Annotation of Transition-Relevance Places and Interruptions for the Description of Turn-Taking in Conversations in French Media Content

Rémi Uro[1,2,3], **Marie Tahon**[3], **Jane Wottawa**[3],
David Doukhan[1], **Albert Rilliard**[2,4], **Antoine Laurent**[3]
[1]Institut National de l'Audiovisuel, [2]Université Paris-Saclay, CNRS, LISN,
[3]LIUM, Le Mans Université, [4]Universidade Federal do Rio de Janeiro.
{ruro,ddoukhan}@ina.fr, albert.rilliard@lisn.fr
{marie.tahon,jane.wottawa,antoine.laurent}@univ-lemans.fr

Abstract

Few speech resources describe interruption phenomena, especially for TV and media content. The description of these phenomena may vary across authors: it thus leaves room for improved annotation protocols. We present an annotation of Transition-Relevance Places (TRP) and Floor-Taking event types on an existing French TV and Radio broadcast corpus to facilitate studies of interruptions and turn-taking. Each speaker change is annotated with the presence or absence of a TRP, and a classification of the next-speaker floor-taking as *Smooth*, *Backchannel* or different types of turn violations (cooperative or competitive, successful or attempted interruption). An inter-rater agreement analysis shows such annotations' moderate to substantial reliability. The inter-annotator agreement for TRP annotation reaches κ=0.75, κ=0.56 for Backchannel and κ=0.5 for the Interruption/non-interruption distinction. More precise differences linked to cooperative or competitive behaviors lead to lower agreements. These results underline the importance of low-level features like TRP to derive a classification of turn changes that would be less subject to interpretation. The analysis of the presence of overlapping speech highlights the existence of interruptions without overlaps and smooth transitions with overlaps. These annotations are available at https://lium.univ-lemans.fr/corpus-allies/.

Keywords: Spoken interaction, Media, TV, Radio, Transition-Relevance Places, Turn Taking, Interruption, Backchannel, Overlapping Speech, Simultaneous Speech

1. Introduction

Interruptions constitute an important aspect of social interactions. While interruption-related incivilities are often evoked in public debates (Bennet, 2015), quantifying these phenomena in media is an essential but complex issue. For example, two studies of the same French 2007 Presidential debate (Sandré, 2009; Constantin de Chanay and Kerbrat-Orecchioni, 2010), that focused on interruptions between the second turn candidates reached opposite conclusions on who committed most turn-tacking violations. While both works were manually produced by experts, large differences may be found in the end: this shows the complexity of the interruption phenomena and the need for other approaches of quantifiable robustness. To study speaker changes in French TV and Radio conversations, we present a set of annotations of an existing French speech corpus (Larcher Anthony et al., 2021) in terms of Transition-Relevance Places (TRPs) (Sacks et al., 1974) and types of transitions between turns. One aim is to determine whether human TRP annotation is more reliable than the direct annotation of various interruption types and whether detecting TRPs automatically can help analyze interruptions in general.

TRPs are defined by Sacks et al. (1974) as the end of a Turn-Constructional Unit –i.e., a place where the speaker may change. Building upon this concept, Levinson (1983) proposes that taking the floor outside a TRP constitutes a *violation* and may be considered an interruption. Wells and Macfarlane (1998) state that TRPs must be identifiable by all participants in a conversation to allow for a smooth transition.

Overlapping speech correlates with interruptions (Makri-Tsilipakou, 1994), but also commonly occurs in anticipated turn-taking (Heldner and Edlund, 2010; Adda-Decker et al., 2008; Gravano and Hirschberg, 2012). Levinson and Torreira (2015) show that a floor-taking action needs to be planned to minimize floor transfer offsets: this requires the next speaker to anticipate TRPs. There shall thus be cues to the future end of a speech turn. Grosjean (1983) shows that human listeners can predict a sentence's remaining number of words. Swerts and Geluykens (1993); Swerts (1997); Moneglia and Raso (2014); Gambi et al. (2015), among others, showed there are cues to the terminality of an utterance; the exact nature of those cues –lexical, prosodic– are debated.

Several corpora propose annotations of turn-taking events. Gravano and Hirschberg (2011) offer an annotation scheme used in Gravano

LREC-COLING 2024, pages 1225–1232
20-25 May, 2024.

and Hirschberg (2012) and Brusco and Gravano (2023); this scheme is based on the notion of Inter-Pausal Units (IPU) and uses six transition types on task-oriented dialogues: *Smooth Switch*, *Overlap*, *Pause Interruption*, *Simple Interruption*, *Butting-in*, *Hold*. Hara et al. (2019) present an annotation scheme of TRPs on simulated human-robot dialogues. In ten Bosch et al. (2004), the authors classify utterances as continuation, interruption, or turn change. Adda-Decker et al. (2008) focuses on overlapping speech segments from French political interviews to annotate them as *interruptions*, *backchannels*, *anticipated turn-taking*, and *complementary*.

We are not aware of an available corpus of French broadcast media that focuses on interruptions and not only on overlapped speech. Lebourdais (2023) reports it is difficult to annotate different categories of interruptions. We thus focus here on terminality and its relationship to the type of floor-taking events. This paper presents a dataset of annotations for TRPs and types of floor-taking events made on French TV and Radio content.

2. Method

2.1. Data

2.1.1. Allies clean

The ALLIES corpus (Tahon et al., 2024) is a French meta-corpus whose publication is planned in 2024 and designed to gather and extend previous French data collected for diarization and transcription evaluation campaigns. It consists of 328 hours of audio extracted from 1998 to 2020 in 1048 shows. In terms of duration, the proportion of overlaps fluctuates widely between broadcast news and debates. ALLIES-clean is a subset of ALLIES, composed of 23 shows at the time of the annotation presented in this paper. The description of each show is reported in Appendix A. The dataset presented here is part of a collective effort to produce richly annotated speech resources to help develop better automatic processing.

2.1.2. Selection of speech samples to be annotated

Samples of speaker change with and without overlapping speech were selected. Each sample comprises two or three intervals, depending on whether overlapping speech was included. Samples without overlapping speech comprise one segment[1] before and after speaker changes. For overlapping speech, we only kept turns shorter than 2 seconds and with only two speakers to capture backchannels, apparent interruptions, and anticipated starts. Longer overlapped speech samples were excluded to maximize speaker changes and avoid segments with cumbersome structures and unintelligible speech (Gay, 2023). The samples with overlapped speech are composed of 3 intervals: the part of the segment preceding the overlap, the overlap, and the segment following the overlap.

A total of 2041 samples were extracted, among which 1064 contain an overlapped speech interval. The samples last from 0.48s to 31.27s, with a mean duration of 5s. The overlapping speech intervals range from 0.1s to 1.99s, with a mean duration of 0.65s and 80% lasting less than 1s.

2.1.3. Speaker categories

Table 1 presents the distribution of speakers found in the selected speech samples. The gender and role of the 128 unique speakers, including only 38 women, were manually annotated using categories inspired from Doukhan et al. (2020) to help the description of turn-taking dynamics. The categories of Anchors and Invited Journalists were separated to differentiate the ones in charge of managing the conversation from those invited to give their opinions. The Experts, Politicians, and Celebrities categories are individuals used to appearing in the media, whereas Witnesses (whose profession may be unknown) are not. Speech Turns from 2 Witnesses and 6 Celebrities not speaking in French are followed by the corresponding French version uttered by a Translator. The most represented roles are male Anchor Journalists, followed by male Witnesses and Politicians.

Role	**Women**	**Men**
Anchor Journalist (AJ)	12	24
Invited Journalist (IJ)	1	7
Witness (W)	11	18
Politician (P)	4	15
Celebrity (C)	2	14
Expert (E)	2	12
Translator (T)	2	4
Total	38	90

Table 1: Distribution of the retained role categories observed in the corpus, by gender.

2.2. Annotation

A paid linguistics student annotated the speech samples. The estimated annotation time for each sample was around 30 seconds. The annotator was instructed to identify two phenomena: the TRPs and the types of floor-taking events. The

[1] defined as "sequences containing complete words which are syntactically and semantically coherent" in the Allies annotation guidelines(Larcher Anthony et al., 2021)

types of floor-taking events were inspired by existing classifications of interruptions (Yang, 2001; Li, 2001; Makri-Tsilipakou, 1994). The annotator had to determine for each spoken interval if it ended with a TRP or not and to qualify each floor-taking event as either **Smooth**, *Backchannel* (abbr. **Back**), *Competitive Interruption* (abbr. **Compl**), *Cooperative interruption* (abbr. **Coopl**), *Competitive interruption attempt* (abbr. **ComplA**) or *Cooperative interruption attempt* (abbr. **CooplA**). An interruption attempt happens when the next speaker fails to take the floor from the initial speaker. The Cooperative / Competitive distinction was inspired from Yang (2001): Cooperative violations are defined as a violative floor-taking event aiming to help the current speaker or asking for or giving clarification, whereas a Competitive violation happens when the interrupter tries to steal the turn. A *Backchannel* is defined by Traverso (2005) as a short intervention that does not contribute to the thematic development of the conversation.

If a span was deemed irrelevant (e.g., no speech, no floor-taking event, translation) or could not be annotated (i.e., a sample artificially cut before the end of a turn or without speech), the field would be left empty. The annotator was provided with a Praat (Boersma and Weenink, 2023) interface, with the mandatory fields pre-filled with all possible labels. The first (resp. second) group of three tiers corresponds to the speaker's name, the TRP annotation, and the Floor-taking event type of the first (resp. second) speaker; the seventh tier is for possible comments of the annotator. The complete annotation guidelines are reported in Appendix B, and a screen capture of the prefilled Praat interface can be seen in Appendix C.

3. Results

3.1. Annotations distribution

Of the 3070 annotations of TRP, TRPs are present in 1151 intervals including 23% with overlapped speech, and absent in 1919 intervals including 39% with overlapped speech. Table 2 presents the distributions of 1898 floor-taking annotations obtained from 1991 speech samples. Annotation statistics are detailed according to their realization context (within/without overlap). 50 irrelevant speech samples (e.g., ALLIES-clean annotation errors) were discarded from the TRPs and the turn-taking annotation process.

Let us note that 25% of the transitions annotated as smooth occur within overlapping speech intervals, which may be related to turn-end anticipation as described in e.g., Grosjean (1983); Bögels and Levinson (2017). Likewise, 19% of labeled violations were realized without being associated with overlapping speech.

Label	Number	Within overlap
Back	524	80%
Smooth	929	25%
Compl	265	79%
Coopl	37	78%
CooplA	21	86%
ComplA	122	89%
Total	1898	1015

Table 2: Number of floor-taking labels and percentage within overlapping speech

3.2. Interaction pattern

Table 3 lists the five most frequent annotation patterns out of the 161 distinct annotation sequences in the corpus. They correspond to smooth transitions, backchannels, anticipated turn-taking, and successful and aborted competitive interruptions. While these five patterns represent 69% of the annotations, each of the remaining patterns' frequencies represent less than 2%.

Sequence	%
[S1: TRP] [S2: Smooth]	32%
[S1: nTRP] [S1: nTRP | S2: Back]	17%
[S1: nTRP] [S1: TRP | S2: Smooth]	8%
[S1: nTRP] [S1: nTRP | S2: Compl]	7%
[S1: nTRP] [S1: nTRP | S2: ComplA]	5%

Table 3: Top 5 sequences of labels forming the most common interaction patterns, and their proportion in the corpus

3.3. Speakers

Journalists are the category that interrupts the most. Of the 437 turns labeled as violative (i.e., something other than Smooth or Back), 301 are produced by journalists. Interestingly, one-third of those journalists are not presenters (Anchor) but are invited.

The two main interruptive patterns –non-terminal turns with an interruption or interruption attempt– mostly present interruptions by anchors, between invited journalists, or between experts. It is also interesting to note that while witnesses are rarely interrupted by journalists (6 in total), journalists seem to let themselves be interrupted by witnesses (1 interruption attempt and 16 actual interruptions). This is primarily true for shows where journalists interview people unaccustomed to speaking to the media. Gender does not seem relevant in turn-taking dynamics in this corpus. About 25% of male-male, male-female, female-male dyads are interruptions, and 30% of female-female ones. However, there are only 23 female-

female dyads vs. 1588 male-male dyads; the percentage difference is likely irrelevant because of the gender imbalance.

3.4. Inter-rater agreement

338 samples were selected as described for multi-rater annotation to evaluate the quality and reproducibility of the proposed annotation scheme. This subset was annotated by two of the authors: one speech sciences computer scientist and one phonetician. Examples to be annotated by multiple raters were selected to encompass all possible label sequences as well as maximize the number of examples from the most represented sequences in the main annotation to obtain a precise agreement for those sequences. All three raters agreed on 74% of examples for TRP classification. When looking only at the five most represented categories, comprising 195 annotated intervals (i.e., 70% of the intervals), the agreement is 88%, showing that the main categories are the most robust ones.

A substantial inter-annotator agreement was observed for first speaker TRP annotations (Fleiss' Kappa=0.75 (Fleiss, 1971)), showing the reliability of our proposed annotation scheme. Table 4 shows the agreements for each possible label for the type of floor-taking event and their grouping in the following meta-labels: **Inter**: *CoopI*+*CompI*; **Coop**: *CoopI*+*CoopIA*; **Comp**: *CompI*+*CompIA*; **Attempt**: *CoopIA*+*CompIA*; **Violation**: All of the above. It indicates that Backchannels, Smooth transitions, and Violations –i.e., taking the floor when it's not expected– (especially competitive ones, which are more represented than cooperative ones) are reliably annotated. However, the type of violation seems more complicated to classify. Those findings are consistent with results reported in Lebourdais (2023) and Adda-Decker et al. (2008).

	overlap		no overlap		all	
	κ	n	κ	n	κ	n
Back	.56	87	.57	16	.60	103
Smooth	.42	65	.66	84	.62	149
CoopI	.13	34	.33	10	.18	44
CompI	.48	43	.51	18	.49	61
CoopIA	.10	30	.08	8	.11	38
CompIA	.09	50	.34	10	.16	60
Inter	.51	63	.54	25	.50	88
Coop	.22	55	.21	18	.23	73
Comp	.36	81	.61	23	.44	104
Attempt	.16	70	.22	18	.19	88
Violation	.41	110	.66	32	.51	142

Table 4: Fleiss' Kappa values for each label

Table 5 shows the confusion between labels. The *Main* column states the number of phenomena annotated with each label by the main annotator, and the other columns the corresponding phenomena annotated by the two other annotators. If there is little confusion between Smooth or Backchannel vs. others labels, we also observe that apart from Competitive Interruptions, the violative labels are more prone to confusion.

4. Discussion

We proposed an annotation of TRPs and floor-taking events classification on French media. Out of the 2041 labeled samples, the most commonly observed interaction patterns correspond to smooth speaker changes, backchannels, competitive interruptions -successful or not- and anticipated starts. An analysis of the inter-rater agreement shows that the manual annotation of TRPs is reliable. If classifying different types of interruptions is more variable across annotators, the general concept of interruption itself seems to be identifiable. An in-depth analysis of randomly selected examples of the five main patterns was conducted to determine possible reasons for annotation disagreements. It highlights that when all annotators agree, the samples are textbook examples of TRP respect or violation. Disagreements seem to be related to the attention paid to non-verbal characteristics –underlining the importance of prosody for TRPs–, and the complexity of some samples that escape explicit criteria.

For instance, cooperative and competitive interruptions may be used depending on the perception of the listener and the interpretation of the speakers' communicative aim. Although in our dataset, gender doesn't seem to play a role in turn-taking dynamics –the primary factor being the role– it is important to remember that this corpus has a highly imbalanced gender distribution. The effect of gender is thus to be considered with caution.

One important takeaway of this analysis is that about a quarter of smooth transitions occur with overlapped speech, while a fifth of the violative turn-taking occurs without overlap. This added to the fact that TRP annotations are more reliable than turn-taking events, hinting that analyzing TRPs may be a good proxy to study interruptions, even without overlapping speech. The annotations of this corpus are available at https://lium.univ-lemans.fr/corpus-allies/.

5. Acknowledgements

This work has been partially funded by the French National Research Agency (project Gender Equality Monitor - ANR-19-CE38-0012).

The authors would also like to thank the main annotator for this study.

	Main	Secondary annotators (percentage)						
	Number	Backchannel	Smooth	CoopI	CompI	CoopIA	CompIA	Not relevant
Backchannel	80	**69.4**	8.1	0.6	1.9	6.9	10.00	3.1
Smooth	126	6.8	**72.6**	4.0	3.2	4.8	4.8	4.0
CoopI	13	11.5	15.4	**34.6**	7.7	11.5	19.23	0.00
CompI	39	5.1	16.7	15.4	**51.3**	1.3	10.3	0.00
CoopIA	9	16.7	16.7	22.2	16.7	**22.2**	5.56	0.00
CompIA	23	15.2	17.4	13.0	17.4	10.9	**26.1**	0.00
Not relevant	48	14.6	15.6	12.5	4.2	3.1	5.2	**44.8**

Table 5: Confusion table between the main and the secondary annotators

6. Bibliographical References

Martine Adda-Decker, Claude Barras, Gilles Adda, Patrick Paroubek, Philippe Boula de Mareuil, and Benoit Habert. 2008. Annotation and analysis of overlapping speech in political interviews. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*.

Jessica Bennet. 2015. How not to be 'manterrupted' in meetings. *Time*.

Paul Boersma and David Weenink. 2023. Praat: doing phonetics by computer (version 6.3.10).

Pablo Brusco and Agustín Gravano. 2023. Automatic offline annotation of turn-taking transitions in task-oriented dialogue. *Computer Speech Language*, 78:101462.

Sara Bögels and Stephen C. Levinson. 2017. The brain behind the response: Insights into turn-taking in conversation from neuroimaging. *Research on Language and Social Interaction*, 50(1):71–89.

Hugues Constantin de Chanay and Catherine Kerbrat-Orecchioni. 2010. Les interruptions dans les débats médiatiques□: une stratégie interactionnelle. *Pratiques. Linguistique, littérature, didactique*, (147–148147–148):83–104.

David Doukhan, Cécile Méadel, and Coulomb-Gully Marlène. 2020. En période de coronavirus, la parole d'autorité dans l'info télé reste largement masculine. *La revue des médias*.

Joseph L. Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382.

Chiara Gambi, Torsten Jachmann, and Maria Staudte. 2015. The role of prosody and gaze in turn-end anticipation. *Proceedings of the 37th Annual Conference of the Cognitive Science Society*.

Déborah Gay. 2023. Radio libre (skyrock) ou la libre antenne comme dispositif d'encadrement d'une sexualité hétéronormative. *Réseaux*, N° 237(1):37–65.

Agustín Gravano and Julia Hirschberg. 2011. Turn-taking cues in task-oriented dialogue. *Computer Speech Language*, 25(3):601–634.

Agustín Gravano and Julia Hirschberg. 2012. A corpus-based study of interruptions in spoken dialogue. In *Interspeech 2012*, page 855–858. ISCA.

François Grosjean. 1983. How long is the sentence? prediction and prosody in the on-line processing of language. *Linguistics*, 21(3).

Kohei Hara, Koji Inoue, Katsuya Takanashi, and Tatsuya Kawahara. 2019. Turn-taking prediction based on detection of transition relevance place. In *INTERSPEECH*, pages 4170–4174.

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

Martin Lebourdais. 2023. *Interactions entre locuteurs : de la détection de la parole superposée à la détection des interruptions*. Ph.D. thesis, Le Mans Université.

Stephen C. Levinson. 1983. *Pragmatics*. Cambridge [England]; New York: Cambridge University Press.

Stephen C. Levinson and Francisco Torreira. 2015. Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, 6:731.

Han Z. Li. 2001. Cooperative and intrusive interruptions in inter- and intracultural dyadic discourse. *Journal of Language and Social Psychology*, 20(3):259–284.

Marianthi Makri-Tsilipakou. 1994. Interruption revisited: Affiliative vs. disaffiliative intervention. *Journal of Pragmatics - J PRAGMATICS*, 21:401–426.

Massimo Moneglia and Tommaso Raso. 2014. *Appendix: Notes on the Language into Act Theory*, volume 61, page 468–495. John Benjamins Publishing Company, Amsterdam.

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

Marion Sandré. 2009. Analyse d'un dysfonctionnement interactionnel – l'interruption – dans le débat de l'entre-deux-tours de l'élection présidentielle de 2007. *Mots. Les langages du politique*, (8989):69–81.

Marc Swerts. 1997. Prosodic features at discourse boundaries of different strength. *The Journal of the Acoustical Society of America*, 101:514–21.

Marc Swerts and Ronald Geluykens. 1993. The prosody of information units in spontaneous monologue. *Phonetica*, 50(3):189–196.

Marie Tahon, Anthony Larcher, Martin Lebourdais, Fethi Bougares, Anna Silnova, and Pablo Gimeno. 2024. Allies: a speech corpus for segmentation, speaker diarization speech recognition and speaker change detection. In *LREC*.

Louis ten Bosch, Nelleke Oostdijk, and Jan Peter de Ruiter. 2004. Turn-taking in social talk dialogues: temporal, formal, and functional aspects. In *Proc. 9th Conference on Speech and Computer (SPECOM 2004)*, pages 454–461.

Véronique Traverso. 2005. *L'analyse des conversations*. Armand Colin.

Bill Wells and Sarah Macfarlane. 1998. Prosody as an interactional resource: Turn-projection and overlap. *Language and Speech*, 41(3–4):265–294.

Li-chiung Yang. 2001. Visualizing spoken discourse: prosodic form and discourse functions of interruptions. In *Proceedings of the Second SIGdial Workshop on Discourse and Dialogue -*, volume 16, page 1–10, Aalborg, Denmark. Association for Computational Linguistics.

7. Language Resource References

Larcher Anthony, Mehrish Ambuj, Tahon Marie, Meignier Sylvain and Carrive Jean, Doukhan David, Galibert Olivier and Evans Nicholas. 2021. *Speaker Embedding For Diarization Of Broadcast Data In The ALLIES Challenge*.

A. Description of shows

Table 6 shows the description of each show with the number of separate broadcasts, the number of extracted samples, the total duration of those samples, the number of unique speakers and the type of show.

B. Annotation guidelines

For each of the following, if it is not possible to annotate a phenomenon (e.g. no speech), leave empty or fill in "NA".

The TextGrid files are prefilled with tiers 1 to 3 corresponding to Speaker 1, tiers 4 to 6 to Speaker 2 and tier 7 is to be used for comments:

- Tiers 1 and 4: prefilled with the speakers name, do not change them
- Tiers 2 and 5: fill with terminality annotation for the relevant speaker
- Tiers 3 and 6: fill with floor-taking event type for the relevant speaker
- Tier 7: use for comments

Terminality: Annotate with "Term" if the segment ends with a TRP, "NonTerm" if it doesn't. Do not annotate the last segment, as it may have been artificially cut.

Turn-taking event type: Annotate with the following categories:

- Smooth: Smooth
- Back: Backchannel
- CompI: Competitive interruption
- CoopI: Cooperative interruption
- CompIA: Competitive interruption attempt
- CoopIA: Cooperative interruption attempt

Do not annotate the first segment, as it lacks previous context.

Definitions:

- Smooth: Speaker takes the floor when expected, at a TRP
- Backchannel: Short verbal regulator signifying attention; may or may not be during an overlapping speech segment
- Interruption: Speaker to take the floor outside a TRP
- Competitive: Speaker wants to take the floor for themself

Show	Channel	Nb Broad.	Samples	Dur.	Nb Spk.	Type
BFMStory	BFMTV	2	204	1550	15	News + itw
CaVousRegarde	LCP	4	282	4860	24	News + itw
CultureEtVous	BFMTV	2	8	25	11	Culture magazine
DEBATE	France Inter	2	147	567	17	Casual itw
EntreLesLignes	LCP	3	312	2173	8	News + debate
LaPlaceDuVillage	TV8	2	803	2569	14	Casual itw
PileEtFace	LCP	2	153	1069	6	Debate
PlaneteShowbiz	BFMTV	2	12	47	13	Culture magazine
TopQuestions	LCP	2	8	44	7	Parliament questions
fm	France Inter	2	112	501	37	News + itw

Table 6: Description of each show with the broadcasting channel, the number of separate broadcasts, number of extracted samples, total duration of the samples in second, number of speakers and type of show

- Cooperative: Speaker wants to help, add or ask for clarifications
- Attempt: Speaker does not keep the floor after interrupting the other

Special cases:

- Music, noise: specify in the comments field
- Segmentation error: flag as Invalid if not deemed usable, else specify in the comments field
- A segment contains only laughs, breathing, silences: leave empty and specify in the comments fields
- Speaker change is not an interaction (e.g. speech playback, translation): annotate TRPs but not floor-taking event type; specify in the comments field

C. Prefilled Praat interface

Figure 1 show a screen capture of a prefilled Praat Textgrid opened in Praat, for a sample with an overlapped speech interval.

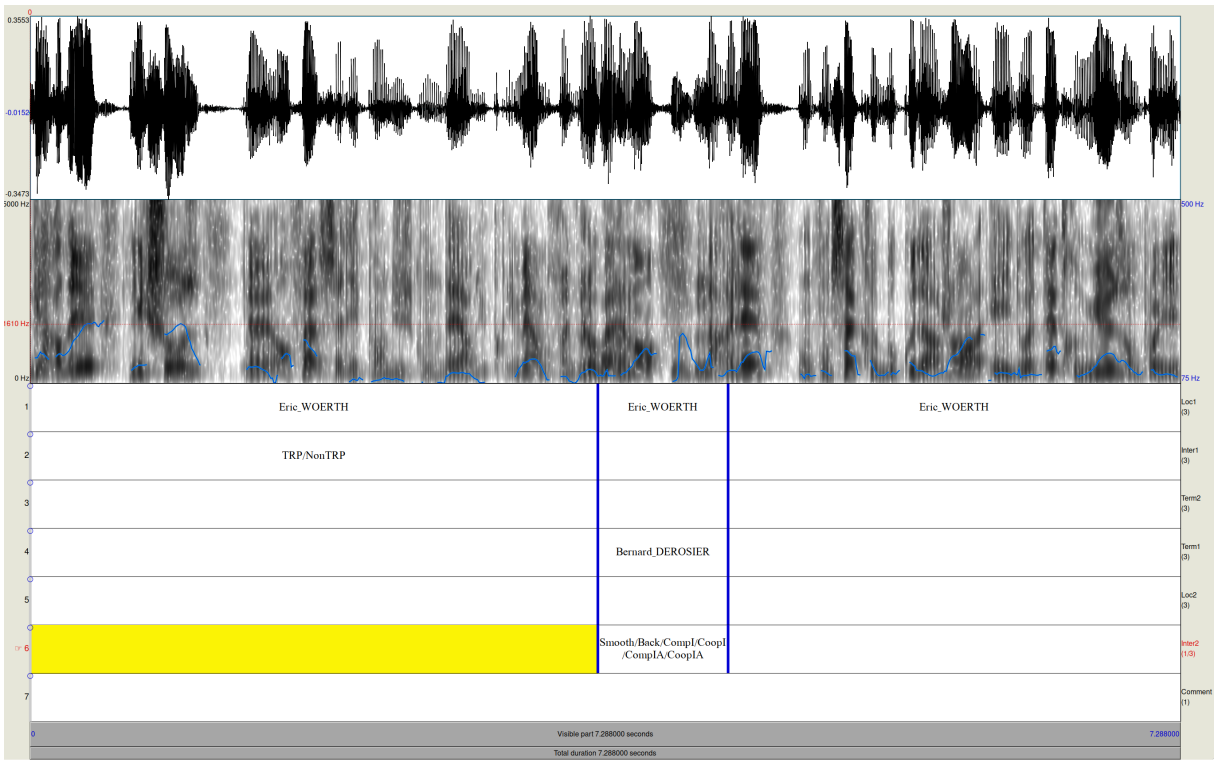

Figure 1: Screenshot of a pre-filled Praat TextGrid, showing a sample with S1 speaking in all three intervals and an overlapped utterance of S2 in the second interval only.