

# Functions of Silences towards Information Flow in Spoken Conversation

Shammur Absar Chowdhury and Evgeny A. Stepanov and Morena Danieli  
and Giuseppe Riccardi

Signals and Interactive Systems Lab  
Department of Information Engineering and Computer Science  
University of Trento, Italy

## Abstract

Silence is an integral part of the most frequent turn-taking phenomena in spoken conversations. Silence is sized and placed within the conversation flow and it is coordinated by the speakers along with the other speech acts. The objective of this analytical study is twofold: to explore the functions of silence with duration of one second and above, towards information flow in a dyadic conversation utilizing the sequences of dialog acts present in the turns surrounding the silence itself; and to design a feature space useful for clustering the silences using a hierarchical concept formation algorithm. The resulting clusters are manually grouped into functional categories based on their similarities. It is observed that the silence plays an important role in response preparation, also can indicate speakers' hesitation or indecisiveness. It is also observed that sometimes long silences can be used deliberately to get a forced response from another speaker thus making silence a multi-functional and an important catalyst towards information flow.

## 1 Introduction

Silence is a multifaceted natural phenomenon in human conversations that carries information rich in meaning and function. Even though “silence” is generally defined as the absence of speech (Jaworski, 1993) or a break in a conversation flow, its occurrence has the power to deliver a message, as well as trigger human response similar to any other conversational behavior. Silence in human conversations provides insights into the thought process, emotion, and attitude (Richmond et al.,

1991) among others. At the same time, silence is used to convey power (dominance) (Saunders, 1985; Tannen, 1990), respect, and manage conflicts.

Along with speech, silence is an integral part of human interaction, and the two complement and provide information about each other. In the words of Bruneau (Bruneau, 1973):

*“Silence is to speech as the white of this paper is to this print”*

– Thomas J Bruneau.

Given that the reasons for silence are limitless, it also has many functions. One function is “eloquent silences” that includes the use of silence in the funeral, at religious ceremonies, as a legal privilege, or in response to a rhetorical question (Ephratt, 2008). Apart from this, silence can be used to indicate topic avoidance, lack of information to provide response, agreement, disagreement, anger, frustration, uncertainty, hesitancy and others.

Over the years, researchers have studied silence with respect to, but are not limited to, the location of silence in a conversation (Richmond et al., 1991; Jensen, 1973) or its role in a conversation (Cappella, 1980; Zimmermann and West, 1996; McLaughlin and Cody, 1982) or how its duration changes with different emotions (Alam et al., 2016). Silence has also been studied as a method for non-verbal communication (Kogure, 2007; Bruneau, 2008) and its practices in different cultures (Richmond et al., 1991), or in different contexts. It has also been observed as a powerful tool for conflict-management (Oduro-Frimpong, 2007), and within the context of psychotherapy (Frankel et al., 2006; Gale and Sanchez, 2005; Ladany et al., 2004; Ronningstam, 2006).

Unlike research on speech, the studies on silence are either definitional (theoretical) or de-

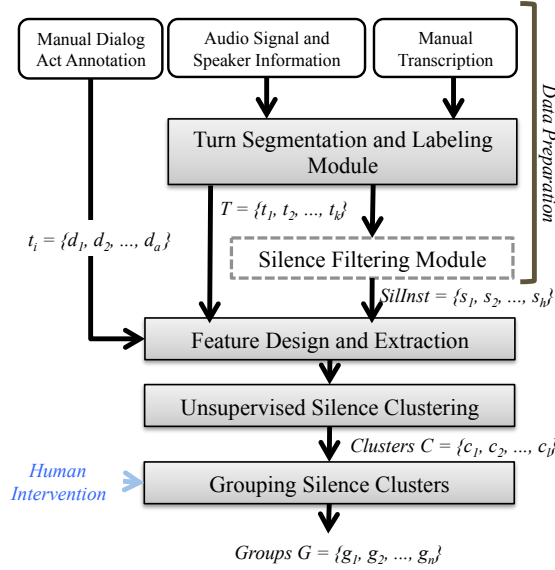


Figure 1: Framework for categorizing functions of long silences.

scriptive. Even within speech research communities, there are very few studies that have analyzed function of silence in a methodological manner. Generally, in a dialog system, silence is not acknowledged as a form of interaction, but rather its function in a conversation is seen as a “pause” or a “gap”. Whereas speech is viewed as the primary carrier of information. Thus, a further study of silence and its functions is important, as silence *often does serve as a message*, or at least as *means that offers contextual cues to the surrounding speech*.

Therefore, the goal of this paper is to analyze the function of long silences<sup>1</sup> occurring between and within speakers in dyadic spoken conversations. Our focus is on understanding the perceived reasons of such functions towards the information flow in spoken conversations.

The paper is organized as follows. An overview of the experimental methodology used in this study is given in Section 2. We present an analysis of our findings in Section 3 and provide conclusions in Section 4.

## 2 Methodology

The methodology followed for grouping functions of the long silences is shown in Figure 1. The pipeline includes a data preparation phase, which involves extraction and selection of the long silence instances; followed by the feature design and

<sup>1</sup>In our study, we defined long silences as having the duration greater and equal to 1 second.

extraction phase. The next phase in the pipeline is unsupervised clustering of the selected silence instances, which are later grouped into hierarchical clusters for the analysis of their functions.

### 2.1 Data Preparation

For this study, we have used conversations from the SISL Human-Human Conversational Discourse Corpus. The data is a subset of a large Italian call-center corpus where call center agents are engaged in conversations with real customers. The customers are calling to solve some specific problem or seek information. The inbound Italian phone conversations are recorded on two separate audio channels with a quality of 16 bits, 8kHz sample rate. The collected conversations ( $\approx 10K$ ) have an average duration of  $396.6 \pm 197.9$  seconds.

To analyze the role of silence in information flow of the conversation, we have selected 10 conversations that contain manual dialog act annotations (Chowdhury et al., 2016b) following dialog Act Markup Language (DiAML) (Bunt et al., 2010, 2012) annotation scheme. The details of the dimensions and the communicative functions considered for the annotation are given in Table 1. The dimensions such as: Discourse Structuring, Speech and Turn Management dimensions are mapped to the tag *Other*, as they are very infrequent.

#### 2.1.1 Extraction of Silence

Silence positions, as well as turn types, such as speakers’ turns, overlapping turns, are extracted using the *turn segmentation and labeling system* (Chowdhury et al., 2016a). The input to the system is the audio of the conversation, the manual transcription and speaker information.

The forced-aligned transcription is obtained using an in-domain Automatic Speech Recognition (ASR) (Chowdhury et al., 2014). Lexical information from these forced-aligned transcripts is used to extract turn-taking sequences. The pipeline uses the time aligned output as tokens to create Inter-Pausal Units (IPUs) for each input channel. IPUs are defined as segments of consecutive tokens with no less than 50 ms gaps in-between. Using the time information of inter-IPUs and intra-IPUs, we then define **steady conversation segments** where each segment maintains a steady time-line for both interlocutor channels. The labels of each silence segment are then defined by a set of rules as follows:

- Pause ( $P$ ): Gaps between the turns of the same speaker with no less than 0.5 second.  $P_A$  and  $P_C$  represent agent and customer’s pauses respectively.
- Lapse between speakers ( $L_B$ ): Floor switches between the speakers with a silence duration of 2 seconds or more.
- Lapse within speaker ( $L_W$ ): Gaps between the same speakers’ turns with a silence duration of 2 seconds or more.
- Switch ( $SS$ ): Floor switches between the speakers with silence duration of less than 2 seconds or with overlapping frames not more than 20 ms. This category is also known as *gaps*.

The labeled turn sequences are then used to select silence instances for the analysis.

### 2.1.2 Silence Filtering

From the 10 conversations we have extracted 433 instances of silences with the duration greater or equal to 1 second. The instances are categorized into two groups:

- Between-Speaker Silences ( $B$ ): These instances of silence include gaps between different speaker turns that are greater or equals to 1 second.  $B = \{S_l, L_B\}$ , where  $S_l$  stands for gaps  $\geq 1$  second and  $< 2$  seconds whereas  $L_B$  are lapse between speakers  $\geq 2$  seconds.
- Within-Speaker Silences ( $W$ ): These instances of silence include pauses between the same speaker’s turns that are greater or equals to 1 second.  $W = \{P_l, L_W\}$ , where  $P_l$  stands for pauses  $\geq 1$  second and  $< 2$  seconds whereas  $L_W$  are lapse within speaker  $\geq 2$  seconds.

For the initial analysis, the instances of long silences that occur after or before overlapping speech (61 silence instances) are ignored. As a result, the analysis is performed on 372 instances.

## 2.2 Feature Design and Extraction

Even though silence is an inherently valueless phenomena that possesses no function on its own, individual instances of silence gain its meaning and function from the surrounding context. Consequently, modeling functions of silences requires conceptualization of the context and features that capture it. Dialog acts carry specific communicative functions such as question, answer, expression

of agreement, disagreement, etc. Since dialog acts are assigned to the speech segments (turns) that surround the long silences, they provide the information that could be used to model the context of silence instances.

Table 1: Core dimensions and communicative functions from ISO 24617-2 standard considered for dialog act annotation.

Dimension	Comm.Function	Group
General (Task)	<i>Information Transfer Functions</i>	
	Question Set Question Choice Question Propositional Question Check Question	Information Seeking
	Inform Answer Confirm Disconfirm Agreement Disagreement Correction	Information Providing
	<i>Action Discussion Functions</i>	
	Offer Promise Address Request Accept Request Decline Request Address Suggest Accept Suggest Decline Suggest	Commissives
	Suggest Request Instruct Address Offer Accept Offer Decline Offer	Directives
	<i>Time Management</i>	
	<i>Auto-Feedback</i>	
	<i>Allo-Feedback</i>	
	<i>Social Obligations Management</i>	
	Stalling, Pausing	
	Positive, Negative	
	Positive, Negative, Feedback Elicitation	
	Initial-Greeting, Return-Greeting Initial-Self-Intro, Return-Self-Intro Apology, Accept-Apology Thanking, Accept-Thanking Initial-Goodbye, Return-Goodbye	

The dialog act dimensions and communicative functions listed in Table 1 are used as features for the analysis of between and within speaker silence instances. Each turn preceding or following a silence is transformed into a feature vector using one-hot representation for dialog acts.

The vectors encode information such as the following. Feedback, a joined dimension of auto-feedback and allo-feedback, ( $fb$ ) =  $\{0, 1\}$ , where  $fb=0$  represent the absence of feedback dialog acts in the turn and vice-versa. Similarly, the vector also includes other dialog act dimensions like Time Management ( $tm$ ), and Social Obligations Management ( $s$ ). The General dimension is split into two: (a) information seeking ( $q$ ) and (b) information providing and action discussion functions

(*ac*). The motivation behind such a split is to distinguish between information seeking dialog acts which impose an expected pattern on its recipient, i.e an obligation to provide the requested information.

Since according to the DiaML annotation standard a turn can contain several dialog acts, the vector representation specifically encodes the last dialog act of the preceding turn (*lact*) and the first dialog act of the turn following the long silence (*fact*). Both *lact* and *fact*  $\in \chi$ , where as  $\chi = \{Ac, Q, F, TimeM, Ap, Thank, Int, Other, None\}$ . In the set,  $\chi$ , *Ac* represents communicative functions from *information providing* and *action discussion* functions; *Q* represents *Information Seeking* functions; *F* represents *Feedback (auto-feedback and allo-feedback)* functions; *Apo* represents *apology and accept-apology* functions; *Thank* represents *thanking and accept-thanking*; *Int* represents *initial and return greetings, self-introductions, and goodbyes*; *Other* represents all the dialog acts not used for the analysis. *None*, on the other hand, indicates absence of dialog acts.

The feature vectors of preceding, pr ( $|pr|=6$ ) turn, and succeeding (following), su ( $|su|=6$ ) turn, are merged to represent a silence instance for categorization ( $|sil| = 6 * 2 = 12$ ).

### 2.3 Unsupervised Annotation of Silence Function

The described representation of silence instances is applied for clustering using Cobweb clustering algorithm (Fisher, 1987) – a well-known concept formation system designed to model human concept learning. Cobweb constructs clusters using “concept hierarchy” that optimally and incrementally accounts for the observed regularities on a set of instances. In other words, given a set of silence instances, Cobweb discovers a classification scheme that covers the patterns with respect to provided feature vectors. Instead of forming concepts at a single level of abstraction, Cobweb groups instances into a classification tree where leaves represent similar instances, and internal nodes represent broad concepts. The generality of a broader concept increases as the nodes approach the root of the tree. Each cluster is characterized with a probabilistic description.

The classification tree is constructed incrementally inserting the instances into the tree one by one. When adding an instance, the algorithm tra-

verses the tree top-down starting from the root of the tree. At each node, there are four possible operations: (a) insert (b) create (c) merge and (d) split. These operations are selected with respect to the highest category utility (*CU*) metric (Gluck and Corter, 1985). The metric is derived from the categorization studies in cognitive psychology and is shown in Equation 1.

Category utility, *CU*, attempts to maximize both (a) the probability of the instances in the same category to have feature values in common; and (b) the probability of the instances in different categories to have different feature values.

$$CU(C_l) = \sum_i \sum_j (Pr[f_i = v_{ij}|C_l]^2 - Pr[f_i = v_{ij}]^2) \quad (1)$$

In the equation,  $Pr[f_i = v_{ij}]$  represents the marginal probability that feature  $f_i$  has value  $v_{ij}$ , whereas  $Pr[f_i = v_{ij}|C_l]$  represents the conditional probability that feature  $f_i$  has value  $v_{ij}$ , given the instance belongs in cluster  $C_l$ .  $CU(C_l)$  estimates the quality of individual cluster.

To measure the quality of overall clustering of the silences, we calculate the average category utility function  $CU(C_1, C_2, \dots, C_k)$ , as shown in Equation 2.

$$CU(C_1, C_2, \dots, C_k) = \frac{1}{k} \left( \sum_l Pr[C_l] \right) \quad (2)$$

In the equation,  $k$  is the total number of categories. The overfitting is controlled by  $\frac{1}{k}$ .

Therefore, for each set ( $B, W$ ), we applied Cobweb clustering algorithm implemented in (Hall et al., 2009) with acuity  $A = 1.0$  and cutoff threshold of  $C = 0.0028$ .

## 3 Analysis

### 3.1 Resulting Clusters

For between-speakers silences ( $B$ ), we have obtained 24 leave clusters, whereas for within-speaker silences ( $W$ ), we have obtained 26 leave clusters. The distribution of dialog act sequences in each cluster is given in Tables 2 and 3.

### 3.2 Categorization of Silence Functions

Assuming that each cluster represents a specific function of a silence, the clusters are manually grouped with respect to their parents in the classification tree. The manual grouping of silence clusters is performed considering conversation scenarios surrounding the silence events. For instance,

Table 2: Preceding (column two) and succeeding (column three) turn communicative function sequences for each clusters and their frequency inside parenthesis for between speaker silences. The first column of the table represents the classification tree’s leaf id of the corresponding cluster.

Id	Preceding turn dialog acts	Succeeding turn dialog acts
2	question(19); checkquestion(9); inform question(2); inform checkquestion(2); inform autopo- sitive question(1); choicequestion(1); au- topositive checkquestion(1)	answer(12); confirm(11); inform(3); answer in- form(3); disconfirm(2); confirm inform(2); dis- confirm answer(1); answer request(1)
3	question(2); initialselfintroduction initialgreeting returnselfintroduction question(1); initialselfin- troduction initialgreeting initialselfintroduction question(1); inform checkquestion(1); choice- question(1)	other(2); autopo- sitive(2); autopo- sitive return- greeting stalling inform(1); allopositive(1)
5	question(2)	stalling answer(2)
6	question(1)	stalling checkquestion(1)
8	initialgreeting initialselfintroduction question(2)	returngreeting returnselfintroduction answer in- form(1); returngreeting inform(1)
9	initialselfintroduction question(1)	returngreeting returnselfintroduction(1)
11	inform(20); request(6); confirm(2); answer(2); suggest(1); stalling request(1); offer(1); ini- tialgreeting initialselfintroduction request(1); in- form none inform(1); answer request(1); answer autopo- sitive inform(1); agreement(1); addressre- quest(1)	inform(22); acceptrequest inform(4); inform in- form(3); confirm(3); acceptrequest(2); inform question(1); answer request(1); agreement(1); addressrequest(1); acceptoffer inform stalling(1)
13	autopo- sitive(16); allopositive(1)	inform(15); inform request(1); correction(1)
15	other(6)	inform(5); suggest(1)
17	answer thanking(1)	inform(1)
18	pausing(2); stalling(1); inform stalling(1)	inform(2); confirm(1); answer(1)
19	allopositive none(1)	inform inform(1)
22	inform(23); answer(2); request(1); correction(1); confirm(1); acceptrequest inform(1)	autopo- sitive(19); autopo- sitive inform(4); au- topositive question(3); autopo- sitive checkques- tion(2); allopositive(1)
25	allopositive(1)	autopo- sitive(1)
29	pausing(1)	autopo- sitive(1)
31	inform(10); answer(4); confirm(3); request(1); disconfirm(1); correction(1)	question(12); checkquestion(6); question in- form(1); question checkquestion(1)
33	autopo- sitive(2)	question(2)
34	autopo- sitive(1)	question acceptthanking(1)
37	inform(2); confirm(2); offer(1)	pausing(4); stalling(1)
38	inform(1)	none(1)
43	other(5)	other(5)
45	other(1)	returnselfintroduction(1)
46	initialgreeting initialselfintroduction question other(1)	returngreeting(1)
47	inform(2); request(1); other inform(1); de- clinerequest(1); answer(1); acceptrequest(1)	other(5); other stalling(1); other other ques- tion(1)



Table 3: Preceding (column two) and succeeding (column three) turn communicative function sequences for each clusters and their frequency inside parenthesis for within speakers silences. The first column of the table represents the classification tree’s leaf id of the corresponding cluster.

Id	Preceding turn dialog acts	Succeeding turn dialog acts
2	inform(95); answer(6); request(3); stalling inform(2); inform inform(2); correction(2); question request(1); offer(1); inform request(1); confirm(1)	inform(90); request(5); answer(5); inform inform(4); offer(2); inform stalling(2); inform question(2); suggest(1); inform stalling inform stalling(1); correction(1); addressrequest(1)
3	none(1)	inform(1)
7	pausing(2)	question(1); checkquestion(1)
8	autopositive(1)	question(1)
9	question(8); checkquestion(3); inform question(1)	question(8); checkquestion(3); question inform(1)
10	question(1)	other(1)
11	question(1)	pausing(1)
12	question(1)	autopositive autopositive(1)
14	other(1)	apology inform(1)
15	other(3)	other(3)
16	other(1)	autopositive inform(1)
19	pausing(1)	pausing(1)
20	inform stalling(1)	stalling(1)
21	autopositive pausing(1)	pausing autopositive inform(1)
22	stalling(1)	other inform(1)
23	autopositive(1)	other(1)
24	autopositive(5); autopositive autopositive(1)	autopositive(4); autopositive thanking(1); autopositive question(1)
25	autopositive(1)	stalling inform(1)
29	inform none(1)	none inform(1)
33	stalling(1); pausing(1); other stalling(1)	inform(3)
34	autopositive(5)	inform(4); inform autopositive question(1)
36	question(4)	inform(3); inform inform(1)
37	other(1)	inform(1)
39	inform(9)	stalling inform(7); stalling(2)
40	inform(7)	question(4); question inform(2); choicequestion(1)
41	inform(2); agreement Null inform(1)	autopositive(3)

in a conversation a participant may expect an answer to a question or a contribution from another speaker that might yield a long silence due to the time required to prepare an answer. It might take long to get the information to the query or simply be an act of noncompliance. This long silence period is considered as a failure to contribute to an ongoing conversation. To repair this speakers may use strategies such as repeating the query, changing the topic, or ask for more time to respond.

Below we give example scenarios observed in the silence cluster groups:

The Between-Speaker Silence cluster groups are:

- A mode of response preparation (*RP*): In this group, there can be two subcategories based on the type of response given by the speaker after the silence. The subcategories are:
  - Response to the previous turn’s question in the form of information that includes an answer to the question, a feedback, or asking for

more time to answer. This category includes clusters **RP1**={2, 3, 5, 6, 8, 9}.

- A response can also be a question to the information/feedback provided in the previous turn. This category includes clusters **RP2**={31, 33, 34}.

- A mode of information flow (*IF*): These silences can either be a: 1) conversational silences where both speakers are exchanging information or feedback 2) forced silences (deliberate<sup>2</sup>), where the current speaker is using a silence as a tool to force the interlocutor to respond. The member clusters of this group are **IF**={11, 13, 15, 17, 18, 19, 22, 25, 29, 37, 38}.
- Silences in Other categories (*B – Oth*): These are the silences which are motivated by factors, such as discourse structuring, not considered in the study. This group includes clusters

<sup>2</sup>These silence instances are usually longer. For this study the threshold for this type of silences is  $\geq 2$  seconds.

**B-Oth**={43, 45, 46, 47}.

The above-mentioned categories are presented in Examples 1 and 2. In Example 1, we observe that the caller is asking the call center operator a reason behind an action, and the act is followed by a long silence of 1.41 seconds. After the interval, the operator is passing some information regarding the earlier query by the caller. From the operational point of view, the interval might have been used to either acquire information or to structure it. Similarly, in *RP2* scenario in Example 1, after the operator informs that the ‘electric power’ will not be activated, the caller is taking a long silence of 1.38 seconds to respond to the given information, asking another question. This silence could have been again used for preparing the answer, or it might be the time taken by the responding speaker to compose the next action. In Example 2, we present a scenario where the silence category *IF* is used deliberately to force another speaker to reply.

The silence in both examples may play other cognitive functions such as controlling emotional attitudes. However, as the focus of this study is to understand the function of long silences in the information flow, these cognitive functions are not considered.

**Example 1. Example of silence category *RP* : *RP1***

caller: al distacco perfetto ora eh eh  
su che base mi perdoni  
caller: *the complete interruption ... perfect! now*  
*ehm ehm due to what reason, excuse me?*  
(1.14) Category - *RP1*  
operator: ah ascolti qui ci sono una  
serie di fatture malgrado  
operator: *Listen (please) we have here a number of*  
*unpaid bill in spite of*

**Example of silence category *RP* : *RP2***

operator: la luce non gliela riprist  
non viene ripristinata  
operator: *the electric power will not be reactiv will*  
*not be reactivated*  
(1.38) Category - *RP2*  
caller: ma cosa devo pagare se io ho  
gi congruato tutto con  
trecentoquarantacinque euro mi  
perdoni cosa devo pagare la  
caller: *but what do I have to pay if I have already*  
*paid 345 euros I beg you pardon but what do I*  
*need to pay the*

**Example 2. Example of silence category *IF***

caller: [lei deve fare una cosina  
lei ha un delle]  
caller: [You have to do a small thing you have  
some]  
operator: [per e se]  
operator: [but and if]

caller: belle schermate a disposizione  
mi deve aprire la mia ehe il mio  
fax inviato il ventitr zero otto  
duemiladodici cortesemente  
caller: *beautiful screens available you have to open*  
*my own and you will find my fax sent on 23rd of*  
*August 2016*  
(2.12) Category - *IF* (deliberate  
silence)  
operator: vediamo subito  
operator: *let us see immediately*

The Within-Speaker Silence cluster groups are:

- Organizational silence (*CS*): The long pause used for the purpose of organizing the information flow in the conversation. This group contains clusters of silences where a speaker is providing information and the silence between turns can either be a time taken to think, find information, or to compose and plan the next turn. **CS**={2, 3, 19, 20, 21, 22, 23, 24, 25, 29, 33, 34, 39}.
- Indecision or Hesitation silence (*H*): In this group of silences, speaker is either confused about some information, needs clarification, or have some queries. The member clusters of this group are **H**={7, 8, 9, 10, 11, 12, 36, 40}.
- Silences in Other categories (*W – Oth*): These are the silences which are motivated by other factors, not considered for the present study. This group includes clusters **W-Oth**={14, 15, 16, 37, 41}.

**Example 3. Example of silence category *CS* and *H***

caller: non riesco a parl devo  
parlarle ho parlato con cinque  
suoi colleghi e mi hanno  
chiamato due consulenti  
caller: *I cannot tal ... I need to talk ... I talked with*  
*five colleagues of you and two consultants called*  
*me*  
(1.16) Category - *CS*  
caller: io oggi pomeriggio devo andare  
dall avvocato per denunziarvi  
per diecimila euro al giorno di  
danni che mi avete arrecato da  
stamattina  
caller: *this afternoon I will go my lawyer for*  
*sueing you due to ten thousand euros in damage*  
*per day due to this morning (power) interruption*  
(1.65) Category - *CS*  
caller: ehe perch io ho gi pagato  
tutto nel  
caller: *ehm because I already paid all what I due*  
caller: senso che tutte queste  
bollette sono state congruagiate  
con una di  
trecentoquarantacinque euro  
incluso

caller: *because all these bills were paid with another one of 345 euros including*  
 caller: *il mese di luglio e agosto*  
 caller: *the months of July and August (as well)*  
 (1.57) Category - *CS*  
 caller: *ehe avevo gi chiarito il (.) primo distacco l abbiamo sospeso mi hanno richiamato perch non trovate una vostra lettera di risposta*  
 caller: *and I already told this when (.) there was the first interruption (that) was suspended they called me because you are not able to find a reply letter from you*  
 (1.01) Category - *H*  
 caller: *ora devo (.) parlare con lei o devo parlare con qualcuno sopra di lei mi perdoni se sono abbastanza*  
 caller: *now (I) have (.) to call with you or have (I) to call with you boss? sorry but (I) am enough*

In Example 3, we present dialog scenarios with assigned categories. It is observed that the top three long silence intervals are used either to plan the next turn or to take the time to think. Whereas in the last (shortest) silence of 1.01 second, before threatening the operator, the caller either hesitates, feels bad, or is not sure whether a threat is going to work.

The duration distribution statistics for each category of silence functions are presented in Tables 4 and 5. For between-speaker silence categories, in Table 4, it is observed that median duration of silence category *RP2* along with *B-Oth* are longer compared to *RP1* and *IF*. As for within-speaker silence categories, it is observed that median duration of *H* categories is longer than *CS*. The observation is explained as the speakers might need more time to take the next turn when s/he is facing indecision, hesitation, or need clarification about something.

Table 4: Statistics of between-speaker long silences categories.

<i>Between-Speaker Silence</i>	<i>RP1</i>	<i>RP2</i>	<i>IF</i>	<i>B-Oth</i>
1st Quadrant	1.21	1.33	1.27	1.36
Median	1.37	1.76	1.59	1.96
3rd Quadrant	1.62	2.67	2.13	2.93
No. Instances	47	23	107	12
<b>Total</b>	<i>189</i>			

## 4 Conclusion

The main focus of this analytical study is to explore the functions of long silence within and between speakers towards the information flow in a conversation. In an attempt to find such functions, the study utilizes the sequences of dialog act tags

Table 5: Statistics of within-speaker long silences categories.

<b>Within-Speaker Silence</b>	<b>CS</b>	<b>H</b>	<b>W-Oth</b>
1st Quadrant	1.13	1.10	1.32
Median	1.36	1.42	1.63
3rd Quadrant	1.76	2.63	2.06
No. Instances	145	29	9
<b>Total</b>	<i>183</i>		

present in the left and right context (concerning speaker turns) surrounding the silence itself, and designs feature vector to represent a long silence. These designed feature vectors are later used to cluster silences using a well-known hierarchical concept formation system (Cobweb), which is designed to model different aspects of human concept learning. Following the clustering, we have manually grouped the clusters into functional categories and have studied their significance, and duration distribution.

The functions of silence we observe vary from response preparation to hesitation to asking some queries. It is also observed that sometimes these long silences are used deliberately to get a forced response from another speaker. It can also indicate the indecisiveness of the current speaker.

Even though most of the research from speech communities ignores the silences, our observation shows that by considering the function of long silences, we can better understand the information flow in the conversation. As silences do contribute to explaining the information presented by the speech signals. Silence also has the potential to explain long term behavioral traits and short term states.

This study is our first attempt to analyse, understand and group functions of long silence in dyadic conversations. The observed functions, such as hesitations, are also related to another speech phenomenon – disfluencies. In future work we plan to address the relationship between speech disfluencies and long silences. This analysis will help us to understand the factors and contexts that represent cues of the silence function which is indeed necessary to design computational models for such a simple yet informative event of human conversation.

## References

Firoj Alam, Shammur Absar Chowdhury, Morena Danieli, and Giuseppe Riccardi. 2016. How inter-



- locutors coordinate with each other within emotional segments? In *COLING: International Conference on Computational Linguistics*.
- Thomas J Bruneau. 1973. Communicative silences: Forms and functions. *Journal of Communication* 23(1):17–46.
- Thomas J Bruneau. 2008. How americans use silence and silences to communicate. *China Media Research* 4(2).
- Harry Bunt, Jan Alexandersson, Jean Carletta, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Kiyong Lee, Volha Petukhova, Andrei Popescu-Belis, Laurent Romary, et al. 2010. Towards an ISO standard for dialogue act annotation. In *Seventh conference on International Language Resources and Evaluation (LREC'10)*.
- Harry Bunt, Jan Alexandersson, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Volha Petukhova, Andrei Popescu-Belis, and David R Traum. 2012. ISO 24617-2: A semantically-based standard for dialogue annotation. In *LREC*. pages 430–437.
- Joseph N Cappella. 1980. Talk and silence sequences in informal conversations ii. *Human Communication Research* 6(2):130–145.
- Shammur A. Chowdhury, Giuseppe Riccardi, and Firoj Alam. 2014. Unsupervised recognition and clustering of speech overlaps in spoken conversations. In *Proc. of Workshop on Speech, Language and Audio in Multimedia*.
- Shammur Absar Chowdhury, Evgeny Stepanov, and Giuseppe Riccardi. 2016a. Predicting user satisfaction from turn-taking in spoken conversations. In *Proc. of INTERSPEECH*.
- Shammur Absar Chowdhury, Evgeny A. Stepanov, and Giuseppe Riccardi. 2016b. Transfer of corpus-specific dialogue act annotation to iso standard: Is it worth it ? In *Proc. of LREC*.
- Michal Ephratt. 2008. The functions of silence. *Journal of Pragmatics* 40(11):1909 – 1938.
- D. H. Fisher. 1987. Knowledge acquisition via incremental conceptual clustering. *Machine Learning* 2:139–172.
- Ze'ev Frankel, Heidi M Levitt, David M Murray, Leslie S Greenberg, and Lynne Angus. 2006. Assessing silent processes in psychotherapy: An empirically derived categorization system and sampling strategy. *Psychotherapy Research* 16(5):627–638.
- John Gale and Beatriz Sanchez. 2005. The meaning and function of silence in psychotherapy with particular reference to a therapeutic community treatment programme. *Psychoanalytic Psychotherapy* 19(3):205–220.
- M. A. Gluck and J. E. Corter. 1985. *Information, uncertainty, and the utility of categories*. Proceedings Seventh Annual Conference of the Cognitive Science Society. Lawrence Erlbaum Associates.
- Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten. 2009. The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter* 11(1):10–18.
- A. Jaworski. 1993. *The power of silence: social and pragmatic perspectives*. Language and language behaviors. Sage. <https://books.google.mu/books?id=0NFoAAAAIAAJ>.
- J Vernon Jensen. 1973. Communicative functions of silence. *ETC: A Review of General Semantics* pages 249–257.
- Masato Kogure. 2007. Nodding and smiling in silence during the loop sequence of backchannels in japanese conversation. *Journal of Pragmatics* 39(7):1275–1289.
- Nicholas Ladany, Clara E Hill, Barbara J Thompson, and Karen M O'Brien. 2004. Therapist perspectives on using silence in therapy: A qualitative study. *Counselling and Psychotherapy Research* 4(1):80–89.
- Margaret L McLaughlin and Michael J Cody. 1982. Awkward silences: Behavioral antecedents and consequences of the conversational lapse. *Human communication research* 8(4):299–316.
- Joseph Oduro-Frimpong. 2007. Semiotic silence: Its use as a conflictmanagement strategy in intimate relationships. *Semiotica* 2007(167):283–308.
- Virginia P Richmond, James C McCroskey, and Steven K Payne. 1991. *Nonverbal behavior in interpersonal relations*. Prentice Hall Englewood Cliffs, NJ.
- Elsa Ronningstam. 2006. Cultural function and psychological transformation in psychoanalysis and psychoanalytic psychotherapy. *The International Journal of Psychoanalysis* 87(5):1277–1295.
- George R Saunders. 1985. Silence and noise as emotion management styles: An italian case. *Perspectives on silence* pages 165–83.
- Deborah Tannen. 1990. Silence as conflict management in fiction and drama: Pinter's betrayal and a short story, great wits. *Conflict talk: Sociolinguistic investigations of arguments and conversations*, ed. AD Grimshaw pages 260–279.
- Don H Zimmermann and Candace West. 1996. Sex roles, interruptions and silences in conversation. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4* pages 211–236.