Silent pauses as clarification trigger

Loredana Schettino
Università degli Studi
di Salerno
Italy
lschettino@unisa.it

Maria Di Maro
Università degli Studi
di Napoli "Federico II"
Italy
maria.dimaro2@unina.it

Francesco Cutugno
Università degli Studi
di Napoli "Federico II"
Italy
cutugno@unina.it

Abstract

Among possible pragmatic feedback an interlocutor can use to acknowledge the degree of understanding of an utterance, clarification requests (CRs) are to be considered. The functional role of CRs can furthermore be expressed via silent pauses - or failed turn-giving moves - which express an understanding problem and are solved through a *clarify* speech act. In this work, we therefore hypothesise that some silent pauses, in specific conditions, may also have an interactional role which is interpreted by the speaker as a clarification need.

1 Introduction

While conversing, interlocutors constantly need to signal or check the understanding of the latest utterance. If the information acquired as input is clear for the hearer, the conversation can go on, otherwise CRs can be used (Ginzburg, 1998). CRs are anaphoric feedback initiated when the processing of a preceding utterance occur (Purver and Richard, 2004), in order to acknowledge - in a metacommunicative way - the degree of understanding of the input (Ginzburg and Macura, 2005). The perlocutive effect of such requests is another peculiar speech act, called *clarify*, described as *additional information to own previous or other speaker's utterance* (Traum and Hinkelman, 1992; Savy, 2010). According to speech communication research, a similar effect could be exerted by means of another conversational tool, i.e., silence.

Since the major study by Sacks and his colleagues (1974), the communicative role of silences has been variously investigated, thus overcoming the idea of silence as a mere absence of speech.

Delving into the linguistic functions of *eloquent silence*, Ephratt (2008) defines it as discourse marker (Schiffrin, 1987) which activates the addressee to take the conversational floor.

Most commonly, speakers change involving a slight gap (about 150–250 ms) may be considered *smooth transitions*. (Jefferson, 1984; Schegloff, 2000). As a matter of fact, gap duration may vary due to the characteristics of the specific communicative situation (i.e., competitive or cooperative conversation, stress, cognitive workload, eye contact between speakers, (Heldner and Edlund, 2010)), but when reaching about 1s duration, inter-turn silences are most likely perceived as a cue for troubles in conversation (Jefferson, 1989; Roberts et al., 2006). In a recent study, Chowdhury and his colleagues (2017) explored the functions of long silences (1 second and above) occurring between and within speakers and found them to vary from response processing, to discourse structuring, to hesitation due to troubles in information processing and need for clarification (Ephratt, 2008; Chowdhury et al., 2017).

Different terms have been used in literature to refer to silences within and/or between turns (Heldner and Edlund, 2010). In our account, in line with Heldner and Edlund (2010), we refer to acoustic silences preceding a speaker change as *gaps* and to silences occurring within speech by the same speaker as *pause*.

Given this framework, this study is meant to investigate the interactional role of within speaker silent pauses interpreted by the speaker as a clarification need, following a speech act requiring a response and preceding a reformulation of the prior utterance so as to make it more understandable. This way, we aim at obtaining a more clear picture of the formal instruments (explicit and/or implicit), speakers may rely on to carry out the same pragmatic function, specifically CRs.

2 Methodology

In order to verify the hypothesis, we analysed 8 Italian dialogues from the CLIPS Corpus (Savy and Cutugno, 2009)[1], comprising 16 different speakers (10 from Naples, 6 from Rome), accomplishing two different tasks, i.e., 6 map-tasks (MT) and 2 image comparison (TD), for about 1,5 h interaction. What is important to underline is that during the exchange, the interlocutors were not allowed to see each other. This means that all the pragmatic functions expressed during the interaction are either linguistic or para-linguistic.

Table 1 provides an overview of the duration of the analysed dialogues, turns/minutes rate was also considered so as to compare speaker change trends.

Dialogue	Duration (min)	turns/min
MTA01 N	05:42	23,3
MTA02 N	17:30	22,1
MTB01 N	9:05	16,8
MTA01 R	9:28	20,7
MTA02 R	18:48	17,0
MTB04 R	10:01	14,9
TDA01 N	14:18	25,2
TDA02 N	10:16	22,4
mean	11:53	20,3
std.dev.	4,2	3,4

Table 1: Duration and turns/minutes rate of the analysed CLIPS dialogues.

For each dialogue, we firstly annotated CRs in the form of questions, here called *explicit* CRs. These can have different formulations (i.e. polar questions, wh-questions, alternative questions) and express specific communicative problems, such as acoustic, lexical, syntactic, logical, and inferential problems (Rodríguez and Schlangen, 2004).

On a second annotation level, gaps, and pauses were considered. Among the latter, pauses carrying out structuring (STR) and interactional (INT) functions were selected. That is to say, pauses playing a structuring role in the discourse on both syntactic and intonation level (Boomer and Dittmann, 1962; Origlia et al., 2019; Cataldo et al., in press), and an interactional role, working as a signal of discourse processing and/or planning within the interaction (e.g. turn-taking devices, signal of trouble in discourse, (Bazzanella, 2006)).

[1]available at www.clips.unina.it

Each STR-INT pause was annotated along with its left and right span. Specifically, the two parts of the turn, separated by the silence, were annotated according to their speech act. For the pragmatic annotation, we used the annotation scheme Pr.A.T.I.D - Pragmatic Annotation Tool for Italian Dialogues (De Leo and Savy, 2007; Savy, 2010) - from which we used the *final moves* tags, i.e., *action_directive*, *open_option*, *explain*, *check*, *query_w*, *query_y/n*, *info_request*, *align*, *over*, *acknowledgment*, *repeat/rephrase*, *fatic*, *continue*, *not_ready*, *clarify*, *reply_y/n*, *reply_w*, *reply*, *object*, *hold*, *correct*. This pragmatic annotation was useful to determine when a STR-INT pause was perceived as a signal of problems due to clarification need (CR-pauses) or not (other-STR-INT pauses), and in which pragmatic context it mostly occurs.

It was then introduced the concept of *implicit* CR defined as a less frequent and marked CR formulation w.r.t. the above mentioned explicit one. Specifically, it refers to pauses occurring in "failed turn-giving scenarios", where the turn given was not taken by the interlocutor, interpreted by the speaker as a request to get more clarifications, although not specifying which problem occurred.

Other measurements obtained from the corpus analysis are the number of clarification needs on the number of turns per dialogue, the number of all pauses and gaps, and duration of STR-INT pauses (CRs and other STR-INT pauses) compared to the duration of between speakers gaps and other within speaker pauses. In the next section (Section 3), these measurements are presented and discussed.

3 Results and Discussion

General data on CRs' occurrence and incidence in the analysed speech are reported in the following table (Table 2). In about 1,5 h a total amount of 168 CRs (comprising both explicit and implicit ones) occurred, with an incidence on the number of turns varying from one CR each 6,6 turns to one CR each 24,1 turns.

As expected, explicit CRs (135 occurrences) are more frequently used than implicit CRs (33 occurrences), as they express interactional troubles concerned with information processing difficulties which are not yet verbalised and are solved by the speaker before the hearer manages to explicitly express them. Table 3 points out how the frequency of implicit CRs depends on the interlocutors: some

of them prefer to verbalise the problem, some of them are not able to verbalise it as the speaker tries to solve it before it is verbalised, and some other interactions are simply more successful.

The analysis of silences duration confirmed the tendency described in Jefferson (1984; 1989) and Schegloff (2000), as intra-turn silences (pauses), and successful inter-turn silences (gaps) are shorter, whereas STR-INT pauses resulting from failed inter-turn silences, expressing some kind of troubles, are longer (Table 4).

dialogue	CRs	turns/CRs
MTA01 N	9	14,8
MTA02 N	58	6,7
MTB01 N	19	8,1
MTA01 R	26	7,5
MTA02 R	22	14,5
MTB04 R	9	16,6
TDA01 N	15	24,1
TDA02 N	15	15,3
mean	21,6	13,4
std.dev.	*15,8*	*5,8*

Table 2: Number of CR occurrences and turns/CRs rate per dialogue.

Dialogue	explicit_CR	implicit_CR
MTA01 N	78%	22%
MTA02 N	91%	9%
MTB01 N	74%	26%
MTA01 R	69%	31%
MTA02 R	73%	27%
MTB04 R	89%	11%
TDA01 N	80%	20%
TDA02 N	73%	27%
mean	78%	22%
std.dev.	*0,1*	*0,1*

Table 3: Percentages of explicit CRs and implicit CRs per dialogue.

Type of Silence	Duration (mean sec.)
gap	0,52
pause	0,46
STR-INT pause	1,46

Table 4: Average duration of different types of silences in CLIPS.

However, STR-INT pauses can be perceived as a sign of different kinds of troubles. What differentiate them is the speech act that precedes and, more importantly, follows. Figure 1 shows how STR-INT pauses functioning as implicit CRs are mostly preceded by an explanation and followed by a clarification, indicating that the explanation not being clear caused the hearer not to take the given turn with the speaker continuing with a clarification. On the other hand, the other STR-INT pauses are mostly followed by detailed explanations or other kinds of speech acts, such as align (i.e., *Did you get it?*), questions (i.e., *Can you tell me what you see?*), and check (i.e., *Do you have this woman in the small display?*).
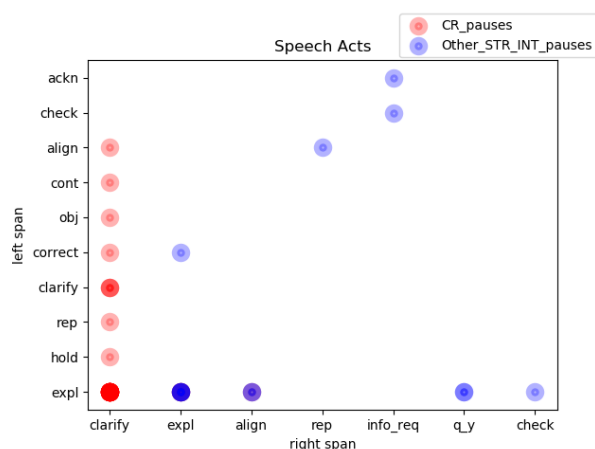

Figure 1: Speech acts occurring before (left-span) and after (right-span) implicit CRs (red spots) and other STR_INT pauses (blue spots). Speech acts frequency for each type of STR_INT pauses (CRs_pauses and other_STR_INT_pauses) is here represented as colour intensity.

4 Conclusions

This study has investigated silent pauses in task-oriented dialogues. More specifically, silent pauses playing both a structuring and interactional role were analysed in order to verify and describe in which conditions they work as CRs. The presence of troubles in conversation due to understanding problems was found to be signalled in two different ways, though explicit requests (explicit CRs) and implicitly through structuring and interactional silent pauses (implicit CRs), significantly longer than the other pauses. The latter covered a smaller, though pretty consistent, amount of CRs.

The analysis we carried out has provided preliminary results to be tested on a larger corpus. How-

ever, these first considerations contribute to better understanding interactional dynamics and which different formal ways may carry out the functional role of CRs in conversation.

Acknowledgments

References

Carla Bazzanella. 2006. Discourse markers in italian: towards a'compositional'meaning.

Donald S Boomer and Allen T Dittmann. 1962. Hesitation pauses and juncture pauses in speech. *Language and speech*, 5(4):215–220.

Violetta Cataldo, Loredana Schettino, Renata Savy, Isabella Poggi, Antonio Origlia, Alessandro Ansani, Isora Sessa, and Alessandra Chiera. in press. Phonetic and functional features of pauses, and concurrent gestures, in tourist guides' speech. In *Studi AISV*.

Shammur Absar Chowdhury, Evgeny Stepanov, Morena Danieli, and Giuseppe Riccardi. 2017. Functions of silences towards information flow in spoken conversation. In *Proceedings of the Workshop on Speech-Centric Natural Language Processing*, pages 1–9.

Simona De Leo and Renata Savy. 2007. Specifiche per l'etichettatura pragmatica dei testi in pra.t.i.d. [online] Parlaritaliano.it.

Michal Ephratt. 2008. The functions of silence. *Journal of pragmatics*, 40(11):1909–1938.

Jonathan Ginzburg. 1998. Clarifying utterances. In *Proc. of the twente workshop on the formal semantics and pragmatics of dialogues. Universiteit Twente, Faculteit Informatica, Enschede*, pages 11–30.

Jonathan Ginzburg and Zoran Macura. 2005. The emergence of metacommunicative interaction: some theory, some practice. In *Symposium on the Emergence and Evolution of Linguistic Communication (EELC'05)*, page 35. Citeseer.

Mattias Heldner and Jens Edlund. 2010. Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, 38(4):555–568.

Gail Jefferson. 1984. Notes on some orderlinesses of overlap onset. *Discourse analysis and natural rhetoric*, 500:11–38.

Gail Jefferson. 1989. Preliminary notes on a possible metrics which provides a 'standard maximum' silence of approximately one second in conversation. *Conversation: An interdisciplinary perspective*, 3:166.

Antonio Origlia, Renata Savy, Violetta Cataldo, Loredana Schettino, Alessandro Ansani, Isora Sessa, Alessandra Chiera, and Isabella Poggi. 2019. Human, all too human: Towards a disfluent virtual tourist guide. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, pages 393–399.

Matthew Purver and John Richard. 2004. *The theory and use of clarification requests in dialogue*. Ph.D. thesis, University of London.

Felicia Roberts, Alexander L Francis, and Melanie Morgan. 2006. The interaction of inter-turn silence with prosodic cues in listener perceptions of "trouble" in conversation. *Speech communication*, 48(9):1079–1093.

Kepa Joseba Rodríguez and David Schlangen. 2004. Form, intonation and function of clarification requests in german task-oriented spoken dialogues. In *Proceedings of Catalog (the 8th workshop on the semantics and pragmatics of dialogue; SemDial04)*, pages 101–108.

Harvey Sacks, Emanuel Schegloff, and Gail Jefferson. 1974. The simplest systematics for the organization of turn-taking for conversations. *Language*, 50(4):696–735.

Renata Savy. 2010. Pr. a. ti. d: A coding scheme for pragmatic annotation of dialogues. In *LREC*, pages 2141–2148.

Renata Savy and Francesco Cutugno. 2009. Diatopic, diamesic and diaphasic variations in spoken italian. In *Proceedings of the 5th Corpus Linguistics Conference: CL2009*, pages 20–23.

Emanuel A Schegloff. 2000. Overlapping talk and the organization of turn-taking for conversation. *Language in society*, 29(1):1–63.

Deborah Schiffrin. 1987. *Discourse markers*. 5. Cambridge University Press.

David R Traum and Elizabeth A Hinkelman. 1992. Conversation acts in task-oriented spoken dialogue. *Computational intelligence*, 8(3):575–599.