

# Investigations of Conversational Implicatures and Turn-taking

**Mateo Jaramillo**

MSc. Logic

ILLC

Universiteit van Amsterdam

mcjaramillo97@gmail.com

**Paul Stöckle**

MSc. Logic

ILLC

Universiteit van Amsterdam

mail@paulstoeckle.de

## Abstract

With the rise of more advanced corpora which contain time stamp information of turns, NLP studies have been able to focus on the phenomena of turn-taking, e.g. how pause/overlap duration of utterances vary under different circumstances. Conversational implicatures convey responses to questions without explicitly stating a polar answer. Our investigations examine the consequences implicatures have on turn-taking. Our research is motivated by results produced by [Stivers et al. \(2009\)](#) that show a tendency to minimize both overlaps and long pauses. We show that the use of implicatures correlates with a minimization of turn-taking time of the subsequent turn. I.e., we show that pause durations are on average shorter following an implicit response rather than an explicit response, and that implicit responses are in much fewer cases followed by a relatively long pause ( $\geq 0.6$  s) than in comparison to an explicit response. Our results reinforce well establish pause duration results by [Stivers et al. \(2009\)](#) and [Kendrick and Torreira \(2015\)](#).

## 1 Introduction

Recent research gives evidence that interlocutors tend to minimize the time between turn taking within conversations ([Stivers et al., 2009](#)). Despite this observation, there are phenomena concerning the use of implicatures in which it is not obviously clear if interlocutors consequently minimize or elongate their turn taking time.

One such phenomena is the broad use of indirect responses to polar questions using implicatures ([Hockey et al., 1997](#)). Polar questions warrant a simple positive or negative response (*yes/no*), but the use of conversational implicatures as a response typically gives the listener a non-explicit response that provides context, either situational

or cultural, allowing them to deduce the intended meaning ([de Marneffe et al., 2010](#)).

The following is an example of an implicit answer to a polar question. While the expected response to the said question is an affirmative or a negative, the response given is instead a statement which allows the listener to infer a negative reply.

**Polar Question:** Do you want to go for a walk?

**Implicit Answer:** I already went this morning.

One might hypothesize that an implicit response increases the turn taking time of the following turn, arguing that in order to produce and convey a reply to the implicit response, the listener must first deduce its intended meaning which generates an additional effort in comparison to explicit responses and therefore correlates with an additional demand for time. If this is the case we cannot generalize the tendency of minimizing turn taking time as a principle of conversations.

On the other hand, one could argue as well that the use of indirect responses is on a par with such a principle. In most languages, the standard time needed to generate an utterance and begin the muscular action of speaking is 600 ms ([Indefrey, 2011](#)). In spite of this, the average time between turns in conversation ranges from 200-300 ms ([Kendrick and Torreira, 2015](#)). This indicates that interlocutors begin planning their response before the previous turn has ended. Since a direct *yes/no* response to a polar question is of much shorter duration than a common implicit response, one might argue that the longer implicit response provides the listener with more time to plan her answer and therefore prevents some pause within the conversation a direct answer might cause due to a lack of planning time.

The research presented in this paper investigates what effect the use of non-explicit responses has

on turn taking time in order to reveal whether or not one of the above theses can be supported. We investigate data provided through the *Switchboard* corpus (Godfrey et al., 1992). The corpus provides annotations to determine whether an utterance is a polar question as well as whether a response is explicit or implicit. We look at partial conversations each consisting in an initial polar question, a response turn, and a follow-up turn and compare the turn-taking time of the follow-up turn regarding the type of response (explicit/implicit) given to the initial question.

Section 3 describes our methods and procedures for procuring the data. Using the annotations and POS tags of the corpus we are able to identify polar questions, explicit responses, and determine the pause durations. Table 3 provides examples of these tags. Section 4 presents our initial results, where the counts of each type of response can be visualized as well as the connections between implicit responses and follow-up pause durations. Taking these results, we investigate our initial competing hypotheses regarding the effects of conversational implicatures. Section 5 provides analysis of the results as well as comparing them to existing results from Stivers et al. (2009) and Kendrick and Torreira (2015).

## 2 Related Works

There are universal conversational tendencies, despite differing languages. Stivers et al. (2009) shows that conversational participants avoid both overlapping utterances and pauses between utterances. Similarly, ten Bosch et al. (2005) presents results that show a form of pause-duration matching between interlocutors. While this result tends towards research regarding communication accommodation and alignment, they also conclude that the character of the dialogue determines pause-durations. We use this assumption from basic conversation to determine whether there is a significant pause duration or overlap when dealing with polar questions and implicit responses.

George and Mamidi (2019) have identified nine different types of conversational implicature. These range from scalar responses to cultural idioms as responses, and there is a potential difference in the way interlocutors respond upon hearing these types of implicatures. Due to the nature of the *Switchboard* corpus, which transcribes phone conversations between interlocutors on a more formal level

than familiar, the presence of implicatures such as Pope Questions <sup>1</sup>, tautologies <sup>2</sup>, and idioms <sup>3</sup> are rare. We expect the majority of implied responses to instead be scalar, generalised, particularised, or relevance responses. These types of implicatures use relevant contextual clues and typically provide either a more generalized response or particularised response (eg. Question: Did you call the landlord? Answer: I called her assistant. Meaning ‘No’) Further analysis of these types of implicatures and their meaning can be found in de Marneffe et al. (2010).

Kendrick and Torreira (2015) investigates the connection between positive and negative responses with pause duration. They report that when answering a question with a negative response, the responder takes more time which in turn allows the listener to interpret the polarity of the response before the utterance has even been made. Thus, the listener is able to predict the response and begin formulating a follow-up before the previous turn has ended. *Switchboard* corpus also contains annotations for both the explicit and implicit responses, tagging them as either positive or negative which allows us to determine if there is a connection between the polarity of the response and the pause duration preceding both the response and the follow-up.

## 3 Data and Methods

**Switchboard Corpus** The *Switchboard* corpus contains the transcripts of 642 time-annotated two person conversations conducted over the telephone. Each transcript consists of an indexed list of utterances, where each utterance is annotated with a POS tag. The tags the corpus provides for annotation and which were particularly needed for our investigations are illustrated in Table 2. Furthermore, each utterance is indexed by the turn to which it corresponds, where a turn is a contiguous series of utterances uttered by one speaker. Each turn is time-annotated, i.e. each turn is annotated with the start-time (the time at which its first token starts) and the end-time (the time at which its last token ends). This turn-based time annotation is

<sup>1</sup>eg. Question: Are you hungry?  
Answer: Is the pope Catholic? (Meaning ‘Yes’)

<sup>2</sup>eg. Question: Is New York pizza better than Chicago pizza?  
Answer: Pizza is pizza. (Meaning equal preference)

<sup>3</sup>eg. Question: Aren’t you sad?  
Answer: It is useless to cry over spilt milk. (Meaning ‘No’)

Response Type	Polar Question	Response	Follow-up
Explicit	You don't have any children? Have you ever been to New Mexico? Is this your first baby?	No. Uh, no actually. Uh-huh.	But you're married? Oh. Have you crocheted or knitted any baby clothes or baby blankets?
Implicit	Have you ever served on a jury before? Isn't she fabulous? You do it yourself?	Uh, I have twice. Oh, she is wonderful. I do it myself.	Oh, okay. Oh, once you st-# Yeah.

Table 1: Example triples from the *Switchboard* corpus. (# represents a speech cut-off)

Speech act	Tag	Response
Polar Question	qy	-
Yes answer	nn	explicit
No answer	ny	explicit
Affirm. non-yes answer	na,ny^e	implicit
Neg. non-no answer	ng,nn^e	implicit

Table 2: Annotations made in the dataset that allowed for efficient identification of polar questions as well as implicit and explicit responses.

the only time annotation the corpus provides, so it does not provide time information of particular utterances in the case a turn consists of more than one utterance.

**Data processing** In order to investigate the effects of the usage of implicit and explicit responses on the turn-taking time<sup>4</sup> of the following turn, we only considered those question-response-follow-up turn-triples where the question, response and follow-up utterance appeared within a sequence. The reason for this restriction is that we want to exclude cases in which there exists cross-talk in between utterances which might distort the measuring of the particular turn-duration and turn-taking-time and additionally makes manual checking of the triples more difficult. In this manner we isolated those triples of turns where a first turn ends with an utterance which is annotated as a polar question, the second turn consists in only one utterance which is annotated as a response, and a third turn which must not be further specified. In order to identify explicit and implicit responses we followed the transcription of the corpus and counted every utterance which was tagged as a yes/no answer as

<sup>4</sup>The turn-taking time describes the pause duration between the end of one turn and the start of a second turn and might be negative if there is an overlap between both terms.

explicit and every utterance which was tagged as an affirmative non-yes answer or a negative non-no answer as implicit, as illustrated in Table 2. Following this method we could isolate 335 triples with an explicit response in the middle turn and 135 triples with an implicit response in the middle turn. Even though, the corpus comes with the mentioned speech act annotations, we manually checked that the isolated triples describe reasonable parts of conversation. Examples triples pulled directly from the corpus which we use to analyze can be seen in Table 1.

Given those triples, we calculated the pause duration between the response and the follow-up utterance by simply subtracting the end-time of the response from the start-time of the follow-up turn. Since the middle turn consists only in the response utterance, the response duration could be calculated easily by subtracting the start-time of the corresponding follow-up from its end-time.

## 4 Results

Given the simple nature of an explicit polar response (*yes/no*), explicit responses had an average length of 1.5 tokens while the implicit responses had an average length of 6.6 tokens. Explicit responses typically consisted of single token responses such as ‘yes’ and ‘no’, or ‘huh-uh’ (no) and ‘uh-huh’ (yes). Slightly longer utterances also appeared such as ‘yeah, actually’ which count as explicit responses. Similarly to the number of tokens per utterance, the time taken to utter the response of explicit responses is significantly shorter than the implicit counter-part.

This result is to be expected as it is natural to suppose that an explicit response to a polar question could usually be uttered with one word, while an implicit response is by nature a short phrase or

a full sentence. Correspondingly, it is natural to suppose that with an increasing number of words within an utterance the average duration of the utterance increases as well.

The average pause duration between the response utterance and the follow-up utterance is 260 *ms*. This result fits well with the average turn-taking time of 200-300 *ms* predicted by Kendrick and Torreira (2015). Comparing the turn-taking time following an explicit response with that following an implicit response, we get an average duration of 313 *ms* in the explicit case compared to 132 *ms* in the implicit case. This is illustrated in Table 3, as well. Figure 1 compares both types of responses as well as plotting them with respect to their actual utterance duration and the pause duration that precedes it. Similar to results produced by Stivers et al. (2009), for both the explicit response follow-ups and the implicit response follow-ups, the majority of pause durations are clustered between -0.1 and 0.3. Results can be seen in the histograms of Figure 2 and Figure 3 for follow-ups to both explicit and implicit responses, respectively. Furthermore we could obtain that in the case of an explicit response the duration of the subsequent pause is in 19.4% of the isolated triples longer or equal than 0.6 *s*, while in the case of an implicit response the pause is only in 5.9% of the triples longer than or equal to 0.6 *s*. Supposing a pause with a duration that lasts for at least 0.6 *s* as a long pause, there are relatively significantly fewer long pauses following an implicit response.

Out of the 335 explicit responses, 278 of them were affirmative answers. The implicit counterparts follow a similar structure in which out of the 135 implicit triples we identified 105 positive responses and 30 negative response.

## 5 Analysis

While we initially had two seemingly competing hypotheses that proposed that either the use of implicatures would cause the follow-up to occur later as it would require more thought or that the use of implicatures would allow more time resulting in a shorter pause duration. As stated in Section 4, our results mimic results produced by Stivers et al. (2009) which show that interlocutors minimize overlaps as well long pauses (exceeding 0.6 *s*). The

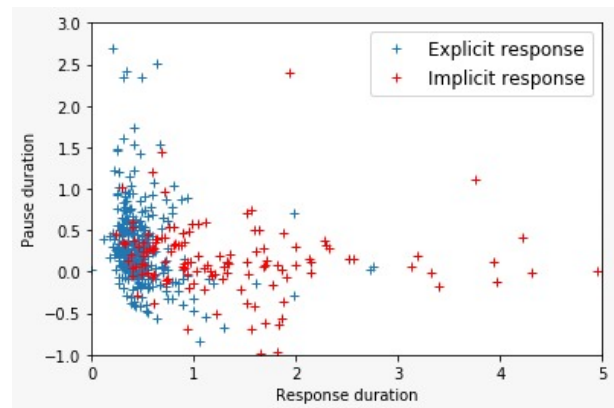


Figure 1: Scatter gram: Plots both implicit and explicit responses to their preceding pause duration and the time taken to utter said response.

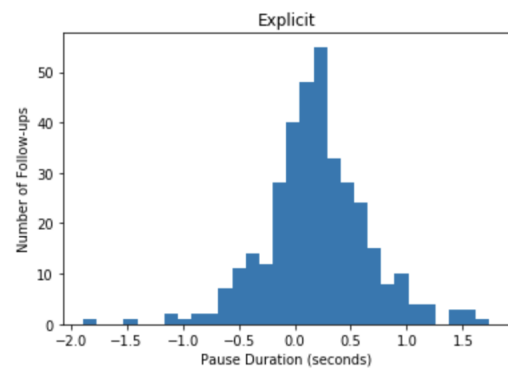


Figure 2: Histogram: Frequency of pause duration lengths between explicit responses and follow-up utterances.

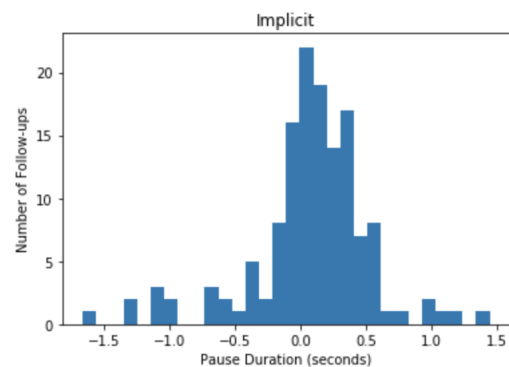


Figure 3: Histogram: Frequency of pause duration lengths between implicit responses and follow-up utterances.



Turn Type	Amount	Response Duration (Std)	Pause Duration (Std)
<b>All</b>	470	812 <i>ms</i> (1024)	260 <i>ms</i> (498)
<b>Explicit all</b>	335	480 <i>ms</i> (294)	313 <i>ms</i> (501)
<b>Explicit aff.</b>	278	459 <i>ms</i> (256)	295 <i>ms</i> (446)
<b>Explicit neg.</b>	57	587 <i>ms</i> (416)	401 <i>ms</i> (703)
<b>Implicit all</b>	135	1636 <i>ms</i> (1578)	132 <i>ms</i> (467)
<b>Implicit aff.</b>	105	1547 <i>ms</i> (1633)	140 <i>ms</i> (491)
<b>Implicit neg.</b>	30	1947 <i>ms</i> (1321)	103 <i>ms</i> (370)

Table 3: Results of Explicit response triples compared to Implicit response triples. ‘aff’ responses are marked as positive or affirmative replies and ‘neg’ responses are marked as negative replies. Column 1 is the average follow-up time (st. dev). Column 2 is average response length of answers.

results visualized in Figure 2 and Figure 3 show there is a slight skew toward positive pause durations. For implicit responses, 65% of the follow-up pause durations were positive, likewise for explicit responses, 71% were positive. The use of interruptions has been studied to represent two actions in a conversation, to begin a new turn or to cut-off an existing turn (Lestary et al., 2018). Interruptions are typically a sign of lively conversation, as interlocutors want to chime in and contribute to what is being said while establishing conversation membership (Eecke and Fernández, 2016). However, the sociolinguistic aspects of interruptions and conversations are beyond the scope of this paper. The important factor to note is the tendency away from overlaps.

Table 3 shows a difference of over 100 *ms* between explicit affirmative follow-up pause durations and explicit negative follow-up pause durations. Yet, this result is inverted for the implicit counterpart as implicit affirmative follow-up pauses are longer on average than implicit negative follow-up pauses by 37 *ms*. Regardless of the positive and negative distinction, implicit responses are preceded by shorter follow-up pauses compared to the explicit counterparts. This data supports our hypothesis that implicit responses offer the listener more time to formulate their follow-up.

Our results clearly reveal that the use of implicit responses as responses to polar questions results in a decrease in turn-taking time in comparison to the use of explicit responses. Even though the difference in the average pause duration between the implicit and the explicit case is only 181 *ms*, the all in all amount of pauses in the implicit case is less than in the explicit and the amount of large pauses (duration  $\geq 0.6$  s) is much higher in the explicit case than in the implicit (19.4% vs. 5.9%).

## 6 Conclusion

Through our investigations into the use of implicatures in the *Switchboard* corpus, we are able to present data that supports the hypothesis that implicatures and subsequently longer responses to polar questions present the listener with more time to internally generate a follow-up. Our research was initially motivated by results regarding pause durations between interlocutor turns that showed a tendency for minimization of these pauses. We had two initial contradicting hypotheses stating that the use of implicatures would result in either longer pause durations as they required the use of deduction to understand or shorter pause durations because implicit responses are longer in duration and would allow the listener more time to formulate a response during the previous turn. Our results presented are in favour of the latter hypothesis.

Further research could investigate the connection between implicatures and information comprehension. As our results signify quicker follow-up times, they do not make the distinction as to whether these follow-ups are confirmations (such as ‘okay.’) or clarification questions (such as ‘Really?’). Likewise, given an available corpus, it would be beneficial to investigate conversations which take place face-to-face. The social interruption tendency and ability to perceive the other interlocutors might produce interesting results with respect to turn-taking.

## Acknowledgments

Our initial project began as an investigation into the use of hesitations and filled pauses within the the *Spoken British National Corpus* Love et al. (2017). Upon recommendation from Raquel Fernández we moved to the *Switchboard* corpus which made our investigations much easier, so we thank her for that.

Paul Stöckle worked on the majority of the code

related to obtaining the data as well as providing insights for the overall direction of the project. Mateo Jaramillo worked on supplementary code and calculations as well as writing many of the sections for the report. Visit <https://github.com/P451M/CDM-Project.git> for all related code.

Tanya Stivers, Nicholas J Enfield, Penelope Brown, Christina Englert, Makoto Hayashi, Trine Heine-mann, Gertie Hoymann, Federico Rossano, Jan Peter De Ruiter, Kyung-Eun Yoon, et al. 2009. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26):10587–10592.

## References

- Louis ten Bosch, Nelleke Oostdijk, and Lou Boves. 2005. On temporal aspects of turn taking in conversational dialogues. *Speech Communication*, 47(1-2):80–86.
- Paul Van Eecke and Raquel Fernández. 2016. On the influence of gender on interruptions in multiparty dialogue. In *INTERSPEECH*.
- Elizabeth George and Radhika Mamidi. 2019. Conversational implicatures in english dialogue: Annotated dataset.
- J. J. Godfrey, E. C. Holliman, and J. McDaniel. 1992. Switchboard: telephone speech corpus for research and development. In *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 517–520 vol.1.
- Beth Ann Hockey, Deborah Rossen-Knill, Beverly Spejewski, Matthew Stone, and Stephen Isard. 1997. Can you predict responses to yes/no questions? yes, no, and stuff. In *Fifth european conference on speech communication and technology*.
- Peter Indefrey. 2011. [The spatial and temporal signatures of word production components: A critical update](#). *Frontiers in Psychology*, 2:255.
- Kobin H. Kendrick and Francisco Torreira. 2015. The timing and construction of preference: A quantitative study. *Discourse Processes*, 52(4):255–289.
- Agustina Lestary, Ninuk Krismanti, and Yulieda Hermaniar. 2018. Interruptions and silences in conversations: A turn-taking analysis. *PAROLE: Journal of Linguistics and Education*, 7(2).
- Robbie Love, Claire Dembry, Andrew Hardie, Vaclav Brezina, and Tony McEnery. 2017. The spoken bnc2014: Designing and building a spoken corpus of everyday language. *International Journal of Corpus Linguistics*, 22(3):319–344.
- Marie-Catherine de Marneffe, Christopher Manning, and Christopher Potts. 2010. "Was it good? It was provocative" Learning the meaning of scalar adjectives. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 167–176.