

Mapping Meetings Project: Group Interaction Labeling Guide

Rebecca Bates Elizabeth Willingham Chad Kuyper Patrick Menning

August 24, 2005: DRAFT

Contents

1	Introduction	2
2	Quick Reference Information	2
2.1	Terminology	2
2.2	Meeting Act Tagset	3
3	Labeling of Meetings	4
3.1	Basic Format of MAs and Labels	4
3.2	Segmentation of Interaction Regions	4
4	Meeting Act Tag Descriptions	7
4.1	Group 1: Administration and Planning	7
4.2	Group 2: Decision Making	8
4.2.1	Negotiation	8
4.2.2	Voting	9
4.3	Group 3: Types of Discussion	9
4.3.1	Reporting	9
4.3.2	Brainstorming	9
4.3.3	Other Discussion	15
4.4	Group 4: Commentary and Breaks	15
4.4.1	Commentary	16
4.4.2	Breaks	16
4.4.3	Outside Topic	17
4.5	Group 5: Humor	17
4.6	Diacritics	17
5	Label File Formats	19
5.1	Header Information	23
5.2	Tags and Label Order	23
5.3	Comment Columns	23
5.4	Meeting Notes	23
A	Labeled Meeting Samples	25
A.1	Meeting Act Label Example	25
A.2	Meeting Act Label Example with DAs and Text	27
B	Dialog Act Labels	49
C	Index of Tags	49

1 Introduction

The Meeting Corpus is a collection of over 90 recorded and transcribed technical meetings [4]. The meetings were regularly scheduled meetings of various groups meeting for different purposes. The conversations are natural in that the meetings would have happened even if they were not being used for data collection. Seventy-five of the meetings are transcribed with dialog act labels describing sentence-like activity [6]. This manual describes the labeling of the Meeting Corpus for group interaction using meeting acts, descriptors of which high-level meeting function is occurring. Interaction segments are at least one utterance long and could be as long as an entire meeting.

Meeting acts are intermediate labels between dialog acts and topic information. Many of the meetings are segmented by topic information, and these topic labels could be used as a required boundary for group interaction. Multi-level labeling for topics has been used at various sites. We adapt this idea to allow MA labels to indicate interaction at up to two levels (with an additional level available to indicate humor).

The interaction segments used here are related to the meeting actions described in [3], actions such as “monologue,” “presentation,” “white-board,” or “consensus” are used to script recorded meetings. An important difference is that these labels focus on verbal interaction and do not report physical activity (e.g., whiteboard). Similarly, labeling of “hot spots” in meetings that describe locations of agreement/disagreement, floor grabbing, or topic shift [5, 7] is related to the style functions labeled here.

This manual describes both the tagset used to label group interaction in meetings and how to label the meetings. There are complete descriptions of each tag along with examples from the Meeting Corpus. The guide is modeled after the *Meeting Recorder Project: Dialog Act Labeling Guide* [2].

There are three appendices in this guide. The first is a complete sample of a labeled meeting. The second is a list of the dialog act labels and basic descriptions for labelers who have no experience with the DA labeling system. The final appendix is an index of the tags used in the meetings.

2 Quick Reference Information

2.1 Terminology

The terminology presented in [2] is re-presented here, since the underlying work is very similar. We also present new terminology used for this work.

- utterance: A segment of speech occupying one line in the transcript by a single speaker which is prosodically and/or syntactically significant within the conversational context. In DA labeled data, utterances are indicated with start and stop times. has
- speech: A group of successive utterances or successive portions of an utterance.
- turn: The period during which a speaker has the floor.
- interaction segment: A region of speech that is identified by the type of group action and/or function in the context of the overall meeting.
- dialog act (DA): The tag or sequence of tags pertaining to the function of an utterance or portion of an utterance. Each DA contains at least one tag and may also contain one or more diacritics, depending upon the nature of the utterance. (Further details can be found in [2] and Appendix B.
- meeting act (MA): The tag or sequence of tags pertaining to the function of a group interaction segment or portion of a segment.
- tag: The individual component(s) of a DA or MA label. The sets of tags for DAs and MAs are distinct.

- **label:** The entire set of DAs and/or other tags applicable to an utterance or the entire set of DAs, MAs, and/or other tags applicable to an interaction segment.
- **level 1:** The level for the primary MA tag which represents the basic form of an interaction segment (e.g., reporting, negotiation, planning, etc.)
- **level 2:** A nested level for secondary MA tags which represents secondary form(s) of an interaction segment. The tag set is the same as that used for level-1 M-labels. Each MA contains at least one level-1 tag and may contain one level-2 tag and/or diacritics, depending upon the nature of the utterance.
- **diacritic:** The tag which represents a function or characteristic of an interaction segment and is appended to the primary tag (e.g., long-term, short-term, agreement, disagreement, visual aid used, etc.)

2.2 Meeting Act Tagset

This section summarizes the tags used for labeling meeting acts. Detailed descriptions of the tags can be found in Section 4. Tags shared by level-1 and level-2 labels are introduced, followed by the diacritics that can be used to modify tags.

Group 1: Administration and Planning

- **G** Agenda work
- **P** Planning of future work

Group 2: Decision Making

- **N** Negotiation
- **E** Voting

Group 3: Types of Discussion

- **R** Reporting
- **Z** Brainstorming
- **T** Other discussion

Group 4: Commentary and Breaks

- **C** Commentary
- **BE** Technical or equipment testing breaks-
- **BO** Breaks: interrupts, filler talk, off-topic self-talk, etc.
- **OT** Outside topic

Group 5: Humor

- **H** Joking, sarcasm, use of humor

Diacritics:

- **A** agreeing
- **D** disagreeing

- **FA** ends in agreement
- **FD** ends in disagreement
- **FT** ends in tabling (no decision made)
- **I** consists primarily of an individual speaker
- **M** consists primarily of multiple speakers
- **L** long-term (typically for planning)
- **ST** short-term (typically for planning)
- **S** suggest alternate
- **V** use of visual aids
- **W** extremely formal in style
- **Y** summary or recap of previous knowledge

3 Labeling of Meetings

3.1 Basic Format of MAs and Labels

The basic format for a label is

<level-1 tag> [^diacritic] [<level-2 tag> [^diacritic]].

A single level-1 tag is the only required tag for an MA. The level-1 tag can be modified by one or more diacritics, and/or it can have a level-2 tag attached to it. The level-2 tag can also be modified by one or more diacritic(s).

While labels can be generated in two passes (with the second using the audio tracks), the first release of labeled data comes from the transcripts alone. This first pass requires reading the transcripts and their associated dialog act labels. A sample of the text seen by labelers is shown in Figure 1. While the first pass results in useful labels, we have a plan for a second pass of labeling that will incorporate audio information. Since the labels from text alone are likely to bias the results using audio, an audio pass may need to be completely independent. A dependent second pass would be performed using the first pass of labels and the audio for the meetings. Tag boundaries can be shifted or labels changed based on new information from changes in the audio signal. The primary reasons to listen to the audio would be to confirm diacritics and to adjust segment boundaries. An additional issue that is likely to be clarified when using audio information is the labeling of humor. Since sarcasm can be hard to distinguish through text alone, an audio pass will allow labelers the chance to do more accurate humor labeling.

ASCII text files are generated for each meeting. Each new MA tag is placed on a line, preceded by the start and end times. The start and end times correspond to utterance times in the dialog act transcriptions. A sample of a file generated by labelers is shown in Figure 2. Future examples of tags in this document will be a merging of the word- and MA-level transcriptions.

3.2 Segmentation of Interaction Regions

The segmentation of meetings into interaction regions is not always clear-cut. Transitional speech activity at interaction segment boundaries, which may be analogous to disfluent words at the beginning of a sentence unit, could belong to either region if participants are clarifying consensus on the previous topic or if they are negotiating agreement about which topic to discuss next. Utterance boundaries define the candidate segment boundaries; i.e., a new MA cannot occur in the middle of a DA. Segmentation is currently an

2057.68	2059.01	Bdb001-c3	s and they'll just be done sequentially .
2059.68	2060.04	Bdb001-c1	s^bk o_k .
2060.25	2060.45	Bdb001-c3	fh so ==
2062.08	2062.52	Bdb001-cB	s.%-- so i guess ==
2062.52	2062.76	Bdb001-cB	s^bk yeah .
2062.76	2066.54	Bdb001-cB	s^co if - if you and don can - if you can show him the p_file stuff and see .
2065.99	2066.25	Bdb001-c3	s^aa sure .
2066.54	2068.2	Bdb001-cB	s^df so this would be like for the f_zero .
2066.81	2070.2	Bdb001-c3	s i mean if you do man p_file or apropos p_file you'll see a lot .
2067.05	2067.47	Bdb001-c2	b true .
2069.19	2070.87	Bdb001-c2	s^ng i've used the p_file i think .
2070.87	2076.03	Bdb001-c2	s^bd i've looked at it at least briefly i think when we were doing s- - something .
2074.17	2075.67	Bdb001-c1	qw what does the p_ stand for anyway ?
2075.65	2076.45	Bdb001-c3	s^no i have no idea .
2078.26	2078.54	Bdb001-c2	s^t1 oh in there .
2080.04	2081.43	Bdb001-c3	s^df i didn't de- - i didn't develop it .
2081.43	2083.12	Bdb001-c3	s^df you know it was - i think it was dave johnson .
2084.02	2086	Bdb001-c3	s^df so it's all part of the quicknet library .
2086	2087.28	Bdb001-c3	s^e it has all the utilities for it .
2088.36	2090.4	Bdb001-c1	s^ar s^nd no p_files were around way before quicknet .
2090.87	2091.24	Bdb001-c3	qy^bh^rt oh were they ?
2090.95	2094.67	Bdb001-c1	s^df p_files were - were around when - w- - with um rap .
2094.85	2095.37	Bdb001-c4	s^aa mm-hmm .
2095.41	2095.81	Bdb001-c1	qy^d^g^rt right ?
2095.54	2096.99	Bdb001-cB	s^j it's like the history of icsi .
2096.56	2097.64	Bdb001-c1	s^bu you worked with p_files .
2097.15	2097.24	Bdb001-cB	% like ==
2097.58	2097.98	Bdb001-c3	b mm-hmm .
2097.67	2098.12	Bdb001-c4	s^ar no .
2098.17	2099.05	Bdb001-c1	s i worked with p_files .
2098.55	2098.79	Bdb001-cB	qy^bh^rt yeah ?
2101.06	2102.32	Bdb001-c4	s^no i don't remember what the p_ is though .
2103.66	2104.05	Bdb001-c1	s^ar no .
2105.12	2106.19	Bdb001-c3	s.%-- but there are ni- - they're ==

Figure 1: Sample of DA-labeled text used by MA labelers with utterance start and stop times, speaker and meeting ID labels, DA labels, and utterance text.

START	STOP	LABEL	COMMENTS
623.923	721.119	T	
647.871	668.128	T T	really strong topic break
721.119	842.702	Z	?
842.566	998.624	N^FT	overlap. (poss top. break)
999.352	1047.88	R^M	
1048.49	1134.22	N^FT	
1132.99	1210.29	R	overlap
1209.54	1236.98	B0	overlap
1240.74	1372.18	T	
1372.64	1641.71	T^V	
1400.37	1407	T^V R	
1473.23	1518.33	T^V R	
1641.04	1649.48	N^FA	
1649.81	1805.54	T	
1805.41	1823.59	P	overlap
1824.13	1858.79	P^S	
1858.79	1870.7	C	
1870.75	1907.68	T	
1908.02	1954.52	N^FT	
1956.05	2028.94	T	
2028.94	2030.9	G	
2032.61	2069.26	P^L	
2070.1	2182.13	T	
2183.64	2200.42	C	
2201.81	2250.85	T	

Figure 2: Sample of an MA-labeled meeting with interaction segment start and stop times and MA labels.

1874.63	1875.98	Bmr015-c1	s.%-- i mean i think if you ==
1876.15	1878.51	Bmr015-c1	qw does i- - does it - why would it correspond to pitch ?
1879.58	1879.82	Bmr015-c3	z yeah .
1880.13	1881.3	Bmr015-c3	z i - i - i not sure .
1881.37	1882.62	Bmr015-c3	z i - i - i try to - to ==
1882.77	1883.06	Bmr015-c1	b yeah .
1883.49	1888.51	Bmr015-c3	z when first i - i calculate eh using the f.f.t ==
1887.23	1888.13	Bmr015-c8	s <huh> di- - digital camera .
1888.94	1890.01	Bmr015-c3	z the - the ==
1890.62	1891.22	Bmr015-c2	s^t1 keep forgetting .
1891.49	1896	Bmr015-c3	z i get the - the spectrum ==
1897.04	1897.36	Bmr015-c1	b yeah .
1897.86	1898.78	Bmr015-c3	z and i represent ==
1900.8	1901.63	Bmr015-c3	z all the frequency .
1902.12	1902.39	Bmr015-c1	b yeah .
1902.2	1902.96	Bmr015-c3	z and - when ou- ==
1902.96	1904.88	Bmr015-c3	z i obtained the instantaneous frequency .
1904.97	1907.35	Bmr015-c3	z and i change the - the - the ==
1909.54	1910.25	Bmr015-c3	z using the ==
1911.74	1913.38	Bmr015-c3	z instantaneous frequency here .

Figure 3: Sample of intertwined meeting acts from BMR015, in which the speakers c8 and c2, who rarely speak in the rest of the meeting, are having a brief external conversation (**BO**). Only these two utterances belong in that MA, while all of the surrounding utterances belong in another (**R**).

active part of the discussion of “how to label.” Any additional passes will focus on the area of segmentation. Given this issue, all test data will be labeled by at least two labelers so that “truth” can be represented by multiple labelers. The issue of MA segmentation is related to topic segmentation. Given multi-level topic labels (generated at CSLI at Stanford University [1]), there is already an interesting interplay between MAs and topic labels. The initial plan is that high-level topic labels will indicate a change in MA, even if the MA style does not change, since a high-level topic change is a substantive change in interaction.

Additionally, the opportunity for overlapping MAs exists with meetings more than with two-person dialogs. Multiple topic threads can be intertwined, with secondary threads being entirely surrounded by primary threads. One example is shown in Figure 3. As topics are changing, the start and stop points may overlap as well, similar to overlapping sentences or dialog acts.

4 Meeting Act Tag Descriptions

There are five major groups of MA tags: administration and planning, decision making, discussion, humor, and breaks and commentary. This section provides a detailed description of each MA tag. Diacritics are described at the end of the section.

4.1 Group 1: Administration and Planning

Primary Labels:

- **G** Agenda work
- **P** Planning of future work

Table 1: An excerpt of transcripts from BED009 with the agenda planning label (**G**) and from BMR002 with the planning label (**P**).

G -Start	G -Stop	Spkr	Utterances
221.373	228.101	c5	so we're gonna try to finish by five so people who want to can go hear nancy chang's talk uh - downstairs .
227.218	227.678	c0	hmm .
229.032	233.562	c5	and you guys are g- - giving talks on tomorrow and wednesday lunch times .
233.037	233.367	c0	yes .
233.389	233.629	c3	mmm .
233.562	234.262	c5	right ? that's great .
P -Start	P -Stop	Spkr	Utterances
535.379	537.089	c1	anyway \$—\$ so why did i call you all here ?
535.394	535.754	c2	@reject@ .
536.962	537.412	c3	yes .
537.474	540.554	c1	um - \$—\$ what i'd like to talk about is um - the transcription .
537.938	538.388	c3	oh you did .
540.554	559.1	c1	so as part of this um - project we need to transcribe these meetings . um - to do training and test and so on . and uh - so we need word transcripts and uh - speaker change identification . so speaker i.d . and uh - so i'd sort of like to know who's going to do that .
			...
	568.473		[End of segment]

In meetings, there are often multiple levels of planning. Here, we label regions of **agenda** building or discussion, thus marking a very local level of planning. We exclude from this label single utterances whose purpose is simply to move along the agenda rather than mapping out the agenda. These utterances are often indicative of topic shift and may be indicative of an MA change but do not constitute a **G** label. An example might be the phrase from BRM025 “As a somewhat segue into the next topic...” We also note **planning** of future work or activities. Planning can be modified with diacritics that show agreement or disagreement as well as short- and long-term planning. Short-term planning may involve what will happen in the next week or specific activities that will be accomplished by attendees. Long-term planning typically involves larger directions and might not involve discussion of specific activities. In many cases, it will not be clear from the meeting being annotated whether plans are short- or long-term. In these cases, the label will simply be **P**. See Table 1 for examples of **G** and **P** labels.

4.2 Group 2: Decision Making

Negotiation and voting are the two labels used to indicate decision-making actions. Some portions of negotiation will have aspects of discussion, but the goal of reaching a decision places negotiation in this group.

4.2.1 Negotiation

Primary Label: **N**

The label **negotiation** is used when multiple talkers are working to reach a decision. As level-2 tags, negotiation periods may or may not be connected to planning or discussion regions. An example of this

label as a level-2 tag would be when speakers are negotiating which talker next gets the floor, or negotiating the best way to approach a research problem.

Negotiation regions have a required diacritic labeling the final consensus action, with available diacritics for agreement, disagreement, or a tabling of the negotiation. A tabled negotiation is used to indicate that no final decision was made, whether or not the issue was formally tabled. Negotiation differs from other discussion labels (such as discussion or brainstorming) in that there is an expected final action. See Table 2 for examples.

Required diacritics (labeling final consensus action):

- **^FA** ends in agreement
- **^FD** ends in disagreement
- **^FT** ends in tabling (no decision made)

4.2.2 Voting

Primary Label: **E**

This label is available for decision making that requires some sort of vote or accounting of assessment. This would be separate from the final points of negotiation unless a tally is taken or responses comes from every participant. It can be used as a level-2 tag to indicate this when coupled with negotiation. An example of this tag used as a level-2 tag is in Table 3.

4.3 Group 3: Types of Discussion

There are multiple types of discussion that can happen in a meeting. Here, they are divided into three subgroups: reporting, brainstorming, and other discussion. While negotiation has aspects of discussion, the goal of negotiation is to reach a decision, so it is included in Group 2. While most types of discussion will have one of the first two labels, we include the third so that there is a label that can be used when neither of the others fit exactly.

4.3.1 Reporting

Primary Label: **R**

The label **reporting** is used when either primarily one talker or a team focussed on disseminating the same topic is giving information to the others in the meeting. Clarification questions may be asked, but the meeting remains in this phase. If the type of interaction moves towards discussion but does not change topic or the person who is leading the meeting (the reporter), then the segment does not change. When there is discussion in direct response to the reporting, a level-2 tag for discussion (see below) should be used to note the change. When the reporter changes, a new reporting label should be assigned. See Tables 4 and 5 for examples.

4.3.2 Brainstorming

Primary Label: **Z**

The MA **brainstorming** describes meeting activity that generates ideas or possibilities. Brainstorming may describe the decision of what should happen, but could also be a level-2 label for how or when things might happen – for example, in the case of agenda planning. Negotiating about the generated ideas would be a new MA. Within a negotiation segment, a level-2 tag of brainstorming could be used when a series of alternate suggestions are made in response to negotiation. Diacritics indicating the generation of ideas by a single individual or multiple speakers should accompany the brainstorming MA. See Table 6 for examples.

Table 2: An excerpt of transcriptions from BNS003 with the level-1 label of negotiation (N^{FT}). The labels in the table are the level-2 labels. The final consensus action is to table the discussion, so the region has the required diacritic of FT . There are multiple level-2 tags that occur throughout this excerpt. The level-1 tag does not change.

N^{D}^{FT} -Start	N^{D}^{FT} -Stop	Spkr	Utterances
1268.82	1269.57	c4	everybody agree ?
1270.39	1271.71	c1	um i don't know .
1270.57	1270.87	c2	o_k .
N^{S}^{FT} -Start	N^{S}^{FT} -Stop	Spkr	Utterances
1272.13	1291.7	c1	maybe it could be a good thing if we have such an event button that um somewhere beneath this button or wherever that there's a date when the last update was . so that you know it's worth to click this button to see there's a new um announcement or not .
1293.32	1293.85	c0	yeah o_k .
1293.36	1300.28	c1	s- - so otherwise you are often clicking this button . and you get your old information or even an - uh empty page .
1300.92	1304.1	c0	yeah but this will be then down on the button of the page anyways .
1304.44	1307.68	c4	yeah maybe we can have updated on you know the button itself .
1304.89	1305.03	c1	so ==
N^{A}^{FT} -Start	N^{A}^{FT} -Stop	Spkr	Utterances
1307.17	1308.39	c1	yeah . something like that .
1308.18	1312.51	c4	and then updated and the current date when - when something happened .
1310.15	1311.73	c1	last updated . yeah .
N^{D}^{FT} -Start	N^{D}^{FT} -Stop	Spkr	Utterances
1313.17	1314.32	c4	i don't know how we deal with that .
1313.31	1313.45	c0	yep .
1314.8	1315.68	c4	is it another thing we have to ?
1315.01	1315.23	c0	yep .
1316.69	1321.66	c1	so it's just a su- - suggestion . i don't know how easy it is to - to implement something like that .
1322.58	1322.84	c4	yeah .
1322.6	1323.22	c0	huh . o_k .
N^{FT} -Start	N^{FT} -Stop	Spkr	Utterances
1323.73	1324.29	c4	o_k you ==

Table 3: An excerpt of transcripts from BTR002 with the voting label (**E** in the form $\mathbf{T}^{\wedge}\mathbf{V}|\mathbf{E}^{\wedge}\mathbf{V}$).

E-Start	E-Stop	Spkr	Utterances
1569.68	1574.66	c7	is it - is it the - so if we had a show of hands how many people would vote for the sun keyboard versus the ?==
1572.02	1572.28	c3	o.k .
1574.66	1580.82	c7	so let's put it the other way . how many people would not vote for the sun keyboard ? would vote for the other one . the p_c - p_c keyboard .
1579.36	1579.81	c6	which is the ==
1580.08	1580.73	c5	the p_c ?
1580.36	1581.09	c3	the p_c one ?
1581	1581.98	c5	i would vote for the p_c .
1581.16	1582.55	c6	the p_c one is the amaretto one . right ?
1582.85	1584.54	c7	i don't know . i don- - i'm not sure . it's the ==
1584.1	1584.54	c5	yeah .
1584.45	1584.63	c9	huh .
1584.7	1586.42	c3	i- - i'm not sure which one is which .
1584.87	1585.27	c5	or no .
1586.14	1587.76	c9	yeah . i just get used to it .
1586.92	1587.22	c7	uh-oh .
1587.97	1588.33	c9	so ==
1588.23	1588.73	c3	yeah .
1588.61	1589.77	c5	i like the linguine keyboard .
1588.69	1589.27	c4	yeah .
1588.73	1593.48	c3	i mean like when i first had started working i would have definitely wanted the p_c keyboard .
1590.51	1592.06	c7	the linguine is - o.k .
1590.61	1591.4	c5	that's all i know .
1593.48	1595.19	c3	but now i'm - i'm just used to it .
1594.76	1595.62	c7	you're multi-keyb- ==
1594.89	1598.87	c6	i prefer th- - i prefer the one on linguine or basmati .
1595.62	1595.94	c9	yeah .
1598.87	1599.6	c6	i don't know what it's called .
1599.52	1600.33	c7	o.k let's do it that way .
1599.7	1600.1	c4	mm-hmm .
			...
			[Segment continues.]
			...
1642.91	1656.25	c7	who - who would prefer the keyboard on linguine ? o.k we got two votes for linguine . and who would prefer amaretto ? amaretto . so we got more - more multi-keyboardal people .
1654.41	1654.83	c4	ready .
1655.87	1657	c9	i just have no preference .
1657.12	1657.74	c7	well that's lovely .
1657.28	1658.49	c4	yeah it doesn't really matter .
1657.96	1658.44	c7	same here ?
1658.03	1658.45	c3	yeah .
1658.03	1658.17	c9	yeah .
1658.73	1659.18	c7	excellent .
1658.87	1659.16	c4	yeah .

Table 4: An excerpt of transcriptions from BED009 with the reporting label (**R**).

R -Start	R -Stop	Spkr	Utterances
601.594	611.83	c1	o_k so == facing to - to what we've sort of been doing here == um == well for one thing we're also using this room to collect data .
612.46	613.31	c0	yeah obviously .
612.53	621.147	c1	um - um - not this type of data . no not meeting data but sort of - sort ah - our version of a wizard experiment such ==
616.49	617.06	c0	oh o_k .
621.687	625.117	c1	not like the ones in munich but pretty close to it .
624.77	625.01	c0	mm-hmm .
625.577	630.425	c1	the major difference to the munich ones is that we do it via the telephone .
631.263	631.383	c0	o_k .
631.48	638.424	c1	even though all the recording is done here . and so it's a - sort of a computer call system that gives you tourist information .
637.048	637.218	c0	mm-hmm .
638.424	654.44	c1	tells you how to get places . and it breaks halfway through the experiment and a human operator comes on . and part of that is sort of trying to find out whether people change their linguistic verbal behavior when first thinking they speak to a machine and then to a human .
652.869	653.069	c0	yeah .
655.61	673.09	c1	and we're setting it up so that we can - we hope to implant certain intentions in people . for example um - we have first looked at a simple sentence that how do i get to the powder-tower . o_k so you have the - castle of heidelberg .
672.386	672.636	c0	o_k .
673.09	675.04	c1	and there is a tower and it's called powder-tower .
673.84	674.54	c0	oh o_k . yeah .
			...
	921.31		[End of segment]

Table 5: An excerpt of transcriptions from BED009 with the two-level label of reporting with additional discussion (**R|T**).

R T-Start	R T-Stop	Spkr	Utterances
921.254	929.09	c5	no no . it's a good time to pause . i s- - i see questions on peoples' faces . so why don't ?== let's - let's - let's hear .
927.12	936.52	c0	oh- - . well the obvious one would be if - if you envision this as a module within smartkom where exactly would that sit ?
937.24	943.17	c1	um - so far i've thought of it as sort of adding it onto the modeler - knowledge module .
937.59	938.68	c0	that's the d- ==
942.945	943.435	c3	hmm .
943.091	943.711	c0	o_k yeah .
943.17	947.87	c1	so this is one that already adds additional information to the ==
944.168	946.264	c0	makes perfect sense . yes .
946.45	947.2	c3	hmm - ah .
948.55	955.97	c1	but it could sit anywhere in the attention-recognition . i mean - this is what attention-recognition literally sort of can ==
955.58	956.36	c0	well it's supposed to do .
956.24	956.53	c3	mmm .
956.44	956.6	c0	yeah .
956.596	957.706	c5	that's what it should do . right . yeah .
957.585	958.005	c0	yeah . yeah .
958.092	958.262	c3	huh .
958.906	959.086	c1	yeah .
960.86	980.007	c0	well f- - from my understanding of what the people at phillips were originally trying to do doesn't seem to quite fit into smartkom currently . so what they're really doing right now is only selecting among the alternatives - the hypothesis that they're given enriched by the domain knowledge and the um - discourse modeler and so on .
975.566	981.073	c1	yeah . yeah .
980.89	987.86	c0	so if - if this is additional information that could be merged in by them . and then it would be available to action planning and - and others .
987.39	987.51	c1	yeah .
	1152.17		... [End of segment]

Table 6: An excerpt of transcriptions from BMR002 with the label of brainstorming (Z^M). The diacritic showing that multiple speakers are talking is attached.

Z^M -Start	Z^M -Stop	Spkr	Utterances
718.559	724.695	c3	however if we have a transcribing machine i would accept it . the other thing is - you know having - getting hold of a transcriber . i think that cogsci has one .
725.005	725.495	c1	mm-hmm .
725.017	725.297	c0	yeah .
726.02	727.29	c3	and i could probably borrow one .
726.164	726.724	c0	well one th- ==
726.173	727.433	c2	i - i - i actually have one .
727.38	728.09	c3	you have one yourself ?
728.172	729.032	c2	well sarah has one .
728.403	728.743	c3	oh i see .
729.032	731.152	c2	because she does a lot of transcribing of interviews .
730.5	730.89	c3	oh super .
732.093	738.294	c2	but - the crazy thing is we'd have to convert the s- - we'd have to dump the audio files onto cassette to use it .
732.856	734.942	c0	the other thing is is that we might be == oh .
737.976	738.846	c3	oh i see . o_k .
738.778	744.623	c2	i mean given that we've already got them digitally it would be == you know it's quite tempting to use some kind of uh - software thing .
744.624	745.004	c3	oh o_k .
745.154	747.604	c2	but - but then there's the question of the ergonomic interface that ==
745.235	745.645	c3	i kn- ==
747.621	750.171	c1	oh i actually have some foot pedals for the p_c .
749.114	753.414	c0	well that's what i was gonna say is that adam has foot pedals that we could probably try and rig something up to .
750.51	750.82	c2	o_k .
753.669	754.329	c3	oh that'd be super .
754.11	754.82	c2	what do they connect to ?
755.287	755.787	c1	uh - keyboard . and they generate keystrokes .
757.643	760.943	c2	oh that's totally cool . now how - how do you - how do you decide which keystrokes they generate ?
760.917	763.857	c1	um - there's a little program that downloads into it . um ==
763.367	763.437	c2	hmm .
763.372	763.702	c3	this is great .
			...
	831.822		[End of segment]

Table 7: An excerpt of transcriptions from BED011 with the label of other discussion (**T**).

T-Start	T-Stop	Spkr	Utterances
1491.39	1521.49	c3	but anyway so um - if you - if you can think of a way to uh - point us to th- - to interesting things sort of as you're doing this or - or something . uh - make your - make notes or something that - that this is you know - something worth looking at . and other than that yeah i guess we'll just have to uh - listen == although i guess it's only ten minutes each . right ? roughly .
1521.05	1524.34	c4	well i guess . i'm not sure how long it's actually going to take .
1523.55	1530.29	c2	the reading task is a lot shorter . that was cut by fifty percent . - and the reading == nobody's interested in that except for the speech people .
1529.13	1530.31	c3	right . no we don't care about that at all .
1530.91	1534.23	c2	so == it's actually like five minutes dialogue .
1534.5	1538.46	c3	i - b - my guess is it's gonna be ten . people ==
1538.32	1539.12	c2	ten minutes is long .
1539.64	1542.24	c3	i understand . but people - people - you know uh ==
1541.65	1542.98	c4	it feels like a long time . but ==
1542.88	1543.24	c2	yeah .
1543.47	1543.72	c3	yeah .
1543.5	1547.78	c2	it feels like forever when you're doing it . but then it turns out to be three minutes and forty five seconds .
1545.53	1548.01	c3	yeah . could be .
1547.82	1548.15	c1	b yeah .
1548.28	1553.75	c3	o.k . i was thinking people would you know - hesitate and whatever . whatever it is we'll - we'll deal with it .
1553.89	1554.39	c2	yeah it's not ==
1555.03	1559.12	c3	o.k so that'll be - that'll be - um - on - on the web page .
1555.34	1559.77	c2	and it's fun . o.k .
1559.63	1560.02	c3	that's great .
			...
	1600.46		[End of segment]

4.3.3 Other Discussion

Primary Label: **T**

The label **discussion** indicates balanced conversation between multiple talkers. New information could be brought in by any talker in the meeting, as opposed to reporting where new information is delivered by an individual or a team. Presenting pros and cons of ideas that have already been generated but not actually negotiating a decision would be labeled as discussion. (Since **T** is assumed to involve multiple speakers, it is not necessary to include the diacritic for multiple speakers.)

An example of the discussion marker as a level-2 label would be for discussion following a report that is in direct response to the reporting action. This could include clarification, where the responses may come from the reporter(s) or from other meeting attendees. See Table 7 for examples.

4.4 Group 4: Commentary and Breaks

This grouping describes meeting acts that do not typically produce action or decisions. The first MA is commentary, typically asides not directed to the entire meeting. The second set is for breaks in the meeting

Table 8: An excerpt of transcriptions from BMR002 with the label of commentary (**C**).

C-Start	C-Stop	Spkr	Utterances
2464.88	2467.92	c3	i do find myself whispering . i mean - i suspect - my normal s- ==
2467.01	2467.3	c1	why ?
2467.92	2468.7	c3	i don't know why .
2469.06	2470.51	c1	i had noticed that you were talking quietly .
2469.11	2469.84	c3	it's bizarre .
2470.51	2471.8	c1	and i was actually going to ask you .
2470.57	2472.65	c3	yeah this is not really my normal pattern .
2471.8	2475.35	c1	i was going to ask you to speak up . but i didn't want you to feel self conscious .
2475.55	2478.55	c3	i think the problem is that i'm afraid that i'm talking into someone's ear .
2478.94	2479.21	c1	oh .
2479	2485.75	c3	i really do . so it's - so it's like == and i and i realize now that you have volume uh - control . but uh == yeah .
2486.1	2489.46	c2	i don't feel like you're speaking unnaturally quietly .
2488.96	2489.34	c3	oh good .
2489.56	2491.7	c2	i mean - i feel like it's all quite normal .
2489.71	2490.25	c3	okay . good .
2491.2	2491.87	c0	well i hear what you're saying .
2492.41	2492.67	c3	i'm glad .
2492.72	2492.9	c2	yeah .
2494.78	2494.9	c3	good .
2495.16	2498.99	c2	eric's really got it easy because he's got the lapel mike which is so unobtrusive .
			...
	2547.77		[End of segment]

to deal with technical issues or to deal with interruptions. The third is an outside topic marker used to indicate when portions of the meetings are used for outside recording work.

Level 1 tags should not be interrupted by breaks. If the primary function and topic continue, a level 2 tag indicating a break should be used. Level 2 labels can be interrupted by a break to avoid needing a third level of labeling.

4.4.1 Commentary

Primary Label: **C**

The label **commentary** is used for commentary about meeting action or for asides. These would be apart from the regular or current discussion. Self-talk about wording or correction of previous speech would also be considered commentary. Commentary that is about the topic rather than about the meeting action would not be labeled as commentary unless it becomes an extended branch of the conversation. It would then take on the enveloping region label as a level 2 tag. See Tables 8 and 9 for examples.

4.4.2 Breaks

Primary Labels:

- **BE** Technical or equipment testing breaks

Table 9: An excerpt of transcriptions from BNS003 with the label of commentary (**C**).

C-Start	C-Stop	Spkr	Utterances
439.139	440.019	c4	is it alumni - alumni ?
440.162	440.292	c4	i'm ==
440.308	440.678	c0	alumni .
440.914	441.244	c4	alumni .
441.745	443.125	c4	i'm not sure how to pronounce this word .
442.385	443.415	cA	aber das wird jetzt geaendert .
444.101	444.281	c4	huh ?
444.511	445.211	cA	das wird geaendert .
446.724	448.024	c4	and um ==
448.637	450.107	c4	but um ==

- **BO** Breaks: interrupts, filler talk, off-topic self-talk, etc

The two **break** labels are used for speech that is apart from the focus of the meeting. **BE** is used when there is discussion of technical issues. In the ICSI Meeting set, this is primarily about the recording equipment but would also include discussion about projectors, lighting, furniture, etc. **BO** is used to mark breaks in the meeting for what would otherwise be “third-party” talk. These include interrupts, filler talk, and off-topic self-talk that is not commentary on the meeting or the process. Breaks can be used as level-2 labels when the meeting act remains the same before and after the break region. See Table 10 for examples.

4.4.3 Outside Topic

Primary Label: **OT**

The **outside topic** label is primarily used for DIGIT_TASK labels, a separate recording task that occurred during the meeting recorder project. It can also be used as a level-2 tag for meeting acts that are primarily focussed on setting up the outside topic region. An example could be **P| OT** for time spent planning how an outside topic segment will be done. See Table 11 for examples of **OT** as a level-1 tag as well as a level-2 tag.

4.5 Group 5: Humor

Primary Label: **H**

We decided to do a parallel labeling of humor, in part because labelers were distracted by humorous sections which were not necessarily connected to functional regions within the meetings. This means regions can have another level of labels. While humor is sometimes labeled with the DA “j” for joke, it was actually easier for the labelers to note when something was humorous than to not label anything. Because humor can play many roles within meetings, a marking of humorous sections may be useful for future work. In some cases, a humor label can be seen as a disruption to the current MA segment. In others, a humorous portion comprises a new MA. Examples will be provided in future drafts.

4.6 Diacritics

Diacritics are modifiers that can be used for any of the primary MA tags at either level-1 or level-2. They should be used to indicate general aspects of the MA segment. If the segment is generally filled with agreement or disagreement, then **A** or **D** would be used to indicate that. The negotiation label has a required diacritic (applied to the entire label) that indicates what the final action is (agreement,

Table 10: Excerpts of transcriptions from BMR002 with the labels for equipment-related break (**BE**) and other break (**BO**).

BE-Start	BE-Stop	Spkr	Utterances
3.553	236.174	c1	o_k . so ==
4.715	5.545	c2	i think that's recording now .
6.459	11.146	c1	o_k i'm on - i'm talking on mike two . so == which channel is that ?
12.461	13.221	c2	um - do it again .
13.444	15.224	c1	mike number two i'm talking right now on .
15.448	17.546	c2	that's uh - that's in channel one .
15.829	19.856	c1	mike number two ? channel one ? is that channel one or channel zero ?
19.93	20.96	c2	channel one . let me make sure .
20.466	20.926	c0	channel one .
			...
	236.174		[End of segment]
BO-Start	BO-Stop	Spkr	Utterances
1209.54	1212.52	c3	can i be excused ? and i'll go down and get my stuff . because
			um - i - i do think ==
1212.08	1213.59	c2	you don't == just take the thing with you .
1214.46	1215.58	c1	th- - i- - it's wireless .
1214.47	1215.19	c3	are you serious ?
1215.58	1215.96	c1	i mean it won't ==
1215.82	1217.9	c2	well i don't know what will happen . but you may as well .
			...
	1236.98		[End of segment]

Table 11: Outside Topic labels from BMR002 as level-1 and level-2 labels.

T OT-Start	T OT-Stop	Spkr	Utterances
291.882	296.502	c1	so if you could just fill out the form . and then later
			on we'll read it with pauses between .
296.965	297.135	c3	good .
297.589	298.899	c1	and we're session two .
302.416	303.036	c0	do you want to read next ?
302.476	303.197	c3	shall i - n- ?==
303.036	303.446	c0	or d- ?==
303.539	304.219	c3	i'd like to hear one ==
303.549	304.749	c1	let her fill it out and go ahead and read .
304.839	305.679	c3	yeah i'd like to hear one more .
305.084	313.273	c0	uh == okay . um - this is eric on mike one . the wireless lapel .
			channel zero . um ==
OT-Start	OT-Stop	Spkr	Utterances
313.722	363.134	c0	DIGIT_TASK

Table 12: An excerpt from BSR001 with a level-2 humor label.

R H-Start	R H-Stop	Spkr	Utterances
2011.6	2012.35	c0	delete that part .
2011.79	2015.31	c4	none of that . n- - no eating during meetings .
2015.4	2016.69	cA	right right right .
2015.55	2016.82	c8	yes . or drinking .
2016.38	2018.24	c4	maybe salivating if you're thinking about the food . but it ==
2019.64	2022.68	c8	y- - getting - getting coffee on these is definitely a bad thing .
2019.72	2020.95	c2	not with these around .
2019.88	2020.22	c4	um ==

disagreement or no decision made). **V**, the diacritic for the use of visual aids, can indicate slides, whiteboard use, handouts, and gestures. The key reason to include this diacritic is to indicate that the audio signal does not contain all of the information from the meeting at that point. **Y** is used to indicate summarization. Summarization can be used for information from earlier meetings, from the current meeting, or from outside sources. There is typically checking for agreement amongst participants during the summary. Examples of diacritic usage can be found in Appendix A (**I**, **L**, **M**, **V**, and **Y**) and other examples in the text (**D**, **FT**, **M**, **S**). Examples of the remaining diacritics ((**FA**, **FD**, **ST**, **W**,) are in Tables 13, 14, and 15.

- **A** agreeing
- **D** disagreeing
- **FA** ends in agreement
- **FD** ends in disagreement
- **FT** ends in tabling (no decision made)
- **I** consists primarily of an individual speaker
- **M** consists primarily of multiple speakers
- **L** long-term (typically for planning)
- **ST** short-term (typically for planning)
- **S** suggest alternate
- **V** Use of visual aids.
- **W** extremely formal in style
- **Y** summary or recap of previous knowledge

5 Label File Formats

There are four key areas of the ascii label files: header information, the tag with start and stop times, tag-level comments and meeting notes. See Appendix A.1 for a complete example.

Table 13: Examples of diacritic usage for **FA** from BMR002 and **ST** from BED011.

Meeting BMR002: Complete N ^{FA} Example			
Start	Stop	Spkr	Utterances
595.155	595.654	c0	uh - right .
596.022	598.762	c1	and so one of the first questions is do you need help .
596.049	596.269	c0	o.k .
599.594	600.064	c2	well hang on .
600.636	600.786	c0	oh .
600.717	601.667	c1	that's not the first question .
600.929	602.561	c2	well - well i guess .
601.133	601.503	c0	o.k .
603.432	607.862	c2	do you - so do you ?== what are - what are we talking about ? that you're gonna - you're gonna do transcriptions ? you're going to listen to these things and ==
607.121	614.136	c3	i'm going to provide - i'm going to be == once i'm given the data i will provide w- - word-level transcripts with uh - indication of speakers .
612.414	612.824	c2	o.k .
614.358	619.845	c3	and um == i guess speaker change in any case . so speaker i_d . and words .
620.002	622.422	c0	o.k so you're gonna do the actual transcription yourself ?
622.613	622.803	c3	yeah .
622.954	623.174	c0	o.k .
Meeting BED011: Complete P ST Example			
Start	Stop	Spkr	Utterances
586.71	589.62	c2	and we're gonna start tomorrow at three ? four ? one ?
589.605	593.535	c4	tomorrow == well we don't know for sure . because we don't know whether that person is coming or not .
593.362	593.642	c2	o.k .
593.535	593.695	c4	but ==
593.872	605.081	c2	around four-ish . and um - we're still l- - looking for a room on the sixth floor because they stole away that conference room um - behind our backs . but ==
605.6	610.74	c3	well there are these uh - uh == oh i see we have to == yeah it's tricky . we'll - let's - let - we'll do that off-line o.k .
609.62	615.91	c2	yeah but i - i - i - it's happening . david and - and jane and - and lila are working on that as we speak .
612.86	613.27	c3	o.k .

Table 14: Examples of diacritic usage for **FD** from BUW001. The labeler called this section **T|N^FD** with a comment that it was tabled to avoid conflict. This probably explains why **FT** is more commonly used than **FD**.

Start	Stop	Spkr	Utterances
2752.62	2755.79	cA	absolutely so i think we're gonna have to start with keywords .
2752.97	2753.25	cB	mm-hmm .
2755.79	2759.22	cA	and - and if someone becomes more interested we could work our way up .
2757.02	2760.83	c3	i- i'm == i- i'm not so sure i agree with that .
2759.32	2759.58	cA	but ==
2759.46	2760.5	c1	it - but it may well ==
2760.76	2761.05	cA	really ?
2760.93	2764.71	c3	because - uh b- - because it depends on uh what our goal is . if our goal is wizard of oz-ish ==
2765.1	2765.68	cA	oh that's true .
2765.38	2768.41	c3	we might want to know what is it that people would really like to know about this data .
2768.2	2768.47	c8	yeah .
2768.41	2771.09	c3	and if it's - if - if it's something that we don't know how to do yet th- - great .
2768.83	2769.03	c8	yeah .
2769.15	2769.54	c0	right .
2771.09	2773.57	c3	that's you know research project for year four or something .
2772.09	2772.83	c8	research yeah .
2772.55	2772.91	c0	mm-hmm .
2773.57	2773.75	c3	you know .
2774.21	2774.6	c0	mm-hmm .
2775.25	2780.06	cA	yeah i was thinking about wizard of oz . but it requires the wizard to know all about the meetings .
2778	2779.22	c8	we'd have to listen to all the data .
2782.42	2786.13	c3	um == well not - maybe not true wizard of oz .
2784.95	2786.61	cA	so == oh yeah i - i understand .
2786.13	2787.15	c3	because people are too ==
2787.53	2788.51	c8	well just imagine if ==
2787.94	2790.51	c3	uh aware of what's going on . but - but just ==
2790.65	2793.98	c8	<yeah> get people to ask questions that they def- - the machine definitely can't answer at the moment .
2793.79	2794.09	cA	yep .
2794.08	2794.22	c8	but ==
2794.31	2795.6	c3	yeah . w- - just what would you like to know .
2794.58	2794.69	c8	yeah .
2794.65	2797.48	cA	but that - neither could anyone else though is what uh my point is .
2797.16	2797.74	c8	yes .
2798.52	2842.68		<i>Meeting continues...</i>
2846.18	2846.37	c1	so ==
2847.39	2847.6	c3	hmm .
2848.75	2850.46	cB	yeah . well but - you know but you could get again @@ .
2849.42	2852.37	c0	well we're piloting . we'll just do it and see what happens .
2850.98	2852.34	c1	yeah . yeah .
2852.05	2852.34	cB	yeah .
2852.35	2858.54	c3	yeah . i usually don't remember my action items . but - i'd - i ==

Table 15: Examples of diacritic usage for **W** from BNS002. The principal speaker talks for over three minutes before an interruption. The labeler called the first section where c4 speaks $\mathbf{R}^{\wedge}\mathbf{V}^{\wedge}\mathbf{W}$. After this portion, the speaker begins calling on people who have indicated they have questions. The last two turns in the example are labeled $\mathbf{R}^{\wedge}\mathbf{V}^{\wedge}\mathbf{W}|\mathbf{T}$.

Start	Stop	Spkr	Utterances
165.079	380.826	c4	thank you . well eh i - i have prepared eh a really short talk about some of the topics . i - i have been doing some w- - some research work uh before coming here . and eh well mainly i will present . uh here you have more or less the sketch of the presentation . i will present some ideas and some concepts about what ad-hoc networks are . about the issues in media access control in wireless networks . also about routing in such kind of networks . and eh i will make some comments also about uh the current standards of wireless local area networks . and finally i will present the blueprint of - of my research project or my research uh proposal here . so the first thing is eh what an ad-hoc network is . you know uh an ad-hoc network is made of wireless nodes . these nodes of course can be mobile nodes . and eh in this network there is no network infrastructure . so eh the only active thing we have is the mobile nodes . and eh == end-to-end communication is sometimes uh done by direct communication from one node to the other . because they can talk to each other . but sometimes it is not possible this direct communication because nodes end - uh source or - and destination nodes are um - are too far away . and in this case it can be done communication can be done by uh using multi-hub routing . so == an intermediate or a set of intermediate nodes will route or will forward the packet from the sender to the destination . so == uh in a sense what we have in this kind of networks is that uh nodes are as both arced as both uh endpoints and also routers of other neighboring nodes . and this makes a uh difference from other kind of network . and well regarding the purposes of what can this kind of network be used for eh we can say that well cellular networks are infrastructure-based networks . in this kind of ad-hoc networks we don't have anything uh with the exception of the network nodes . so this uh network can be useful for eh uh occasions where no available infrastructure is present . like for example in - in - eh after some kind of - of - eh um catastrophe- - catastrophic event . or also when - even when there is a network available we don't want for different reasons to use the available network . this latt- - later case can be for example the use in military systems . where maybe we don't want to use the other - the enemy network infrastructure for evident reasons .
380.957	381.177	c4	yeah .
381.296	399.816	c3	hmm ? but you can also i believe uh to have an ad-hoc network where one node is then connecting to the outer world . for instance what i have in mind you have a car where you plug in different devices and they talk to each other for instance . but one device decided to go to the g_s_m network or g_p_s network or whatever kind of things .
399.971	400.711	c4	that's o_k .

5.1 Header Information

Each file begins with one line of global labeling information. This includes the meeting name (e.g., BSR001), the date (typically in the format “12 May 2005”), the word “LABELER” followed by the initials of the labeler. This is followed by a column header line:

```
START    STOP    LABEL    COMMENTS
```

5.2 Tags and Label Order

The lines following the header lines each describe one distinctly labeled region. They are ordered by their start times, which is key since embedded labels (i.e., when a level-2 label is used) are contained within the start and stop times of the associated level-1 label. Start or end times can be the same for level-2 labels associated with level-1 labels but if the level-1 label ends, so does the level-2 label. However, a new level-1 tag can begin before the ending of the current level-1 tag. Since one speaker can be finishing up a meeting act while another starts the next one, this allows for overlapping. Start times can be used to generate sequencing information, while stop times, when compared with start times, indicate overlap.

5.3 Comment Columns

The fourth column of a tag line is for comments about that particular tag or region of the meeting. These are used to indicate any labeler uncertainty as well as to indicate problems that exist at that point in the transcription. They are frequently used to indicate overlapping MA regions.

5.4 Meeting Notes

Comments about the overall meeting may be placed at the bottom of the file with at least one blank line between the tag lines and the notes. The word “NOTE:” is the first word for each line. When there are meeting notes, this also suggests the fact that the user should look at the comment columns. Example notes show problems in transcriptions or recordings, such as “missing speaker.”

References

- [1] CALO topic labeling and meeting annotation resources, 2005. Stanford University CSLI. Information available at <http://cujo.stanford.edu/twiki/bin/view/Public/SemlabProjects>.
- [2] R. Dhillon et al. Meeting recorder project: Dialog act labeling guide. Technical Report TR-04-002, ICSI, 2004.
- [3] I. McCowan et al. Modeling human interaction in meetings. In *Proc. ICASSP*, pages IV-748-751, 2003.
- [4] N. Morgan et al. The Meeting Project at ICSI. In *Notebook Proceedings of HLT (Human Language Technology) Conference*, pages 29-36, 2001.
- [5] N. Morgan et al. Meetings about meetings: Research at ICSI on speech in multiparty conversations. In *Proc. ICASSP*, pages IV-740-743, 2003.
- [6] E. Shriberg et al. The ICSI Meeting Recorder Dialog Act (MRDA) Corpus. In *Proc. HLT-NAACL SIGDIAL Workshop*, 2004.
- [7] B. Wrede and E. Shriberg. The relationship between dialogue acts and hot spots in meetings. In *Proceedings of IEEE Workshop ASRU*, 2003.

A Labeled Meeting Samples

There are two portions of this appendix. The first is a MA labeled file for BMR-002. The second is a labeled version of BMR025 that includes the full text as well as the DA and MA labels.

A.1 Meeting Act Label Example

This section contains an example of labels for one meeting. Level-2 labels are indicated with a |. Note that the start and end times for labels with level-2 labels are typically within a larger level-1 segment. When the start time follows the previous stop time, a new level-1 segment begins. Otherwise, the level-1 label stays in effect.

MA Labels for BMR-002			
Header: BMR002 7 May 2005 EAW			
START	STOP	LABEL	COMMENTS
3.553	206.876	BE	
208	217.235	R	
217.574	236.174	BE	
236.756	245.386	G	
245.552	249.995	T OT	
250.454	287.406	OT	
291.882	313.273	T OT	
313.722	363.134	OT	
365.231	373.323	T OT	
374.14	415.203	OT	
416.9336	451.369	T OT	
452.053	492.82	OT	
494.218	495.032	T OT	
495.311	503.31	P OT	
504.044	534.984	T OT	
535.379	568.473	G	
569.571	595.654	C	
595.155	623.174	N^FA	
623.923	721.119	T	
647.871	668.128	T T	really strong topic break
721.119	808.414	Z	
758.653	780.07	Z T	
809.019	842.702	T	
842.566	924.286	N^FT	overlap. (poss top. break)
914.353	924.286	N^FT C	
924.286	998.624	T^Y	
999.352	1047.88	R^M	
1048.49	1134.22	N^FT	
1132.99	1210.29	R	overlap
1209.54	1236.98	BO	overlap
1240.74	1372.18	T	

MA Labels for BMR-002			
START	STOP	LABEL	COMMENTS
1372.64	1638.76	T^V	
1400.37	1407	T^V R^Y	
1473.23	1518.33	T^V R^Y	
1534.92	1558.54	T^V P	
1640.62	1641.71	C	OVERLAP
1641.04	1649.48	N^FA	
1649.81	1805.54	T^V	
1707.8	1710.76	T^V C	
1805.41	1823.59	P	OVERLAP
1824.13	1856.45	P^S	
1858.79	1870.7	C	
1870.75	1907.68	T	
1908.02	1954.52	N^FT	
1956.05	2028.94	T	
2028.94	2030.9	G	
2032.61	2069.26	P^L	
2070.1	2182.13	T	
2183.64	2200.42	C	
2201.81	2249.55	T	
2249.55	2261.64	N^FA	
2269.17	2361.64	N^FA T	
2366.54	2367.01	G	
2368.82	2418.47	P	
2418.22	2465	T	OVERLAP
2464.88	2531.22	C	
2531.87	2534.97	G	
2536.64	2541.52	BE	
2543.07	2549.26	T OT	
2547.59	2580.94	OT	OVERLAP
2587.32	2592.12	T OT	
2592.97	2638.17	OT	
2640.41	2640.83	T OT	
2642.11	2685.56	OT	
2688.55	2689.46	T OT	
2689.9	2695.44	OT	
2696.14	2696.44	T OT	
2696.82	2732.07	OT	
2734.57	2736.29	T OT	
2741.24	2745.29	BE	

A.2 Meeting Act Label Example with DAs and Text

This section contains the full text of a meeting jointly labeled by the three labelers. The first column shows the level-1 MA tags, with the start time being the same as the current utterance start time. The end time for a tag is the stop time of the utterance before a new level-1 tag. The second column is for the level-2 tags. Level-2 tags do not span level-1 changes so if there is a new level-1 tag, both the previous level-1 and any existing level-2 tags end. A level-2 tag may be embedded in a level-1 tag. In this case, the table below shows the flag “end” to show the end of the level-2 tag, with the stop time coinciding with the stop time of the current utterance. In the distribution files, the tags appear as shown in the previous section.

Header: BMR025 Labeled 8 MAY 2005 PKM CJK EAW

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
BO		0.04	0.11	c4	z	yeah .
G		1.733	2.823	c1	fh	um - so ==
		3.6	7.19	c1	s^t	if we can't we can't but uh we're gonna try to make this an abbreviated meeting .
		7.19	11.95	c1	s^df	cuz the - the next - next occupants were pushing for it .
		12	12.28	c1	fh	so ==
		13.432	13.952	c1	fh	um .
		17.72	17.95	c1	fh	so ==
		19.02	26.937	c1	s^t	agenda is - according to this is transcription status darpa demos x_m_l tools disks backups et cetera and crosspads.
		29.11	31.03	ca	qy^rt^t	does anyone have anything to add to the agenda ?
		33.288	33.598	c1	s^bk^rt	o.k .
		35.706	36.526	c1	qy^rt^t	should we just go in order ?
		36.616	37.576	c1	s^e^rt	transcription status .
		37.576	38.526	c1	s	who's - that's probably you .
R		38.618	39.458	c0	s^na	i can do that quickly .
		39.458	41.398	c0	s^rt	um i hired several more transcribers .
		41.398	42.628	c0	s	they're making great progress .
		41.859	42.259	c1	qy^bu^d^rt	seven ?
		42.628	43.938	c0	s^bc^nd	seve- - several several .
		43.61	43.9	c1	s^bk	oh .
		46.403	47.153	c0	fh	and uh ==
		47.891	48.671	c0	fh	and uh ==
		49.73	52.87	c0	s	uh i've been uh finishing up the uh double checking .
		53.511	54.811	c0	s	i hoped to have had that done by today .
		54.811	55.911	c0	s	but it's gonna take one more week .
T		57.03	57.43	ca	fg	um ==
		58.43	66.45	ca	qy^rt	as a somewhat segue into the next topic um - could i get a hold of uh - the data even if it's not really corrected yet ?
		58.471	58.751	c3	%-	i g- ==
		66.45	8.2	ca	s^df^e	just so i can get the data formats .
		68.81	71	ca	s^	df^e and make sure the information retrieval stuff is working .
		71.03	71.34	c0	s^aa	certainly .
		71.39	72.76	c0	s	yeah i mean it's in the same place it's been .
		71.78	72.26	ca	qy.%-	so can you just ?==
		72.88	73.23	ca	qh^rt	oh it is ?
		73.23	73.44	ca	s^bk	o.k .
		73.39	74.1	c0	s^aa	uh-huh no change .
		73.44	76.62	ca	qy^bu^d^rt	just - so transcripts is the sub-directory ?
		75.524	77.36	c0	h s^aa^rt	uh yes uh-huh .
		77.15	77.37	ca	s^bk	o.k .
		77.982	80.072	ca	s^cc	so i'll - i'll probably just make some copies of those .
		80.382	81.842	ca	s^e	rather than use the ones that are there .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		83.33	83.65	c0	s^bk^rt	o.k .
		83.45	88.92	ca	fh s^co	um and then just - we'll have to remember to delete them once the corrections are made .
		90.489	90.849	c0	s^bk	o.k .
R		92.394	93.164	c1	fg %-	o.k wh- ==
		92.57	95.02	c3	s	i also got anot- - a short remark to the transcription .
		95.02	99.02	c3	s	i've uh just processed the first five e_d_u meetings .
		99.02	100.49	c3	s	and they are chunked up .
		100.49	103.43	c3	s^rt	so they would - they probably can be sent to i_b_m .
		104.139	104.959	c3	s^e^rt	whenever they want them .
	T	105.92	107.16	c5	s.%-	well the second one of those ==
		106.013	106.293	c2	s^bk	cool .
		107.13	108.82	c3	s^aa s^2	yep it's already at i_b_m .
		107.479	108.509	c5	s^m	is already at i_b_m .
		108.82	109.62	c3	s.%-	but the other ones ==
		109.044	111.074	c5	s	that's the one that we're waiting to hear from them on .
		109.62	109.78	c3	b	yeah .
		111.027	111.147	c3	b	yeah .
		111.7	111.91	c5	b	yeah .
		111.85	112.19	c0	b	o.k .
		112.535	114.225	c0	s^bu.%-	these are separate from the ones that ==
		112.57	112.92	c5	s.%-	as soon as ==
		114.435	114.905	c0	s.%-	i mean these are ==
		114.63	115.44	c5	s	they're the i_b_m set .
		115.727	116.207	ca	s.%	it's this one .
		115.824	116.014	c3	s^aa	yep .
		116.36	116.84	c0	b	excellent .
		116.987	117.167	c5	b	yeah .
		117.11	117.3	c0	b	good .
	BE	117.59	118.16	ca	qy^rt^t1	is my mike on ?
	T	117.78	121.7	c5	s^rt	and so as soon as we hear from brian that this one is o.k .
		118.16	118.38	ca	b	yeah .
		121.7	122.92	c5	s	and we get the transcript back .
		122.92	129.389	c5	s	and we find out that hopefully there are no problems matching up the transcript with what we gave them .
		130.45	133.98	c5	s^cc	then uh we'll be ready to go and we'll just send them the next four as a big batch .
		134.318	134.678	c0	b	excellent .
		134.81	135.77	c5	s	and let them work on that .
		136.12	139.26	ca	qy^bu^d^rt	and so we're doing those as disjoint from the ones we're transcribing here ?
		139.58	139.72	c5	s^aa	yes .
		139.91	140.3	ca	s^bk	o.k good .
		140.13	140.58	c5	s^aa	exactly .
		140.58	141.72	c5	s	we're sort of doing things in parallel .
		141.72	143.33	c5	s^df	that way we can get as much done a- - at once .
		142.91	144.02	ca	s^na	yeah i think that's the right way to do it .
		143.91	144.09	c5	s^aa	yeah .
		144.17	146.06	ca	s^e	especially for the information retrieval stuff .
	G	148.891	150.401	ca	qy^d^rt^t	anything else on transcription status ?
		150.217	150.437	c0	s^ar	hm-mmm .
		151.62	151.8	ca	s^bk^rt	o.k .
	R	152.36	152.93	c1	s^rt^t	darpa demos .
		152.93	154.67	c1	s^e^rt	we had the submeeting the other day .
		154.56	155.7	ca	fg %-	right which uh ==
		156.82	161.38	ca	s	so i've been working on using the thisl tools to do information retrieval on meeting data .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		161.908	165.822	ca	s	and the thisl tools are - there're two sets there's a back-end and a front-end .
		165.942	167.312	ca	s	so the front-end is the user interface .
		167.312	170.532	ca	s	and the back-end is the indexing tool and the querying tool .
		171.4	176.37	ca	s	and so i've written some tools to convert everything into the right for- - file formats .
		176.774	181.7	ca	s	and the command line version of the indexing and the querying is now working .
		182.369	188.493	ca	s	so at least on the one meeting that i had the transcript for - uh conveniently you can now do information retrieval on it .
		188.593	193.003	ca	s^	e do - type in a - a string and get back a list of start-end times for the meeting .
		193.479	194.389	ca	s^	e uh of hits .
		193.57	196.28	c5	qw	what - what kind of uh - what does that look like the string that you type in ?
		196.28	198.23	c5	qr.%-	what are you - are you - are they keywords or are they ?==
		197.91	198.34	ca	s^aa^m	keywords .
		198.46	198.77	c5	s^bk	o.k .
		199.431	199.691	c5	s^	bk i see .
		199.82	206.376	ca	s right	and so - and then it munges it to pass it to the thisl i_r which uses an s_g_m_l-like format for everything .
		206.73	207.22	c5	s^bk	i see .
		207.427	209.21	c1	qy ^rt	and then - does it play something back ?
		209.21	210.54	c1	qr	or that's something you're having to program ?
		209.533	215.043	ca	s^na	um right now i have a tool that will do that on a command line using our standard tools .
		215.199	215.639	c1	b	yeah .
		215.475	218.935	ca	s.%-	but my intention is to do a prettier user interface based either ==
P		219.44	222.99	ca	s^cs^t:qw	so - so that's the other thing i wanted to discuss is well what should we do for the user interface .
		222.99	225.706	ca	s	we have two tools that have already been written .
		225.796	228.266	ca	fh s	um the softsound guys did a web-based one .
		228.689	229.069	c1	b	mm-hmm .
		229.001	231.331	ca	fh s	um which i haven't used haven't looked at .
		231.331	232.681	ca	s	dan says it's pretty good .
		232.962	233.272	c1	b	mm-hmm .
		233.196	235.906	ca	s	but it does mean you need to be running a web server .
		236.418	236.768	c1	b	mm-hmm .
		236.76	238.69	ca	s	and so it - it's pretty big and complex .
		239.424	241.674	ca	fh s	uh and it would be difficult to port to windows .
		241.674	243.894	ca	s^df	because it means porting the web server to windows .
		244.49	244.93	c1	b	mm-hmm .
		244.967	250.614	ca	s	uh the other option is dan did the tcl-t.k thisl gui front-end for broadcast news .
		249.616	249.906	c1	b	yeah .
		250.614	251.584	ca	s^ba	which i think looks great .
		251.584	252.654	ca	s^ba	i think that's a nice demo .
		253.177	255.757	ca	s	um and that would be much easier to port to windows .
P end		256.07	257.49	ca	s^cs	and so i think that's the way we should go .
		258.666	259.796	c0	qy	i - can i ask a question ?
		259.612	259.982	ca	s^aa	mm-hmm .
		260.006	265.59	c0	fh s:s	so um as it stands within the - the channeltrans interface it's possible to do a find and a play .
		266.3	67.75	c0	s	you can find a searched string and play .
		267.75	274.352	c0	qo^bu.%-	so e- - are you - so you're adding like um - i don't know uh are they fuzzy matches or are they uh ?==

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		273.632	276.301	ca	s	it's a sort of standard text-retrieval-based .
		277.32	280.62	ca	s	so it's uh term frequency inverse document frequency scoring .
		281.036	281.466	c0	s^bk^rt	o_k .
		281.395	284.589	ca	fh s	um and then there are all sorts of metrics for spacing .
		284.589	286.679	ca	s^e	how far apart they have to be and things like that .
		287.343	288.233	ca	s.%-	so it - it's ==
		288.337	291.77	c0	s^bu	it's a lot more sophisticated than the uh - the basically windows-based .
		289.366	293.922	ca	s^na	i- - it's like doing a google query or anyth- - anything else like that .
		294.271	297.481	ca	s	so i- - it uses - so it pr- - produces an index ahead of time .
		294.367	294.677	c0	s^bk^rt	o_k .
		297.481	299.991	ca	s	so you don't - you're not doing a linear search through all the documents .
		300.515	304.405	ca	s	cuz you can imagine if - with - if we have the sixty hours' worth you do - wouldn't wanna do a search .
		304.281	304.591	c0	s^aa	hm-mmm .
		305.07	306.44	ca	s	um you have to do preindexing .
		305.396	305.606	c0	b	good .
		306.44	308.08	ca	s	and so that - these tools do all that .
		309.161	312.211	ca	s.%-	and so the work to get the front-end to work would be porting it ==
		312.529	320.372	ca	fh s	well uh to get it to work on the unix systems our side is just rewriting them and modifying them to work for meetings .
		320.552	327.963	ca	s	so that it understands that they're different speakers and that it's one big audio file instead of a bunch of little ones and just sorta things like that .
		320.903	321.283	c1	b	mm-hmm .
		323.332	323.672	c1	b	mm-hmm .
		327.075	327.345	c9	b	mm-hmm .
T		328.28	330.6	c5	qw	so what does the user see as the result of the query ?
		331.23	332.2	ca	qw^bu	on which tool ?
		332.722	333.132	c5	s	thisl .
		334.04	337.31	ca	s	the thisl gui tool which is the one that dan wrote tcl-t.k .
		335.161	335.431	c5	b	yeah .
		337.795	342.113	ca	fh s	um you type in a query and then you get back a list of hits .
		342.456	344.076	ca	s	and you can type on them and listen to them .
		344.839	346.249	ca	s^bsc	click on them rather with a mouse .
		345.74	345.97	c5	s^bk	ah .
		346.63	349	c5	qx^bu.%-:s	so if you typed in small heads or something you could ?==
		346.731	346.901	c1	b	mmm .
		348.652	349.702	ca	s^aa %-	right you'd get ==
		349.347	355.311	c5	s^2	get back a uh uh - something that would let you click and listen to some audio where that phrase had occurred .
		352.556	353.006	ca	s.%-	something ==
		354.08	356.3	ca	s:s^j	you - you'd get to listen to beep .
		355.311	355.551	c5	%-	or some- ==
		356.38	357.3	c1	s^ba^j	that was a really good look .
		357.5	359.93	c1	s^ba.%-	it's too bad that that couldn't come into the ==
		358.92	360	ca	s^j	you couldn't get a video .
		360.54	361.62	c9	s^co	guess who i practice on .
C		364	366.29	c0	s	at some point we're gonna have to say what that private joke is .
		366.29	367.16	c0	s	that keeps coming up .
		366.37	366.61	c1	s^aa	yeah .
		367.974	369.054	c1	s^ar	and then again maybe not .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
	T	369.054	370.914	c1	fh	so uh ==
		371.59	372.64	c1	s	yeah that soun- - that sounds reasonable .
		372.64	377.68	c1	s	yeah it loo- - it - my - my recollection of it is it's - it's a pretty reasonable - uh demo sort of format .
		372.893	373.153	ca	s^bk	right .
		376.35	376.95	c5	s^ba	yeah that sounds good .
		376.74	380.161	ca	s	and so i think there'd be minimal effort to get it to work minimally .
		377.24	378.01	c5	s^ba	that sounds really neat .
		380.161	384.511	ca	s^cs:s	and then we'd wanna add things like query by speaker and by meeting and all that sort of stuff .
		384.971	387.651	ca	s	um dave gelbart expressed some interest in working on that .
		387.651	388.691	ca	s^cc	so i'll work with him on it .
		389.078	390.248	ca	s	and it - it's looking pretty good .
T		390.408	392.638	ca	s^e	you know the fact that i got the query system working .
		393.193	396.333	ca	s^ba^cs	so if we wanna just do a video-based one i think that'll be easy .
		396.693	397.063	c1	b	mm-hmm .
		396.811	399.051	ca	s	if we wanna get it to windows it's gonna be a little more work .
		399.051	404.63	ca	s^df.%-	because the thisl i_r - the information retrieval tool's - um ==
		404.84	406.93	ca	s^df	i had difficulty just compiling them on solaris .
		407.303	407.673	c1	b	mm-hmm .
		407.46	410.797	ca	s	so getting them to compile on windows might be challenging .
		411.231	411.581	c1	b	mm-hmm .
		411.61	416.805	c5	s.%-	but you were saying that - that the uh - that there's that set of tools uh cygnus tools that ==
		411.812	412.082	ca	fh	so ==
		416.732	417.732	ca	s^na	it certainly helps .
		418.11	418.43	c5	b	uh-huh .
		418.202	420.502	ca	s	um i mean without those i wouldn't even attempt it .
		420.67	420.98	c5	s^bk	yeah .
		420.88	421.34	c1	b	mm-hmm .
		421.222	425.472	ca	s	but what those - they - what those do is provide sort of a b_s_d compatibility layer .
		425.943	428.323	ca	s	so that the normal unix function calls all work .
		426.153	426.663	c1	b	mm-hmm .
		427.58	428.09	c1	b	mm-hmm .
		428.998	429.358	ca	fh	um ==
		429.341	430.241	c5	s^bu.%-	and you have to have all the o- ==
		429.89	433.48	ca	s	but the problem is that - that the thisl tools didn't use anything like autoconf .
		434.423	440.499	ca	s	and so you have the normal porting problems of different header files and th- - some things are defined and some things aren't .
		441.092	444.6	ca	s	and uh - different compiler work-arounds and so on .
		444.6	453.701	ca	s	so the fact that - um it took me a day to get it c- - to compile under solaris means it's probably gonna take me s- - significantly more than that to get it to compile under windows .
		453.8	456.01	c1	qo^cs	how about having it run under free b_s_d ?
		453.941	454.461	c4	s.%-	well what you need ==
		456.09	457.59	ca	s^ba^na	free b_s_d would probably be easier .
		457.98	463.02	c4	s:s	all you need to do is say to dan gee it would be nice if this worked under autoconf and it'll be done in a day .
		463.05	463.46	ca	s^aa	that's true .
		463.26	463.54	c4	qy^d^g^rt	right ?
		463.719	463.839	c3	b	uh ==
		465.39	466.47	ca	s^cc	actually you know i should check .
		466.47	469.08	ca	s^df	because he did port it to sprachcore .
		469.491	469.711	c4	s^aa	right .
		469.526	470.766	ca	s	so he might have done that already .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		470.75	472.54	c4	s	i - i - i wouldn't be surprised .
		473.09	473.38	c1	fg	so ==
		473.49	474.14	ca	s^cc	i'll check at that .
		473.94	474.57	c5	qw	how does it play ?
		474	474.22	c4	s.%-	what i ==
		474.352	479.542	c1	s^co	but it would - what would serve - would serve both purposes is if you contact him and ask him if he's already done it .
		479.378	479.998	ca	s^aa	yeah right .
		479.938	481.278	c1	s	if he has then you learn .
		481.608	482.828	c1	s	if he hasn't then he'll do it .
		482.9	483.19	ca	s^aa	right .
	C	485.96	486.22	c0	b	wow .
		486.71	488.08	c5	s^j	i hope he never listens to these meetings .
	C end	488.52	489.03	ca	s^aa.%	that's right .
		490.744	497.29	ca	fg s.%-	so and i've been corresponding with dan and also with uh - uh - softsound guy uh ==
		491.66	492.14	c0	b	it's amazing .
		493.283	493.683	c1	b	yeah .
	C	498.41	499.18	ca	s	blanking on his name .
		499.731	500.341	c5	qy^bu^d^rt	tony robinson ?
		500.01	500.53	c1	qy^bu^d^rt	tony robinson ?
		500.8	501.31	ca	qy^no^rt	do i mean tony ?
		501.31	501.89	ca	s^aa	i guess i do .
		502.034	502.244	c1	b	yeah .
		502.317	503.607	ca	s^e	or s- - or steve renals .
		502.39	503	c4	s^2	james christie .
		504.029	504.809	ca	qw	which one do i mean ?
		504.16	505.46	c1	s	steve renal- - steve renals .
		504.69	505.92	c4	s^nd	steve renals is not softsound .
		505.92	506.15	c4	qy^d^g^rt	is he ?
		506.83	507.02	c1	s^ar	no .
		507.935	508.855	ca	s^bd	my brain is not working .
		508.855	510.495	ca	s^no	i don't remember who i've been corresponding with .
		508.97	509.46	c1	s^bk	o.k .
		509.73	512.834	c4	s	steve wro- - i- - it's ste- - steve renals wrote thisl i.r .
		512.89	513.85	ca	s	then it's steve renals .
	C end	514	514.24	c4	s^bk	o.k .
		514.259	514.689	c1	s^bk	oh o.k .
		514.311	518.111	ca	s	so uh just getting documentation and uh and f- - and formats .
		516.89	517.08	c1	b	yeah .
		518.111	519.331	ca	s^ba	so that's all going pretty well .
		519.331	520.311	ca	s	i think we'll be o.k with that .
		519.607	519.777	c4	b	right .
		519.68	520.32	c1	s.%-	assuming we're ==
		519.7	522.81	c5	qo	what about issues of playing sound files between the two platforms ?
		523.37	523.86	ca	h	um ==
		525.57	526.14	ca	s.%-	we have ==
		526.743	527.753	ca	s^ba	well that's a good point too .
	Z^I	527.99	529.83	c4	s	here's a - here's a crazy idea actually .
		528.64	529.01	ca	s^no	i don't know .
		530.546	534.646	c4	qh^cs	why don't you try and merge transcriber and thisl i.r ?
		535.02	535.99	ca	s.%-	well this is one of the reasons ==
		535.113	536.543	c4	s^df	they're both tcl interfaces .
		536.491	540.471	ca	s^r	this is the - one of the reasons that i'm gonna have uh dave gelbart - gelbart .
		540.903	543.023	ca	s^ba	having him volunteer to work on it is a really good thing .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		543.445	545.665	ca	s^df	because he's worked on the transcriber stuff .
		545.83	546.03	c4	s^aa	right .
		546.072	547.862	ca	s^df	and he's more familiar with tcl-t_k than i am .
		546.94	549.87	c4	s	and then you get - they - then you get the windows media playing for free .
		550.23	552.73	ca	s^ng	well - that's snack not - not transcriber .
		553.6	558.94	c4	s^aa s^df	right but the point is that the transcriber uses snack and then you can - but you can use a - a lot of the same functionality .
		558.879	563.239	ca	s^aa s	yeah yeah i mean i - i think thisl - thisl gui probably uses snack .
		558.94	559.22	c4	%-	and it's ==
		564.4	567.02	ca	s	and so my intention was just to base it on that .
		569.33	569.51	c4	s^bk	yeah .
		569.901	576.1	c4	s^cs	well my thought was is that it would be nice - it would be nice to have the running transcripts - um eh you know from speaker to speaker .
		570.48	571.13	ca	s.%-	and if it doesn't ==
		576.75	576.98	c4	qy^d^g	right ?
		577.11	580.56	c4	qy^rt	do you have - you have you know a speaker mark here and a speaker mark here ?
		580.183	583.233	ca	s^bk s^cc	right we'll have to figure out a user interface for that so .
		582.42	586.255	c4	fg s^cs	right well that - eh my thought was if you had like multitrans or whatever do it .
		586.83	587.18	c4	s^e	or whatever .
		587.052	587.582	ca	s^am	yeah .
		588.748	594.6	ca	s	it might be fairly difficult to get that to work in the little short segments we'd be talking about .
		594.6	596.866	ca	s^e	and having the search tools and so on .
		596.866	597.826	ca	s^cc	we - we can look into it .
		597.866	598.076	ca	fh	but ==
		597.971	598.171	c4	b	yeah .
T		599.237	603.507	c1	s.%-	the thing i was asking about with um free b_s_d is that it might be easier to ==
		603.949	609.178	c1	s.%-	get powerpoint shows running in free b_s_d than to get this other package running in ==
		609.891	613.631	ca	s^cc	yeah i mean we have to - i have to sit down and try it before i make too many judgments .
		613.809	614.099	c1	b	yeah .
		613.811	614.401	ca	fh	so uh ==
		615.83	616.27	ca	fh	um ==
		616.53	623.32	ca	s	my experience with the gnu compatibility library is really - it's just as hard and just as easy to port to any system .
		624.18	624.32	ca	qy^d^f^g	right ?
		624.32	625.81	ca	s	the windows system isn't any harder .
		625.81	627.69	ca	s^df	because it - it looks like a b_s_d system .
		627.905	628.285	c1	b	mm-hmm .
		628.47	632.71	ca	s	it's just you know - just like all of them the include files are a little different and the function calls are a little different .
		632.86	633.17	c1	b	right .
		633.261	635.221	ca	s	so i - it might be a little easier .
		635.221	636.511	ca	s^e	but it's not gonna be a lot easier .
		637.18	637.48	c1	b	o.k .
		638.959	640.939	c1	s^tc	so there was that demo which was one of the main ones .
		640.939	643.57	c1	s	{and} then we talked about - um some other stuff .
		643.57	648.47	c1	s	which would basically be - um showing off the - the transcriber interface itself .
		648.47	651.62	c1	s^cs	and as you say maybe we could even merge those in some sense .
		651.62	652.98	c1	fh	but - but um ==

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		653.96	661.18	c1	fh s	uh - and part of that was showing off what the speech-non-
		659.89	660.07	c3	b	- uh nonspeech stuff that thilo has done s- - looks like .
		661.462	662.512	c0	qy	yeah .
		662.456	662.746	ca	s^aa^rt	can i ask one more thing about thisl ?
		662.844	666.524	c0	qy^bu^d^rt.%-	mm-hmm .
		666.467	666.727	ca	s^aa^rt	so with the i_r stuff then you end up with a somewhat
		666.987	667.297	ca	s^2	prioritized um ?==
		667.08	667.42	c0	s^bk	mm-hmm .
		667.54	667.92	c0	s^bk^r	ranked .
		668.863	669.013	c0	b	excellent .
		669.18	679.334	c9	s^cs	excellent .
		679.736	680.136	c1	b^rt	yeah .
		680.56	683	c9	fh s^cs^e	so another idea i w- - t- - had just now actually for the
		683	688.36	c9	s^cs^e	demo was whether it might be of interest to sh- - to show
		689.33	689.67	c9	fh	some of the prosody - uh work that don's been doing .
		690.36	693.15	c9	s^cs	mm-hmm .
		693.22	696.18	c9	s	um actually show some of the features .
		696.7	697.07	c1	s.%-	and then show for instance a task like finding sentence
		696.729	698.599	c9	s^no	boundaries or finding turn boundaries .
		697.07	700.87	c1	s^na	um ==
		700.87	702.36	c1	s^df	you know you can show that graphically sort of what the
		702.47	708.57	c1	s	features are doing .
		702.87	703.21	c9	b	it - you know it doesn't work great but it's definitely giving
		709.16	710.74	c1	s^	us something .
		709.69	711.44	c9	s	well i think {it} ==
		710.74	711	c1	fh	i don't know if that would be of interest or not .
		711.44	713.12	c9	qr^rt	at <at> the very least we're gonna want something
		712.52	713.87	c2	h s	illustrative with that .
		714.18	714.61	c2	fh	cuz i'm gonna want to talk about it .
		715.741	716.771	c2	fh	and so i- - if there's something that shows it graphically it's
		717.736	723.05	c2	s^cc	much better than me just having a bullet point .
		720.03	720.34	c9	s^aa.%-	yeah .
P^L		722.869	729.419	c1	s	e pointing at something i don't know much about .
		723.126	723.376	c9	%-	i mean you're looking at this now .
		730.168	731.688	c1	s	so ==
		730.863	732.363	c9	qr^bu	are you looking at waves or matlab ?
		731.16	731.64	c2	s.%- ooo	um yeah i'm starting to .
		732.466	732.976	ca	s	and ==
		732.768	733.878	c9	s^bk qy^bu^d^rt	um ==
		733.42	733.69	c1	%-	yeah we can probably find some examples of different type
		734.341	734.591	c9	s^bk	of prosodic events going on .
		735.2	736.5	ca	s^fa s^bsc	yeah def- ==
		735.33	736.88	c1	s	s- - so when we here were having this demo meeting what
		735.43	735.79	c9	s^	we're sort of coming up with is that we wanna have all
		735.95	736.36	c2	b	these pieces together to first order by the end of the month .
						i- ==
						and then that'll give us a week or so .
						oh the end of this month or next month ?
						the end of ==
						this month .
						oh you mean like today ?
						ju- ==
						oh .
						oh sorry next month .
						june june june .
						bk^m next month .
						yeah .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		736.12	736.35	c9	s^fa	sorry .
		736.5	737.58	ca	s	today isn't june first .
		737.541	738.221	c5	s	there's another one .
		737.58	737.79	ca	qy^d^g	is it ?
		738.68	739.1	c4	b	exactly .
		738.703	745.834	c1	fg s	uh that'll - that'll give us - that'll give us a week or so to uh - to port things over to my laptop and make sure that works .
		740.245	740.475	c9	s^fa^r	<o.k> sorry .
		745.8	747.11	c9	s.%-	i think - i mean eh where ==
		745.834	745.894	c1	b	yeah .
		746.4	747.63	c2	s^bk s	yeah i mean i'll be here .
T		747.75	750.71	c9	s^cs.%-	yeah if d- - if don can sort of talk to whoever's ==
		750.71	753.89	c9	s^df	cuz we're doing this anyway as part of our - you know the research .
		751.069	751.279	c1	b	yeah .
		753.89	755.62	c9	s^e	visualizing what these features are doing .
		755.14	755.32	c1	b	yeah .
		755.62	756.38	c9	s.%-	and so either ==
		756.697	757.797	c9	s	it might not be integrated .
		757.797	760.047	c9	s	but it - it could potentially be in it .
		761.327	761.887	c9	s	could find some .
		761.47	763.13	c1	s	yeah well this is to an audience of researchers .
		763.13	763.9	c1	%-	so i mean you know ==
		763.96	766.47	c1	s	to let s- - the goal is to let them know what it is we're doing .
		764.558	765.408	c9	s	i mean it's different .
		766.47	766.87	c1	%-	so that's ==
		766.5	768.39	c9	s	i don't think anyone has done this on meeting data .
		768.39	769.93	c9	s^ba	so it might be neat you know .
		769.43	769.67	c1	b	yeah .
		771.654	771.804	c1	b	good .
G		773.88	774.23	c1	s^rt^t^tc	done with that .
		775.47	776.18	c1	qy^d^rt^t^tc	x.m.l tools ?
R		776.53	777.03	ca	fh	um ==
		777.83	779.93	ca	s	so i've been doing a bunch of x.m.l tools .
		779.93	783.735	ca	s	where you - we're sort of moving to x.m.l as the general format for everything .
		783.735	785.145	ca	s^ba	and i think that's definitely the right way to go .
		785.145	790.265	ca	s^df	because there are a lot of tools that let you do extraction and reformatting of x.m.l tools .
		790.803	791.203	ca	fh	um ==
		792.002	796.512	ca	s^cs	so yet again we should probably meet to talk about transcription formats in x.m.l .
		796.512	799.042	ca	s^df	because i'm not particularly happy with what we have now .
		799.525	800.805	ca	s	i mean it works with transcriber .
		800.805	803.135	ca	s	but it - it's a pain to use it in other tools .
		803.515	805.285	ca	s^df	uh because it doesn't mark start and end .
		806	807.75	c5	qy^d.%-	start and end of each ?==
		806.169	806.409	ca	h	uh ==
		806.556	806.686	c3	b	yeah .
		807.7	808.23	ca	s^2	utterance .
		808.322	808.822	c5	s^2^bk	utterance .
		808.822	809.392	c5	qy^bu^d	just marks ?
		809.11	810.54	ca	s	so it's implicit in - in there .
		810.54	812.33	ca	s	but you have to do a lot of processing to get it .
		810.68	810.97	c5	b	yeah .
		811.93	812.47	c5	s^aa	right .
		812.53	812.83	c5	s^aa^r	right .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		812.898	816.258	ca	s^cs	and so - and also i'd like to do the indirect time line business .
		816.721	820.131	ca	s	um but regardless i mean w- - that's something that you me and jane can talk about later .
		820.501	825.34	ca	fh s	um - but i've installed x_m_l tools of various sorts in various languages .
		825.34	831.009	ca	s	and so if people are interested in doing - extracting any information from any of these files .
		831.009	835.413	ca	s^e	either uh information on users because the user database is that way .
		835.852	842.688	ca	s^e	i'm converting the key files to x_m_l so that you can extract m- - uh various inf- - uh sorted information on individual meetings .
		839.512	839.822	c2	s^ba	cool .
		840.796	840.986	c3	b	yeah .
		843.224	844.524	ca	s^e	and then also the transcripts .
		844.664	845.784	ca	s^co	and so l- - just let me know there .
		845.784	847.214	ca	s	it's mostly java and perl .
		847.717	850.317	ca	s^cs	but we can get other languages too if - if that's desirable .
T^V		850.98	854.82	c9	qy^rt	oh quick question on that is - do we have the - the seat information ?
		855.63	857.24	c9	qy^e^rt	in - in the key files now ?
		856.24	856.45	c0	s^aa^rt	mm-hmm .
		856.7	859.54	ca	s^aap.%-	the seat information is on the key files for the ones which ==
		857.349	857.469	c0	b	ah ==
		858.04	858.58	c9	s^bk.%-	oh in ==
		860.39	861.11	ca	s	it's been recorded .
		860.83	861.32	c9	s^2	for the new one- - .
		861.32	861.76	c9	s^bk	o_k .
		861.42	861.62	ca	s^aa	yeah .
		861.94	862.44	c1	qy^bu^d^rt	seat ?
		862.02	862.27	c9	s^	ba great .
		862.86	863.5	c9	s^aa	sea- - yeah .
		863.262	863.942	ca	s	where - where you're sitting .
		864.17	866.36	c1	s^bk s	oh not - not the quality or anything no .
		864.93	865.05	c3	%-	n- ==
		865.98	866.26	ca	s^aa	right .
		866.55	867.05	c1	s^bk	o_k i see .
	BO	866.68	867.86	ca	s^ba^j	it's pretty soft and squishy .
		867.795	867.995	c1	b	yeah .
		868.962	869.112	c9	b	alright .
		869.391	869.541	c1	b	yeah .
		870.399	870.829	c1	fg	o_k .
		870.8	871.87	ca	s^j	oh but that might just be me .
		870.876	871.136	c9	b	o_k .
		872.28	872.46	ca	fh	um ==
		873.392	873.802	c9	b	alright .
		874.001	875.831	c1	s^j	that's more seat information than we wanted .
		874.874	875.564	c9	s^bd	never mind .
	BO end	875.663	876.923	c4	b	hmm ==
		876.196	881.136	c9	s:qr^rt	i'm just trying to figure out you know when morgan's voice appears on someone's microphone are they next to him or are they across from him .
	BO	877.819	879.139	ca	s^cs^j	maybe we should bleep that out .
		877.951	878.301	c1	b	yeah .
	BO end	879.178	879.768	c1	b	mmm yeah .
		881.22	881.58	c5	s^co	wait a minute .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		881.38	881.65	ca	s^aa	right .
		881.58	883.65	c5	qw^rt	how - how w- - eh where is it in the key file ?
		881.854	882.104	c1	b	yeah .
		884.23	885.05	ca	s^rt	the square bracket .
		885.86	887.683	c9	s	cuz i mean i haven't been putting it in .
		887.683	888.573	c9	s.%-	and - in by ==
		888.16	889.18	ca	s^bu	you haven't been putting it in .
		889.18	889.43	c9	s^aa	right .
		889.67	890.44	c9	s^r	i have not .
		889.74	890.09	c0	%-	well bu- ==
		890.07	890.85	ca	s^bk	oh o_k .
		890.62	891.87	c0	qy^bu^rt	isn't it always on the digits ?
		890.63	890.94	c9	fh	and ==
		890.7	891.61	c1	s^bu	some of these are missing .
		891.61	891.96	c1	qy^rt	aren't they ?
		892.64	894.1	c0	qy^bu^r^rt	isn't it always on the digits forms ?
		893.038	894.178	c1	s^t1.%-	some fall out of ==
		893.29	893.75	c9	%-	well it- ==
		893.54	895.75	ca	s^aa s^cs	yeah so we can go back and fill them in for the ones we have .
		894.193	894.323	c2	b	ooo .
		895.76	897	c9	s	i mean they're on th- - right - these .
		897	899.41	c9	s	but i just hadn't ever been putting it in the key files .
		898.46	899.44	c5	s.%-	yeah i - i never ==
		899.41	900.63	c9	s	and i don't think chuck was either .
		900.53	902.49	c5	s	i never knew we were supposed to put it in the key file .
		900.63	900.85	c9	%-	cuz ==
		901.93	903.3	ca	s	i had told you guys about it .
		903.12	904.29	c9	s^bk s^fa	oh so we're both sorry .
		903.3	903.51	ca	fh	but ==
		903.46	903.87	c5	qh^rt	oh really ?
		904.29	904.63	c9	fh	so ==
T		904.648	908.728	ca	s^df	i mean this is why i wanna use a g- - a tool to do it rather than the plain text .
		906.18	906.41	c9	b	o_k .
		908.728	911.238	ca	s^df	because with the plain text it's very easy to skip those things .
		909.14	909.6	c5	s^bk	o_k .
		912.19	912.43	c9	s^bk	o_k .
		912.5	912.75	ca	fh	so ==
		913.836	915.556	ca	s.%-	um if you use the edit-key ==
C		916.55	917.2	ca	s^bsc^rt	or key-edit .
		917.2	918.402	ca	s^rt	i think it's edit-key command .
C end		917.687	918.297	c3	s^aa^m	edit-key .
		918.547	919.367	ca	qy^rt	did i show you guys that ?
		919.716	919.876	c3	s^aa	yep .
		919.797	920.567	ca	s	i did show it to you .
		920.14	920.79	c5	s^na	you mentioned it .
		920.567	924.155	ca	s:s	but i think you both said no you'll just use text file .
		920.79	921.34	c5	s^aa	yeah yeah .
		924.05	924.28	c5	s^2	text .
		924.155	926.945	ca	fh s	um it has it in there a place to fill it in .
		926.398	926.848	c9	s^bk	o_k .
		927.06	927.75	c5	s^bk	o_k .
		929.27	932.03	ca	s	yeah and so if you don't fill it in you're not gonna get it in the meetings .
		929.77	930.08	c9	%-	so if ==
		932.29	932.64	ca	fh	so ==
		932.53	934.61	c9	s^df	right well i - i just realized i hadn't been doing it .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		934.167	934.417	ca	b	yep .
		934.61	935.7	c9	%-	and probably - so ==
T		936.67	939.96	ca	s^rt	yeah and then the other thing also that thilo noticed is on the microphone .
		936.697	937.137	c2	%-	u- ==
		940.102	943.192	ca	s^rt:s	on channel zero it says hand-held mike or crown mike .
		942.205	942.385	c9	b	yeah .
		943.197	943.577	c9	s^aa	right .
		943.405	944.705	ca	s^co	you actually have to say which one .
		944.66	945	c9	s^aa	i know .
		945	946.35	c9	s.%-	yeah i usually delete the ==
		945.155	945.345	ca	fh	so ==
		946.658	946.968	c9	%-	i don't ==
		946.83	948.07	c5	s^bk s	oh o_k i didn't do that either .
		947.258	948.098	c9	s.%-	maybe i forgot to d- ==
		947.54	947.77	c3	b	yeah .
		950.04	951.38	c5	s	takes me no time at all to edit these .
		950.37	951.07	c9	s.%-	but it's almost ==
		951.495	951.815	c9	fg	yeah .
		951.66	953.19	ca	s.%-	yeah that's cuz you kn- ==
		952.33	953.09	c5	s^2^j	i'm not doing anything .
		953.19	953.39	ca	s	i - i know why .
		955.168	957.658	c9	s:qw	and i was - i was looking at chuck's like oh what did chuck do .
		957.658	958.368	c9	s:s^cc	o_k i'll do that .
		958.558	958.678	c9	fh	so ==
		959.73	967.16	ca	s:s	and then uh also in a couple of places instead of filling the participants under participants they were filled in under description .
		960.56	961.14	c1	s^bk	ah o_k .
		967.67	967.94	c9	fg	uh ==
		967.951	969.141	ca	s	and so that's also a problem .
		970.82	971.31	ca	s^tc	so anyway .
		972.41	973.32	c9	s^cc	we will do better .
		972.8	973.2	ca	s^tc	that's it .
R		974.82	978.24	ca	fg s^rt	oh uh also i'm working on another version of this tool .
		978.808	982.835	ca	s^e	the - the one that shows up here that will flash yellow if the mike isn't connected .
		983.44	984.88	ca	s	and it's not quite ready to go yet .
		984.91	988.406	ca	s	because um it's hard to tell whether the mike's connected or not .
		988.496	993.72	ca	s^rt	because the best quality ones - the crown ones - are about the same level if they're off .
		994.012	996.362	ca	s^e	and no one's o- - off or if they're on and no one's talking .
		996.45	996.7	c2	b	huh .
		997.51	1000.47	ca	s	um these - these ones they are much easier .
		1000.61	1001.58	ca	s^df	there's a bigger difference .
		1002.25	1003.25	ca	s	so i'm working on that .
		1003.25	1004.35	ca	s	and it - it sorta works .
		1004.35	1005.91	ca	s^cc	and so eventually we will change to that .
		1005.91	1008.55	ca	s^df	and then you'll be able to see graphically if your mike is dropping in or out .
		1009.63	1011.79	c2	qy^rt	will that also include like batteries dying ?
		1011.65	1011.82	ca	s^aa	yep .
		1012.04	1014.44	c2	s^bu	just a- - any time the mike's putting out zeros basically .
		1013.1	1013.28	ca	s^aa	yep .
		1014.27	1014.44	ca	s^aa^r	yep .
		1015.29	1016.88	c5	s.%-	but with the screensaver kicking in it ==
		1015.35	1015.65	ca	%-	now ==
		1015.79	1015.98	c3	fg	but ==

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1015.98	1016.48	c3	b	y- - yeah .
		1016.98	1018.43	ca	s^cc	well i'll turn off the screensaver too .
		1017.09	1017.4	c2	b	yeah .
		1017.39	1017.55	c3	b	yeah .
		1018.74	1019	c2	s^fe	oops !
		1019.33	1022.6	ca	fh s	um the other thing is as i've said before it is actually on .
		1019.6	1020.34	c2	s	speaking of which .
		1023.03	1023.47	ca	s^e	the thing .
		1023.47	1024.57	ca	s	there's a little level meter .
		1024.86	1026.57	ca	s	but of course no one ever pays attention to it .
		1026.57	1030.48	ca	s	so i think having it on the screen is more easy to notice .
Z		1030.86	1033.6	c0	s^cs	it would be nice if - if these had little light indicators .
		1033.6	1034.68	c0	s^e.%-	little l_e_ds for ==
		1034.11	1034.59	ca	s^2^j	uh buzzer .
		1035.15	1035.79	c0	s^aa^j^r	yeah a buzzer .
		1035.99	1036.92	ca	s^fe	bamp bamp !
		1037.4	1038.62	c1	s^j	small shocks .
		1038.75	1038.99	c0	s^aa	yeah .
		1039.4	1039.9	c0	%-	actually ==
		1040.14	1041	c1	%-	administered to the ==
		1041.25	1041.55	c1	fg	o_k .
		1042.64	1043.91	ca	qy^d^rt^t^tc	o_k disk backup et cetera ?
Z end		1042.67	1042.93	c1	fg	oh ==
		1044.93	1051.23	ca	s	um i spoke with dave johnson about putting all the meeting recorder stuff on non-backed-up disk to save the overhead of backup .
		1051.67	1054.04	ca	s:s	and he pretty much said yeah you could do that if you want .
		1054.33	1055.56	ca	s^ba	but he thought it was a bad idea .
		1056.2	1063.23	ca	s^ba^rt	in fact what he said is doing the manual one - doing uh n_w archive to copy it - is a good idea .
		1063.52	1065.68	ca	s	and we should do that and have it backed up .
		1066.65	1069.96	ca	s	he w- - he's a firm believer in - in lots of different modalities of backup .
		1070.35	1071.7	ca	s	i mean his point was well taken .
		1071.85	1073.44	ca	s^df	this data cannot be recovered .
		1073.78	1074.06	c9	s^aa	yeah .
		1074.4	1079.52	ca	s	and so if a mistake is made and we lose the backup we should have the archive .
		1079.52	1082.48	ca	s	and if then a mistake is made and we lose the archive we should have the backup .
		1082.73	1087.99	c1	s.%-	well i guess it is true that even with something that's backed up it's not gonna ==
		1088.51	1093.55	c1	s	if it's stationary it's not going to go through the increment- - it's not gonna burden things in the incremental backups .
		1092.87	1094.51	ca	s^2	just - just the monthly full .
		1093.38	1093.55	c0	b	hmm ==
		1094.85	1096.44	c1	s^aa s	yeah so the monthly full will be a bear .
		1096.44	1096.54	c1	fh	but ==
		1096.6	1102.28	ca	s	yeah but he said that - that we sh- - shouldn't worry too much about that .
		1102.42	1104.1	ca	s^df	that we're getting a new backup system .
		1104.1	1108.56	ca	s	and we're far enough away from saturation on full backups that it's w- - probably o_k .
		1106.48	1106.68	c1	qy^bh	really ?
		1108.96	1119.43	ca	fh s	and uh so the only issue here is the timing between getting more disks and uh - recording meetings .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1120.09	1129.31	c1	s^bu	so i guess the idea is that we would be reserving the non-backed-up space for things that took less than twenty-four hours to recreate or something like that .
		1129.13	1133.34	ca	s^na	things that are recreatable easily and also - yeah basically things that are recreatable .
		1129.31	1129.54	c1	qy^d^g^rt	right ?
		1130.87	1131.26	c1	b	yeah .
		1132.67	1133.04	c1	b	yeah .
		1134.06	1135.75	ca	s	the expanded files and things like that .
		1137.06	1137.53	c1	b	o.k .
		1137.18	1138.4	ca	s	they take up a lot more room anyway .
		1139.19	1139.51	c1	b	yeah .
		1140.47	1141.55	ca	s	uh but we do need more disk .
		1142.23	1143.8	c1	s^bk	so we can get more disk .
		1143.67	1143.88	ca	s^aa	yeah .
		1143.8	1144.48	c1	fh	yeah - so .
		1145.21	1146.23	ca	s	and i - i think i agree with him .
		1146.23	1150.18	ca	s	i mean his point was well taken that if we lose one of these we cannot get it back .
		1150.6	1150.93	c1	s^bk	o.k .
	C	1154.37	1156.11	ca	s^t	i don't think there was any other et cetera there .
	C end	1157.67	1161	c1	s.%-	well i was allowing someone else to come up with something related that they had uh ==
	T	1163.11	1164.9	c4	s^bu	i thought you guys were gonna burn c.ds .
		1165.07	1168.39	ca	h s^ng	um unfortunately - we could burn c.ds but first of all it's a pain .
		1168.81	1169	c4	b	yeah .
		1168.87	1171.82	ca	s^df^rt	because you have to copy it down to the p.c and then burn it .
		1171.82	1173.38	ca	s^e^rt	and that's a multi-step procedure .
		1173.76	1178.55	ca	s	and second of all the - the write-once burners as opposed to a professional press don't last .
		1179.08	1179.28	c4	b	yeah .
		1179.35	1181.86	ca	s^ba	so i think burning them for distribution is fine .
		1182.19	1184.3	ca	s^ba	but burning them for backup is not a good idea .
		1184.44	1184.79	c4	s^bk	i see .
		1185.63	1188.56	ca	s^df	cuz th- - they - they fail after a couple years .
		1185.79	1186.03	c4	b	o.k .
		1188.92	1189.12	c4	b	alright .
	G	1191.16	1194	c0	s^t^tc	i do have uh uh - it's a different topic can i add one top- - topic .
		1194	1194.39	c0	qy^rt	we have time ?
	T	1194.71	1195.46	c0	s.%-	i wanted to ask ==
		1195.65	1201.63	c0	s^bu^rt	i know that uh that thilo you were um - bringing the channeltrans interface onto the windows machine .
		1202.46	1203.89	c3	s	yeah it's - it - basically it's done .
		1202.57	1203.41	c0	s.%-	and i wanted to know is th- ==
		1203.89	1204.01	c3	b	yeah .
		1204.28	1205.91	c0	s^bk^m	it's all done that's g- - wonderful great .
		1205.05	1205.17	c3	b	yeah .
		1206.21	1209.93	ca	s	yes since tcl-t.k runs on it basically things'll just work .
		1209.05	1209.26	c3	%-	yeah it ==
		1210.12	1213.79	c3	s	yeah it was just a problem with the snack version and the transcriber version .
		1213.79	1215.68	c3	fh	but it's solved .
		1215.92	1216.15	c3	fh	so ==
		1216.39	1218.11	c0	qy.%-	does - and that - does that mean i ?==
		1218.58	1219.56	c0	s	maybe i should know this but i don't .
		1219.56	1225.25	c0	qy^bu.%-	does this mean that the - that this could be por- - uh ported to a think-pad note- - or some other type of uh ?==

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1225.58	1228.26	c3	s^aa	yeah basically uh i did install it on my laptop .
		1228.26	1228.94	c3	fh	and - yeah .
		1229.42	1229.72	c0	s^bk	wonderful .
		1229.55	1229.9	c3	s	it worked .
		1230.23	1230.6	c0	s^bk^r	wonderful .
		1230.26	1230.54	c1	b	hmm !
		1231.22	1231.42	c1	s^ba	good .
		1233.83	1234.56	c1	qo^d^rt^t^tc	crosspads ?
		1234.82	1238.13	ca	s	uh {i} got an email from uh - james landay .
		1238.13	1242.1	ca	s^e:qy	who basically said if you're not using them could you return them .
		1242.44	1248.73	ca	s	so he said he doesn't need them he just periodically w- - at the end of each term sends out email to everyone who was recorded as having them .
		1249.17	1251.1	ca	s	and asks them if they're still using them .
		1251.09	1252.5	c1	s^bu	so we've never used them .
		1252.58	1253.29	c0	s^nd	we used them once .
		1253.42	1254.72	ca	s^	nd we - we used them a couple times .
		1253.46	1253.8	c1	qy^bu^d^rt	once ?
		1254	1254.29	c0	s^aa	mm-hmm .
		1254.46	1254.82	c5	qy^bu^d^rt	them ?
		1254.72	1254.99	ca	fh	but ==
		1254.78	1255.26	c0	s^aa^m	couple times .
		1254.82	1255.48	c5	qy^bu^d^e^rt	there's more than one ?
		1255.88	1256.53	ca	s^aa s^na	yeah we have two .
		1256.04	1256.2	c0	s^aa	yeah .
		1256.09	1256.45	c1	%-	i- ==
		1257.42	1257.84	ca	fg	um ==
		1258.6	1258.86	c1	fg	but ==
		1259.12	1260.77	ca	s.%-	my opinion on it is ==
		1260.77	1262.23	ca	s	first i never take notes anyway .
		1262.54	1263.81	ca	s	so i'm not gonna use it .
		1264.76	1268.29	ca	fh s	um and second it's another level of infrastructure that we have to deal with .
		1269.23	1274.61	c0	s^ba	and i have - uh so my - my feeling on it is that i think in principle it's a really nice idea .
		1274.61	1279.17	c0	s^ba	and you have the time tags which makes it better tha- - than just taking ra- - raw notes .
		1279.25	1284.2	c0	s	on the other hand i - the down side for me was that i think the pen is really noisy .
		1284.2	1285.96	c0	s:s	so you have ka- - kaplunk kaplunk kaplunk .
		1286.1	1287.94	c0	s.%-	and i - and i don't know if it's audible on the ==
		1288.26	1290.19	c0	s	but i - i sort of thought that was a disadvantage .
		1290.33	1291.12	c0	s	i do take notes .
		1291.12	1292.99	c0	s^cs	i mean i could be taking notes on these things .
		1292.99	1296.57	c0	s	and i guess the plus with the crosspads would be the time markings .
		1296.57	1296.72	c0	fh	but ==
		1297.27	1297.64	c0	s^no	i don't know .
		1298.18	1299.35	c3	fg qw^rt	uh what is a crosspad ?
		1300.27	1308.06	c1	s	so it's - it's um - it's a regular pad just a regular pad of paper but there's this pen which indicates position .
		1302.6	1302.94	c2	s^ft^t3	thank you .
		1308.75	1312.1	c1	s	and so you have time and position stuff stored .
		1308.83	1309.18	c3	s^bk	o_k .
		1312.1	1315.12	c1	s	so that you can - you have a record of whatever it is you've written .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1315.41	1315.66	c3	s^bk	o.k .
		1316.24	1317.26	ca	s	and then you can download it .
		1317.26	1320.55	ca	s	and they have o_c_r and searching and all sorts of things .
		1319.24	1319.52	c3	s^bk	o.k .
		1320.14	1320.32	c3	b	o.k .
		1321.43	1323.94	ca	s	so i- - if you take notes it's a great little device .
		1321.93	1322.54	c0	%-	could - mm-hmm ==
		1324.91	1326.18	ca	s	but i don't take notes .
		1326.25	1332.41	c1	s	and one of the reasons that it was brought up originally was because uh we were interested in - in higher-level things .
		1326.5	1326.78	ca	fh	so ==
		1332.41	1336.75	c1	s	not just the you know microphone stuff but also summarization and so forth .
		1336.75	1343.93	c1	s:qw	and the question is if you were going to go to some gold standard of what wa- - what was it that happened in the meeting - you know where would it come from .
		1344.11	1347.83	c1	fh s	and um i think that was one of the things .
		1344.36	1344.55	c3	b	yeah .
		1346.7	1346.91	ca	s^aa	yep .
		1347.04	1347.41	c9	s^aa	yeah .
		1347.83	1348.16	c1	qy^d^g^rt	right ?
		1348.16	1350.01	c1	s^ba	and so the - it seemed like a neat idea .
		1350.01	1351.44	c1	s^rt	we'll have a - you know have a scribe .
		1351.94	1355.69	c1	s	have somebody uh take good notes and then that's part of the record of the meeting .
		1356.15	1358.32	c1	s.%-	and then we did it once or twice and we sort of ==
		1357.4	1359.11	ca	s^aa s	yep and then just sort of died out .
		1358.84	1360.03	c1	s^j	probably chose the wrong scribe .
		1360.03	1361.63	c1	s.%-	but it was - it's uh ==
		1360.4	1361.17	ca	s^aa	yeah that's right .
		1360.48	1360.84	c9	%-	i mean ==
		1361.47	1362.31	c0	s	well i did it one time .
		1362.31	1363.16	c0	fh	but um ==
		1362.31	1362.57	ca	b	yep .
		1362.64	1362.9	c1	b	yeah .
		1363.6	1371.25	c0	s^cs	u- - but i guess the - the other thing i'm thinking is if we wanted that kind of thing i wonder if we'd lose that much by having someone be a scribe .
		1371.25	1373.72	c0	s^cs^e	by listening to the tape - to the recording afterwards .
		1374	1376.73	c0	s^e	and taking notes in some other interface .
		1376.7	1378.19	c5	s	i mean we're transcribing it anyways .
		1378.19	1378.98	c5	qw^rt	why do we need notes ?
		1379.32	1380.39	c0	s^nd	oh it's la- - it's useful .
		1379.48	1380.28	ca	s^ng	because that's summary .
		1380.39	1382.18	c0	s^ng	have a summary and high points .
		1380.42	1380.87	c9	%-	i think ==
		1380.73	1381.33	c1	s^m	summary .
		1381.3	1381.87	c9	%-	there's also ==
		1382.5	1383.37	c9	s.%-	there's this use that ==
		1383.99	1385.2	c5	s^cs	summarize it from the transcription .
		1384.14	1384.28	c9	%-	the ==
		1385.31	1387.6	c9	qo	well what if you're sitting there and you just wanna make an x_ ?
		1387.6	1388.69	c9	s	and you don't wanna take notes .
		1388.69	1389.72	c9	s.%-	and you're - you just wanna ==
		1389.96	1390.18	c5	s^2^rt	doodle .
		1390.11	1393.1	c9	s	get the summary of the transcript from this time location .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1393.1	1396.04	c9	s	like - you know and - and then while you're bored you don't do anything .
		1396.04	1398.24	c9	s	and once in a while maybe there's a joke and you put a x_ .
		1398.94	1399.19	c9	fh	<and> but ==
		1399.35	1404.1	c9	s^bs^cs	in - in other words you can use that just to highlight times in a very simple way .
Z^I		1404.1	1407.68	c9	s	also with - i was thinking - and i know morgan disagrees with me on this .
		1407.68	1415.9	c9	s^rt	but - suppose you have a group in here and you wanna let them note whenever they think there might be something later that they might not wanna distribute in terms of content .
		1416.65	1419.51	c9	s^cs	they could just sort of make an x_ near that point .
		1419.51	1426.49	c9	s^e	or a question mark that sort of alerts them that when they get the transcript back they c- - could get some red flags in that transcript region .
		1426.49	1427.69	c9	s^e	and they can then look at it .
		1427.69	1427.99	c9	fh	so ==
		1428.35	1429.59	c9	s	i know we haven't been using it .
		1429.62	1435.29	c9	s^cs	but i w- - i can imagine it being useful just for sort of marking time periods .
		1434.4	1434.59	ca	s^aa	right .
		1435.29	1437.44	c9	s^e	which you then get back in a transcript .
		1436.88	1437.03	c0	b	well .
P		1437.12	1444.95	c1	s^cs	i guess - so {w-} - you know what - what makes one think i- - is maybe we should actually schedule some periods where people go over something later .
		1437.44	1437.74	c9	fh	so ==
P end		1444.95	1448.61	c1	s	and - and - and put some kind of summary or something .
		1448.83	1454.74	c1	fh s^cs	uh you know some - there'd be some scribe who would actually listen w- - who'd agreed to actually listen to the whole thing .
		1455.06	1458.63	c1	s	not transcribe it but just sort of write down things that struck them as important .
		1459.11	1465.75	c1	s	but then you don't - you don't have the time reference - uh that you'd have if you had it live .
		1462.77	1463.15	c9	s^aa	right .
		1465.9	1469.52	c9	s	and you don't have a lot of other cues that might be useful .
		1467.84	1468.2	c1	s^am	yeah .
T		1468.59	1471.54	c5	qw	how do you synchronize the time in the crosspad and the time of the recording ?
		1469.88	1470.06	c9	fh	so ==
T^Y		1470.88	1473.17	ca	s	i mean that was one of the issues we talked about originally .
		1473.17	1477.85	ca	s	and that- - that's w- - part of the difficulty is that we need an infrastructure for using the time - the crosspads .
		1477.85	1479.61	ca	s.%-	and so that means synchronizing the time ==
		1478.99	1479.46	c9	b	{oh} .
		1480.92	1482.4	ca	s	you know you want it pretty close .
		1480.95	1481.21	c0	b	mm-hmm .
		1482.66	1485.87	ca	s	and there's a fair amount of skew because it's a hand-held unit with a battery .
		1486.17	1486.98	c0	fg s.%-	well when - when i d- ==
		1486.54	1486.93	ca	%-	and so you ==
		1487.48	1487.71	c0	b	o.k .
		1487.89	1492.28	ca	s^co	so you have to synchronize at the beginning of each meeting all the pads that are being used .
		1493.08	1494.89	ca	s^df	so that it's synchronized with the time on that .
		1494.89	1497.99	ca	s^e	and then you have to download to an application .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1497.99	1503.82	ca	s^e	and then you have to figure out what the data formats are and convert it over if you wanna do anything with this information .
	T^Y end	1504.64	1506.55	ca	s.%-	and so there's a lot of infrastructure which ==
		1504.84	1505.62	c0	b	w- - mm-hmm .
		1505.11	1505.25	c4	qw.%-	why ?==
		1506.86	1507.83	c0	s^cs^nd	there is an alternative .
		1507.15	1507.84	ca	%-	unless someone ==
		1508.3	1509.18	c0	s^r	there is an alternative .
		1509.18	1513.25	c0	s^na	i mean it's still there's uh - you know your point stands about there be - needing to be an infrastructure .
		1513.25	1516.65	c0	s	but it doesn't have to be synchronized with the little clock's timer on it .
		1516.65	1520.1	c0	s	you c- - i mean i - when i - when i did it i synchronized it by voice .
		1520.1	1522.41	c0	s^e	by whispering one two three four onto the microphone .
		1521.51	1521.76	ca	b	hmm ==
		1522.77	1523.94	c0	fh s^f	and uh you know .
		1524.64	1526.5	ca	s	well but then there's the infrastructure at the other end .
		1525.22	1525.38	c4	b	right .
		1526.5	1528.81	ca	s^e	which someone has to listen to that and find that point .
		1528.94	1529.62	c0	s^na	yeah it's transcribed .
		1529.26	1529.97	ca	s^e	and then mark it .
		1529.62	1530.4	c0	s	it's in the transcript .
		1530.56	1530.9	ca	fh	so ==
		1530.76	1530.93	c0	b	yeah .
		1532.04	1532.83	c0	s^r	well it's in the transcript .
		1532.23	1535.25	c9	qy^rt	juh; well could we keep one of these things for another year ?
		1535.25	1536.8	c9	s^df.%-	would h- - i mean is there a big cau- ==
		1536.24	1537.97	ca	s^na	we can keep all - both of them for the whole whole year .
		1537.17	1538.82	c9	s^df.%-	just - just in case we ==
		1537.97	1538.59	ca	s.%-	i mean it's just ==
		1539.22	1543.01	c9	s	even maybe some of the transcribers who might be wanting to annotate .
		1543.76	1546.73	c9	s	uh f- - just there's a bunch of things that might be neat to do .
		1546.78	1551.68	c9	s	but i - it might not be the case that we can actually synchronize them and then do all the infrastructure .
		1551.68	1553.35	c9	s^cs	but we could at least try it out .
		1553.58	1564.83	c1	fg s	well one thing that we might try um is on some set of meetings - some collection of meetings - maybe e_d_u is the right one or maybe something else .
		1565.58	1570.41	c1	s^cs	we - we get somebody to buy into the idea of doing this as part of the task .
		1570.41	1570.86	c1	%-	i mean ==
		1570.71	1570.93	c9	s^aa	right .
		1571.38	1579.31	c1	s	uh part of the reason - i think part of the reason that adam was so interested in uh the speechcorder sort of f- - idea from the beginning is he said from the beginning he hated taking notes .
		1579.06	1579.3	ca	s^aa	yep .
		1579.31	1579.45	c1	fh	and ==
		1580.03	1580.63	c1	s^e	and so forth .
		1580.63	1582.99	c1	fh s	so and - and jane is more into it .
		1582.99	1584.35	c1	fh	but eh uh ==
		1584.63	1586.66	c1	s^no	you know i don't know if you wanna really do - do this all the time .
		1586.66	1589.68	c1	s^cs	so i think the thing is to - to get someone to actually buy into it .
		1589.68	1592.25	c1	s^cs^e	and have at least some series of meetings where we do it .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1593.23	1593.81	c1	fh	um ==
		1594.48	1596.4	c1	s	and if so it's probably worth having one .
		1596.71	1599.36	c1	s.%-	the p- - the - the problem with the - the more extended view ==
		1599.36	1603.47	c1	s^e.%-	all these other you know with uh quibbling about particular applications of it ==
		1603.47	1608.73	c1	s^e	is that it looks like it's hard to get people to um - uh - routinely use it .
		1608.73	1610.55	c1	s	i mean it just hasn't happened anyway .
		1611.03	1613.23	c1	s^cs.%-	but maybe if we can get a person to ==
		1612.46	1616.21	c9	s	yeah i don't think it has to be part of a- - what everybody does in a meeting .
		1616.21	1620.38	c9	s^ba	but it might be a useful neat part of the project .
		1620.38	1625.38	c9	s^df	that we can you know show off as a mechanism for synchronizing events in time .
		1625.93	1627.82	c9	s^df^e	that happen that you just wanna make a note of .
		1627.82	1629.4	c9	s	like what jane was talking about .
		1629.77	1632.39	c9	s^e	with some later browsing just - just as a convenience .
		1632.8	1636.3	c9	s	even if it's not a full-blown note taking substitute .
Z^M		1636.25	1640.92	c4	s^cs	well if you wanted to do that maybe the right architecture for it is to get a p_d_a with a wireless card .
		1641.32	1645.99	c4	s	and - and that way you can synchronize very easily with the - the - the meeting .
		1645.99	1649.28	c4	s^df	because you'll be synchroni- - you can synchronize with the - the linux server .
		1650.77	1651.99	c4	fh	and uh ==
		1651.55	1653.03	c9	qw.%-	so what kind of input would you be ?==
		1653.26	1655.01	c4	s.%-	so - so i mean if you're not worried about ==
		1653.93	1654.38	ca	s^cs	buttons .
		1654.8	1656.54	c9	qy^d.%-	you'd just be pressing like a - a ?==
		1656.2	1657.28	c4	s^cs	well - well you have a p_d_a .
		1657.28	1660.4	c4	s^cs^e	and may- - and you could have the same sort of x_ interface or whatever .
		1660.4	1663.13	c4	s	i mean you'd have to do a little eh - a little bit of coding to do it .
		1660.84	1661.07	c9	b	mm-hmm .
		1663.97	1665.35	c4	s.%-	but you could imagine ==
		1664.4	1664.68	c9	s^aa	yeah .
		1665.01	1665.42	c9	s^ba	that be good .
		1665.64	1670.23	c4	s	i mean if - if all you really wanted was - you didn't want this secondary note-taking channel .
		1670.23	1673.68	c4	s	but just sort of being able to use m- - markers of some sort .
		1674.17	1676.85	c4	s^cs	a p_d_a with a l- - a wireless card would be the - probably the right way to go .
		1677.87	1679.1	c4	s^cs	i mean even buttons you could do .
		1679.68	1680.04	c4	s^e	sort of .
		1680.04	1680.73	c4	s	i mean as you said .
		1680.19	1684.09	ca	s^cs	i mean for what - what you've been describing buttons would be even more convenient than anything else .
		1682.32	1682.59	c9	%-	m- ==
		1682.83	1683.21	c9	s^aa	right .
		1684.09	1684.25	ca	qy^d^g^rt	right ?
		1684.16	1685.45	c9	s^ba	that would be fine too .
		1684.25	1684.71	ca	s.%-	you have the ==
		1684.49	1684.69	c4	s^aa	right .
		1685.45	1688.05	c9	s	i mean i don't have you know grandiose ideas in mind .
		1688.05	1691.88	c9	s:s	but i'm just sort of thinking well we've - we're getting into the next year now .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1691.88	1694.45	c9	s	and we have a lot of these things worked out at - in terms of the speech .
		1694.45	1696.53	c9	s	maybe somebody will be interested in this .
		1696.53	1696.71	c9	fh	and ==
		1697.15	1698.35	c0	s^ba	i like this p_d_a idea .
		1698.69	1698.95	c0	b	yeah .
		1698.96	1700.68	c1	s^na	yeah i do like the idea of having a couple buttons .
		1699.02	1699.63	ca	s.%-	well i'm sure there would ==
		1699.2	1699.62	c9	b	yeah .
		1700.68	1702.36	c1	s^e:s^fe	where like one - one button was uh-oh .
		1700.96	1701.29	c0	b	yeah .
		1702.74	1704.34	c1	s^e:s^ba	and then another button was that's great .
		1704.34	1705.91	c1	s.%-	and another button that's f- ==
		1704.34	1707.4	c9	s^2:s	or like this is my i'm supposed to do this kind of button .
		1705.13	1705.43	c0	b	yeah .
		1707.4	1708.7	c9	s.%-	like i better remember to ==
		1708.42	1709.05	ca	s^2	action item .
		1709.18	1709.62	c9	s^aa	yeah .
		1709.72	1710.51	c9	s^e	something like that .
		1710.51	1710.72	c9	%-	or ==
T		1711.52	1713.67	ca	s^ba	i mean i think the crosspad idea is a good one .
		1711.77	1712.04	c0	%-	and then ==
		1712.19	1712.44	c0	b	uh-huh .
		1714.44	1716.55	ca	s^df	it's just a question of getting people to use it .
		1716.55	1720.17	ca	s^df	and getting the infrastructure set up in such a way that it's not a lot of extra work .
		1720.69	1722.67	ca	s^df	i mean that's part of the reason why it hasn't happened .
		1722.67	1726.19	ca	s^e	is that {it} - it's been a lot of extra work for me .
		1723.56	1723.92	c9	b	yeah .
		1725.74	1725.95	c1	b	right .
		1726.68	1726.86	ca	fh	and ==
		1726.86	1727.83	c0	s^nd	well and not just for you .
		1727.85	1728.1	c9	%-	w- ==
		1728	1733.11	c0	s^ng	but it's also it has this problem of having to go from an analog to a d- - a digital record too .
		1733.11	1733.79	c0	qy^g^rt	doesn't it ?
		1733.75	1738.01	ca	s^	aap well it's digital but it's in a format that is not particularly standard .
		1733.79	1734.02	c0	%-	i mean ==
		1736.81	1742.74	c0	s	but i mean say if i- - if - if you're writing - if you're writing notes in it does - it - it can't do handwriting recognition .
		1742.74	1743.03	c0	qy^d^g^rt	right ?
		1742.99	1747.53	c1	s^aa s^na	no no but it's just - it's just storing the pixel informa- - position information .
		1745.93	1746.27	c0	s^bk	o_k .
		1747.53	1748.08	c1	s	it's all digital .
		1749.34	1756.82	c0	s.%-	i - i guess what i'm thinking is that the p_d_a solution you h- - you have it already without needing to go from the pixelization to a - to a ==
		1757.73	1757.89	c0	%-	i mean ==
		1757.73	1757.98	c1	s^aa	right .
		1758.65	1759.06	c1	s.%-	you don't have to ==
		1758.67	1760.28	c4	s^na	the transfer function is less errorful .
		1760.56	1760.78	c4	s^aa	yes .
		1760.68	1761.64	c9	s^aa	yeah yeah .
		1760.86	1761.01	c1	s^aa	yeah .
		1761	1761.69	c0	s^ba	oh nicely put .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1761.88	1762.63	c1	b	yeah yeah .
		1762.49	1767.82	c9	s	well it also - it's maybe realistic cuz people are supposed to be bringing their p_d_as to the meeting eventually .
		1767.82	1768.1	c9	qy^d^g^rt	right ?
		1768.21	1769.33	c9	s^df.%-	that's why we have this little ==
		1769.83	1772.21	c9	s	i don't know what - i don't wanna cause more work for anyone .
		1772.21	1775.96	c9	s	but i can imagine some interesting things that you could do with it .
		1775.96	1777.57	c9	s.%-	and so if we don't have to return it ==
		1777.78	1779.51	c9	s.%-	and we can keep it for a year ==
		1779.66	1779.97	c9	s^no	i don't know .
		1780.27	1782.69	ca	fg s	well w- - we don't - we certainly don't have to return it .
		1782.69	1787.1	ca	s^df	as i said all - all he said is that if you're not using it could you return it if you are using it feel free to keep it .
		1788.05	1789.84	ca	s	the point is that we haven't used it at all .
		1790.1	1791.27	ca	s:qy	and are we going to .
		1791.46	1792.13	c1	s.%-	so we have no ==
		1792.13	1794.66	c1	s^co	but - uh by i - i would suggest you return one .
		1795.02	1795.28	ca	s^bk	o.k .
		1795.18	1795.48	c9	s^aa	yeah .
		1795.62	1797.9	c1	s^df	because we - we you know we - we haven't used it at all .
		1796.61	1797	c9	%-	we c- ==
		1798.01	1800.53	c1	s	we have some aspirations of using them .
		1798.12	1799.3	c9	s^cs	one would probably be fine .
		1800.31	1802	c9	s^cs	maybe we could do like a student project .
		1800.53	1800.77	c1	fh	and ==
		1802	1807.05	c9	s^ba	you know maybe someone who wants to do this as their main like s- - project for something would be cool .
		1804.88	1805.12	c1	b	yeah .
		1807.17	1807.39	ca	b	yep .
		1808.71	1812.13	ca	s^cs	i mean if we had them out and sitting on the table people might use them a little more .
BO		1812.48	1816.18	c1	s^cs^j	maybe jeremy could sit in some meetings and press a button when there - when - when somebody laughed .
		1812.65	1813.52	ca	s.%-	although there is a little ==
		1816.98	1818.88	c9	fg s.%-	well i'm - yeah that's not a bad ==
		1817.88	1818.49	c1	b	yeah yeah .
		1819.14	1819.43	c1	b	yeah .
		1819.27	1824.18	c9	s	jeremy's gonna be an - he's a new student starting on modeling brea- - breath and laughter actually .
		1824.18	1825.55	c9	s^rt	which sounds funny .
		1825.55	1826.65	c9	s^ba^rt	but i think it should be cool .
		1826.75	1826.86	c9	fh	so ==
		1826.81	1827.17	c1	b	yeah .
		1826.95	1827.76	ca	s^j	sounds breathy to me .
		1828.69	1828.78	c9	b	o.k .
		1829.03	1829.49	ca	s.%-	breath and lau- ==
		1829.09	1829.54	c9	s^j	ha-ha-ha .
		1829.49	1829.95	ca	s^j^m	ha-ha-ha-ha .
		1830.95	1831.52	ca	s^j^r	ha-ha-ha-ha .
		1832.36	1832.91	c1	s^fe	well dear !
		1833.71	1834.14	ca	fg	um ==
		1835.21	1835.4	c1	b	hmm .
T		1836.38	1837.36	ca	s	that reminded me of something .
		1837.44	1837.75	ca	s^bd	oh well .
		1837.75	1838.17	ca	s^bd	too late .
		1838.52	1839.16	ca	s^bd	it slipped out .
		1839.67	1839.96	c1	fg	o.k .

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1839.68	1839.91	ca	fg	oh !
		1839.79	1840.73	c9	qy^rt^t3	you're - you're gonna tease me ?
		1839.91	1840.32	ca	s	equipment .
		1840.73	1840.98	c9	s^bk	o_k .
		1841.1	1841.59	ca	%-	ordered ==
		1841.59	1842.74	ca	s	uh well i'm always gonna do that .
		1842.84	1847.02	ca	fh s	w- - uh we ordered uh more wireless .
		1847.12	1849.19	ca	s	and so they should be coming in at some point .
		1848.56	1848.89	c9	s^ba	great .
		1849.64	1853.05	ca	s^cc	and then at the same time i'll probably rewire the room as per jane's suggestion .
		1853.05	1859.9	ca	s^df	so that uh the first n_ channels are wireless - eh are the m- - the close-talking - and the next n_ are far-field .
BO		1861.22	1861.82	c1	s	you know what he means .
		1861.82	1862.93	c1	qh^rt	but isn't that funny sounding ?
		1862.93	1864.55	c1	s:s	we ordered more wireless .
		1864.55	1865.89	c1	s	it's like wires are the things .
		1865.89	1869.52	c1	s	so you're wiring - you're - you're - you - we're - we ordered more absence of the thing .
		1871.89	1875.45	c9	s^j	that's a very philosophical statement from morgan .
		1874.48	1875.63	ca	s^j	wired less wired more .
		1875.45	1877.77	c9	s	i just - it's sort of a anachronism .
		1877.77	1878.54	c9	s.%-	i mean it's like ==
		1878.94	1879.41	c9	s	it's great .
G		1880.61	1881.46	ca	qy^rt^t^tc	should we do digits ?
		1880.75	1881.12	c1	fg	anyway .
		1881.46	1882.17	ca	qy^rt^t^tc	do we have anything else ?
		1882.41	1882.61	c1	s^aa	yeah .
		1883.61	1883.7	c0	b	o_k .
R		1884.11	1893.79	c1	s	i mean there's - there's all this stuff going on uh between uh andreas and - and - and dave and chuck and others with various kinds of runs .
		1893.79	1896.21	c1	fh s^e	uh um recognition runs .
		1896.21	1897.92	c1	s	trying to figure things out about the features .
		1897.92	1899.98	c1	s	but it's - it's all sort of in process .
		1899.98	1901.16	c1	s	so there's not much to say .
		1901.84	1902.2	c1	s^e	right now .
G		1904.7	1908.98	c1	fh qh^cs^t	uh why don't we start with our - our esteemed guest ?
		1909.22	1909.72	c4	fg	o_k .
		1910.16	1910.65	c4	fh	alright .
T	OT	1910.97	1913.44	ca	s^co.%-	so just the transcript number and then the - then the ==
		1911.59	1911.84	c4	%-	this is ==
		1913.72	1914.99	c4	s^aa s^na	yes this is number two for me today .
		1915.26	1917.05	c1	s.%-	see all you have to do is go away to move way up in the ==
		1916.26	1916.56	c4	b	oh .
		1916.94	1918.14	c9	s^cs	we could do simultaneous .
		1918.9	1919.57	c9	s^df	initiate him .
		1920.21	1921.01	c1	s^na	we - we could .
		1920.55	1921.52	ca	qy^rt	should we do simultaneous ?
		1921.24	1925.6	c9	fg qy	well i'm just thinking are you gonna try to save the data before this next group comes in ?
		1921.91	1922.13	c1	s^aa	yeah .
		1925.96	1926.14	ca	s^aa	yeah .
		1926.2	1926.75	ca	s^aa	absolutely .
		1926.83	1928.73	c9	s^cs^r^rt	yeah so we might wanna do it simultaneous .
		1927.99	1929.01	ca	s^na	i mean you hav- - sorta have to .
		1929.3	1930.04	c9	s^aa %-	right so ==

L-1	L-2	START	STOP	SPKR	DA	UTTERANCE
		1929.76	1931.91	c1	fg s^co	well o_k so let's do one of those simultaneous ones .
		1930.54	1933.02	c9	s	so we might n- - we might need to do that actually .
		1931.87	1932.13	ca	b	o_k .
		1932.12	1932.81	c1	s^	ba that sounds good .
		1932.97	1933.13	c4	b	o_k .
		1933.51	1934.07	c1	qy^d^rt	everybody ready ?
		1934.16	1934.29	c0	s^aa	yeah .
		1934.98	1935.3	c1	s	a one .
		1935.12	1937.27	c9	s^co	you have to plug your ears by the way uh eric .
		1936.69	1937.36	ca	s^na	well i have to .
		1937.27	1937.44	c9	%-	or ==
		1937.36	1938.44	ca	s^no	i don't know about other people .
		1937.44	1939.12	c9	s^df	or you start laughing .
		1937.49	1938.37	c4	s^bk	o_k alright .
		1937.64	1938.37	c3	s^nd	you don't have to .
		1939.27	1939.72	c1	fg	o_k .
		1940.11	1942.4	c1	s	a one and a two and a three .
OT		1942.77	1974.23	ca		DIGIT_TASK
		1943.01	1977.36	c1		DIGIT_TASK
		1943.16	1980.52	c3		DIGIT_TASK
		1943.18	1977.73	c5		DIGIT_TASK
		1943.22	1977.67	c2		DIGIT_TASK
		1943.39	1973.63	c9		DIGIT_TASK
		1943.59	1981.39	c0		DIGIT_TASK
		1943.76	1978.71	c4		DIGIT_TASK
BE		1984.55	1986.66	c1	z	o_k babble take five .

B Dialog Act Labels

To be completed. See [2] for further information.

C Index of Tags

To be completed.