Cultural differences of gaps and overlapping speech in political interviews

Vered Silber-Varod¹, Oliver Niebuhr², Loredana Schettino³, Plinio A. Barbosa⁴

¹TAD Center for AI and Data Science, Tel Aviv University, Tel Aviv, Israel
²University of Southern Denmark (SDU), Sonderborg, Denmark
³University of Naples Federico II, Naples, Italy
⁴University of Campinas, Campinas, Brazil

vereds@tauex.tau.ac.il, olhi@sdu.dk, lscbettino@unisa.it, pabarbsa.unicampbr@gmail.com

Abstract

This cross-language study aims to examine both expected and unexpected communication phenomena of silence intervals and overlapping speech events, respectively. Specifically, we investigate these interaction patterns in relation to speakers' self pauses, considering their relative frequency and duration. Our research holds significance in offering cross-linguistic insights, which we accomplish by analyzing the behavior of politicians and TV hosts during political interview panels in four languages: Brazilian Portuguese, German, Hebrew, and Italian. Our findings reveal linguistic (and potentially cultural) differences in the timing of turn taking in the interviews. Notably, Israeli and Italian politicians exhibit quicker responses and comments compared to their counterparts in German and Brazilian Portuguese languages, who demonstrate a tendency for larger gaps before a turn is taken. Moreover, Hebrew and Italian speakers engage in more frequent overlapping speech, interrupting their interlocutors' turns, in contrast to German and Brazilian Portuguese speakers. Building upon these results, we propose silence-based definitions for two conversational styles: "high involvement" languages and "high considerateness" languages (Tannen, 1994), also conceptualized as "word cultures" and "silence cultures" by Agliati et al. (2005).

Index Terms: pauses, gaps, overlapping speech, turn taking, political interviews, Hebrew, Brazilian Portuguese, German, Italian

1. Introduction

Overlapping speech, also known as simultaneous speech or cross-talk [5], occurs when two or more speakers talk at the same time, with their speech partially or fully overlapping. Speakers' overlaps can make it challenging for listeners to follow the conversation and understand each speaker's message, hence "violating" the interaction. However, research has shown that underlyingly even overlaps are systematic and give rise to intense co-attention and orderliness [6]. In a cross-linguistic study by Stivers et al. [7] that explores the timing of turn-taking in ordinary conversations, the authors highlight a common tendency in speaker changes to avoid overlaps on the one hand but minimize silences on the other. However, the study concerns a specific type of change, viz. that after polar questions. Cross-linguistic comparisons of turn timing considering different contexts are still rare. Moreover, probably also due to technical challenges posed by the automatic detection of overlapping speech, quantitative studies on the absence of silence and on overlapping speech did not penetrate traditional discourse analysis research, in general, and the study of communicative dynamics, e.g., in political interviews, in particular. Starting with the ideal speech situation doctrine of the German philosopher Jürgen Habermas [8], which is based on rational consensus between participants and is managed by basic, implied rules, e.g., that good communication involves not only conveying information but also being able to listen and respond effectively to others, our main concern in this paper is to examine the implementation of this ideal situation in real situations, and to investigate whether the comparison of the turn-taking dynamics in four languages (i.e., Hebrew, Italian, German, and Brazilian Portuguese) might reveal differences which could reflect culture-specific behavior. For example, Tannen coined the term "high involvement style" for a conversational Jewish style (specifically New York Jewish speakers) in her works on silences, or the absence thereof, in conversation [9]. One of the aspects of this style is what she called "cooperative overlaps". This terminology differs from that used by researchers who have linked interruptions in conversation with power relations and gaining dominance. Her argument is that what has been the stereotype of New York Jews as pushy and aggressive should rather be interpreted as high involvement as opposed to another style, which she called "high considerateness" that is used, e.g., by California non-Jewish speakers. Building on these ideas, our goal is to ask whether the four languages could be classified as "high considerateness" languages or "high involvement" languages [9]. Not only that, the ideal speech situation presupposes equal participation opportunity. However, it is common to identify political interviews as an arena of explicit and implicit struggles of power relations [10].

This paper aims to compare the durations of between-speakers silences (termed gaps) and the frequency of occurrence overlapping speech in standard varieties of four different languages: Brazilian Portuguese (BP), German, Hebrew, and Italian. Gaps and overlaps are two realizations of breaks in the conversation. The duration of the first and the frequencies of the second can shed light on personal and cultural communication styles [1]. It is our assumption here that differences between languages in the timing of turn-taking reflect cultural patterns.

In turn-taking events, expected silences are also known as Floor Transfer Offset (FTO), or "phonic gaps“, which are defined as the amount of time between the end of one turn and the beginning of the next [2], [3], [4]. FTO can be informative since it can actually take on a negative value (overlapping speech), it can be zero (no gap), or it can take on a positive value (henceforth, a gap).
Political interviews are an intriguing specific context for discourse and conversation analysis, and beyond the influence on public opinion, they are assumed to also have a linguistic impact on the audience. Televised political interviews are designed to produce face-to-face confrontational and challenging interactions between journalists and politicians, while both parties attempt to generate discourse for the “overhearing audience” [11, 12]. The interviewers are responsible for the structure and content of the interview (by posing questions), while the interviewees’ task is to answer those questions while sending their messages to the audience watching the broadcast. Studies showed that politicians are rarely giving direct answers to questions they are asked, which leads to “vagueness, evasiveness, or equivocal communication style…” [11]. Regarding turn taking strategies, it is suggested in the literature that speakers who are under high mental stressed, such as politicians during their campaigns, tend to shorten their response time [3], which means shorter gaps, and use other strategies of turn taking such as hedges and interruptions, irrespective of the relative communicative skills of politicians or interviewers [10].

1.1 Research questions
Following the above research goals, we ask the following questions:
1. Are there differences in the duration of gaps between the languages?
2. Are there differences in the frequency of occurrence and duration of overlaps between the languages?
3. Do languages with short gaps also have relatively more overlaps?
4. Do languages with short gaps also have relatively longer overlaps?

2. Material and Method

In our analysis of gaps and overlaps, we follow Feldner and Edlund’s [3] terminology of silence intervals: Pauses are within-speaker silence intervals; Gaps are between-speakers silence intervals; Overlaps are synchronous speech towards turn taking. Zero-gaps are gaps spanning between 0 to 70 ms (similar to the no-gap no-overlap cases of [3]). The upper threshold of zero-gaps was set in a bottom-up approach, post manual labeling of the authors. When measuring the labeled cases, we found that all brief pauses are under 80 ms, yet few gaps were measured between 70-80 ms. Therefore, we have decided to set the upper threshold of 70 ms for zero-gaps.

2.1 Data

The dataset is composed of 12 TV interviews with 12 well-known (male) politicians from four countries: Brazil, Germany, Israel, and Italy (three politicians per language). Each interview was led by two interviewers, one senior and more dominant and the other a junior and less active. Altogether 29 speakers; six different Brazilian interviewers; four German interviewers; five Israelis; and two Italian interviewers led the respective TV interviews. The interviews are from the years 2019-2021.1

The total duration of our dataset is about eight hours and 45 minutes. Altogether, our dataset consists of 7,609 pauses (within-speaker silences), 1,100 overlapping events, 536 gaps (between-speakers silent breaks), and 139 zero-gaps. Due to limited space, we will not present the zero-gaps analysis.

2.2 Labeling

The labeling process involved assigning specific labels to the roles of the speakers involved in the interviews. The politician was labeled as “1,” the senior interviewer as “2,” and the junior interviewer as “3.” Each period of silence was labeled using two digits, indicating the speaker before and after the silent interval. For instance, “11” represented a pause by the politician. The label “12” indicated a gap between the politician’s turn (“1”) and a subsequent turn by the senior interviewer (“2”), etc.

Additionally, overlaps were labeled to identify instances where one speaker began a turn before another speaker had finished. For example, “1 in 2” denoted that the politician started speaking before the senior interviewer had concluded his/her turn.

During the analysis, we categorized the gaps based on the turn-takers (i.e., the second digit label) and we attributed the overlap to the speaker who interrupted (indicated by the label “1” in the case of “1 in 2”). However, in situations where it was challenging to determine who interrupted whom, such as when two speakers started talking simultaneously, we used a deviating labeling approach that only indicated the involved speakers. For instance, if the three participants overlap, the label would be overlap123.

2.3 Statistical analysis

For statistical analyses, linear mixed models were used with factors referred to below. For paired comparisons, the t-test with the Satterthwaite’s method was used. In all cases the significant level was 5%.

3. Results

We first present the distribution of pauses, gaps, and overlaps in each language. Figure 1 shows that there are relatively more pauses than gaps and overlaps for the speakers of all four involved languages. These findings reflect the characteristics of political interviews, wherein it is expected that politicians provide thorough and detailed responses. Furthermore, it is plausible that politicians have been trained to articulate their ideas at length, aiming to capture attention and establish their authority [13]. Yet, differences between the languages emerge as well; gaps are more frequent in the German data than in the Hebrew, Italian and BP data; overlaps in Hebrew and Italian considerably outnumber those in German and, even more, in BP. The findings also show that most pauses belong to the politicians (who presumably speak more than the interviewers).

In the following we will focus on the politicians’ behavior and look at how fast politicians take their turns after the interviewers finished their own turns (i.e., gap durations), and how many times the politicians were taking turns before the interviewers finished their own turns (i.e., number of overlapping events). As we explained above, we expect to find more overlapping speech in languages with shorter gap durations.

---

1 Detailed information about the corpus can be found here: https://drive.google.com/file/d/1BT-U407oN1WckjXhDnZy3p2uUVB59V9/view?usp=sharing
3.1 Gaps

3.1.1 Gap frequency per speaker (= turn-taker) and language

As expected, gaps in political interviews are quite scarce (less than 3 per minute), probably due to the specific genre where the journalists wish (or were instructed) to let the politician speak as much as possible. As illustrated in Figure 2, there were more gaps in the Israeli interviews compared to the other languages. A chi-square test of independence showed that the variables Language and Turn-taker (the speaker who took the turn – the politician, or one of the interviewers), are significantly interrelated. \( \chi^2(3, N = 2458) = 165.5778, p < 0.0001 \).

3.1.2 Gap duration

To analyze how fast, or slow, the interlocutors take turns, we measured the duration of gaps. Table 1 shows the results of the durations of the two types of gaps where the politician takes the turn after the interviewers (i.e., “21” and “31”). N = 273 and the two types in which the interviewers take the turn after the politician (i.e., “12” and “13”). N = 242. No significant difference was found regarding the durations of the four gap types between politician and an interviewer: “12”, “13”, “21”, and “31”. The factor Language caused significant difference, though. In particular, Brazilian Portuguese is different from Hebrew and Italian but not from German. Table 1 presents the results of a linear mixed model with Speaker as the random factor and Interview and Language as the fixed factors (fit by REML). In this model, BP is compared with the other three languages.

### Table 1: Gap duration differences; Results of the Linear mixed model.

<table>
<thead>
<tr>
<th>Language</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>df</th>
<th>t</th>
<th>p-value</th>
<th>BMI corrected p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>0.017</td>
<td>0.125</td>
<td>21</td>
<td>0.138</td>
<td>0.892</td>
<td>0.891</td>
</tr>
<tr>
<td>Hebrew</td>
<td>-0.547</td>
<td>0.1150</td>
<td>18</td>
<td>-4.845</td>
<td>0.0001</td>
<td>0.000***</td>
</tr>
<tr>
<td>Italian</td>
<td>-0.660</td>
<td>0.135</td>
<td>22</td>
<td>-4.891</td>
<td>0.000***</td>
<td>0.000***</td>
</tr>
</tbody>
</table>

Figure 3 presents the duration of the gaps “21” and “31”, when politicians take the turn after the interviewers. Following [3], [14], and [15], the duration values were transformed to \( \log(10) \) values in order to shift the strongly right-skewed distribution of the duration (ms) towards a normal distribution. Gaps are on average longer in BP and German and shorter in Hebrew and Italian.

3.2 Overlapping speech

3.2.1 Overlap frequency per speaker (= turn-taker) and language

In Figure 4, we present a bubble diagram of the average relative frequency of overlaps per interview (represented by the politicians’ label on the circle). The relative frequencies were calculated as the number of overlapping events divided by the interview’s total duration. The average frequency ranges from 2 (in Ciro’s interview) to 17.62 (in Netanyahu’s interview).
3.2.2 Overlap duration

Compared to the within-speaker pauses, gaps are somewhat shorter, and overlaps are on average longer (Figure 5). Longer durations of overlapping events were found across languages (Figure 6).

![Figure 5: The global average duration of overlaps, pauses, and gaps.]

A linear mixed model with the Speaker as random factor and Interview and Language as fixed factors yielded significant results. In the analyzed data, overlaps (mean: 756 ms) are on average significantly longer than gaps (mean: 439 ms, Est.: 317.3, SE: 70.860, df: 2.01, t: 4.479, p: 0.04).

![Figure 6: A box plot of duration (ms) pauses, gaps, and overlaps per language.]

The overlap results reveal a different language order than for the gaps: BP overlaps are the longest, Italian follows, also with higher SD, then German and, finally, Hebrew overlaps are shortest, differing significantly from all three languages (Est.: -297.1, SE: 43.7, df: 46.4, t: -6.7, p: 1.79e-08).

4. Discussion and Conclusions

In this study, we investigated the frequency and duration of gaps and overlaps, two turn-taking strategies, within the context of political interviews conducted in four different languages. By analyzing these fundamental conversation events quantitatively, we aimed to identify language-specific characteristics that could indicate a preference for a high involvement style or a high considerateness style [9]. Our findings revealed notable differences in the duration of gaps observed when politicians took turns after the interviewers (Question #1). Specifically, Israeli and Italian politicians were observed to be quicker in taking turns compared to their counterparts from the German and BP languages, who exhibited a slower pace in turn-taking. Furthermore, our analysis also identified disparities in the duration of overlapping events (Question #2). On average, the longest overlaps were observed in BP, followed by Italian and German, while Hebrew exhibited the shortest overlap durations.

With respect to the relationship between gaps and overlaps (third and fourth questions in 1.1 above), our study yielded noteworthy findings. We discovered that languages characterized by shorter gap durations tended to exhibit a higher frequency of overlaps (Question #3). Put differently, the duration of positive gaps can serve as a proxy for the occurrence of negative gaps (overlaps). However, our analysis revealed that languages featuring shorter gaps do not necessarily have longer overlaps (Question #4). Interestingly, this trend was reversed in the case of BP, a language distinguished by longer gaps, extended pauses (as also discussed in [16]), and longer instances of overlap. Additionally, our investigation unveiled that the duration of overlaps displays a distinct pattern of language grouping, potentially indicating a divergent pace in the communication process, distinct from the patterns observed in pauses and gaps, and this is seemingly the case of BP, in particular.

Our findings suggest that Hebrew and Italian can be classified as “high involvement” languages [9], characterized by short gaps and a substantial amount of overlapping speech. Conversely, German and BP can be described as “high considerateness” languages, as they exhibit longer gaps and fewer instances of overlaps. These interpretations align with and provide further insights into the observations made by [17], highlighting significant cultural variation in turn-taking, silence, and language usage, distinguishing between “word cultures” and “silence cultures”. The former are characterized by rapid and frequent exchanges of speech, with generally short pauses (both filled and empty) during standard conversations, quick turn-taking, and brief transitional periods. In such cultures, silence may carry negative connotations of indifference, resistance, mistrust, and even hostility in certain contexts. In contrast, “silence cultures” are characterized by longer gaps between turns, indicating thoughtful contemplation and deliberation. Consequently, the pace of conversation tends to be relatively slower, and silence is viewed as a manifestation of trust, confidence, agreement, and harmony.

While our study deviates from previous research that has primarily focused on linguistic cues for identifying turn transition opportunities [18], our findings contribute to explaining stereotypes surrounding the communication cultures of the four languages examined. This spans from the procedural communication culture observed in German to what we interpret as an emotional and, in some cases, argumentative communication culture in Israeli and Italian contexts [19].

5. Acknowledgements

We wish to thank the anonymous reviewers whose suggestions helped us improve our paper. We also want to thank Dr. Aya Vituri from the TAD center for AI and Data Science at Tel Aviv University for her contribution to the statistical analysis. For the 2nd author please see https://oliverniebuhr.com/conflict-of-interest.html.
6. References


