# Analysis of silences in unbalanced dialogues: the effect of genre and role

*Vered Silber-Varod [1] and Anat Lerner [2]*
*[1]The Research Center for Innovation in Learning Technologies, The Open University of Israel*
*[2]Department of Mathematics and Computer Science, The Open University of Israel*

## Abstract

This study examines the diversity of silences in unbalanced dialogues, i.e. dialogues between speakers with different participation levels: responder and reporter. We examined two genres: therapeutic sessions and private dialogues that are based on this responder-reporter structure. When looking at silences versus speech ratios, we found no differences between the genres nor between the roles. However, when grouping the silences by their types: Pauses (intra-speaker silences), gaps (inter-speakers' silences) and silences that occur in the vicinity of speech overlaps, we found that the silence duration of pauses are role dependent in both genres, while the silence duration of gaps were found genre dependent, but not role dependent. Moreover, speech rate was not found genre dependent. It seems that although silences in unbalanced dialogues vary considerably, genre and speaker's role are influential.

## Introduction

Silences are dealt here in the context of prosodic characteristics of speech genre and speakers' role in dialogues. We argue that studying silence types in a quantitative manner can shed light on the study of linguistic, paralinguistic and extra-linguistic aspects of silence. This approach of quantitative analyzes of silences goes back as early as 1939 (Ephratt, 2011:2294), and is carried out mainly in order to analyze the ratios of speech and non-speech, in isolation or in relation to personality variables. Indeed, such studies produced quantitative predictions, such as the "constant ratio" between vocalization and silence in spontaneous speech (Crown & Feldstein, 1991) and "standard maximum" silence in conversation (Jefferson, 1989). By taking this challenge, we declare that we will not consider the linguistic content that surrounds the silence (Wooffitt & Holt, 2010), nor will we try to understand other dimensions in silence research, such as the communicative functions of silence (Bruneau, 1973; Ephratt, 2008; 2011), or the interpretation of silence in the pragmatic sense (inter alia, Kurzon, 2007; Sacks, Schegloff & Jefferson, 1974). The examined two types of silences that this research is focused on are inter-speaker gaps and intra-speaker pauses. These

two terms were mentioned in (Sacks, Schegloff & Jefferson, 1974), who defined pauses as "intra-turn silence (not at a transition-relevance place)", and gaps as "silence after a possible completion point". While pauses are relevant to cognitive processing and are related to effect, style and linguistic structures (Ramanarayanan et al., 2009; Zellner, 1994), gaps are interactive event and were found influential by different types of turns (Tannen, 2000). As Ephrat (2014) claims, "the most prominent case is silence as a discourse marker for turn taking". (Ephrat, 2014:127). Meaning, silence as a discourse marker for turn taking is a direct communicative act that the speaker operates in order to pass the turn from himself to his interlocutor(s).

In this study, we chose dialogues in which one of the speakers is a "reporter", the one who shares a story or a personal problem, while the other is a "responder", the one who listens and comments, hence – unbalanced dialogues.

The contribution of the present study is in the typology of silence types in dialogues with regard to the speakers who speak before and after the silence (reporter, responder or both together; AKA overlaps). We assume that this typology can shed light on broader research on dialogue styles and genres. For example, on the notion of "who owns the gap" – the preceding speaker, who initiated it, or the following speaker, who broke it? For example, in ANALOR tool (Lacheret-Dujour & Victorri, 2002), the silent pause is attached to the preceding speech turn. We further believe that the current study can augment automatic classification of dialogue styles and identifying speakers' role.

## Corpus and method

The examined corpus consists of eight dialogues from two genres of speech. Both genres belong to the spontaneous spectrum: Four private face-to-face dialogues, taken from the Corpus of Spoken Israeli Hebrew (CoSIH) and four therapeutic sessions (for details see Lerner et al. 2016). All participants are fluent Hebrew speakers. All dialogues have the following common feature: in each of them there is one core speaker – the reporter who shares and tells a personal story, and the other participant – the responder, who listens and responds.

The spontaneous dialogues varied in their durations from 7 to 27 minutes. The therapeutic sessions consisted of 22–26.5 minutes.

The segmentation and annotation procedure was carried out manually by a phonetician using PRAAT textgrid tool (Boersma & Weenink, 2015). Each interval was labeled with one of the following four parameters:

*Silence* (non-speech communicative event): Acoustic silence (labeled by a hashtag - #) in the present study is when both speakers are not involved in any vocal production. According to this definition, inhales, exhales, sniffs, sighs, tutting (tsk sound articulation) and coughs were treated as silences; *Reporter*: The reporter's speech intervals (including hesitations and creaky voice); *Responder*: The responder's speech intervals (including hesitations and creaky voice); and *Overlap*: When the two speakers talk simultaneously.

Figure 1 illustrates the labeling scheme at the bottom tier, which is derived from the top and mid tiers, each of which consists of the annotations ascribed to one of the interlocutors in the dialogue.

We then distinguished between three classes of silences: *Pauses* – refer to intra-speaker silences. Pause's duration minimum threshold was set as 100 milliseconds, following Silber-Varod (2013). *Gaps* – refer to inter-speakers* silences (following Edlund, Heldner & Hirschberg (2009) and Heldner & Edlund (2010)), and *Pausal Interruption Silences* (henceforth, PIS) that interact with simultaneous speech (following Bruneau (1973:28–36) and Ephrat (2014:24)). Lapses ("extended silences at transition-relevance places" (Sacks, Schegloff & Jefferson, 1974:715, n. 26) above 10 seconds were omitted from the calculations. This was done since these cases do not reflect pauses and gaps, as argued in Bruneau (1973).
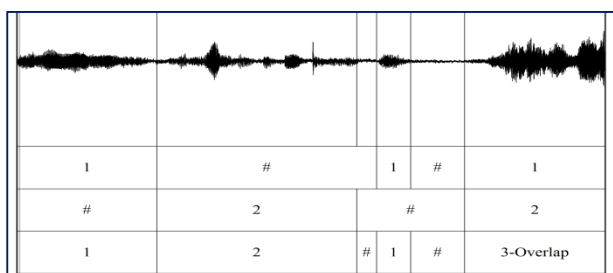


*Figure 1. Annotation scheme via three tiers: Reporter-1 (top tier), responder-2 (middle tier) and the derived annotation (bottom tier).*

## Results

We first measured the relative distribution of the four major interactional parameters in each dialogue (Figure 2). On average, silences compose 35.8% of a dialogue. It is evident that although the therapeutic sessions were taken under specific psychological paradigm, they are varied in terms of the ratios of the relative amount of silence in each dialogue (range from 29% of a session to 47%, similar to the range in the private dialogues).
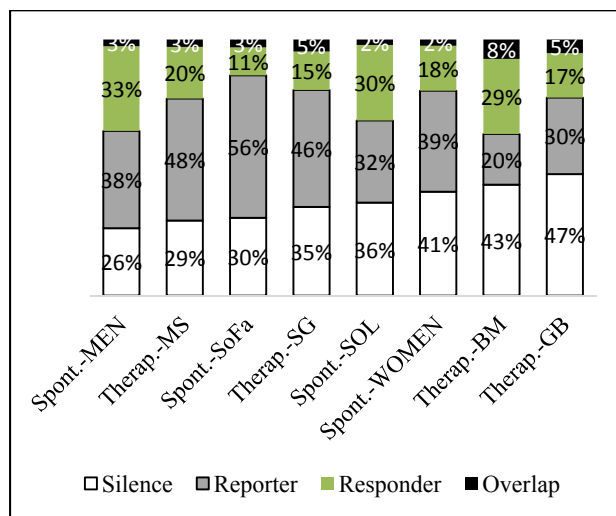


*Figure 2: The durational distribution (%) of the four labels in each dialogue: Silences, Reporter's speech, Responder's speech and overlaps.*

We then divided the pauses (1421 cases) and gaps (720 cases), and measured their duration. Due to the strongly skewed of the distributions to the left, for the duration of both pauses and gaps, the arithmetic values were transformed to geometric ones, as suggested by Campione and Veronis (2002). This suggests that all statistical calculations, that assume normal distributions, should be performed in the logarithmic domain. To this end, the input to all the t-tests is the natural logarithmic (ln) values of the duration rather than the duration (ms) values. Results show that the duration of most of the pauses is between 300 ms to 500 ms (299 cases). Average pause duration is 1114 ms; Most gaps are between 200 ms to 400 ms (164 cases). Average gap duration is 860 ms. The median pause value is 675 ms and the median gap value is 626 ms. In a *t*-test, the difference between pauses and gaps was found significant ($p = 0.012813$.).

We then tagged all the 2,287 silence intervals below 10 seconds according to the preceding and following participant. Reporter was tagged as 1; Responder as 2; an overlap as 3; and silences by the sign #. The number to the left of the hashtag # represents the participant who breaks his speech and the number to the right represents the one who follows the silence: one of the speakers or an overlap. For example, a silence interval "1#2" means that the reporter (1) spoke before the silence and the responder (2) after the silence (means, this is a gap). This process resulted in nine different

tags, as shown in Table 1. Finding shows that reporter's pauses are the most frequent (954 cases), followed by 467 cases of responder's pauses. Both 1#2 and 2#1 gaps have similar frequency (360 cases), while overlaps rarely occur. As to the distribution among the two genres; Pauses were found in similar distribution – 62%–63%, while gaps were found more in private dialogues (34%) than in therapeutic sessions (29%).

In Table 1 we present the average duration of each of the nine silence types, from the shortest "1#1" pause type (with an average of 1.086 s) to the longest "2#3" PIS type (with an average of 2.388 s). Several differences emerge from the average data: The average duration of reporters' pauses (1#1) is the shortest. The average responders' pause (2#2) is above 400 ms longer than the reporters' pause. The average duration of the 1#2 gap is 100 ms shorter than the 2#1 gap. Pausal interruptions (PIS) are the longest type of silences, except for 1#3 (i.e., reporter # overlap) type. All averages are above the 'tolerance interval' length "of approximately one second" (Jefferson, 1989:170), which is suggested to mark the normatively acceptable length of absence of talk in conversational interaction. Standard deviation values indicate the variance of silences in spontaneous speech. The general finding here is that, the silence length is affected by the event (speaker or overlap) prior to the silence more than by the event that follows the silence: the three cases of 1#? are the shortest; the three cases of 2#? are longer; and the three cases of 3#? are even longer, although the gap 2#3 (representing responder's break and both speakers initiate the next turn synchronically) is the longest.

*Table 1. The average (and standard deviation) duration of the nine silence types.*

| Silence (#) types | Tags | Average duration (seconds) and standard deviation |
|---|---|---|
| Pause | 1#1 | 1.086 (sd = 1.417) |
| Gap | 1#2 | 1.202 (sd = 1.684) |
| PIS | 1#3 | 1.261 (sd = 3.929) |
| Gap | 2#1 | 1.368 (sd = 2.464) |
| Pause | 2#2 | 1.523 (sd = 1.915) |
| PIS | 3#2 | 1.579 (sd = 1.792) |
| PIS | 3#1 | 1.720 (sd = 2.117) |
| PIS | 3#3 | 1.795 (sd = 1.930) |
| PIS | 2#3 | 2.388 (sd = 4.710) |

Next, we compared the duration of the pauses between the two speakers, in the two genres. A *t*-test was carried out in order to measure the significance of the difference between intra-reporters' pauses ("1#1") and intra-responder pauses ("2#2"), and it was found that in both genres, the responder's pauses ("2#2") are significantly different from the reporter's ones

("1#1"): $t = 4.261$; $p = 2.2E-05$. For Private: $t = 2.657$; $p = 0.008$. For Therapeutic: $t = 3.027$; $p = 0.002$. Due to the unique setting of the therapeutic sessions, where each intern served once as a client and once as a therapist, we measured pause difference between "1#1" and "2#2" of the same speaker in *two different* sessions. For three speakers (S, G and M) results are statistically significant: for S: $t = 2.128$; $p = 0.034$; for G: $t = 5.478$; $p = 0.00001$; and for M: $t = 2.206$; $p = 0.028$. For speaker B, the difference was found not significant ($t = 0.128$; $p = 0.897$).

As to the gap duration between every turn taking – reporter-responder ("1#2") and responder-reporter ("2#1"), the difference was found not significant in both genres together: $t = 0.679$; $p = 0.497$. The result is not significant at p < 0.05 in Private: $t = 0.75$; $p = 0.451$, and in Therapy: $t = 1.493$; $p = 0.135$).

Last, we measured the differences of the average gaps and pauses durations for the two genres. It was found that the average durations of gaps are different, while pauses are almost identical in the two genres (Figure 3). The difference between the intra-speaker average pause durations in both genres was found statistically not significant for both roles: For 1#1: $t = 1.552$; $p = 0.120$; for 2#2: $t = 0.336$; $p = 0.736$. While the inter-speaker gaps were found significantly different: For "1#2": $t = 3.589$; $p = 0.0003$; for "2#1": $t = 5.652$; $p < 0.0001$.
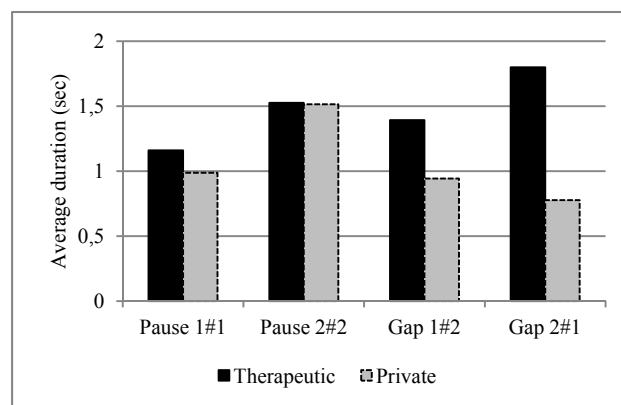


*Figure 3. The average duration of gaps and pauses in the two genres.*

We further tried to examine if the genre affects the speaking rate. We calculated the speech rates per dialogue in two steps. First, we created the graph of speech intervals' onset time and then ran linear regression to evaluate the speaking rate (as a function of speech intervals' onset rate). The speaking rate was then calculated as the slope of the linear regression evaluator (Table 2). Higher values indicate slower speaking rate (i.e., longer time laps between two speech intervals).

*Table 2. Speaking rates of each dialogue (slope of the estimated linear regression).*

| Dialogue | Speaking rate |
|---|---|
| SOL | 1.818 |
| SG | 2.002 |
| MS | 2.095 |
| SOFA | 2.314 |
| MEN | 2.591 |
| WOMEN | 2.628 |
| GB | 2.894 |
| BM | 3.304 |

## Discussion

The present study suggests a basic method to learn about similarities and differences between genres and speaker's role, not only regarding silence types, but also to other parameters of the interaction.

As expected, there are more gaps (indication to turn-takings), in the private dialogues compared to therapeutic session: Gaps distribution ratios were found in private dialogues more than in the therapeutic sessions. On the other hand, pauses distribution ratios were found similar in both genre.

As to the durational distribution, we found the duration of pauses role dependent – reporters' pauses are shorter than responders', in both genres, while the duration of gaps was found genre dependent. These findings suggest that the socio-linguistic aspect of role affects not only speech acoustics (Lerner et al., 2016), but also silence duration. Moreover, speech rate was not found genre dependent. This calls for more inspection on the issue of silence variations, and the linguistics, as well as non-linguistics, parameters that affects silence behavior. As suggested by Jefferson (1989), it seems that although silences in unbalanced spontaneous dialogues vary considerably compared to silences in reading, this does not mean that in spontaneous speech sheer cognitive demands determine the silences. This complex form of silence variability should be further investigated by itself, and by its interface to other prosodic parameters and to speech interval size.

## References

Boersma P. & D. Weenink, 2015. Praat: doing phonetics by computer [Computer program]. Version 5.4.06, retrieved 21 February 2015 from http://www.praat.org/

Bruneau, T. J. 1973. Communicative silences: Forms and functions. *Journal of Communication* 23(1):17–46.

Campione, E. & J. Véronis. 2002. A large-scale multilingual study of silent pause duration. *Proceedings of the first international conference on speech prosody (Speech Prosody 2002)*, 11–13 April 2002, Aix-en-Provence, France, 199–202.

The Corpus of Spoken Israeli Hebrew (CoSIH) http://humanities.tau.ac.il/~cosih

Crown, C. L. & S. Feldstein. 1991. The perception of speech rate from the sound-silence patterns of monologues. *Journal of Psycholinguistic Research* 20(1):47–63.

Edlund, J., M. Heldner & J. Hirschberg. 2009, January. Pause and gap length in face-to-face interaction. *Proceedings of Interspeech 2009, 2009*, 6–10 September 2009, Brighton, UK, 2779–2782.

Ephratt, M. 2014. *When Silence Speaks: Silence as Verbal Means of Expression from a Linguistic Point of View*, Magnes Press. (Hebrew)

Ephratt, M. 2011. Linguistic, paralinguistic and extralinguistic speech and silence. *Journal of pragmatics* 43(9):2286–2307.

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

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

Jefferson, G. 1989. Preliminary notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation. In D. Roger & P. Bull (eds.), *Conversation: An interdisciplinary perspective* (vol. 3), Multilingual matters.

Kurzon, D. 2007. Towards a typology of silence. *Journal of Pragmatics* 39(10):1673–1688.

Lacheret-Dujour, A. & B. Victorri. 2002. La période intonative comme unite d'analyse pour l'étude du français parlé, modélisation prosodiques et enjeux linguistiques. *Verbum* 24(1–2):55–72.

Lerner, A., V. Silber-Varod, F. Batista & H. Moniz. 2016. In search of the role's footprints in client-therapist dialogues. *Proceedings of Speech Prosody 2016 (SP2016)*, 31 May – 3 June 2016, Boston, USA, 400–404.

Ramanarayanan, V., E. Bresch, D. Byrd, L. Goldstein & S. S. Narayanan. 2009. Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation. *The Journal of the Acoustical Society of America* 126(5):EL160–EL165.

Sacks, H., E. A. Schegloff & G. Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 696–735.

Silber-Varod, V. 2013. *The SpeeCHain Perspective: Form and Function of Prosodic Boundary Tones in Spontaneous Spoken Hebrew*, Lambert Academic Publishing (LAP).

Tannen, D. 2000. "Don't Just Sit There-Interrupt!" Pacing and Pausing in Conversational Style. *American Speech* 75(4):393–395.

Wooffitt, R. & N. Holt. 2010. Silence and its organization in the pragmatics of introspection. *Discourse Studies* 12(3):379–406.

Zellner, B. 1994. Pauses and the temporal structure of speech. In E. Keller (ed.) *Fundamentals of speech synthesis and speech recognition*. Chichester: John Wiley, 41–62.
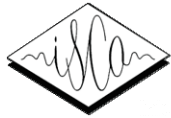
# Proceedings of

# DiSS 2017

## The 8th Workshop on Disfluency in Spontaneous Speech

**KTH Royal Institute of Technology**
**Stockholm, Sweden**
**18–19 August 2017**

## TMH-QPSR
### Volume 58(1)



**Edited by**
**Robert Eklund & Ralph Rose**