Overview of the data

As mentioned earlier, the present study builds on Kosmala and Crible's (2021) earlier study of the distribution of filled pauses in the same data. Their main results are presented here, as a reminder. Overall, 666 filled pauses were extracted from the data, amounting to a rate of 6.8 per hundred words during presentations ($N = 385$), and 4.2 ($N = 281$) during conversations, which is a significant difference ($LL = 47.02, p < .001$). In addition, filled pauses were found to be of longer duration during presentations ($M = 415$ ms, $SD = 240$ ms) than conversations ($M = 343$ ms, $SD = 341$ ms) ($\beta = -71.97, SE = 25.98, t = -2.770, p < .01$), and more instances of the nasal variant (*eum*) were found in class (23% 88/385) than in conversation (8% 22/281); $z = 5.12, p < .001$). Finally, more instances of filled pauses in utterance-initial position were found in class (40% 156/385) than in conversation (24% 67/281; $z = 4.44, p < .001$). As the authors argued, such distributional differences ultimately reflected differences in speech genre: while a class presentation requires clear discourse boundaries to help the audience follow, a face-to-face interaction relies on more intersubjective mechanisms, which inevitably has an effect on filled pauses' uses and behavior. The aim of this paper is not to linger on their pattern of distribution, but rather to explore the interplay between filled pauses, gesture, and gaze, to include a different set of variables, which have received less attention in the literature.

5.2. Gesture and gaze behavior during filled pauses: differences between the two settings

Table 1 shows the proportion of filled pauses during gestural phases, from rest position to retraction. The results reveal no significant relation between gesture phasing and filled pauses production across the two situations ($\chi^2 (1, N = 666) = 5.6, p < .05$), as a majority of them were produced while the hands were in rest position in both conditions ($N = 233/385$ for the

class presentations and $N = 187/281$ for the conversation, which amounts to about 60-65% of the time).

Table 1. Proportion of filled pauses during gestural phases (raw and relative values)

Gest. Phase	Class	Conversation	Total	Z (p)
rest	233 (61%)	187 (66%)	420 (63%)	-0.57 (.05)
stroke	56 (14%)	45 (16%)	101 (15%)	1.10 (.2)
preparation	28 (7%)	11 (4%)	39 (5%)	1.79 (.07)
hold	48 (12%)	27 (10%)	75 (12%)	-1.36 (.1)
retraction	20 (6%)	11 (4%)	31 (5%)	0.7 (.4)
Total	385	281	666	

Only about 15% of the filled pauses were accompanied by a stroke in the two situations, which suggests little gestural activity during filled pauses, in line with previous work (see Christenfeld, Schatcher & Bilous 1991). However, it is also relevant to note that about 30% and 19% of filled pauses ($N= 96/385$ in class and $N =49/281$ in conversation) occurred during other phases of gesture, such as hold, preparation, and retraction, suggesting a certain relationship between speech suspension and gesture suspension (marked by holds and retractions) as well as speech preparation (with the role of filled pauses as *planners*) and gesture preparation (with the preparation phase). This is further exemplified in the qualitative analyses in the following section.

Table 2. Proportion of gesture types ($N=102$) during filled pauses ($N= 666$)⁵

	class	conversation	Z (p)
referential gestures	1% (5)	6% (17)	-3.14 (0.006)
discursive gestures	7% (26)	3% (9)	0.03 (0.04)
interactive gestures	1% (5)	4% (12)	-2.42 (0.01)
word-searching gestures	5% (21)	3% (7)	1.85 (0.06)

⁵ It should be noted that the present analyses do not compare nasalised filled pauses with oral ones because of the sample size ($N= 102$) which is too small to conduct statistical analyses.

Table 2 displays the proportion of completed gesture strokes during filled pauses across the two situations. While only 102 instances of filled pauses were accompanied by gestures, which represents a small proportion overall (only 102 out of 666 in total) it is still relevant to examine which specific types of gestures are most often found, in line with previous research (e.g. Graziano & Gullberg, 2018). As the Table shows, a slightly higher proportion of referential as well as interactive gestures were found in the conversational setting, but more discursive gestures were found during the class presentations. This further illustrates situational differences: while a class presentation requires structure and clarity, a conversation relies more on intersubjective processes, where interactants may wish to (co)-construct meaning with referential gestures, or deal with turn-taking. These differences are illustrated in the following examples, taken from the data:

Example 1 – face-to-face interaction

In Example 1, one of the speakers (pseudonymized as David) is talking about the TV Show *The Crown*. The filled pause is marked in bold, along with other vocal markers (a lengthening and an unfilled pause).

DAVID: c'est sur la reine Elisabeth et là c'est complètemen:**ent euh (0.705)** assez plat.

*It's on Queen Elizabeth and it's completely:**y euh (0.705)** quite flat.*



Fig. 1. Palm Open Down Gesture

As Fig. 1 shows, David executed a lateral movement away to both sides with his two open hands, palms facing down, to metaphorically and iconically represent the abstract concept of

being flat, or dull. This type of gesture is considered referential, as it is related to referential content. This type of gesture is quite different from discursive gestures, which do not relate to content, but are rather used to structure and segment speech, as illustrated in Example 2:

Example 2 – class presentation

In Example 2, the student (pseudonymized as Laura) is presenting a specific scene from a play, and she is producing a series of filled pauses, transcribed in the following utterance.

LAURA: hhh. donc euh par exemple euh on peut le voir euh dans l'Acte 1 scene 8.

hhh. so euh for example euh we can see it euh in Act 1, scene 8.



Fig. 2. Palm-Open Up Gesture

As Fig. 2 further shows, Laura is also producing a series of palm-presentation gestures (Kendon 2004: 264), where the Open Hand Supine is “presented” or “displayed” into the frontal space to introduce elements in discourse. In this example specifically, each gesture stroke coincided with the production of filled pauses.

So far, Examples 1 and 2 suggest situational differences in gestural behavior, through the lens of filled pauses, which deserves more attention. While these examples focused more specifically on the timing of gestures relative to filled pauses, more detailed contextualized analyses are provided in the next section, so as to shed more light on the contextual and interactional nature of filled pauses.

Further significant differences between the two situations were detected in gaze behavior, as illustrated in Table 3.

Table 3. Proportion of filled pauses during changes in gaze direction

	class	conversation	Total	Z (p)
away	23 (5%)	128 (46%)	151	-12.1 (<0.0002)
in different directions	18 (5%)	25 (9%)	43	-2.8 (0.02)
towards interlocutor	57 (15%)	124 (44%)	181	-8.31 (< 0.0002)
towards paper	287 (75%)	4 (1%)	291	N/A
Total	385	381	666	

Overall, a significant relationship is found between these two variables ($\chi^2 (1, N= 666) = 371.04, p < 0.00001$) with a proportion of about 46% of filled pauses that occurred during gaze aversions during conversations, as opposed to only 6% in class presentations. This is a striking result, as it is very common to withdraw one's gaze when engaged in an interactional practice, so as to display a state of disengagement, a word search, or the end of a sequence, among other actions (e.g. Goodwin & Goodwin 1986; Kendon 1967; Rossano 2013; Streeck 2014). This finding may reveal that looking away is a pattern of gazing more commonly found in face-to-face conversations than formal class presentations⁶. Once again, this sheds light on the multifunctionality of filled pauses, which can serve self-oriented or discourse-oriented functions, as well as interaction-oriented ones. As Table 3 further shows, the students' gazes were mostly directed towards their notes when they produced filled pauses during their presentations (about 75%), which suggests very little interaction with the audience. Once more, these results highlight situational differences, which are further illustrated in the following qualitative analyses, based on three extracts from the data.

6. Qualitative analyses

Following the work of Tottie (2011, 2014, 2016) the aim of the present research is to explore the pragmatic dimension of filled pauses across two different settings, by adopting a multimodal perspective. The following analyses focus on three specific functions served by

⁶ While this study only targets gazing behavior during filled pauses specifically, Kosmala (2021c) found a similar tendency in the whole data, regardless of whether they co-occurred with filled pauses or not.

filled pauses, which have been documented in the literature, mainly (1) planning and word searching (Tottie 2014, 2016 ; Jehoul, Brône & Feyaerts 2016; Finlayson & Corley 2012), (2) stalling and time-buying (Hiecke 1981; Brennan & Schober 2001) and (3) turn-taking/holding (Beňuš 2009; Kjellmer 2003). The following analyses will show how these practices are not only performed vocally, but through the means of visible actions, such as gesture and gaze. All the excerpts are transcribed with a gestural notation system, based on Kendon (2004). The transcription conventions are listed in the Appendix.

6.1. Filled pauses as planners: evidence from the two settings

As mentioned in the introduction, filled pauses have been given the label “planner” by Tottie (2014, 2016) and Jucker (2015) to highlight their planning function, based on the empirical evidence that they are very often used to plan upcoming material in discourse. The two following excerpts will reveal how the notions of planning, projection, or preparation, which are closely associated with filled pauses, may further be reflected in multiple modalities.

Excerpt 3 – Class presentation (Laura)

In Excerpt 3, Laura is giving an oral assignment in front of the class, and is introducing characters from the exposition scene in a French play.

Transcription

~ : preparation phase ; *** gesture stroke , -.- retraction , *** gesture hold

- 1 donc hhh. euh elles interagissent dès le début avec les personnages principaux
so hhh. they interact right from the beginning with the main characters
- 2 hhh. (0.400) **e:et** eum (1.313)

~~~~~\*\*\*\*\*

((moves both her hands in preparation then brings them together, palms pressed against one another))



par cette idée de personnage exposante.

*with this idea of the character as an exhibitor*

((brings both her lax flat hands forward, palm up, bent fingers + retraction, then repeats the same gesture))

3 eum déjà je voudrais établir un point.

*eum first I'd like to make a point.*

((hands return to rest position + initiates a similar gesture))

#### Excerpt 4 – Conversation between Jenny and Alex

In Excerpt 4, Jenny is about to tell a story to her conversational partner Alex about a friend of a friend who encountered a famous French actor at a bar.

#### Transcription

1 \*JEN: hhh. ma:ais euh la dernière fois j'étais au café (il) y'a une pote (laughs) qui me raconte qu'elle avait une pote (laughs) complètement bourrée qui était dans un bar.

*hhh. bu:ut last time I was at a cafe and a friend (laughs) told me that she had a friend (laughs) who was completely drunk and who was at a bar.*

2 \*JEN: et elle était en fait **euh** assise en face de Louis Garrel.

*And she was in fact **euh** sitting opposite Louis Garrel.*

~ ~ ~ \*\*\*\*\*

((moves her vertical left palm open hand in preparation + extends her arm towards ALEX))



When a manual action unfolds, it is usually marked by a preparatory movement phase where the hand(s) move from a resting position (i.e. the lap) to prepare the execution of the gesture stroke, and this is known as the *preparation phase*. In Excerpts 3 and 4, the preparation of the gesture stroke was found to coincide with the production of a filled pause (marked in bold in both transcripts). In Excerpt 3, Laura produced a nasalized filled pause in utterance-initial position, coupled with other vocal markers (two unfilled pauses and a prolongation), so as to project and index her next piece of discourse that was potentially relevant for her presentation; as she executed this action, she also moved both her hands together from her desk in preparation, and brought them together, with her palms pressed against another. This type of gesture seems to embody a posture of “getting ready” for her next stretch of discourse, as it is followed by a series of discursive gestures used for presentation and emphasis. She is also looking up, as if she was “collecting her thoughts” and transitioning from one piece of discourse to the next. Similarly, in Excerpt 4, Jenny produces an euh-type filled pause in medial position while gazing away, and slowly moving her left hand and arm forward as she describes that the friend was sitting opposite a famous French actor, Louis Garrel. Therefore, the initiation of this gesture paved the way for the upcoming spatial description of the event, but it also introduced the speaker’s viewpoint, functioning as a cohesive device which created visual coreference in space. This phase of preparation is also manifested in her gaze, which is averted.

In both cases, the notions of preparation, or planning are found in two modalities: in the vocal channel, the filled pauses introduce a new piece of discourse, and in the gestural channel they further project a new gestural action. This further reveals that the same functions served by filled pauses in discourse can also be manifested in gesture, suggesting a tight relationship between the two systems. This leads us to another excerpt, taken from the class presentations.

## 6.2. Filled pauses as time-buying strategies: attending to other relevant activities in multimodal talk

We shall now turn to the role of filled pauses as time-buying strategies, and how they may be used in discourse to buy time in order to plan upcoming material, or attend to other relevant activities in the multimodal talk. Excerpt 5 is taken from the presentation-session, during Dan's assignment, in which he is analyzing a French play and introducing the topic of disobedience.

Excerpt 5 – Dan's presentation

Transcription

1 DAN: hhh. maintenant (je vais) parler un peu plus parler de:e la désobéissance.

*hhh. now I (am going) to talk a little bit more about disobedience*

2 DAN: euh notamment chez Scapin

*euh notably in Scapin*

3 DAN: hhh. euh à l'Acte 1 scene 4

*hhhh. euh at Act 1 scene 4*

4 DAN: **euh** [!] on a:a (0.678) une scène où Scapin tient tête

(0. 492) à:à **euh** (0. 870) [!] à Arg(ante) **eum** à Argant:te ouais.

***euh** [!] we ha:ave (0.678) one scene where Scapin stands up*

(0.492) to:o **eu**h (0.870) [!] to Arg(ante) **eum** to Argant:te yeah.

((Dan is looking for a specific passage from the book))



As the transcription shows, the student is inserting a number of utterance-initial delays in his discourse (on lines 2, 3, and 4) in the form of filled pauses, co-produced with other pauses and vocalizations such as tongue clicks and inbreaths, as he presents a specific scene from the play. However, unlike in a conversational setting, he cannot rely on his interlocutor to co-construct meaning or take the floor; therefore, he relies more significantly on material objects he has at his disposal, mainly his book and his notes, as he begins to search for the right passage from the play that will be relevant to his current argument (lines 4 and 5). These filled pauses in discourse thus mark a shift, or a transition from one presentation-relevant activity (Rendle-Short 2005) to the next, from one participation framework to the next (book-oriented to audience-oriented) in order to find ways to manage his orientation towards his audience, but also his actual live performance, which requires structure and organization. Here the filled pauses serve time-buying functions as Dan is attending to other relevant activities related to the task at hand. This example shows the importance of our embodied material environment, as our voices and bodies do not act on their own, but are in constant interaction with our surrounding space, where we manipulate different objects within the environment (Morgenstern & Boutet forth.; Streeck, Lebaron & Goodwin, 2011; Goodwin 2003). In summary, filled pauses can be used to mark relevant delays in the course of the talk, related to the task at hand, also captured in contextualized and

embodied discourse. This leads us to the final excerpt, where I explore the communicative values of filled pauses in conversational settings in greater detail.

### 6.3. Filled pauses as intersubjective markers used for turn-taking

Excerpt 6 is taken from the conversation-session between the two speakers Paul and Paula, in which they are thinking about a funny anecdote they could share about university. As explained earlier, the participants in the study were given a piece of paper prior to the recordings with a list of topics written on it, but they were free to talk about anything else if they wanted. This short excerpt is taken from the story preface, before Paul initiated the telling of his narrative.

#### Excerpt 6 - Transcription

1 \*PAULA: eum une anecdote t'en as pas une à dire avant que je t'en trouve une?

*eum an anecdote don't you have one to tell me before I find you one?*

2 \*PAULA +< sur l'université

*at university*

3 \*PAUL: **eu:um:m:m** (0.400) [!] mmm:m:mm



((gazes away, shakes his head then extends his left Palm Up Open Hand towards Paula))

((1.358))

4 \*PAUL: ah si j'en ai une elle est très très précise!

*ah right I have one and it's very very specific!*

On line 3, Paul initiates a significant delay in the course of the exchange, marked by a series of vocalizations, one nasalized filled pause, an unfilled pause, a tongue click, and a nasal

sound. These vocalizations take up most of Paul's turn, followed by a lapse in the interaction. The nasalized filled pause is particularly long (about one second), and marks a significant delay in both the acoustic and interactional channel, as the progression of the exchange is momentarily disturbed; however, it also gives an opportunity for Paul and Paula to orient themselves towards a joint thinking activity.

It is also of note that while Paul momentarily delays the course of his vocal utterance with a significantly long filled pause and other vocalizations, he mobilizes a number of visual-gestural resources relevant to the task at hand as well: he is first seen looking up, his body and head not moving, as he displays a current thinking posture (Heller 2021) signaling that he is trying to remember an anecdote he could share with his partner. Then, he shifts from his solitary search activity, and invites Paula to take the floor by extending his right palm open hand towards her, while gazing in her direction. But Paula has still not found an anecdote yet, so she does not take the floor. After a lapse, Paul eventually reclaims his turn, and finds an anecdote he could share, as shown on line 4. These interactional actions took place during the production of the different vocalizations within Paul's turn, and the two speakers relied on each other's facial and bodily behavior to pursue the exchange. This further shows that filled pauses are not only used to signal uncertainty or deal with planning, but can also perform meaningful interactional actions, such as turn-taking and turn-yielding.

## Conclusion

The aim of this paper was to explore the interplay between filled pauses, gestures, and gaze. It supports Tottie's (2011, 2014, 2016) argument that filled pauses are not merely a sign of disfluency, but can be considered as pragmatic markers in discourse, despite their a priori non-lexical status. Based on an earlier corpus study of French students across two communication situations (Kosmala & Crible, 2019), this paper has further confirmed that



filled pauses are strongly shaped by situational and contextual factors, in line with previous usage-based studies. Filled pauses are multifunctional, and serve a variety of pragmatic functions in discourse, such as planning, discourse structuring, turn-taking and the like, and this can further be reflected in their accompanying visual-gestural behavior (gaze direction, body orientation, and manual gestures). Grounded in an interactional and multimodal approach to language, the present study has highlighted the importance of gaze and gesture with respect to filled pause behavior, by combining quantitative treatments and detailed micro-analyses of several excerpts. Even though relatively few gestures actually co-occur with filled pauses (only about 15% of cases), their relationship should not be overlooked, as they can tell us a great deal about the different cognitive and interactive processes underlying the use of filled pauses across settings. Gesture and gaze patterns have been shown to reveal relevant aspects of filled pauses depending on the setting (RQ2): they are more often accompanied by interactive and referential gestures during casual conversations than formal class presentations, and are predominantly accompanied by gaze withdrawals during the class presentations, revealing recurrent multimodal patterns (RQ3). Co-occurring hand gestures may also further demonstrate a particular pragmatic function served by a filled pause based on its configuration and orientation; for instance, in Excerpt 6, the Palm-Up-Open Hand gesture oriented towards the interlocutor revealed a turn-taking function, which was not overtly expressed by the speaker at the speech level (RQ1). Cases of gesture preparation also synchronized with filled pauses (Excerpts 3 and 4) underlying a planning function at the multimodal level.

However, it should be noted that the qualitative analyses presented in this paper have only illustrated a very small proportion of the data under study with only three functions served by filled pauses, which is not at all representative of the whole sample. Therefore,

more work needs to be done on the interplay between filled pauses, gaze, gesture, and discourse in a more extensive study.

For future research, we could further integrate acoustic and prosodic analyses of filled pauses and see how they relate to gaze and gesture, with perhaps specific variants (e.g. the nasal vs oral variants, largely underexplored in this paper due to corpus size) or intonational patterns associated with specific visible bodily features. More recently, some phonetic work has also been done on the perception of filled pauses regarding speakers' performance and charisma (e.g. Niebuhr & Fischer, 2019; Voss & Niebuhr, 2022) which could be supplemented with other visible factors (use of the gaze, body, and the hands). This would support arguments concerning the multimodal nature of filled pauses which rely for their interpretation on a multiplicity of modes and semiotic features. In addition, the data sample used under study is still relatively small compared to most corpus-based studies, as already pointed out in Kosmala & Crible (2021), so the present study should be replicated on a larger dataset and include more languages and speech genres.

Despite its limitations, the present study can be said to demonstrate the interest of multimodal analyses in deepening our understanding of the pragmatic function of filled pauses and their link with interaction and cognition.

## Appendix

### Transcription conventions

| CHAT conventions (MacWhinney, 2000) |                                         |
|-------------------------------------|-----------------------------------------|
| +/                                  | interruption by other participant       |
| +//                                 | self-interruption                       |
| [/]                                 | word repetition                         |
| [//]                                | self-repair                             |
| +...                                | trailing off                            |
| (0.250)                             | unfilled pause (number in milliseconds) |
| wo:rd                               | prolonged vowel or consonant            |

|                                          |                                             |
|------------------------------------------|---------------------------------------------|
| +< >                                     | overlapping talk                            |
| (a)bout                                  | shortenings                                 |
| +/ +“/.                                  | quoted utterance                            |
| xxx                                      | unintelligible words                        |
| <b>CA conventions (Jefferson, 2004)</b>  |                                             |
| [!]                                      | tongue click                                |
| .hhh                                     | inbreath                                    |
| hhh                                      | outbreath                                   |
| *creaky*                                 | creaky voice                                |
| (( ))                                    | description of events, or analyst's comment |
| <b>Gesture annotation (Kendon, 2004)</b> |                                             |
| ~ ~ ~                                    | preparation of gesture stroke               |
| ***                                      | gesture stroke                              |
| ***                                      | hold                                        |
| -.-.-                                    | return to rest position                     |

## References

- Allwood, Jens, Elisabeth Ahlsén, Johan Lund, & Johanna Sundqvist. 2005. Multimodality in own communication management. *In Proceedings from the Second Nordic Conference on Multimodal Communication*.
- Arnold, Jennifer E., Michael K. Tanenhaus, Rebecca J. Altmann, & Maria Fagnano. 2004. The old and thee, uh, new: Disfluency and reference resolution. *Psychological Science* 15(9). SAGE Publications Sage CA: Los Angeles, CA. 578–582.
- Bavelas, Janet Beavin, Nicole Chovil, Linda Coates, & Lori Roe. 1995. Gestures specialized for dialogue. *Personality and Social Psychology Bulletin* 21(4). Sage Publications Sage CA: Thousand Oaks, CA. 394–405.
- Beattie, Geoffrey, & Rima Aboudan. 1994. Gestures, pauses and speech: An experimental investigation of the effects of changing social context on their precise temporal relationships. *Semiotica* 99(3–4). Walter de Gruyter, Berlin/New York. 239–272.
- Beňuš, Štefan. 2009. Variability and Stability in Collaborative Dialogues: Turn-Taking and Filled Pauses. *In Tenth Annual Conference of the International Speech Communication Association*. Brighton, United Kingdom. 796–799.
- Bortfeld, Heather, Silvia D. Leon, Jonathan E. Bloom, Michael F. Schober, & Susan E. Brennan. 2001. Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech* 44(2): 123–147.
- Brennan, Susan E., & Michael F. Schober. 2001. How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language* 44(2). Elsevier: 274–296.

- Brennan, Susan E., & Maurice Williams. 1995. The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language* 34(3): 383–398.
- Chawla, Purnima, & Robert M. Krauss. 1994. Gesture and speech in spontaneous and rehearsed narratives. *Journal of Experimental Social Psychology* 30(6). Elsevier: 580–601.
- Christenfeld, Nicholas, Stanley Schachter, & Frances Bilous. 1991. Filled pauses and gestures: It's not coincidence. *Journal of Psycholinguistic Research* 20(1): 1–10.
- Cienki, Alan. J. 2015. Spoken language usage events. *Language and Cognition* 7. Cambridge University Press: 499–514.
- Cienki, Alan. J. 2004. Bush's and Gore's Language and Gestures in the 2000 US Presidential Debates: A Test Case for Two Models of Metaphors. *Journal of Language and Politics* 3(3). John Benjamins: 409–440.
- Cienki, Alan J. 2012. Usage events of spoken language and the symbolic units we (may) abstract from them. *Cognitive Processes in Language*. Peter Lang: 149–158.
- Clark, Herbert H., & Jean E. Fox Tree 2002. Using uh and um in spontaneous speaking. *Cognition* 84(1): 73–111.
- Clark, Herbert H. 1996. *Using Language*. Cambridge: Cambridge University Press.
- Crible, Ludivine, Amandine Dumont, Iulia Grosman & Ingrid Notarrigo. 2019. (Dis)Fluency across spoken and signed languages: Application of an interoperable annotation scheme. In Fluency and disfluency across languages and language varieties. In Liesbeth Degand, Gaëtanelle Gilquin, and Anne Catherine Simon, (eds). *Corpora and language in use- Proceedings 4*. 17-35. Louvain-la-Neuve: Presses universitaires de Louvain.
- Duez, Danielle. 2001. Signification des hésitations dans la parole spontanée. *Revue Parole*. 17–18.
- Esposito, Anna, & Maria Marinaro. 2007. What pauses can tell us about speech and gesture partnership. In Esposito, Anna, Maja Bratanić, and Eric Keller (eds). *Fundamentals of Verbal and Nonverbal Communication and the Biometric Issue*. 18: 45. IOS Press
- Finlayson, Ian R. & Martin Corley. 2012. Disfluency in dialogue: An intentional signal from the speaker? *Psychonomic Bulletin & Review* 19(5): 921–928.
- Goodwin, Charles. 1981. *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Goodwin, Charles. 2003. The body in action. In Justine Coupland and Richard Gwin (eds.) *Discourse, the body, and identity*. 19–42. London: Palgrave Macmillan
- Goodwin, Marjorie, & Charles Goodwin. 1986. Gesture and coparticipation in the activity of searching for a word. *Semiotica* 62(1–2): 51–76.

- Graziano, Maria, & Marianne Gullberg. 2013. Gesture production and speech fluency in competent speakers and language learners. *TIGER*, Tilburg University, Holanda.
- Graziano, Maria, & Marianne Gullberg. 2018. When speech Stops, Gesture Stops: Evidence from Developmental and Crosslinguistic Comparisons. *Frontiers in Psychology* 9: 879—910.
- Gullberg, Marianne. 2011. Multilingual Multimodality: Communicative Difficulties and Their Solutions in Second-Language Use. In Jürgen Streeck, Charles Goodwin, and C. LeBaron, (eds.) *Embodied Interaction: Language and Body in the Material World*. 137–151. Cambridge, UK: Cambridge University Press.
- Heller, Vivien. 2021. Embodied Displays of “Doing Thinking.” Epistemic and Interactive Functions of Thinking Displays in Children’s Argumentative Activities. *Frontiers in Psychology* 12. Frontiers: 369—400.
- Heritage, John. 2004. Conversation analysis and institutional talk. In David Silverman (eds.) *Qualitative Research: Theory, Method and Practice* (London: Sage, 1997), 161—182
- Hieke, Adolf E. 1981. A content-processing view of hesitation phenomena. *Language and Speech* 24(2): 147–160.
- Jehoul, Annelies, Geert Brône, and Kurt Feytaerts. 2016. Gaze patterns and filled pauses: Empirical data on the difference between Dutch ‘euh’ and ‘euhm.’ Paper presented at *DGKL/GCLA 7: Cognitive Approaches to Interaction and Language Attitudes*.
- Jucker, Andreas H. 2015. Pragmatics of fiction: Literary uses of uh and um. *Journal of Pragmatics* 86: 63–67.
- Kendon, Adam. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica* 26: 22–63.
- Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kjellmer, Göran. 2003. Hesitation. In defence of er and erm. *English Studies* 84(2): 170–198.
- Kosmala, Loulou. 2020. Euh le saviez-vous ? Le rôle des (dis)fluences en contexte interactionnel : Étude exploratoire et qualitative. *SHS Web of Conferences* 78. EDP Sciences. 2—15.
- Kosmala, Loulou. 2021a. On the specificities of L1 and L2 (dis)fluencies and the interactional multimodal strategies of L2 speakers in tandem interactions. Elena Babatsouli and Martin J. Ball, (eds) *Journal of Monolingual and Bilingual Speech*. 69-101
- Kosmala, Loulou. 2021b. Gestures in fluent and disfluent cycles of speech: What they may tell us about the role of (dis)fluency in L2 discourse. *In Proceedings of DiSS 2021*. Paris 8 University, France. 77- 82

Kosmala, Loulou. 2021c. *A multimodal contrastive study of (dis)fluency across languages and settings: Towards a multidimensional scale of inter-(dis)fluency*. Unpublished PhD thesis, Sorbonne Nouvelle.

Kosmala, Loulou & Ludivine Crible. 2021. The dual status of filled pauses: Evidence from genre, proficiency and co-occurrence. *Language and Speech*. SAGE Publications Sage UK: London, England: 216-239.

Kosmala, Loulou, & Aliyah Morgenstern. 2019. Should “uh” and “um” be categorized as markers of disfluency? The use of fillers in a challenging conversational context. In Liesbeth Degand, Gaëtanelle Gilquin & Anne Catherine Simon (eds.) *Fluency and disfluency across languages and language varieties. Corpora and language in use-proceedings 4*, 67–85. Louvain-la-Neuve: Presses universitaires de Louvain.

MacWhinney, Brian. 2000. The CHILDES Project: Tools for analyzing talk. transcription format and programs. *Psychology Press 1*, 657—677.

McNeill, David. 1985. So you think gestures are nonverbal? *Psychological Review 92*(3): 350—380.

Mondada, Lorenza. 2007. Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies 9*(2): 194–225.

Mondada, Lorenza, 2016. Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics 20*(3). 336–366.

Morgenstern, Aliyah. 2014. Children’s multimodal language development. In Christiane Fäcke (ed.) *Manual of language acquisition*. 123–142. Berlin & Boston: De Gruyter.

Morgenstern, Aliyah, and Dominique Boutet. forth. The orchestration of bodies and artifacts in French family dinners.

Morita, Emi, & Tomoyo Takagi. 2018. Marking “Commitment to Undertaking of the Task at Hand”: Initiating Responses with Eeto in Japanese Conversation. *Journal of Pragmatics 124*. Elsevier: 31–49.

Müller, Cornelia. 1998. *Redebegleitende Gesten: Kulturgeschichte, Theorie, Sprachvergleich*, vol.1. Berlin: Spitz.

Müller, Cornelia. 2014. Gesture as ‘deliberate expressive movement.’ *From Gesture in Conversation to Visible Action as Utterance: Essays in Honor of Adam Kendon*: 127–151.

Niebuhr, Oliver, & Kerstin Fischer .2019. Do not hesitate! – Unless you do it shortly or nasally: How the phonetics of filled pauses determine their subjective frequency and perceived speaker performance. 10th Proc. International Conference of Spoken Language Processing (Interspeech), 544-548.

- Peltonen, Pauliina. 2019. Gestures as fluency-enhancing resources in L2 interaction: A case study on multimodal fluency. In Pekka Lintunen, Maarit Mutta, and Pauliina Peltonen (eds.) *Fluency in L2 Learning and Use* 138—158. Clevedon: Multilingual Matters.
- Rendle-Short, Johanna. 2005. Managing the transitions between talk and silence in the academic monologue. *Research on Language and Social Interaction* 38(2). 179–218.
- Rose, Ralph. 2008. Filled pauses in language teaching: Why and how. *Bulletin of Gunma Prefectural Women's University* 29: 47–64.
- Rossano, Federico. 2013. Gaze in conversation. In Jack Sidnell and Tanya Stivers (eds.) *The handbook of conversation analysis*. 308—330 Oxford: Blackwell .
- Schegloff, Emanuel A. (2010): Some other "uh(m)"s. *Discourse Processes* 47(2). 130-174.
- Schnadt, Michael J., & Martin Corley. 2006. The influence of lexical, conceptual and planning based factors on disfluency production. *Proceedings of the Annual Meeting of the Cognitive Science Society*. 750 – 755.
- Seyfeddinipur, Mandana. 2006. *Disfluency: Interrupting speech and gesture* Unpublished doctoral dissertation, Radboud University.
- Shriberg, Elizabeth E. 1994. *Preliminaries to a theory of speech disfluencies*. Unpublished PhD Thesis, University of California.
- Sloetjes, Han & Peter Wittenburg. 2008. Annotation by Category-ELAN and ISO DCR. In *6th International Conference on Language Resources and Evaluation (LREC 2008)*.
- Smith, Vicki L., & Herbert H. Clark. 1993. On the course of answering questions. *Journal of Memory and Language* 32. 25–38.
- Sterponi, Laura, & Alessandra Fasulo. 2010. “How to Gg on”: Intersubjectivity and progressivity in the communication of a child with autism. *Ethos* 38(1). 116–142.
- Stivers, Tanya. 2015. Coding social interaction: A heretical approach in conversation analysis? *Research on Language and Social Interaction* 48(1). 1–19.
- Stivers, Tanya, and Jack Sidnell. 2005. Introduction: Multimodal interaction. *Semiotica* 2005: 1–20.
- Streeck, Jürgen. 2009. *Gesturecraft: The manu-facture of meaning*, vol.2. Amsterdam/Philadelphia: John Benjamins.
- Streeck, Jürgen 2014. Mutual gaze and recognition. In Mandana Seyfeddinipur and Marianne Gullberg (eds.) *From gesture in conversation to visible action as utterance*. 35–55. Amsterdam/Philadelphia: John Benjamins.

Streeck, Jürgen, Charles Goodwin & Curtis LeBaron. 2011. *Embodied interaction: Language and body in the material world*. Cambridge: Cambridge University Press.

Swerts, Marc. 1998. Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30(4). 485–496.

Tottie, Gunnel. 2011. Uh and um as sociolinguistic markers in British English. *International Journal of Corpus Linguistics* 16(2). 173–197.

Tottie, Gunnel. 2014. On the use of uh and um in American English. *Functions of Language* 21(1). 6–29.

Tottie, Gunnel. 2015. Uh and um in British and American English: Are they words? Evidence from co-occurrence with pauses. In Nathalie Dion, André Lapierre and Rena Torres Cacoullos (eds.) *Linguistic variation: Confronting fact and theory*. 38–55. London/New York: Routledge.

Tottie, Gunnel. 2016. Planning what to say: Uh and um among the pragmatic markers. In Gunther Kaltenböck, Evelien Keizer & Arne Lohmann (eds.) *Outside the clause. Form and function of extra-clausal constituents*. 97–122. Amsterdam/Philadelphia: John Benjamins.

Tottie, Gunnel. 2019. From pause to word: Uh, um and er in written American English. *English Language & Linguistics* 23(1). 105–130. Cambridge: Cambridge University Press.

Voss, Sina, & Niebuhr Oliver. 2022. Beautiful noise? The impact of filled pauses on the perception of speaker charisma. In *Book of Abstracts of the 13th Nordic Prosody Conference*. 14-16. Sonderborg, Denmark.

Yasinnik, Yelena, Stefanie Shattuck-Hufnagel and Nanette Veilleux. 2005. Gesture marking of disfluencies in spontaneous speech. *DiSS- 2005, Disfluency in spontaneous speech*. 183–178.

Yule, George. 1996. *Pragmatics*. Oxford: Oxford University Press.