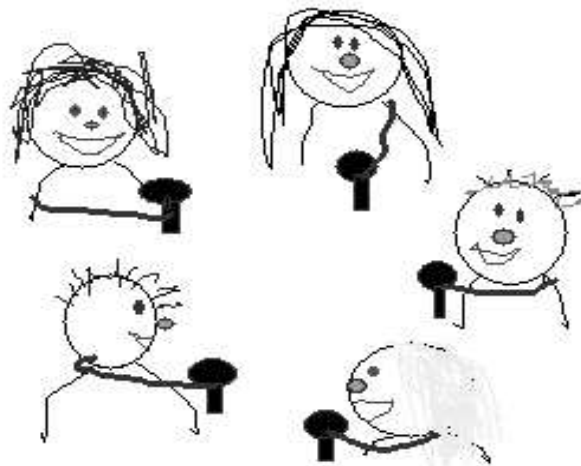


Draft 1: Labeling Guide for Dialogue Act Tags in the Meeting Recorder Meetings



by
Sonali Bhagat
Raj Dhillon
Hannah Carvey
Ashley Krupski
Liz Shriberg
August 8, 2002

Table of Contents

Introduction	4
Section 1: How to Label the Dialogue	5
Part 1: Tiered Labels	5
Part 2: More Than One DA Per Utterance	6
Part 3: Diacritics	7
Part 4: Using TableTrans	8
Section 2: Mapping from Switchboard DAMSL to MRDA	11
Section 3: Listing of Tags	14
Section 4: Explanation of Tags	16
s	17
sj	18
qr, qrr	20-21
qw, qy	21-22
%, %-, %--	23-24
t3, x	24-25
b	26
bk, fg	27-29
fh	29
aa	30
h	31
qh, bh	32
t1	33
aap, arp, nd	34-35
ng, ar	35-36
am, na	36-37
no.	37
co	39
fx, cc	40
cs, oo	41
qo, tc	42
ba, bc	44
bsc, bd, br, bs	45-46
bu	47
df	48
fa	49
by	49
2, e, t	50-51
d, f, g.	51-52
j	53
m, r	54
rt	55
fe, ft, fw	56-57

Section 5: Common Dialogue Act Sequences	58
Section 6: Complex Dialogue Act Sequences	61
Works Cited	63
Index of Tags in Alphabetical Order	64

Introduction

This labeling guide is adopted from work on the DAMSL-Switchboard recordings and the accompanying manual written in 1997. The DAMSL-Switchboard manual for labeling one-on-one phone conversations provided a useful starting point for the types of dialogue acts that arose in the ICSI meeting corpus. The tagset for labeling meetings presented here has been adapted as necessary to include different and new areas which came up due to the multi-speaker setting of meetings.

The pragmatic meanings of utterances were examined and a list of all the possible categories of dialogue acts was compiled. The reasoning behind the tagset presented in this manual reflects the intent to build the capability of automatically annotating meetings with DAs.

Future work involves incorporating information from dialogue act labeling for annotating adjacency pairs and "hot spots" (identifiable areas in segments of speech in which tension, excitement, or memorable correspondence exists between speakers).

Section 1

How to Label Dialogue:

Terminology

Here is some rudimentary terminology used to explain how to label the dialogue:

label = the entire set of dialogue acts applicable to a whole utterance

dialogue act (DA) = the tag or sequence of tags in a tiered form pertaining to an utterance or a portion of an utterance

tag = the individual component(s) of a dialogue act

Examples of Terminology in Three Hypothetical Utterances:

1. s^aa.%- Yeah, ri--
2. s.%-- I think spontaneous speech annotation is fun and ==
3. fg|qy^rt^bu So - so, | are you saying that we should stick to the original implementation?

1. label: <s^aa.%->
 dialogue acts: <s^aa> and <%->
 tags: <s>, <aa>, <%->
2. label: <s.%-->
 dialogue acts: <s>, <%-->
 tags: <s> <%-->
3. label: <fg|qy^rt^bu>
 dialogue acts: <fg>, <qy^rt^bu>
 tags: <fg>, <qy>, <rt>, <bu>

Part 1. Tiered Labels

The first-tier tag represents the basic form of the dialogue act. Sometimes, only one tag is needed to express the DA of an utterance. If more than one tag is necessary, tags are entered in a hierarchical form using carets < ^ > to separate the tiers. Every discernible utterance is labeled as a statement, question, or backchannel using the first-tier tags from Set 1. Once the first-tier has been labeled, the remaining tiers can be labeled in any order. Non-discernible or incomplete utterances are labeled using the tags from Set 2. Only the tags from Set 1 can be used in the first tier; these tags cannot be used elsewhere.

Set 1: First-Tier Tags for Discernible Utterances

s	sj	qh	qo	qr	qrr	qw	qy	b	fg	fh	h
---	----	----	----	----	-----	----	----	---	----	----	---

Set 2: Non-Discernible or Incomplete Utterances

x	%	%-	%--
---	---	----	-----

Part 2. More Than One DA Per Utterance

In some cases, one DA does not suffice to represent an utterance and two DAs are required. There are two ways of using more than one DA per utterance: with a pipe bar < | > or a period < . >.

The pipe bar should be used when sequential portions of the utterance require different characterizations. The pipe bar should be used when an utterance split is not needed but the beginning or ending of the utterance seems to serve a different role than can be characterized by one DA. Often, these brief phrases are Floor Holders or Floor Grabbers. The pipe bar separates these brief phrases from the rest of the utterance. (Splits are made when an utterance has more than one semantically and informational distinct conversational function.)

The pipe bar should be used in both the DA field and the TRANS field (transcription).

Example from Data (Bmr012):

94.861-99.771 c4 fg|s^t um - | everyone should have at least two forms possibly three in front of you depending on who you are .

A period should be used when providing information on the completeness of the utterance. Here are some examples of tiered DA labels and disruption forms separated by periods:

s^ar.%-- qy^d^rt.%- sj.%- s.x

The tags in Set 2 comprise entire DAs and therefore cannot be used as part of tiered DAs. An utterance may be labeled using both a tiered DA and tags from Set 2. The exception to this rule is Indecipherable <%> (page 23) which may not be accompanied by other DAs since <%> indicates that an annotator cannot comprehend the utterance sufficiently enough to label it with another DA.

The distinction between the use of the pipe bar and a period exists in how an utterance

can be divided. When dividing an utterance with a pipe bar, a segment of the utterance on one side of the pipe bar will be annotated with a particular DA that is different from the DA used to annotate the other side. The pipe bar exists as a clear boundary which marks where one DA ends and another begins in a single utterance. With regard to periods, no such clear boundary exists. A disruption form is actually an addition to an utterance as a whole, rather than being a separate segment of an utterance that can be marked with a pipe bar. It is also very difficult to use a pipe bar to mark where an interruption begins or where a speaker abandons his utterance, especially since the tiered DA on one side of the pipe bar may also apply to the other side where the disruption form is marked. In saying that a disruption form is an addition to an utterance as a whole, the reason why disruption forms are not used as tags in tiered DAs is because the tags used in the tiered DAs apply to the entire utterance. Disruption forms, however, usually apply only to the end of the utterance. For this reason and that of the inability to use pipe bars in conjunction with disruption forms, the use of periods with disruption forms is necessitated.

Part 3. Diacritics

Diacritics are used when it is impossible to decide upon a tag to label an utterance due to lack of background information of the conversation or speakers in the meeting. The sixth utterance in the example below implements a diacritic, an asterisk <*>. The asterisk separates the two tags upon which an annotator cannot decide.

Example from Data (Bmr013):

1586.76-1590.2	c8	s^bu?	so i - i understand that's what you were saying about your problem with minimum .
1589.55-1589.7	c1	s^aa	yeah .
1589.83-1590.04	c1	fg	and ==
1590.34-1593.3	c1	s	yeah and - and i had - i had uh - specific problems with ==
1590.41-1590.94	c8	s	i get it .
1592.21-1595.47	c8	s^bs*bu	so new - use ninetieth quartile rather than minimum .
1594.33-1594.48	c1	s^aa	yeah .
1594.98-1595.1	c1	s^aa	yeah .

The background information needed to discern the correct DA can be reported in the comment column. Because this column is usually reserved for personal notes of the annotator, the explanation on the use of the diacritic should be enclosed by an asterisk on the left and right hand side in the comment field.

end times for each utterance and the third column denotes the speaker. The fourth and fifth columns are where the DA and AP labels are entered. (This manual does not provide instructions on how to label AP's.) The comment field is located in the sixth column. This column is primarily for the annotator to write notes regarding an utterance. The last column, under the "Trans" heading, provides the transcript of the utterances.

In order to label a meeting, go to the File Menu and select "Open Annotation File." Another menu will pop up providing three formats to work with. Choose "Table Format." A window will appear with a Feature List and a Delimiter. Simply press "OK." In a few seconds, the segment of the meeting to be transcribed will appear.

The Time-Segmented Portion shows the annotator a series of utterances in chronological order. This portion cannot be modified.

2. TableTrans Features

<i>Command</i>	<i>Action</i>
CHANGING THE TRANSCRIPT	
Ctrl-s	Split the current row at the location of the cursor in the TRANS field.
Ctrl-m	Merge the current row with the next row by the same speaker.
MOVING AROUND WITHIN A FIELD(S)	
Ctrl-f or left-arrow	Move forward one character in a field.
Ctrl-b or right arrow	Move backward one character in a field.
Ctrl-p or up-arrow	Move up to previous row.
Ctrl-n or down-arrow	Move down to next row.
Shift + left-arrow	Move to previous field in the same row.
Shift + right-arrow	Move to next field in the same row.
right-click	(In the Time-Segmented Transcription Window) Opens up Comment Field Window
Ctrl-l	Plays a segment
Ctrl-a	Moves cursor to the beginning of a field
Ctrl-e	Moves cursor to the end of a field

3. Printing Commands

Annotators can print out their comments using the program "csjcomment." In the terminal window, type "csjcomment <csj_file>" where <csj_file> is the name of the ".csj" file to print.

4. How to Play the Sound File

To open up the wave file of a meeting to be labeled, a link command can be made from where the sound files are saved in the annotator's home directory. To create this link, type the following in the terminal:

```
> ln -s <path name>
```

After returning back to the Table Trans interface, go to the File Menu and select "Open Sound File." Browse through the annotator's home directory and open the file.

Section 2

Mapping from SWBD-DAMSL Guide to Meeting Recorder Dialogue Acts (MRDA)

The following table is designed to show the correspondence between Switchboard-DAMSL dialog tags and those used to label Meeting Recorder (MRDA) data. Any tags that are listed in *italics* are based on SWBD-DAMSL tags but have had their meanings slightly altered. To see the complete explanation of every tag, see Section 4. Any **boldface** tags not in the original SWBD-DAMSL manual, and were added specifically for this project. The names of the tags correspond to the MRDA tag names since, often times, the names of the Switchboard-DAMSL tags were not specified.

TAG TITLE	SWBD-DAMSL	MRDA
COMMUNICATION STATUS		
Indecipherable	%	%
Nonspeech	x	x
Interruption	not marked	%-
Abandoned	%-	%--
Self-Talk	t1	t1
3 rd -Party Talk	t3	t3
INFORMATION LEVEL		
Task-Management	t	t
Communication-Management	c	Not marked
STATEMENTS		
<i>Statement</i>	<i>sd</i>	<i>s</i>
<i>Subjective Statement</i>	<i>sv</i>	<i>sj</i>
INFORMATION REQUESTS		
Wh-Question	qw	qw
Y/N Question	qy	qy
Open-Ended Question	qo	qo
Or Question	qr	qr
Or Clause After Y/N Question	qrr	qrr
Rhetorical Question	qh	qh

Wh-Question	qw	qw
Declarative-Question	d	d
Tag Question	g	g
INFLUENCING/COMMITTING SPEAKER-FUTURE-ACTION		
Open-Option	oo	oo
<i>Command</i>	<i>ad</i>	<i>co</i>
<i>Suggestion</i>	<i>co</i>	<i>cs</i>
Commit	cc	cc
OTHER-FORWARD-FUNCTIONS		
Conventional-Opening	fp	not marked
Conventional-Closing	fc	not marked
Topic Change	not marked	tc
Explicit-Performative	fx	fx
Exclamation	fe	fe
Other-Forward-Function	fo	not marked
Thanks	ft	ft
Welcome	fw	fw
Apology	fa	fa
Floor-Holder	not marked	fh
Floor-Grabber	not marked	fg
BACKWARDS-COMMUNICATIVE-FUNCTION		
<i>Accept</i>	<i>aa</i>	<i>aa</i>
Partial Accept	aap	aap
Partial Reject		arp
Maybe	am	am
<i>Reject</i>	<i>ar</i>	<i>ar</i>
Hold	h	h
UNDERSTANDING		
Collaborative-Completion	2	2
Backchannel	b	b
Acknowledgement	bk	bk
Reformulation	bf	bs

Collaborative-Completion	2	2
Appreciation	ba	ba
Sympathy	by	by
Downplayer	bd	bd
Misspeak Correction	bc	bc
Rhetorical-Question Backchannel	bh	bh
Signal Non-Understanding	br	br
Understanding Check	not marked	bu
Defending/Explanation	not marked	df
Misspeak Self-Correction		bsc
ANSWERS		
<i>Expansion/Supporting addition</i>	<i>e</i>	<i>e</i>
<i>Yes answers</i>	<i>ny</i>	<i>aa</i>
<i>No answers</i>	<i>nn</i>	<i>ar</i>
Narrative-affirmative non- yes answers	na	na
Narrative-negative non-no answers	ng	ng
Other answers	no	no
Dispreferred answers	nd	nd
OTHER		
Quoted Material	q	not marked
Humorous Material		j
Continued from previous line	+	not marked
Hedge	h	not marked
Rising Tone		rt

Section 3

Listing of Tags

Tag	Page
Group 1: Statements	
	17-19
s Statement	17
sj Subjective Statement	18
Group 2: Questions	
	20-22
qr Or Question	20
qrr Or Clause After Y/N Question	21
qw Wh-Question	21
qy Y/N Question	22
Group 3: Disruption Forms	
	23-25
% Indecipherable	23
%- Interrupted	23
%-- Abandoned	24
t3 3 rd -Party Talk	24
x Nonspeech	25
Group 4: Short Utterances	
	26-33
b Backchannel	26
bk Acknowledgement	27
fg Floor Grabber	28
fh Floor Holder	29
aa Accept	30
h Hold	31
qh Rhetorical Question	32
bh Rhetorical Question Backchannel	32
t1 Self Talk	33
Group 5: Responses	
	34-38
aap Partial Accept	34
arp Partial Reject	34
nd Dispreferred Answer	35
ng Negative Answer	35
ar Reject	36
am Maybe	36
na Affirmative Answer	37
no Other.	37

Group 6: Forward Functions		39-43
co	Command	39
fx	Explicit-Performative	40
cc	Commitment	40
cs	Suggestion	41
oo	Open-Option	41
qo	Open-ended Question	42
tc	Topic Change	42

Group 7: Backward-Looking Functions		44-49
ba	Assessment/Appreciation	44
bc	Correct-Misspeaking	44
bsc	Self-Correct Misspeaking	45
bd	Downplayer	45
br	Repetition Request	46
bs	Summary	46
bu	Understanding Check	47
df	Defending/Explanation	48
fa	Apology	49
by	Sympathy	49

Group 8: Descriptive Tags		50-55
2	Collaborative Completion	50
e	Elaboration	50
t	About-Task	51
d	Declarative Question	51
f	"Follow Me"	52
g	Tag Question	52
j	Joke	53
m	Mimic	54
r	Repeat	54
rt	Rising Tone	55

Group 9: Conventionalities		56-57
fe	Exclamation	56
ft	Thanks	56
fw	Welcome	57

Section 4

Explanation of Tags

Introduction

This section provides a description of each tag. The tags are categorized into nine groups according to syntactic, semantic, and pragmatic similarities of the utterances they mark. Within each group, the DA tags will often be compared and contrasted with one another for further clarification on the specific usage of each tag. Examples given after the description of a tag are either hypothetical or else taken from transcribed meeting data. Below is the format used for the examples in this manual:

Hypothetical Example:

[speaker, if applicable] [tag] [utterance]

Example from Data ([meeting number]):

[start time] - [end time] [speaker] [tag] [utterance]

Group 1: Statements <s, sj>

S Statement

This tag marks objective statements. A Statement will have one or more of the following characteristics:

1. The utterance is regarded as factual (including utterances that are meant to be interpreted as facts, such as lies and sarcasm).
2. The utterance is unlikely to be disputed.
3. The utterance more often receives backchannels and continuers whereas Subjective Statements <sj> (page 18) are often opinions that may be countered or disputed via further opinions from other speakers.
4. The utterance may indicate the level of expertise of the speaker.
5. The utterance may be very concise, such as those labeled with the tags in Group 4. For instance non one would "dispute" (see characteristic 2) an Acknowledgement <bk> (page 27).
6. The utterance is a narrative in which the speaker is telling a story when the topic is personal. The speaker is likely to use "I" and "we"--the latter with regards to people who are not at the meeting.
7. The utterance may be a declaration of a personal fact. This includes subjective personal statements regarding oneself. The reason such utterances are not labeled with an <sj> is because they cannot be disputed.

The examples below highlight these seven characteristics:

Characteristics 1 and 2

Hypothetical Example:

Sp.1 s Boulder is north of Denver.

Example from Data (Bmr012):

40.009-43.139	c3	sd.x	adam's just trying to generate good uh - data for the recognizer there .
---------------	----	------	--

Characteristics 2 and 4

Example from Data (Bmr012):

62.153-64.053	c3	qw^t3^r	why didn't you get the same results and the unadapted ?
62.79-63.17	c4	%	but ==
64.235-68.995	c0	s^t3	oh because when it estimates the

transformer pro- - produces like a single matrix or something .

Characteristic 5

Example from Data (Bmr013):

102-104	c3	s^bk^tc	okay agenda items
104-105	c3	fh s^t	uh - we have digits.
105-106	c3	qo^t	what else we got ?

Characteristic 6

Example from Data (Bmr013):

1536.33-1550.13	c1	s	and what i did is i used some normalized features which uh - look in- - into the - which is normalized energy uh - energy normalized by the mean over the channels and by the minimum over the other .
-----------------	----	---	--

Characteristic 7

Hypothetical Example:

Sp.1 s I was born in Louisiana.

Other Examples

Example 1 from Data (Bmr012):

67.241-68.771	c5	sd	uh - looks kind of low on channel five.
---------------	----	----	---

Sj Subjective Statement

This tag marks statements that are obviously subjective and/or overt opinions. The default tag to any statement is an <s>, thus <sj> is used only if a statement is clearly a subjective statement. A Subjective Statement will have one or more of the following characteristics:

1. The utterance is subjective. (This does *not* include subjective *personal statements* such as "I feel cold" since such a statement cannot be disputed.)
2. The utterance is an opinion.
3. The utterance can be disputed or rejected.
4. The utterance can be a generally-supposed opinion. In which case, if "we" is used in the utterance, it is used in an impersonal voice.

5. The utterance qualitatively describes the subject at hand.

The examples below highlight some of these characteristics:

Characteristics 1, 2, and 3

Hypothetical Example:

Sp.1 sj That's an ugly sofa.
Sp.2 sj The coffee table isn't so great either.

Characteristic 4

Hypothetical Example:

Sp.1 sj What I find particularly difficult in watching the news is that we see injustice and then the weather and it just switches from one you know topic to the next.

Dialogue acts that should be marked with an <sj> are often signaled by phrases such as:

I think
I believe
It seems
It's my opinion that
I mean
Suppose
Of course
impersonal 'we'
impersonal 'they' as in 'they say it rains a lot there...'

Although the above phrases aid in determining whether something ought to be labeled with an <sj>, they are not a necessary part of an <sj> utterance. Also, utterances such as the following should be labeled as Subjective Statements, rather than as Explicit Performatives <fx> (page 40):

I bet you can't guess.
I bet you he's going to pull a rabbit out of that hat.
I am going to bet you that he'll show up late again.
I will bet you that she'll forget the papers.

These are not explicit performatives since the actual act of betting is not committed as a result of the first two utterances. Instead, using the phrase, "I bet..." either can lead up to a bet or simply acts as a qualifier indicating the degree of certainty the speaker has in his opinion.

Group 2: Questions <qr, qrr, qw, qy>

qr "Or" Questions

"Or" questions offer the listener at least two answers from which to choose. These types of questions are sometimes interrupted and answered as if they are Y/N Questions (page 22), as in Example 2. Other times, a speaker asking the question might trail off as in Example 3. In such a case, the speaker answering the question might respond as if the question were a yes-no question without interrupting the question at all.

Hypothetical example 1:

Sp.A	qr	Are you supposed to bring the ice picks or the headlamps?
Sp.B	s	The headlamps.

Hypothetical example 2:

Sp.A	qr	Are you supposed to bring the ice picks or-
Sp.B	s^ar	No
Sp.B	s	I'm supposed to bring the headlamps.

Hypothetical Example 3:

Sp.A	qr^t.%--	Should we discuss the next topic or ==
Sp.B	s^aa	Yeah, good idea.

Example 4 from Data (Bmr013):

502-506	cB	qr^rt.%--	are they reading actual phone numbers or a - a digit at a time or ?==
---------	----	-----------	---

Example 5 from Data (Bmr012):

1271.55-1277.44	c1	qr	are you - are you using uh - uh - mel cepstrum or p l p over there ?
-----------------	----	----	--

qrr Or-Clause After Y/N Question

This tag marks when a speaker adds an Or-clause to a Y/N question (page 22). A <qrr> is in some ways a "dangling Or-Clause." (Jurafsky 29)

Example from Data (Bmr013):

332-334	c4	fg qy^rt	so are these two separate tasks that can happen ?
335-336	c4	qrr.%-	or do they have to happen at the same time before ?==
336-342	c8	s^ar s	no they don't have - this - you have to enter the data before you do the second task but they don't have to happen at the same time .

qw Wh-Question

Wh-questions begin with one of the following "wh" words:

how
what
where
when
who
why

Example 1 from Data (Bmr012):

62.153-64.053	c3	qw^t3^r	why didn't you get the same results and the unadapted ?
---------------	----	---------	---

Example 2 from Data (Bmr012):

231.944-233.704	c2	qw^t3	i guess - what time do we have to leave ?
-----------------	----	-------	---

Example 3 from Data (Bmr013):

121-127	cB	s^e	and i guess that includes some - the filtering
---------	----	-----	--

129-129	c3	qw^br^rt	for the - the a s i refs too .
130-136	cB	s	filtering for what ? for the references that we need to go from the - the fancy transcripts to the sort of starts to say something that starts with s brain dead .

In the third example, even though the "what" is found at the end of the question, the entire utterance still functions as a <qw> and could be rephrased as "what is the filtering for?"

Example 4 from Data (Bmr012):

61.563-61.713	c0	qw^br^t3	hm?
---------------	----	----------	-----

In the Example 4, the utterance functions as a "what" question, in that "hm?" is synonymous to "what?".

qy Y/N Question

This tag marks yes-no questions if and only if they have the pragmatic force and the syntactic and prosodic indications of a yes-no question (i.e., subject-inversion or question intonation, etc.).

Example 1 from Data (Bmr012):

16.805-17.875	c1	qy^rt	is this channel one ?
---------------	----	-------	-----------------------

Group 3: Disruption Forms <%, %-, %--, t3, x>

% Indecipherable

This tag marks indecipherable speech such as mumbled or muffled words, or utterances that are too difficult to hear on account of the microphone picking up sounds from breathing.

Indecipherable <%> is not to be confused with Nonspeech <x> (page 25). The Nonspeech <x> tag is used for sound segments which are silent or otherwise contain non-vocal sounds such as doors slamming and phones ringing. (<x> does not cover sounds like breathing and sighs, as these are vocal sounds.)

Example from Data (Bmr012):

20.995-21.115 c1 % okay .
(where "Okay" was not discernible.)

% - Interrupted

This tag marks utterances cut short by interruption. This tag is not to be confused with Abandoned <%--> (page 24), which is used when a speaker *intentionally* abandons an utterance.

Example from Data (Bmr012):

326.848-330.118 c0 s.%- and then we noticed that the
reference was always one off with
the ==
330.574-333.004 c5 fg|sj maybe,yeah | that might be - that
might be - that might be my fault .

Example from Data (Bmr012):

136.665-143.767 c8 sj^co%- although on - course on the other -
on the other hand you could - you
view it as the age at the time of the
==
137.83-138.33 c0 % yeah .

138.066-139.376 c2 qy shouldn't it be the other way around?

At first glance, the first utterance in the above example looks similar to an abandoned utterance <%-->. In fact, it is marked as interrupted because the speaker actually is interrupted, even though the speaker doesn't immediately give up the floor.

%-- Abandoned

This tag marks utterances that are abandoned by a speaker. Two cases of Abandoned utterances are where:

1. the speaker trails off
2. the speaker chooses to express his or her own thought with a new utterance without finishing the first one.

In both cases, Floor Holders <fh> (page 29) often occur at the end of the Abandoned utterance.

Example from Data (Bmr012):

8.953-9.803	c1	%--	let's see i should be ==?
13.954-15.514	c4	s^ad	as close to your mouth as you can get it .

t3 3rd-Party Talk

This tag handles side conversations that are not directed towards the main conversation participants. In the example below, two people are discussing adaptation results aside from the main conversation topic of how to wear the lapel microphones.

Example from Data (Bmr012):

45.424-51.184	c4	s^ad	and then also for - for all of them if your boom is adjustable the boom should be towards the corner of your mouth .
46.411-47.581	c0	s^t3	by the way there was a bug .

47.831-51.811	c0	s^aa^t3 s^t3.%--	yeah i- - it wasn't using the proper ==
48.577-49.127	c3	qy^d^bh^rt^t3	oh it was .
51.724-56.624	c4	s^ad	and about a- - uh - a thumb to a thumb and a half distance away from your mouth .
52.971-54.271	c0	s^t3	basically it wasn't adapting anything.
56.624-57.954	c4	s	so about like i'm wearing it now .
57.114-57.804	c3	sj^ba^t3	oh that's interesting .

X Nonspeech

Nonspeech marks any utterance that is or partly consists of non-speech noises such as door slams and phone rings. This tag does not mark Indecipherable <%> (page 23) speech due to muffled speech or mumbling.

Example from Data (Bmr012):

40.009-43.139	c3	s.x	adam's just trying to generate good uh - data for the recognizer there .
---------------	----	-----	--

(In the above example, there were problems with the recording that resulted in clicking sounds throughout the utterance.)

Group 4: Short Utterances

<b, bk, fg, fh, aa, h, qh, bh, t1>

b Backchannel

Common Backchannels:

uhhuh
yeah
right
oh
yes
okay
oh yeah
huh
sure
hm
uh yeah

Utterances which function as backchannels are not made by the speaker who has the floor. Instead, backchannels are utterances made in the background that simply indicate that a listener is following along or at least paying attention. The nature of backchannels does not permit utterances such as "uh," "um," and "well" as being perceived as backchannels since these utterances do not indicate that a speaker is following along, but rather that a speaker has something to say or else is trying to say something. When uttering backchannels, a speaker is not speaking directly to someone in particular or even to anyone at all.

Backchannels are not to be confused with Acknowledgement <bk> (page 27) or Accept <aa> (page 30). In the case of <bk>, a speaker acknowledges another speaker's response, usually after a question and answer sequence. With <aa>, a speaker is providing a "yes" response and is speaking directly to someone else.

Other tags which may be confused with backchannels are Floor-Grabbers <fg> (page 28), Floor-Holder <fh> (page 29), and Holds <h> (page 31). An utterance labeled with any of these three tags indicates that the speaker who made the utterance intends to hold or gain the attention of others. When a backchannel occurs with nothing by the same speaker immediately preceding or following it, the person is not trying to get or maintain the floor. These three tags are used for utterances where the speaker currently holds the attention of others (e.g., <h> and <fh>) or else attempts to do so (e.g., <fg>). Another distinction is that utterances labeled as floor-grabbers generally tend to sound far more assertive than backchannels.

Within a tiered DA, a backchannel tag should never be followed by other subdivisions of backchannels such as:

- Acknowledgement <bk> (page 27)
- Assessment <ba> (page 44)
- Correct-Misspeaking <bc> (page 44)
- Self-Correct Misspeaking <bsc> (page 45)
- Downplayer <bd> (page 45)
- Summary <bs> (page 46)
- Understanding-Check <bu> (page 47)

If an utterance is better specified with one of these other tags, then should not be used. For instance, the DA <b^bk> should never be used.

bk Acknowledgement

The <bk> tag is used to express a speaker's acknowledgment of a previous speaker's utterance. It primarily occurs after a response to a question, proposal, or offer. As opposed to backchannels, <bk> encodes a level of direct communication between speakers, where a speaker who acknowledges a previous speaker's utterance is actually speaking directly to that previous speaker. An Acknowledgment is unlikely to be "disputed" by another speaker and therefore is never labeled as a Subjective <sj> (page 18).

"Oh" often but not always functions as a <bk>. Note that in the second example from the data, the second utterance contains the word "oh" but the "oh" does not function as a <bk> since it is not acknowledging an answer.

Hypothetical Example:

Sp.1	qw	What's wrong with your car ?
Sp.2	s	The battery is dead.
Sp.1	s^bk	Oh, okay.

Example 1 from Data (Bmr012):

58.784-60.504	c3	qw^t3	so why didn't you get the same results and the unadapted ?
62.153-64.053	c3	qw^t3^r	why didn't you get the same results as the unadapted?
64.235-68.995	c0	s^t3	oh because when it estimates the

67.730-69.010	c3	s^bk^t3	transformer pro- - produces like single matrix or something. o- - oh i see.
---------------	----	---------	---

Example 2 from Data (Bmr012):

126.568-131.198	c0	sj^co	instead of date of - uh - put year of birth because age will change but == oh.
132.117-132.347	c4	s^bk	

fg Floor Grabber

The <fg> marks when a speaker has not been speaking and wants to start speaking. Floor Grabbers are often repeated by the speaker to gain attention. Here are some examples of some common Floor Grabbers:

Common Floor Grabbers:

well
but
so
and
yeah

A Floor Grabber is contrasted with a Floor Holder <fh> (page 29) since the latter is used when the speaker *has* been speaking.

Floor Grabbers and Backchannels (page 26) may be confused since they are often composed of the same words. The very intention of a Floor Grabber explains why similar words are used in both dialogue acts. A "successful" floor grabber, that is, one in which the speaker's attempt to gain the floor is accomplished, is one that appears non-intrusive, and therefore, like a backchannel.

Example 1 from Data (Bmr012):

141.602-145.412	c0	s^df^nd	well the thing is if ten years from now you look at this form knowing that ==
144.588-148.398	c8	fg^bk sj^df^nd	yes but what we care about is the age at - at the recording date rather than the ==

Example 2 from Data (Bmr012):

1814.65-1817.01	Bmr012-c2	fg s.%-	well or also for you know - if people are not ==
-----------------	-----------	---------	---

Example 3 from Data (Bmr012):

1816.1-1817.54	Bmr012-c4	sj^co	well we could just hand around the lapel .
----------------	-----------	-------	---

Example 4 from Data (Bmr012):

1822.12-1824.17	Bmr012-c4	fg qy^df	well i mean - is the - is the handheld really any better ?
-----------------	-----------	----------	---

fh Floor Holder

A Floor Holder is used when a speaker is mid speech and uses an utterance like "uh" or "so" as a means to pause and continue holding the floor. Floor Holders are often found at the end of an utterance, thus yielding abandoned utterances.

Common floor holders include:

- so
- and
- or
- um
- uh
- let's see
- and what else
- so anyway

To avert confusion regarding the location of Floor Holders, it ought to be noted that Floor Holders do not occur at the very beginning of a speaker's speech but may occur at the beginning of an utterance. A speaker will not use a Floor Holder when he begins speaking but may use one at the beginning of an utterance which exists within a series of utterances by that very speaker, so long as those utterances are made in succession with no gaps in time occurring among them.

A Floor Holder and a Floor Grabber <fg> (page 28) are contrasted by the fact that a Floor Holder is used when the speaker has been speaking and hesitates and a Floor

Grabber is used when the speaker has not been speaking and wishes to start speaking.

Example 1 from Data (Bmr012):

33.474-36.824	c4	sj^co	it's actually a lot more comfortable then if you try to put it over your temples .
38.044-38.334	c4	fh.%-- so ==	

Example 2 from Data (Bmr012):

114.061-116.811	c4	s^ad	so if you don't know what to put just leave it blank .
120.671-122.251	c4	fh s	um - it's like - i - i designed the form.

aa Accept

This tag is used for utterances which exhibit acceptance or agreement. Generally, a speaker will accept or agree to another speaker's questions, proposals, or statements. The <aa> tag is used to mark short utterances that express acceptance or agreement. Common utterances associated with the <aa> tag are listed below:

yeah
yes
okay
sure
I agree
exactly
definitely
that's true
right

This is not to say that words used in negation cannot be labeled with the <aa> tag. The following example exemplifies this point:

Hypothetical example:

Speaker 1	sd	He didn't go to the party
Speaker 2	sd^aa	No, he didn't

Above, it is seen that, in saying "no," Speaker 2 is not rejecting the statement made by

Speaker 1, but rather is actually agreeing with it.

Utterances marked with the Accept tag <aa> may be confused with Backchannels (page 26) or Acknowledgement <bk> (page 27). The distinction is that <aa> marks a type of answer whereas and <bk> mark continuers/Backchannels.

h Hold

The <h> tag is used when a speaker who is given the floor and is expected to speak "holds-off" prior to making an utterance. A common area where the <h> tag is used occurs when a speaker responding to a question says "uh" or "let's see" prior to answering the question.

Holds occur at the beginning of an utterance, as opposed to floor-holders which occur in the middle or at the end of a speaker's utterance.

Although Holds are similar to Self Talk <t1> (page 33), Holds are distinguished by the fact that they have little semantic content and are intended for the listener to hear them. Holds <h> should not be confused with Floor Holders <fh> (page 29).

Common Holds:

uh
um
let's see
well
so
hmm

Example from Data (Bmr012):

355.015-356.145	c8	s^ad	not an offer, he - he can say something about the glitch .
355.824-357.064	c2	s^h s^bk	um - oh okay .
357.154-359.244	c2	qy.%-	so does this glitch occur at other ?==

qh Rhetorical Question

The tag <qh> marks questions to which no answer is expected. Such questions are used by the speaker for rhetorical effect; they are essentially statements formulated as questions. Although <qh> and Rhetorical Question Backchannels <bh> (page 32) are similar, <bh> lacks semantic content and functions mostly as a continuer.

Example from Data (Bmr012):

486.854-492.724	c0	s	there are a fair number of errors that are you know - where - got the plural s wrong or the inflection on the verb wrong .
491.713-491.893	c4	%	yeah .
491.893,492.493	c4	qh	and who cares ?

bh Rhetorical Question Backchannels

This tag marks continuers/backchannels which take the form of questions and often function as acknowledgments. Although <bh> has the same syntactic form as a Rhetorical Question <qh> (page 32), a <bh> has no semantic content and is simply a device to continue the speech of a previous speaker. Such rhetorical backchannels are often answered to with "yeah." These answers are marked with the Accept <aa> (page 30).

Example 1 from Data (Bmr012):

48.577-49.127	c3	qy^d^bh^rt^t3	oh it was ?
---------------	----	---------------	-------------

Example 2 from Data (Bmr013):

527-530	c3	h s^ar s	uh - i th- - no i think we got under a percent .
531-531	c8	qy^bh	oh really ?

Common Rhetorical Question Backchannels:

Oh really?
Really?
Yeah?
Is that right?

Did you?
Are you?
Was it?
Isn't that interesting?
Isn't that amazing?
Isn't that funny?
You think?
You think so?

t1 Self Talk

This tag is for when a speaker is talking to himself and does not intend for others to hear or respond. Similarly structured utterances with the intention for the listener to hear are Holds <h> (page 31).

Example from Data (Bmr012):

946.44-947.1	c2	sj	so that's good .
947.79-948.89	c2	s^h^fe	and um - let's see .
948.89-951.626	c2	s	there was one more thing i wanted to - to mention .
951.626-953.236	c2	s^t1	i can't remember um ==
956.426-957.216	c2	s^fa^r	sorry can't remember .

Group 5: Responses

<aap, arp, nd, ng, ar, am, na, no>

aap Accept-Part

This tag marks when a speaker explicitly accepts part of a previous speaker's utterance. The emphasis in an <aap> is on the acceptance.

Hypothetical Example:

Sp.1 s^aap|sj^co That is a good idea, | but you could change the font.

Example from Data (Bmr012):

564-567	c8	sj	and it's still like an order of magnitude worse than what humans do .
569-570	c3	sj^aap^j s^aa	when - when they're wide awake yeah .

arp Reject-Part

This tag marks where a speaker explicitly rejects part of a previous person's utterance. The emphasis in an <arp> is on the rejection.

Hypothetical Example:

Sp.1 sj I think ants are great.
Sp.2 sj^arp Except for when they get into picnic food.

Example from Data (Bmr012):

398.311-401.201	sj	i mean - if we know about it then i guess it could always be checked for it .
399.841-401.691	sj^arp	well the acoustic one shouldn't do anything .

nd Answer Dispreferred

This tag marks a statement that acts as an explicit rejection to a previous person's utterance. It is different from a Negative Answer <ng> (page 35) in that an <nd> explicitly rejects whereas an <ng> implicitly rejects a proposal, offer, or a question. The example below highlights this fact.

Example from Data (Bmr012):

138.066-139.376	c2	qy	shouldn't it be the other way around?
141.602-145.412	c0	s^df^nd	well the thing is if ten years from now you look at this form knowing that ==
144.588-148.398	c8	fg^bk sj^df^nd	yes but what we care about is the age at - at the recording date rather than the ==
146.066-147.426	c3	s^nd	but there's no other date on the form.

ng Negative Answer

A Negative Answer *implicitly* rejects a previous person's utterance as opposed to an Answer dispreferred <nd> (page 35) which *explicitly* rejects. Both <ng> and <nd> are similar in that they both do not include the word "no" in the dialogue act. But <ng> differs from <nd> in that, rather than directly responding to the previous utterance, a Negative Answer explains why there a speaker does not favor the previous speaker's suggestion.

In the first example, the speaker is responding to the proposal of having her transcribers enter information for a permission form.

Example from Data (Bmr013):

318-326	c2	s^ng	and that you know - i'm not - that - that one i'm not so sure if it's into the - the things that i wanted to use the hours for .
326-330	c2	s	because the - the time that they'd be spending doing that they wouldn't be able to be putting more words on .

Hypothetical Example:

Sp.1	qy	Did you see "Pocahontas" on the Disney channel last night?
Sp.2	s^nr s^ng	No, I don't have cable.

ar

Reject

This tag marks negatives words such as "no" and other semantic equivalents that reject proposals and statements.

Example from Data (Bmr012):

83.795-84.595	c1	qy^d^rt	you want to close this ?
84.805-85.405	c2	s^ar.%--	no i ==

am

Maybe

This tag marks a response to a question, proposal, or offer that uses the word "maybe" or other words denoting possibility. Sometimes, to express a possibility, a speaker may say "I don't know." Such utterances are marked as Maybe <am> rather than as Other <no> (page 37).

Hypothetical Example 1:

Sp.1	sj^co	I think we'll need to add the second experiment in the report.
Sp.2	b	Yeah.
Sp.2	s^am	I guess so.

Example 2 from Data 1 (Bmr018):

77.631-82.272	c4	s^am	but - uh - channel b is probably liz .
---------------	----	------	--

Example 3 from Data (Bmr013):

645.98-653.181	c3	sj	but it could be that when we're reading digits because it's - it's for such a limited set that maybe - maybe that phenomenon doesn't occur as much .
653.181-653.591	c3	s^no	i don't know .

na A Descriptive/Narrative

This tag marks an utterance that acts as an affirmative response to questions and proposals. Affirmation is *implicitly* stated in the utterance. Because utterances marked with <na> do not contain the word "yes" or any variation thereof, a DA that includes the <na> tag often follows a DA that contains <aa>. If a person responds to a question with "I guess," it should be labeled as Maybe <am> (page 36), not as <na>.

Hypothetical example:

Sp.1	qy	Do you have kids?
Sp.2	s^na	I have three.

Example from Data (Bmr012):

209.361-211.781	c4	qy^co^t^rt	okay so should we do agenda items ?
211.573-213.947	c1	s^h s^na	uh - oh - that's a good idea .

Example from Data (Bmr012):

381.613-383.373	c2	qy^rt,43a	and is it only once that that happens?
383.957-384.837	c4	s^aa s^e^na	right once in the middle .

no Insufficient Knowledge

This tag marks utterances in which the speaker expresses a lack of knowledge with regard to the subject at hand. It is usually in response to a previous utterance made by any speaker including the current one.

The most common expression of an <no> is "I don't know." This is not always a signal for <no>; when this phrase expresses uncertainty or possibility, it is labeled as Maybe <am> (page 36).

Example 1 from Data (Bmr012):

312.632-313.712	c5	qy	was it twenty minutes in ?
313.023-314.003	c8	%	i forgot about that .
313.426-315.481	c2	s^no	if it was twenty minutes in then i don't know .

Example 2 from Data (Bmr012):

1070.52-1079.28	c8	sj^ad	although we should clear it with eric and dan of course but these results are based on data which haven't had the uh - haven't had the chance to be reviewed by the subjects .
1078.15-1078.81	c2	s^bk	that's true .
1079.28-1080.32	c8	s^no (tr)	and i don't know how that stands .

Example from Data (Bmr018):

717.62-720.37	c0	s	but you just can't get the visual display to show quickly .
720.696-721.226	c0	%--	so you have to ==
721.226- 723.406	c0	%--	it takes i don't know three four minutes to ==
723.506-725.476	c0	%--	well i mean it takes - it takes long enough ==
724.71- 726.97	c3	aa	yeah it's very slow to do that .
726.13-727.64	c0	%--	it takes long enough because it has to reload the ==
727.64-729.62	c0	s	i - i don't know exactly what it's doing frankly .

Group 6: Forward Functions

<co, fx, cc, cs, oo, qo, tc>

CO Command

This tag marks imperatives and commands. The syntactic realization of <co> may include questions (e.g. "Do you want to go ahead?") and standard declarative clauses (e.g. "You ought to try using the data from the second set").

Example 1 from Data (Bmr012):

45.424-51.184	c4	s^co	and then also for - for all of them if your boom is adjustable the boom should be towards the corner of your mouth .
---------------	----	------	--

Example 2 from Data (Bmr012):

60.19-62.79	c4	s^co	so- - so jane you could actually do even a little closer to your mouth .
-------------	----	------	--

Example 3 from Data (Bmr012):

102.326-104.551	c4	s^co	and you only have to spea- - fill out the speaker form once .
-----------------	----	------	---

Example 4 from Data (Bmr012):

104.551-105.821	c4	s^co	but everyone does need to do it .
-----------------	----	------	-----------------------------------

Example 5 from Data (Bmr012):

136.656-138.066	c2	s^co	a- - actually wait a minute .
-----------------	----	------	-------------------------------

The following example is much like Example 3 and 4 except that the speaker refers to the forms rather than other meeting participants. Therefore, Example 6 is *not* an example of a command.

Example 6 from Data (Bmr012):

173.653-180.483	c4	s	and so then you also have a digits form which needs to be filled out every time the speaker form only
-----------------	----	---	---

once the digit form every time .

fx Explicit-Performative

This tag marks utterances that are performatives in and of themselves. These do not include the conventional utterances shown in Conventionalities (Group 9).

Some examples:

You're fired.
I wish you all the best.

Explicit performatives are different from Subjective Statements <sj> (page 18) because they are opinions that do not function as actions in themselves. Utterances such as the following should not be labeled as explicit performatives because the betting is not enacted by the utterance itself:

I bet you can't guess.
I bet you he's going to pull a rabbit out of that hat.
I am going to bet you that he'll show up late again.
I will bet you that she'll forget the papers.

Instead, using the phrase, "I bet..." can either lead up to a bet or simply act as a qualifier of how sure the speaker is in his or her opinion; therefore they are Subjective Statements <sj>.

CC Commitments

Commitments explicitly commit the speaker to some future course of action.

Example 1 from Data (Bmr013):

575-580 c8 fg|s^cc okay so | um - what i'll do then is i'll go ahead and enter this data .

They are easily confused with Explicit-Performatives <fx> (page 40) which take place in the present without reference to the future. Suggestions <cs> (page 41) are also easily confused with Commitments. If the speaker politely suggests that he is about to do something and gives the listener a chance to say "no," by default, this is not <cc>, but rather a suggestion <cs>:

s^cs Let me see if that clears up.
s^cs Let me reload and see if your program unfreezes.
s^cs Let me ask, by the way, just for the record.

CS Suggestion

The Suggestion tag marks proposals, offers, advice, and, as the name of the tag denotes, suggestions.

Example from Data (Bmr026):

536.64 - 540.31 It would be cool like to see - to hear it and see it.

If the phrase "excuse me" precedes something the speaker was negotiating permission for (Jurafsky 35), then it should be marked as a Suggestion rather than as an Apology <fa> (page 49).

Hypothetical Example 1:

s^cs Excuse me just a second.

Hypothetical Example 2:

fg|qy^cs^rt Excuse me | may I ask one more question?

Hypothetical Example 3:

s^cs Let me ask one more question.

OO Open Option

This tag marks a statement in which the speaker gives multiple and/or open-ended options to others. It is similar to a Suggestion <cs>

Hypothetical Example:

s^oo "We could have lamb or chicken or..."

Example from Data (Bmr013):

306-309	c8	s	and so it's just you know - typing in name times - time date and so on .
310-317	c8	s^oo	um - which again either they can do but it is you know - firing up an editor or again i can do .
317-318	c8	s^oo	or someone else can do .

qo Open-Ended Question

This tag marks utterances of the "how about you" variety; Open-Ended Questions are the kind of questions which place few syntactic or semantic constraints on the form of the answer (Jurafsky 28).

Example from Data (Bmr012):

224.832-228.262	c4	sj^t	and i'm sure liz and andreas want to talk about recognition results .
229.835-230.345	c4	qo	anything else ?

tc Topic Changer

A topic changer identifies utterances that either begin or end a topic.

Example 1 from Data (Bmr012):

1080.32-1089.21	c8	sj	i mean - if you - if you get fantastic results and it's involving data which - which later end up being lessened by you know - certain elisions then i don't know .
1089.34-1090.33	c8	sj	but i wanted to raise that issue .
1090.46-1090.72	c8	s^tc	that's all .
1090.95-1091.26	c1	fg.%--	well we ==

Example 2 from Data (Bmr013):

102-104	c3	s^bk^tc	okay agenda items
---------	----	---------	-------------------

104-105	c3	fh s^t	uh - we have digits.
105-106	c3	qo^t	what else we got ?

Group 7: Backward-Looking Functions

<ba, bc, bsc, bd, br, bu, df, fa, by>

ba Assessment/Appreciation

Assessment/Appreciations are backchannels/continuers directed at others which function to express slightly more emotional involvement and support than just "uhhuh" (Jurafsky) or the like. A <ba> appends to an <sj> only, not an <s>, since it is by nature an opinion/viewpoint.

Example 1 from Data (Bmr012):

52.971-54.271	c0	s^t3	basically it wasn't adapting anything.
57.114-57.804	c3	sj^ba^t3	oh that's interesting .

Example 2 from Data (Bmr012):

231.944-233.704	c2	qw^t3	i guess - what time do we have to leave ?
234.144-234.774	c2	qy^d^bu^rt^t3	three thirty ?
234.974-235.164	c0	s^aa^t3	yeah .
235.456-236.346	c4	sj^ad	why don't you go first then .
236.927-237.107	c1	s^aa	yeah .
237.267-237.747	c1	sj^ba	good idea .

Example 3 from Data (Bmr012):

421.472-424.142	c2	s	there's things that are l- - in smaller increments than a frame .
424.244-424.974	c8	sj^ba	oh interesting .

bc Correct-Misspeaking

This tag is used when a speaker corrects another speaker's utterance.

Hypothetical Example:

Sp.1	sj	My other son is just as happy as a bed bug.
Sp.2	s^bc	A clam.

bsc Self-Correct-Misspeaking

This marks when a speaker corrects his own error. Sometimes another speaker will initiate the correction, but if the primary speaker is the one making the correction, the utterance is marked as <bsc> rather than <bc>.

Example from Data (Bmr012):

132.336-133.516	c0	s	the year of birth changes -
133.516-134.696	c0	s^bsc	you know - stays the same usually .

bd Downplayer

This tag marks when a speaker is downplaying an apology or a compliment.

Common Downplayers:

That's okay
No
That's all right
Okay
I don't know
Oh well
It's okay
Uhhuh

Example from Data (Bmr012):

960.05-960.79	c8	sj^ba	congratulations .
961.254-964.724	c2	s^bd	well it was i mean - i really didn't do this myself .

br Repetition Request

This tag indicates when a listener could not decipher a previous speaker's utterance and requests that the speaker repeat it.

Hypothetical example:

Sp.1 qy^br Could you repeat that?

Example from Data (Bmr013):

121-127	cB	s^e	and i guess that includes some - the filtering for the - the a s i refs too .
129-129	c3	qw^br^rt	filtering for what ?
130-136	cB	s	for the references that we need to go from the - the fancy transcripts to the sort of starts to say something that starts with s brain dead .

bs Summary

Summary marks when the speaker summarizes a previous utterance or discussion. Sometimes a Summary seems to function as an Understanding Check <bu> (page 47). But if the summary or reformulation does not require validation and is said as fact, mark only as a <bs>.

Example from Data (Bmr013):

1082.53-1087.83	c3	s	but - and it's more accurate than the - than the dictionary or if you've got a pronunciation uh - lexicon that has three or four .
1083.51-1083.92	c8	%.x	the word .
1085.21-1085.42	c4	b	yeah .
1087.83-1090.47	c3	s	this might be have been the fifth one that you tr- - that you pruned or whatever .
1090.47-1093.12	c3	s	so sure .
1091.13-1092.16	c4	s^bs	so it's like a continuum .

bu Understanding Check

This tag marks two incidences:

1. When a speaker checks if he has comprehended the facts correctly.
2. When a speaker checks if he has understood what a previous speaker has said.

The first case can occur independently of a previous speaker's utterance. In the second case, a speaker often proposes a summary or a paraphrase of a previous speaker's speech (Jurafsky 48). Therefore, in the second case, it is a pragmatically construed question rather than an explicit question which translates to "is this an acceptable summary of what you said?" In both cases, if an understanding check statement/question is followed by a Tag Question, the <bu> tag is appended to the understanding check statement/question and not to the Tag Question <g>. The <bu> tags are often followed by utterances from other speakers that verify understanding.

This tag differs from Signal Non-Understanding
 (page 46) in that with a <bu> the speaker wants to know if he or she correctly understands the meaning of a previous speaker's talk. However, <bu> marks a request for a repeat of a previous speaker's utterance simply because the listener wasn't sure if the previous utterance was heard correctly (or at all), not because the listener did not understand the meaning or content of the utterance.

Example 1 of Case 1 from Data (Bmr012):

231.944-233.704	c2	qw^t3	i guess - what time do we have to leave ?
234.144-234.774	c2	qy^d^bu^rt^t3	three thirty ?

Example 2 of Case 2 from Data (Bmr012):

33.474-36.824	c4	sj^co	it's actually a lot more comfortable then if you try to put it over your temples .
34.123-36.013	c5	%	test test test .
38.044-38.334	c4	fh.%-- so ==	
38.482-39.312	c1	qy^d^bu^rt	you do it higher ?

df Defending/Explanation

This tag marks cases in which a person defends his or her own point or downplays it. The intention on the part of the speaker in a <df> utterance is to avoid or counteract a dispreferred response (a request rejection, an offer refusal or a disagreement) after evaluative assessments or comments (Levinson 308). The <df> tags often take the form of alternative suggestions, appearing to lessen the resolution of an utterance without abandoning its main proposition or premise.

Example 1 from Data (Bmr018):

922.580 - 928.120	c0	s^df	there's a reason why I disagree and that is that uh - you - it's very good to have a disassociation between the visual and the audio .
-------------------	----	------	--

Example 3 from Data (Bmr012):

506.032-509.052	c2	s fh	there's no language model for p z m or um ==
515.577-517.617	c4	sj^co^j	maybe we shouldn't say p z m in these meetings .
517.297-519.607	c2	fg s^df	well - well there's all kinds of other stuff like jimlet .

Example 4 from Data (Bmr013):

443-445	c3	%-	so more typical numbers like ==
443-446	c4	sj^bu	although the models weren't that good .
446-446	c4	qy^g	right ?
446-448	c4	qy^g	i mean - the models are pretty crappy ?
448-449	c3	s^aa	you're right .
449-455	c3	s	i think that we could have done better on the models
455-461	c3	s^df	but the thing is that we got - this - this is the kind of typical number for all of the uh - uh - things in this task all of the um - languages .

fa Apology

When a speaker apologizes for something that he or she just did (after a cough, sneeze, or interruption, for example), it is marked as an Apology <fa>. However, if the word "excuse me" occurs when a speaker is negotiating permission to do something in advance of doing it, it is an Suggestion <cs> (page 41).

Example from Data (Bmr012):

948.89-951.626	c2	s	there was one more thing i wanted to - to mention .
951.626-953.236	c2	s^t1 h	i can't remember um ==
956.426-957.216	c2	s^fa^r	sorry can't remember .

by Sympathetic Comment

This tag marks when a speaker is showing sympathy for someone else. It commonly takes the form of "I'm sorry," but does not function as an Apology <fa>. A Sympathetic Comment is often followed by a Downplayer <bd> (page 45).

Hypothetical Example:

Sp.1	s^by	I'm sorry to hear your grandmother is sick.
Sp.2	s^bd	That's all right.

Group 8: Descriptive Tags

<2, e, t, d, f, g, j, m, r, rt>

2 Collaborative Completion

This tag marks when one speaker finishes another speaker's utterance. This is not dependent on the previous speaker's agreement with the utterance completion of the current speaker.

Example from Data (Bmr018):

159.358-163.643	c2	s^nr s	No I mean they're flying up from - from - down from Seattle.
60.961-161.341	c4	s^2	Seattle.

In this example, Speaker c4 says "Seattle" immediately before the original speaker (Speaker c2) finishes their utterance.

Example from Data (Bmr012):

433.904-435.384	c2	s	that's why we only have twenty minutes .
435.384-437.674	c2	s.%-	but there's a significant amount of ==
436.608-437.368	c5	qy^d^rt^2	non zero ?
437.618-442.778	c5	s	um - there are like - more - because there's a lot of zeros i tacked on just because of the way the script ran .

e Elaboration

This tag marks when the current speaker elaborates on a previous utterance by adding finer details (as opposed to simply continuing to speak on the same subject).

Example 1 from Data (Bmr012):

14.119-14.519	c3	s^ad	up high .
14.519-15.369	c3	s^ad^e	high as you can get .

Example 2 from Data (Bmr012):

381.613-383.373	c2	qy^rt	and is it only once that that happens?
383.957-384.837	c4	s^aa s^e^na	right once in the middle .

t About-Task

This tag marks utterances made by a speaker about the meeting agenda, addressing the direction of the meeting conversation.

Hypothetical Example:

s^t Let's get onto the next topic.

Example 1 from Data (Bmr012):

94.861-99.771	c4	fg s^t	um - everyone should have at least two forms possibly three in front of you depending on who you are .
---------------	----	--------	--

Example 2 from Data (Bmr012):

209.2,211.573	c1	fg	so uh -
209.361,211.781	c4	qy^co^t^rt	okay so should we do agenda items ?

d Declarative Question

This tag marks a question that is in the syntactic form of a statement.

Example from Data (Bmr012):

38.482-39.312	c1	qy^d^bu^rt	you do it higher ?
39.312-39.412	c1	%	uhhuh .

f

"Follow Me"

"Follow Me" marks utterances made by a speaker who wants to verify that listeners understand what is being said. Implicitly, the speaker is asking, "Do you follow me?"

Example from Data (Bmr012):

173.653-180.483	c4	s	and so then you also have a digits form which needs to be filled out every time the speaker form only once the digit form every time .
180.483-184.763	c4	s^ad	even if you don't read the digits you have to fill out the digits form so that we know that you were at the meeting .
186.657-186.857	c4	qy^g^f	okay ?
186.857-190.987	c4	s^ad	and then also if you haven't filled one out already you do have to fill out a consent form .
192.207-194.767	c4	s	and that should just be one person whose name i don't know .
197.629-197.829	c4	qy^g^f	okay ?

g

Tag Question

A Tag Question is usually a short question which follows an utterance. The speaker is seeking confirmation of his or her previous utterance.

Common Tag Questions:

Doesn't it?
Won't it?
Correct?
Isn't it?
Right?
Yes?
No?

A Tag Question often follows an Understanding Check <bu> (page 47) as in Example 1 below.

Example 1 from (Bmr012):

92.853-93.433	c1	s^bu	we're recording .
92.951-93.131	c2	%	okay .
92.981-93.241	c4	fg	okay .
93.241-94.531	c4	s^na	so we are recording .
93.433-93.733	c1	qy^g	right ?

Example 2 from Data (Bmr012):

67.241-68.771	c5	s	uh - looks kind of low on channel five.
68.771-70.301	c5	qy^g^rt	no ?

j Jokes

This tag marks utterances in which the speaker is making fun of something someone has said or if the speaker is trying to make a joke.

Example from Data (Bmr012):

498.752,501.182	c2	sj	a lot of the errors i think are out of vocabulary .
501.182,503.472	c2	s	so it's like p z m is three words.
506.032,509.052	c2	s fh	there's no language model for p z m or um ==
508.977,510.707	c4	qy^bu^rt	did you say there's no language for p z m ?
510.782,512.232	c2	%-	no language model i mean - those ==
512.677,513.087	c4	qy.%--	do you mean ?==
513.087,515.177	c4	qy^d^rt	so every time someone says p z m it's an error ?
515.577,517.617	c4	sj^co^j	maybe we shouldn't say p z m in these meetings .
517.297,519.607	c2	fgls^df	well - well there's all kinds of other stuff like jimlet .

m Mimic Other

This tag marks when a speaker mimics another speaker's utterance, often as an indication of confusion or a request for clarification. Mimic Other represents that there is a reason for a repeat, whether the repeat is verbatim or not.

Hypothetical Example 1:

Sp.1	sj^co	I think we should hunt down the jabberwocky with ice picks and headlamps.
Sp.2	s^t1^m^rt	Ice picks? Headlamps?

Example 2 from Data (Bmr012):

287.643-289.093	c2	qy^d	so is that - that's twenty minutes or so ?
290.299-290.769	c4	s^na^m	or so.

Example 3 from Data (Bmr012):

105-106	c3	qo^t	what else we got ?
107-109	c1	s^co^t	new version of the presegmentation .
109-112	c3	s^bk^m^rt,3b	new version of presegmentation .

r Repeat Self

This tag is used when a speaker repeats his or herself. This often occurs when the speaker lacks confidence of his or her utterance or in response to a Repetition Request
 (page 46).

Example from Data (Bmr012):

58.784-60.504	c3	qw^t3	so why didn't you get the same results and the unadapted ?
62.153-64.053	c3	qw^t3^r	why didn't you get the same results and the unadapted ?

rt

Rising Tone

This tag marks rising tone in the speaker's voice at the end of an utterance. This occurs in both questions and statements. Although intonation is not a type of DA, it's an important voice quality that provides useful information for automatic speech recognition.

Example from Data:

Bmr012.1829.59-1837.58 s^rt but hand - handhelds are built wi- - with
sort of uh - anti shock sort of things so that
it - it is less uh - susceptible to hand noises.

Group 9: Conventionalities

<fe, ft, fw>

fe Exclamation

This tag marks an utterance which expresses a speaker's excitement or surprise.

Example from Data (Bmr012):

326.848-330.118	c0	s.%-	and then we noticed that the reference was always one off with the ==
330.758-331.738	c4	s^fe	oh no .

Example from Data (Bmr012):

1613.21-1616.64	s^ad		it - i didn't realize but we also have to get a tuner - the receiver - the other end .
1616.64-1617.87	c4	s	that's uh - four thirty [dollars] .
1618.7-1620.65	c4	fh %-	um - and then also ==
1619.75-1620.02	c3	sj^fe	wow .
1619.96-1620.99	c2	qy^d^rt	for - for each ?

ft Thanks

This tag marks an utterance in which a speaker thanks someone.

Example from Data (Bmr012):

200.767-202.047	c5	qy	do you want this adam ?
203.018-203.378	c4	s^h s^aa	uh - sure .
204.918-205.788	c4	s^ft	thank you .

fw

Welcome

This tag marks any sequence of words that is or functions as "you're welcome." The speaker directs this at a previous person's utterance.

Common Welcomes:

- uhhuh
- okay
- you bet
- no prob
- no problem
- don't mention it
- sure
- sure thing
- you betcha
- yeah
- glad i could be of some help

Section 5

Common Dialogue Act Sequences

S

Example from Data (Bmr012):

94.861-99.771	c4	fg s^t	um - everyone should have at least two forms possibly three in front of you depending on who you are .
95.13-95.23	c0	%	okay .
100.101-102.326	c4	s	<i>um - we - we're doing a new speaker form</i>
102.326-104.551	c4	s^ad	and you only have to spea- - fill out the speaker form once .

fg|s^cs

Hypothetical Example:

Sp.1 fg|s^cs uh - | the font should be bigger

sj^cs

Example from Data (Bmr012):

125.028-126.568	c0	fg qy	um - may i make one suggestion ?
126.568-131.198	c0	sj^cs	<i>instead of age put date of - uh - year of birth because age will change but ==</i>
126.731,126.961	c4	%	sure .
132.117-132.347	c4	s^bk	oh .
132.336-133.516	c0	s	the year of birth changes -
133.516-134.696	c0	s^bsc	you know - stays the same usually .

s^bu

Example from Data (Bmr012):

1019.95-1023.61	c8	s^bu	you're maintaining it in - in a place that wouldn't be publicly readable that - that kind of stuff.
1023.61-1023.79	c8	qy^g	right ?

s^na

Example from Data:

609.401-614.591	c2	fg s^rt.%-	and actually - we actually um - used switchboard telephone bandwidth models which i guess ==
609.47-610.19	c8	s^ba	that's amazing .
609.684-609.954	c1	%	okay .
613.907-614.217	c1	%	yeah .
613.917-614.757	c3	s^na	i was just going to say ==
614.444-616.474	c0	fg s.%-	well that's - those are the only we- - ones there are ==
614.757-615.087	c3	fh	yeah .

s^aa

Example from Data:

872.41-874.05	c3	qy^d^ng	not a j but there seems to be bitter sarcasm,so the real value of the data base is these ?
874.4-874.86	c4	s^aa.%-	yeah abso- ==

qy^g

Example from Data (Bmr012):

66.331-66.901	c5	%--	adam i'm not ==
67.241-68.771	c5	s	uh - looks kind of low on channel five.
68.771-70.301	c5	qy^g^rt	no ?
70.86-72.37	c4	s^ad	channel five s- - speak again .

70.919-71.609	c8	hello .	
71.081-71.511	c5	s^t1^am	maybe not .
72.37-73.61	c4	s^aa	yeah that's alright .

s^bk

Example from Data (Bmr012):

62.153-64.053	c3	qw^t3^r	why didn't you get the same results and the unadpated ?
62.79-63.17	c4	%	but ==
64.235-68.995	c0	s^t3	oh because when it estimates the transformer pro- - produces like a single matrix or something .
67.73-69.01	c3	s^bk^t3	o- - oh oh i see .

Section 6

Complex Dialogue Act Sequences

qy^bu^rt

Example from Data (Bmr013):

998.989-1002.18	c8	s^bu^rt	and in - in fact the inter annotator agreement was not that good .
1001.13-1006.05	c3	s^aa	yeah i mean - it was- ==
1002.18-1002.52	c8	qy^g^rt	right ?
1002.99-1003.94	c8	qy^bu^rt	on the harder ones ?
1005.29-1008.2	cB	s	it depends how you look at it and ==

fg|s^co^j

Example from Data (Bmr012):

8.953-9.803	c1	%--	let's see i should be ==?
13.954-15.514	c4	s^co	as close to your mouth as you can get it .
14.119-14.519	c3	s^co	up high .
14.519-15.369	c3	s^co^e	high as you can get .
15.365-15.425	c1	%	two .
16.022-17.542	cB	%	channel one one one .
16.356-17.526	c8	fg s^co^j	yeah on your upper lip .

fg^bk|sj^df^nd

Example from Data (Bmr012):

141.602-145.412	c0	s^df^nd	well the thing is if ten years from now you look at this form knowing that ==
144.588-148.398	c8	fg^bk sj^df^nd	yes but what we care about is the age at - at the recording date rather than the ==

146.066-147.426 c3 s^nd

but there's no other date on
the form

Works Cited

1. Jurafsky, Dan, Shriberg, Liz, Biasca, Debra. Switchboard SWBD-DAMSL Shallow-discourse-Function Annotation. Shallow-Discourse-Function Annotation University of Colorado at Boulder & SRI International, 1997
2. Levinson, Stephen C. Pragmatics. Cambridge University Press, Cambridge: 1983.

Index

Alphabetical Listing of Tags:

Tag	Page
aa	30
aap	34
am	36
ar	36
arp	34
b	26
ba	44
bc	44
bd	45
bh	32
bk	27
br	46
bs	46
bsc	45
bu	47
by	49
cc	40
co	39
cs	41
d	51
df	48
e	50
f	52
fa	49
fe	56
fg	28
fh	29
ft	56
fw	57
fx	40
g	52
h	31
j	53
m	54
na	37
nd	35
ng	35
no	37
oo	41
qh	32

qo	42
qr	20
qrr	21
qw	21
qy	22
r	54
rt	55
s	17
sj	18
t	51
tc	42
t1	33
t3	24
x	25
2	50
%	23
%-	23
%--	24