Simple Metadata Annotation Specification

Version 6.2 – February 3, 2004

Linguistic Data Consortium

www.ldc.upenn.edu/Projects/MDE

1 Introduction

Working with pre-existing verbatim transcripts of broadcast news and conversational telephone speech, annotators will identify several types of *metadata*. The ultimate goal of metadata annotation is to support the production of transcripts that are maximally readable. This document defines the Simple Metadata (SimpleMDE) annotation task. SimpleMDE identifies a subset of the full range of possible metadata phenomena. Under SimpleMDE, annotators identify fillers, strings of deletable words (deletable regions, or delregs) within edit disfluencies, and SUs ("semantic", "syntactic" or "sentence" units). Transcripts annotated for metadata can be "cleaned up" to enhance readability; for instance, delregs and fillers might be removed and each SU presented as a separate line within the transcript. Each metadata type is described in detail in the sections below.

Two types of data will be annotated: broadcast news and conversational telephone speech. The data types are different from one another in a number of ways. Broadcast news data is single channel data, meaning that all speakers' voices are presented in a single audio file. Broadcast news comprises read speech, formal interview format, man-on-the-street interviews and some spontaneous speech, though usually not conversational. Telephone speech is two channel data, so that each speaker in the phone call is presented on a separate audio signal. Telephone speech is spontaneous, and much of it quite informal and conversational. In some ways broadcast news is simpler for metadata annotation. Annotators will observe fewer fillers and disfluencies. Annotators can also expect to see quite a difference in the prosody and discourse structure of broadcast news data, particularly read speech, when compared to conversational telephone speech. Intonation and breath patterns are likely to be very different in read speech and sentences will be structured differently. This may affect the placement of some SUs.

For both data types, annotators will perform metadata annotation using both the audio file plus a verbatim transcript, executing multiple passes over the data using a tool customized for each specific annotation task. For telephone speech annotators will label one speaker at a time, but will always have access to both audio channels as well as the complete 2-channel transcript. The first metadata annotation pass focuses on filler and delreg annotation, and the second pass identifies SUs. Additional passes over the data that focus on quality control are conducted to improve annotation accuracy and consistency.

1.1 Definitions and concepts

For purposes of SimpleMDE annotation[1], a turn is defined as a portion of speech uttered by a single speaker and bounded by silence from that speaker. The

---

[1] The definition of *turn* used by SimpleMDE was adopted to support annotation consistency by allowing for simple turn-based decision rules. We take no stance about how our definition of turn might map to the various conceptions of turn within the larger research community.

portion of speech contains no noticeable pauses[2]. Turns can be long or short; they might consist of a monosyllable, a phrase, a sentence, a grammatically incomplete utterance. A single turn may consist of an incomplete SU, a single complete SU, or multiple SUs. Turns can overlap, with two or more speakers talking at once, or they may occur sequentially. We say that a turn boundary occurs whenever a single speaker takes a noticeable pause in their speech[3]. Note that isolated speaker or background noise does not count as a turn, and that noises at the edges of turns are not considered part of that turn (see Section X for more about annotating noise).

Occasionally within this document we employ the term utterance, by which we mean a contiguous stream of speech from a given speaker. We use the term utterance when we do not wish to take a stance about whether a given span of speech constitutes a turn, an SU, or some other conceptual construct. Within this document, utterance simply means an unanalyzed stream of speech from one person.

Annotators should note that segment boundaries, or areas of white space separating the speech of multiple speakers in the reference transcript, do not necessarily correspond directly to our notion of turn. Segment boundaries are inserted during the initial transcription process.[4] Typically they appear around regions of silence, and may therefore coincide with a speaker turn as we define it here, but they are also inserted elsewhere for ease of transcription and they show variation from one file (transcriber) to another. In making decisions about MDE annotations, annotators should rely on the concept of turn as defined above.

Related to turn is the concept of control of the floor. When Speaker A is talking to Speaker B and B is listening without attempting to take control of the conversation, Speaker A can be said to have control of the floor. The person controlling the floor bears the burden of moving the discourse along. Change in control of the floor can happen by a number of mechanisms including regular turn-taking and interruptions.

The concepts of turn and control of the floor have bearing on MDE annotation, in particular how to distinguish various phenomena like discourse markers, discourse responses and backchannels; and the assignment of incomplete SUs

---

[2] The exact threshold of "noticeable" is an empirical question that deserves further research, but such research is beyond the scope of our current SimpleMDE annotation task. As a rule of thumb we define noticeable pauses as approximately 0.5 seconds or greater.

[3] Note that in some rare cases, a noticeable pause may occur that does not warrant a turn break, like when a speaker pauses to clear his/her throat or swallow. There may be other outlier cases as well. The 0.5 second threshold is adopted here merely as a convenience, to promote annotation consistency in difficult examples.

[4] MDE transcripts come from a number of different sources, so there is a lot of variability in where segment boundaries can occur.

as distinct from restart disfluencies. We adopt these ten principles in the discussion of MDE annotation in the sections that follow:

1. Only one speaker holds the floor at any given time.
2. A turn is a portion of speech uttered by a single speaker and bounded by silence from that speaker.
3. A speaker's turn may or may not coincide with them holding the floor.
4. A speaker's turn may therefore overlap with that of another speaker, who is holding the floor.
5. Overlaps that cause the other speaker to concede the floor are successful interruptions.
6. Overlaps that do not cause the other speaker to concede the floor are failed interruptions or backchannels.
7. Backchannels are passive contributions to the discourse. Singly or in chains, they constitute an entire speaker turn. They are not interruptions.
8. Discourse markers are active or responsive contributions to the discourse. If they overlap with another speaker's turn, they are interruptions. They can occur in non-overlap contexts, for instance as a signal that the current speaker now controls the floor.
9. Interruptions, whether successful or failed, can terminate in a complete or incomplete SU. If an interruption succeeds, the turn into which it breaks may comprise a complete or incomplete SU.
10. An incomplete SU is defined not on prosodic criteria alone, but syntactically/grammatically.

1.2   A word about notation

Within this document, the following notation conventions are used within the examples.

Fillers and Edit Disfluencies

[text] *	'text' is a delreg and '*' is its interrupt point
text	'text' is the corrected portion of a disfluency
text	'text' is a filler (DM, FP, Aside/Parenthetical or EET)
text	'text' is a discourse response
{text}	'text' is an Aside/Parenthetical filler
te-	'te' is a partial word token
text	'text' is a word relevant to the discussion and italicized for emphasis

SUs
Sentence-level SU Breaks

Symbol	Description	Shorthand
/.	Statement SU break	Statement SU
/?	Question SU break	Question SU
/@	Backchannel SU break	Backchannel SU
/-	Incomplete SU break	Incomplete SU

Sentence-internal SU Breaks

Symbol	Description	Shorthand
/&	Coordination SU break	Coordination SU
/,	Clausal SU break	Clausal SU
∅	No break (where one might be expected)	

2 Fillers

We consider four kinds of fillers: filled pauses (FP), discourse markers (DM), explicit editing terms (EET) and asides/parentheticals (A/P). None of these filler types alters the propositional content of the material into which they are inserted, and their insertion does not depend on the word identities of the surrounding material. Annotating fillers consists of identifying the word(s) that act as a filler and labeling the filler with the appropriate type.[5] Each filler type will be considered in turn below.

2.1 Filled pauses

Filled pauses are hesitation sounds that speakers employ to indicate uncertainty or to maintain control of a conversation while thinking of what to say next. Filled pauses do not add any new information to the conversation (other than to indicate the speaker's hesitance) and they do not alter the meaning of what is uttered. For instance,

> **Um** I do **uh** some **uh** woodworking myself.

Filled pauses can occur anywhere in the stream of speech. For English, the most common filled pauses are as follows:

English Filled Pauses

ah	uh
eh	um
er	

Other sounds or non-lexemes can occasionally be used as a filled pause, and some speakers may adopt an idiosyncratic filled pause noise that does not appear on the above list. Annotators should label all words that function as a filled pause as such.[6]

---

[5] Note that each filler also contains an interruption point at the left edge of the filler's extent. Annotators are not required to explicitly label this filler IP; this information will be added automatically by the annotation tool. Further note that when the interruption point from an edit and filler coincide, they will share the same (single) IP tag, as in this example:
[red] * **uh** blue
Since they are not manually annotated, these left-edge filler IPs are **not** reflected within this document.

[6] The annotation tool pre-identifies the most common FPs (limited to *ah, eh, er, uh, um*) and automatically pre-annotates them as filled pauses. Pre-identified filled pauses are displayed in blue font; annotated filled pauses are displayed with blue underlining. Annotators must verify that each pre-labeled filled pause candidate is actually acting as a filled pause, and must remove the

Be aware that some words that can be used as FPs may have other functions, like question responses and backchannel cues, elsewhere in the discourse. Annotators should label words as filled pauses only when they indicate a speaker's hesitation.

Furthermore, filled pauses can sometimes function like an editing term as well. However, these items will not be doubly-annotated as EETs, and should only be annotated as FPs. (See Section 2.3 for a more complete discussion of EETs).

2.2 Discourse markers

A discourse marker (DM) is a word or phrase that functions primarily as a structuring unit of spoken language. To the listener, a DM signals the speaker's intention to mark a boundary in discourse. Discourse markers are inherently active contributions to the discourse, and signal such activities as a change in the speaker, taking or holding control of the floor, giving up the floor or the beginning of a new topic:

```
That gets on my nerves too.  So anyway, tell me about your new
job.

A: So how do you make this soup?
B: First you take a couple of carrots and chop them. Okay and
then you sauté them in butter.
```

It is nearly impossible to establish an exhaustive list of DMs for a given language due to their wide range of functions and the difficulty of defining them precisely; moreover, discourse markers are subject to much dialectal and individual variation, and novel formations can serve as discourse markers which means that any list quickly becomes out of date. The list below represents some of the most common DMs in American English:

```
actually              now
anyway                okay
and yeah              see
basically             so
I mean                well
let's see             yeah
like                  you know
                      you see
```

This list is not exhaustive. Annotators should tag all words functioning as discourse markers as such.[7]

---

annotation from any non-filled pauses. On occasion, annotators may recognize other words in a transcript that function as a filled pauses. These words should also be manually tagged as filled pauses.

[7] The annotation tool pre-identifies the most common discourse markers (limited to the list above). However, the tool **does not** pre-annotate discourse markers as such. Pre-identified discourse markers are displayed in red font. Annotators must **manually annotate** each of these

Some words that function as discourse markers can also act like backchannels in other contexts. Backchannels, a type of SU, are words or phrases that provide feedback to the dominant speaker by indicating that the non-dominant speaker is still engaged in the conversation (though not actively participating at the moment). In the example below, *okay* is functioning as a backchannel rather than a DM:

```
A: I've lived in Friendship Heights for years.
B: Okay.
A: But I'm thinking of moving a little further out.
```

Many words and phrases that are used as discourse markers also have other literal meanings:

```
Do you know how many minutes we're supposed to talk for?

The situation right now is that we're moving in three weeks.
```

We will label as DM only those instances that function to structure the discourse and do not carry separate meaning. It can sometimes be difficult to distinguish when a word or phrase is functioning as a discourse marker and when it is acting as a content word. In cases of this kind of uncertainty, annotators should leave the word or phrase unannotated.

2.2.1 Discourse Responses

In many cases, discourse markers are used to express a response to what the other speaker is saying in addition to structuring the discourse. We call these special cases discourse responses (DRs) and label them as such[8].   DRs can be single words or phrases, and multiple DRs can occur in succession (in which case each should be separately tagged as a DR). Speakers often use DRs to respond to what the other person has just said, and simultaneously initiate their own attempt to take the floor:

---

pre-identified words, as well as any other word that is acting as a discourse marker. Annotated discourse markers are displayed with red underlining.

[8] Currently, DRs are tagged as an attribute of a discourse marker. DRs seem to occur with some regularity particularly in telephone speech, and they have been adopted for this version of SimpleMDE as a compromise to solving a handful of difficult annotation decisions. It is not yet known how DRs as currently defined map to discourse markers and/or backchannels, or whether in fact they motivate a separate SU type in their own right. It is possible during post-processing to merge DRs with the other discourse markers; to create backchannel SU breaks after DR words; or to create a new SU boundary (SU-response?) after DRs.   However, we believe that the proper treatment of DRs is not yet known, so for the time being we will retain them as attributes of discourse markers. This allows us maximum flexibility while preserving the most information about our annotation decisions. After annotating a sufficient volume of material, we plan to examine all items labeled with the DR attribute and decide how they should ultimately be treated.

```
A:  I think /, he's done a terrible job in that position.  If it
were up to me /, I'd fire him.
B:  See I don't know if I'd go that far.

A:  I think /, they'll capture him sooner or later.
B:  Well Ø I'm not so sure.

add yeah example
```

DRs are not to be confused with direct answers to a question. Direct responses are tagged as statement SUs, not discourse responses. DRs are not only responsive, but also function like regular discourse markers (structuring the discourse), whereas direct responses are only answers to questions.

On the surface, DRs may look a lot like backchannel SUs, which are also responsive in nature. Unlike backchannels, DRs are *active* responses. They not only respond to what the previous speaker is saying, but they also actively structure the discourse, conveying something like: "I'm responding to what you're saying, and I'm also going to say something beyond that". Backchannels, on the other hand, convey something like "I'm still listening, keep on talking". (See Section XX for a full discussion of backchannels.)

In the vast majority of cases, because they act as a response to what the previous speaker has said, DRs occur SU- and turn-initially, or within a string of other DRs.

It can sometimes be difficult to tell DRs and "regular" DMs apart. When in doubt, annotators should label the item with the DR attribute. A full discussion of the distinction between regular discourse markers, discourse responses and backchannel SUs appears in Section 4.4.3.1 below.

2.2.2 You know as generalizer

One common discourse marker that has other functions is *you know. You know* is sometimes employed as a (utterance-final) generalizer, allowing the speaker to extend their specific examples to a more general observation:

```
It's not like Boston, NY, Philly. Or you know...

If the stress is really on she'll break down. And you know...
```

Such cases should be annotated as regular DMs. (See Section 4.6.4.4 for additional discussion of generalizers.)

2.2.3 Like

One particularly difficult word that serves many functions, including acting as a DM, is *like.*   It can act as a preposition, a conjunction, an adverb or adjective, a verb, and even as a substitute for *say* or *said.* The following examples illustrate some of the non-discourse marker functions of *like*.

Like as a preposition meaning 'similar to':

        They're *like* bermuda shorts but a little longer.

Like as a preposition meaning "as if":

        It looks *like* she's wearing bermuda shorts.

Like used with 'to be' as a verb of quotation (instead of *say* or *said*):

        He was *like*, "I'm wearing bermuda shorts."

versus these examples, in which *like* functions as a discourse marker:

        She was **like** wearing bermuda shorts.

        But he played the character in it that was very **like** gross.

        Like when I was **like** in high school and junior high school I used
        to hate it.

In addition to context, prosody (especially the presence of a pause) can help to distinguish cases of DM from non-DM *like*, although some confusing cases will remain, particularly between verb of quotation *like* and discourse marker *like*. Annotators should label only those cases in which *like* is readily identifiable as a discourse marker. This is a tricky decision to make, so the default decision is to leave *like* unannotated. Annotators should use the "difficult decision" label to register cases of uncertainty.

2.2.4 So

So is another item with many different uses in discourse that may be difficult to distinguish from one another. The most common uses of *so* apart from its function as a discourse marker are as a subordinating conjunction:

        We brought out pictures of her grandparents *so* she'll get to know
        them.

and as an adverb meaning, roughly, *therefore:*

        I grew up on a farm *so* I always had outdoor pets.

As a discourse marker, *so* often serves to indicate a topic change:

        I'm not sure, to be honest with you. **So,** your backpacking trip
        through China, that seemed to be an exciting adventure.

Discourse marker *so* can often stand alone as a turn, or stand at the end of an utterance when the speaker trails off (see Section XXX for discussion of SU assignment and fillers)

        A: That was a lot to go through. **So...**
        B: Wow, you must be relieved to be over that.

It is often particularly difficult to distinguish uses of *so* as a conjunction versus *so* as a discourse marker. Trailing *so,* as in the example above, should be annotated as a discourse marker unless there exists compelling evidence that it is behaving as a conjunction (e.g., the speaker continues in the next turn with the rest of a conjoined statement). In other cases of uncertainty, annotators should err on the side of exclusion, and label only those cases in which *so* is readily identifiable as a discourse marker. Cases of uncertainty should also be noted with the "difficult decision" tag (see Section 7).

2.3   Explicit Editing Terms

Explicit editing terms (EET) are fillers that occur within the context of an edit disfluency (See Section 3.2 for a discussion of the structure of edit disfluencies). EETs consist of an overt statement from the speaker recognizing the existence of disfluency. These typically consist of a short phrase such as *I mean, sorry, excuse me, rather,* etc:

```
And when he gets free again he will have no compunction but to
[complete that that same kind of lifestyle] * I mean continue
that same kind of lifestyle.

I think [one of the positive things] * or rather one of the
things that can come out of it is not just discipline.
```

Filled pauses can sometimes function like an editing term as well. However, these items will not be doubly-annotated as EETs, and should only be annotated as FPs.

EETs are quite rare, and annotators should label only those cases in which a word or phrase clearly functions as an EET. Note that they can occur anywhere within the disfluency, including after the correction, and edit disfluencies can contain more than one EET:

```
Three hundred fifty-six residents were [killed] * er injured
rather.
```

EETs are optional elements for all edit disfluencies.

2.4   Asides and parentheticals

Asides occur when the speaker utters a short side comment on a new topic then returns to the main topic being discussed. An aside can also occur when the speaker addresses someone who is not part of the immediate conversation (e.g., when someone else walks into the room during a phone call). Asides are often accompanied by prosodic features like a shift in intonation or the presence of a pause.

Asides can occur anywhere in the utterance. Their extent will be identified and labeled like other filler types.

```
And our neighbors were, {oh I don't know} kind of slimy.

He has now for about {oh gosh, how long has it been} ten years I
guess.

And I couldn't help thinking when that last question {it was a
funny question} came up.
```

Parentheticals are similar to asides in that they are brief remarks that break the flow of the larger utterance, but unlike asides the remark is on the same topic as the larger utterance. In standard writing, parentheticals are often accompanied by dashes or parentheses. They often display similar prosodic features to asides. As with asides, parentheticals are treated as fillers.

```
The head of the United Auto Workers Union responded by calling
the move {his words} nuts.
```

Parentheticals are somewhat common in broadcast news speech, while asides occur more frequently in conversational telephone speech. We do not distinguish between asides and parentheticals for annotation purposes, but group them together as one filler type. [9]

Occasionally a filled pause or discourse marker can occur within the extent of an aside or parenthetical. When this occurs, annotators should separately label the smaller FP or DM as its own filler type, within the extent of the larger filler. For instance, consider this example of a discourse marker occurring within an A/P (DM *like* is underlined here for clarity):

```
I talked about how a lot of the problems they have to overcome uh
{It's a very like complex situation}, to go into space.
```

A/Ps should be used sparingly. They should not be confused with relative clauses that modify a noun phrase (see Section 4.6.4.6 for treatment of relative clauses):

```
And we are presently in receipt of a site permit which will allow us
to have certain emissions up to a certain tonnage.
```

Further, they are distinct from refinements and generalizers (see Section 4.6.4.4) that are used to modify a previous statement:

---

[9] A fuller treatment of asides and parentheticals might distinguish the two types, and might classify them as a type of SU rather than as fillers. Such a treatment is beyond the scope of SimpleMDE. As with other filler types, annotators must identify the full span of text functioning as an A/P, rather than just tagging the end point (as with an SU), so for ease of annotation A/Ps are included with fillers. Asides/parentheticals are not currently evaluated in the MDE Program.

```
Right now I believe there are two policies that the government
uses to modulate immigration, three actually.
```

Some very common words and phrases might be mistaken for A/Ps but should not be tagged as such. These short commentary words typically lack the prosodic features that identify A/Ps. "Say", "I think" and "for example" are common examples of non-asides that should not receive any special markup, e.g.,

```
And when someone is say out of high school

And if he for example wanted to be a ballerina
```

The distinguishing feature of asides and parentheticals is that they break up the stream of discourse. In standard writing, they are often represented with double dashes to set them off from the rest of the utterance. Annotators should label only those cases in which a word or phrase clearly functions as an aside or parenthetical.

See Section 5.5.1 for a discussion of SUs and A/Ps

2.5   Strings of fillers

Speakers often produce multiple contiguous fillers. These should be separately annotated as individual items, rather than grouped together as a single unit. This is true even for strings of fillers of the same type. For some types of multi-word fillers (like discourse markers), it's difficult to know whether a speaker is producing a single or multiple filler instances. When in doubt, annotators should err on the side of separating the fillers into separate units.

```
oh really example
```

3   Edit disfluencies

3.1   Introduction

Edit disfluencies are portions of speech in which a speaker's utterance is not complete and fluent; instead the speaker corrects or alters the utterance, or abandons it entirely and starts over.

3.2   Structure of edit disfluencies

Edit disfluencies have a more complex internal structure than fillers, consisting of the deletable region, interruption point, (optional) explicit editing terms and the correction:

- **Deletable region (delreg)**
  The deletable region comprises the original portion or reparandum in simple edits. This is the speaker's initial attempt to formulate an utterance that exhibits some kind of disfluency, and later gets corrected or (in the case of

restarts) abandoned entirely. In complex edits, the delreg consists of the consecutive string of words drawn from pieces of several component reparanda, all of which may be optionally deleted upon cleanup.

- **Interruption point (IP)**
  The interruption point is the point at which the speaker breaks off the delreg with an explicit editing term (EET), a repetition, revision or restart. All edit disfluencies have at least one interruption point at the right edge of the delreg. Complex edits contain multiple interruption points.

- **Editing Phase**
  The editing phase is an optional components of edit disfluencies. It consists of an overt statement (*I mean, rather,* etc.) from the speaker recognizing the existence of disfluency, and can occur anywhere in the disfluency, including after the correction. The editing phase can be filled by a discourse marker, filled pause or explicit editing term (see Section XX for discussion of these filler types).

- **Correction**
  The correction consists of the portion of the utterance that has been repaired by the speaker and is fluent; the correction will always be retained upon cleanup. Restarts do not contain a correction.

3.3   Annotation of edit disfluencies

Delreg: Annotators select the full extent of the deletable region for each disfluency, and apply the delreg label to those words (shown in [square brackets] within this document). Rules for identifying the full extent of the delreg appear in the sections that follow.

Interruption point: Interruption points at the right edge of a delreg are automatically identified and labeled by the annotation tool. Annotators do not explicitly label such IPs. Multiple IPs within the extent of the deletable region of a complex disfluency are manually labeled by annotators. (IPs are indicated by asterisk * within this document.)

Explicit Editing Term (EET): EETs are categorized as a type of filler; therefore annotation of EETs is detailed in Section 2.3 above.

Correction: For SimpleMDE, annotators do not explicitly tag the corrected portion of edit disfluencies. (Corrections are underlined within this document).

ED type: Within SimpleMDE, annotators do not classify edit disfluencies according to their type. Note that this also means annotators do not distinguish simple edits from complex edits.

3.4 Types of edit disfluencies

Edit disfluencies can be divided into four categories: *repetitions*, in which the speaker repeats the same word, part of a word or sequence of words more than once; *revisions*, in which the speaker explicitly modifies and replaces a portion of the utterance; *restarts*, in which the speaker abandons the utterance s/he's begun and simply starts over; and *complex disfluencies*, which consist of multiple or nested edits.

For purposes of SimpleMDE, annotators do not distinguish between the various types of edit disfluencies. However, for explanatory purposes, an inventory of types is presented in the sections below.

3.4.1 Repetitions

In the case of repetitions, the delreg is repeated in the corrected portion of the utterance. The repeated portion can be a word fragment, a complete word, or more than one word:

```
[He-] * he's really out of line, or at least that's what I was
told

[The thing] * the thing I wanted to cover with you is

[After the] * um after the third paragraph, what did it say next?
```

In cases of repetition, cleanup (removal) of the delreg does not alter the meaning of the utterance at all.

Keep in mind that not all repetitions are disfluencies. Examples like repeated backchannels and the use of repetition for emphasis are entirely fluent and should not be labeled as edit disfluencies:

```
We had a really, really wonderful time.
```

3.4.2 Revisions

Revisions are very similar to repetitions, but the corrected portion that replaces the delreg modifies its meaning, rather than simply repeating it in whole or in part. The revision might change the meaning entirely, or may modify or clarify it only slightly. For instance,

```
Our children like [the dog] * I mean the cat, that the neighbors
got.

Show me flights [from Boston on] * uh from Denver on Monday.

So we built a cradle for it, and [we got th-] * once it was
turned we got one cutout on the saw.

[Do you] * by mistakes do you mean just like honest...
```

3.4.2.1 Revisions with information loss

Another complication that occasionally affects revision disfluencies occurs when information is present in the delreg, but that information is not contained within the revision (i.e., the correction). This most often happens when the delreg contains a full noun or noun phrase, but the corrected portion contains only a pronoun:

> I gave a book [from Susan] * **I mean** <u>by her</u> to Ralph.
>
> I flew [to Dallas] * **I mean** <u>from there</u>, yesterday.

Revision disfluencies of this type are extremely rare, and they will be annotated in the standard way, despite the fact that this may result in a small loss of information upon optional cleanup.

3.4.3 Restarts

In the case of restarts, also known as false starts, a speaker abandons an utterance or constituent, and neither corrects it nor repeats it partially or wholly, but instead restructures the utterance and starts over. A restart is often followed by a noticeable pause and may be accompanied by other prosodic features like resetting of pitch. In restarts, the corrected portion is empty. This means that important information contained within the restart delreg may not be repeated or restated by the speaker later in the discourse.

Most typically, a restart does not add information to the overall discourse. These are cases in which the speaker abandons an utterance or constituent and then starts over, restructuring the original utterance entirely:

> [Do you know whether he wants] * how does he like it to be cooked?
>
> I can't really say [that there should be a] * what type punishment there should be.
>
> You know, [they always end up] * sometimes they change their minds.
>
> [It's also] * I used to live in Georgia.

Sometimes, however, the original utterance contains important information content that may not be presented elsewhere within the discourse:

> A: [I happen to live not too far away from] * well, I've actually worked for the company that has been blamed for the Challenger disaster.
> B: Oh, is that right?

In order to enhance inter-annotator consistency and maximize efficiency, the full extent of all restarts are labeled as delreg, even when they contain additional information that may not be present elsewhere within the discourse.

Note that restarts can be sometimes confused with incomplete SUs. Restarts are limited to those cases where a speaker interrupts him or herself, restructuring the utterance and continuing with the discourse. See Section 4.5.4 for further discussion.

3.4.4 Complex disfluencies

In spontaneous speech, speakers often produce a series of disfluencies in succession. These can be serial, where one occurs directly after another, or they can be nested, where some component of one disfluency itself contains another disfluency.

3.4.4.1 Complex original utterance

In cases where the reparandum, or original utterance, itself contains multiple (serial, adjacent) disfluent utterances, annotators should identify the *maximal extent* of the disfluent portion, starting with the left edge of the first disfluency and continuing to the right edge (IP) of the final disfluency. All of this is tagged as a single delreg, although it contains parts of multiple disfluencies. To indicate that the delreg contains additional internal structure, annotators explicitly identify the interruption points contained within the delreg. The final interruption point at the right edge of the delreg is automatically identified for all disfluencies and is not separately annotated (though it is represented in the examples that follow).

```
Yeah but [the * the big * the b- * the big] * um the betrayal or
whatever she called it

[It's * this is like only like the third or fourth time I've * I
ne- * I'm real bad about] * I never make the phone calls
```

3.4.4.2 Complex correction

In some cases, the corrected portion of one edit disfluency contains another disfluency:

```
                       [---corrected---]
1)    I'm sure [the] * that  [the] * the staff learn what's normal.
                                [disf]

                        [----corrected ----]
2)    I flew [from Dallas] * uh from [Aus-] * Austin this morning.
                               [disf]

                               [----corrected---]
3)    It gets kind of expensive [to renew] * to [renew] * renew it.
                                    [disf ]
```

Annotating overlapping structure within a complex disfluency is beyond the scope of SimpleMDE annotation. Instead, annotators will identify just the deletable region of each disfluency. As a rule of thumb in these cases, annotators should take the rightmost (final) instance of each word within the disfluency as the corrected portion; the other instances of each word are part of the delreg. For SimpleMDE, the examples above are annotated as follows:

```
1)     I'm sure [the] * that  [the] *  the staff learn what's normal.
              delreg         delreg

2)     I flew [from Dallas] * uh from  [Aus-] * Austin this morning.
              ---delreg---   FP        delreg

3)     It gets kind of expensive [to renew] * to [renew] * renew it.
                                   --delreg--       delreg
```

3.4.4.3 Complex disfluency and restarts

In many cases a complex disfluency ends in a restart, or the restart will contain another disfluency within it, e.g.:

```
[I hoped that we could be better than * than um expecting]*,
Well, we should have higher expectations than that.
```

The repetition contained within this restart (than than) cannot be separately annotated as a repetition because of the prohibition on nested disfluencies. In cases like these, annotators should select the maximal extent of the deletable region – in this example, the full extent of the restart – and should indicate any internal interruption points within that extent.[10]

3.4.4.4 Multiple disfluencies in one utterance

When an utterance contains multiple or serial disfluencies, the annotator should apply the rules for annotating edits that are described above: selecting the maximal extent of each delreg; annotating internal interruption points; and treating the rightmost occurrence of each word within the disfluency as the correction.

```
[Did they, * did they] * like on bottles, did they give you [a] *
so many cents back?

[When I] * when I see that money taken out of my paycheck [each *
each week * or each] * every other week, [I] * I really think that
money's history.

Yeah, [I've * I've heard some] * I mean I've heard [sta- * sta-] *
statistics and things.
```

3.5   Disfluencies across segment boundaries

In some cases, disfluencies can occur across same-speaker segment boundaries:

```
A: [I ha-] *
```

---

[10] Note again that the SimpleMDE task is limited to identification of the maximal extent of the deletable sequence, plus any internal interruption points in delregs with multiple IPs. This annotation task does not map directly onto the identification of complex edits, nor does it support identification of the extent of the original portion (reparandum) of an edit disfluency in the case of nested edits.

```
A: I hated that car so
A: it burnt up and I did not set it on fire.
A: I have been accused of that but I did not.
```

These cases should be annotated in the same way as within-segment disfluencies.

3.6   Disfluencies and conjunctions

In many cases, a conjunction will occur at the edge of an edit disfluency. The conjunction should be excluded from the span of the delreg unless the conjunction itself is disfluent (that is, unless the speaker either repeats, revises or restructures the conjunction in the corrected portion of the disfluency). For example, the conjunctions "*but*" and "*because*" (italicized for emphasis) are excluded from the delreg extent in the following examples:

```
I agree, but [I] * I feel that uh a lot of people have gotten
lazy about voting

because [that was]* the person meant to say this.
```

When the conjunction itself is disfluent, it must be contained within the delreg, as in these examples:

```
[and if] * and when he gets free again he will continue that same
kind of lifestyle

[and i-] * but if he's not in the mood you'd better stay away
from him
```

3.7   Disfluencies and fillers

3.7.1   Delregs within the extent of a filler

Deletable regions that occur within the extent of a filler (for instance, a repetition that occurs within an aside) should be treated as a regular delreg:

```
Well [y-] * you know I suppose that's true.
```

3.7.2   Fillers within the extent of a delreg

Fillers that occur within the context of a delreg are to be labeled separately as fillers, but are contained within the extent of the delreg. For instance,

```
[I think uh that] * I think that is a tough one.
```

3.7.3   Fillers at the edge of a delreg

Fillers that occur at the right or left edge of a delreg should **not** be included within the extent of the delreg, but should be separately annotated as fillers. (See Section 4.8.1 for one exception to this rule involving incomplete SUs). For instance,

```
Um [but the job that I ju-] * I had this job that I lost.

[But to just] * Well I don't know exactly.
```

3.7.4 Fillers and complex edits

When a filler occurs between two adjacent deletable regions, the filler is excluded from the span of either delreg. Instead, the two disfluencies should be labeled as two simple delregs, with a filler standing between them:

```
[I] * uh [I don't] * I don't have any cats.
```

NOT   ```[I * uh I don't] * I don't have any cats.```

However, if the filler does not occur between two disfluencies, but rather inside one of them, the filler should be included within the maximal delreg extent:

```
[I * I really uh haven't] * I don't have any cats.
```

3.8 Disfluency conclusions

In labeling difficult examples, annotators should keep in mind that the goal of metadata annotation is to create more readable transcripts while preserving the meaning of the discourse. Annotators should not just read the transcript, but must also listen to the speech, and should use their best judgment about the speaker's intended meaning when determining how to properly label edit disfluencies.

4   SUs

4.1   Introduction

One of the goals of metadata annotation is the identification of all units within the discourse that function to express a complete thought or idea on the part of the speaker. Often times this unit corresponds to a sentence, but sometimes a sentence will contain multiple units; other times a unit is semantically complete but smaller than a sentence. For the purposes of metadata annotation, these elements are called SUs. We will not take a stand on the issue of what SU stands for, but some possibilities include: Sentential Units, Syntactic Units, Semantic Units and Slash Units. As with disfluency annotation, the goal of SU labeling is to improve transcript readability by creating a transcript in which information is presented in small, structured, coherent chunks rather than longer undifferentiated turns or stories.

Every word within the discourse must be assigned to an SU, and all SUs must be classified according to their function within the discourse.   Although SUs typically occur within a single speaker turn, they can occasionally span a turn, particularly in the case of failed interruptions and overlapping speech. SUs can be whole well-formed sentences, phrases or single words. Many short SUs that are not full

sentences or clauses but are nevertheless complete units serve to regulate the discourse.

In identifying SUs, annotators should rely on both the transcript and the audio. Syntactic and prosodic factors can influence the placement of an SU boundary. In read speech, SUs are usually easy to identify and are often accompanied by a pause. However, breath groups may not always correspond to SU boundaries in read speech; sometimes a breath occurs in the middle of an SU. Spontaneous speech presents a more difficult problem, since pauses do not always occur at major constituent boundaries, and other metadata phenomena like fillers and disfluencies can further obscure the presence of SU boundaries. In all cases, annotators should rely primarily on semantic and syntactic information, and secondarily on prosodic information, to inform their decision about where to place the boundary.

The following sections provide guidelines for recognizing the existence of SU boundaries and categorizing them according to their type. Notice that we refer to the SU *boundary* when we are discussing the annotation task. We do so because the annotation task is centered on the detection and classification of the breaks *between* the SUs in a discourse.

During annotation, annotators label only the final word of an SU; the next word uttered by the same speaker will automatically be interpreted as the beginning of a new unit. Every word in the discourse must be assigned to some SU. Each word contained between two SU boundaries is considered *part of* the same SU. Annotators will distinguish two kinds of SUs: sentence-level and sentence-internal.

4.2   Background information for non-linguists

Although deciding between SU types is relatively straightforward, it can be difficult to determine when to begin a new SU boundary and when to place two segments within the same SU. In order to enhance inter-annotator consistency, we will define SUs on the basis of some easily recognized basic syntactic units. While some sentences constitute single and complete SUs, many sentences contain multiple SUs; moreover, in conversational speech, sentences are not always well-formed or easily detected.

Clauses, on the other hand, are relatively easy to identify. A clause is something like an 'atomic sentence'. It is merely a subject plus a single simple predicate. For example, the sentence:

```
If the annotator understands clauses then she sees them both.
```

Consists of the two clauses:

```
the annotator understands clauses
she sees them both
```

We will depend a great deal on the notion of a *clause* in the discussion that follows. For both sentence-internal and sentence-level SU boundaries, the ideal placement of a break is between clauses. In nearly all cases an SU break will exist at the boundary between two clauses; the only remaining question, then, is what type of clause boundary we are dealing with.

The data being annotated for SimpleMDE is quite slippery with regard to the placement of SU boundaries, so it will be important to distinguish between a few parts of clauses as well. Specifically, the sections that follow will make use of the terms clause, predicate and noun phrases in the manner illustrated by the following set of examples:

Clauses (S):

```
Danielle loves annotation projects.
Dalal went to the EARS workshop.
Alex is typing on her laptop.
```

Predicates/Verb phrases (VP):

```
loves annotation projects
went to the EARS workshop
is typing on her laptop
```

Noun phrases/subjects (NP):

```
Danielle
Dalal
Alex
the EARS workshop
her laptop
annotation projects
```

While the examples in this section are very straightforward and not representative of the kinds of data to be annotated for SimpleMDE, the concepts they illustrate will apply to the whole range of data that annotators will encounter during SimpleMDE annotation.

The following sections provide more detailed guidelines for recognizing the existence of SUs and categorizing them according to their type. The examples that follow are all taken from real conversational and broadcast news speech data.

4.3   Sentence-level vs. sentence-internal SUs

Distinguishing sentence-internal SUs from those that occur at sentence boundaries is an important component of the SU tagging task. Sentence-level SU breaks are fundamental and directly support the SU research task.

Sentence-internal breaks are secondary and have been introduced to support inter-annotator consistency.

Sentence-level SU breaks are used to indicate the presence of a main (independent) clause. These independent main clauses can stand alone as a sentence and do not depend directly on the surrounding clauses for their meaning; that is, the proposition asserted by this main clause does not depend on another clause for completion.

Sentence-level SU breaks are inserted between sentence boundaries; within each sentence further internal breaks might also exist. In most cases, sentence-level SUs would be represented in standard writing with end-of-sentence punctuation like periods or question marks. Sentence-level SU breaks may also appear after a short phrase that nonetheless functions as a "complete" sentence (See Section 4.6.1). Sentence-level SUs fall into one of four categories: Statements, Questions, Backchannels and Incomplete-SUs. These categories are described in detail in Section 4.5.

Sentence-internal SU breaks signal units that are smaller than a main clause and cannot stand alone as a complete sentence[11]. These smaller units are typically non-main clauses; they may also be units that are smaller than a clause but still constitute an SU. In many cases, sentence-internal SU breaks would be represented in standard writing with a comma. In all cases, these smaller units cannot stand alone as complete sentence-level SUs because they are conditioned by adjacent clauses that affect their meaning. Sentence-internal SUs can also be used to indicate boundaries between larger clauses containing both subject and predicate, when these clauses depend on one another for their full meaning. A clear example of this is an if-then construction:

```
If you never even see the money, then you don't spend it.
```

We can easily detect the two separate clauses that make up this sentence:

```
you never even see the money
you don't spend it
```

Each of these clauses constitutes a main clause, containing both subject and predicate. However, neither of these clauses can stand on its own and still retain the meaning expressed by the full sentence. The proposition expressed by this sentence relies on both clauses interacting with one another. Therefore, we cannot to draw a sentence-level SU boundary between the two clauses, but we still want to capture the fact that we clearly recognize two separate clauses in this sentence. We rely on a sentence-internal SU boundary to do this.

---

[11] Sentence-internal SU breaks are introduced solely for ease of annotation within the current definition of SimpleMDE. Identification of sentence-internal SUs is not an MDE research task.

Sentence-internal SU breaks fall into two categories. Coordination SU breaks identify two clauses which are not full sentences but which are joined by a coordinating conjunction. Clausal SU breaks identify non-sentence clauses that are joined by subordination and a handful of other phenomena. Specific rules for recognizing and classifying sentence-internal SU breaks are described in Section 4.6.

The goal of SimpleMDE SU annotation is for annotators to recognize every clause in the discourse and to impose upon it the appropriate SU label. By forcing a decision at each clause boundary, we eliminate the more difficult task of identifying both the extent and the type of each discourse-level (sentence-external) SU in isolation. A handful of additional rules are necessary to cover common features in spoken language that don't relate to clause boundaries; these rules are fully described in the sections that follow. For the vast majority of cases, however, the SU annotation task involves recognizing and categorizing clause boundaries.

When deciding between the two types of SU breaks, annotators should look first for evidence of a sentence-level break. Annotators should treat the sentence-level SU as the default break type, and should impose a sentence-internal SU only when no evidence of a sentence-level break exists.

4.4 Sentence-level SUs types

4.4.1 Statement
The statement label is used for a complete SU that functions as a declarative statement. The expected end-of-sentence punctuation for a statement is a period (or exclamation point).

```
It was a strange topic about corruption /.
```

4.4.2 Question
The question label should be used for a complete SU that functions as an interrogative. The expected end-of-sentence punctuation for a question is a question mark; but keep in mind that SUs may be smaller than a sentence.

```
How many hours a week are you supposed to work /?
```

4.4.2.1 Tag questions
A tag question is a phrase added to the end of an utterance that invites the listener to give feedback. Tag questions should be labeled as questions, and should be separated from the previous SU using a clausal break:

```
It seems like winter will never end /, doesn't it /?

You've been working there for years /, haven't you /?
```

Sometimes, discourse marker-type words like *you know* can also be used sentence-finally in a way that is similar to a tag question. In these cases, the tag portion should **not** be labeled as a discourse marker, but instead should be treated like the above examples:

```
examples (2)
```

4.4.2.2 Rhetorical questions
Rhetorical or hypothetical questions are treated just like other questions and receive the Question SU label:

```
Briefly there's personality /. Who can work with who in this new
corporate structure /? There's regulation /. Will the American
regulators allow one combination but not another combination /?
```

4.4.2.3 Question intonation in statements
Some speakers utter statements with rising final intonation so that they sound like questions. When the utterance is clearly functioning as a statement, it should be tagged as such. When the utterance is clearly asking a question, it should be tagged as a question. In cases of uncertainty, annotators should categorize the SU as a statement, and add the "difficult decision" label.

4.4.2.4 Embedded questions
When questions are embedded in larger carrier clauses, the sentence-level SU should be assigned on the basis of the whole sentence, and not according to the embedded question:

```
The principal then asked /, "Does anybody know who that president
was?" /.

And I was like /, "Is this guy a closet fascist?" /.
```

4.4.3 Backchannel
Backchannels are an open class of words that we define by their position and function in the discourse. A backchannel, also known as an acknowledgement or continuer, is a word or phrase that provides feedback to the dominant speaker, indicating that the non-dominant speaker is still engaged in the conversation. Backchannels do **not** signal that the non-dominant speaker is trying to take over control of the floor.

For instance,

```
A: You know it's just been really difficult for me /.
B: Uh-huh /@
A: What with everything happening in the course of two weeks /.

A:  If it happens again /, I'm going to have to say something /.
B:  Yeah /@ yeah /@
A:  Because it's getting out of hand /.
```

The number of potential backchannels is very large and it is nearly impossible to establish an exhaustive list. The following words represent very common backchannel words:

```
hm/hmm                   right
huh                      sure
mm-hm/mm-hmm             yeah, yea
oh                       yep
okay, OK                 yes
really                   uh-huh
```

This list is not exhaustive, and annotators may encounter alternate spellings of the backchannels that do appear on this list. Annotators should tag all words that function as backchannels as such, whether they appear on the list above or not. For instance, in the following example, Speaker B's entire turn consists of a string of backchannel SUs, and contains several words and phrases that do not appear on the above list:

```
A:  And it sounds terrible to hear a little kid singing that song
'Back That Thing Up' /.  They don't even know what they're
talking about /.
B:  Right /@ Right /@ That's true /@ Mm-hm /@
A:  I have little nieces and nephews /.
```

Backchannels may on occasion constitute repetitions of the dominant speaker's words:

```
A: I'm from Pittsburgh /.
B: Pittsburgh /@
A: It's an OK place /.
B: Mhm /@
```

When a turn consists entirely of a single potential backchannel, annotators should consider its function. If the word serves as a response to a direct question, it should be classified as a statement SU:

```
A: How are you doing /?
B: Okay /.
```

If the word provides feedback to the dominant speaker, it is a backchannel:

```
A: I've lived in Friendship Heights for years /.
B: Okay /@
A: But I'm thinking of moving a little further out /.
```

When in doubt, the word should be tagged as a statement, not a backchannel (according to the SU hierarchy, Section 4.7).

When a speaker chains together several backchannels in succession, annotators must tag each backchannel individually as a separate backchannel SU.[12]

4.4.3.1 Distinguishing backchannels from other MDE types

Lexical items that function as backchannels can often serve as discourse markers (and more commonly, as discourse responses). However, backchannels are functionally and structurally distinct from these other phenomena. Discourse markers and discourse responses are inherently *active* contributions to the discourse, and can be used to interrupt the other speaker, to seize control of the floor, to preface a new turn or to indicate the continuation of a turn or retention of control of the floor. (Discourse responses are discourse markers that carry the additional function of responding to the previous speaker's utterance – see Section XX for more on DMs and DRs.)

Backchannels, on the other hand, are passive contributions, and cannot serve these additional functions. They simply provide passive feedback to the dominant speaker, who remains dominant after the backchannel is uttered.

When a potential backchannel word prefaces a longer turn or comes within the extent of the current speaker's turn, it indicates that that speaker is now taking or now retains control of the floor. As such, these words are tagged as discourse markers or discourse responses. DMs and DRs do not invoke a separate SU break:

> ***Right*** Ø so you just bought another Plymouth then **/?**
>
> A: **So** how do you make this soup **/?**
> B: First you take a couple of carrots **/&** and chop them **/.** **Okay** Ø and then you sauté them in butter **/.**
>
> ***Oh*** Ø I didn't realize **/,** that's what you were talking about **/.**
>
> My family has always had cars like that **/,** in that **yeah** Ø my father has always had a VW bus ever since about nineteen sixty **/.**

Backchannels cannot preface an interruption, they do not signal that the speaker is taking control of the floor, and they cannot appear within a longer string of content words. They appear only as separate turns or embedded within a turn consisting solely of other backchannels.

The following table encapsulates the functions and structure of backchannels, DRs and DMs:

	Passive	Restricted position (turn-initial only)	Responsive

---

[12] Note that this represents a change from V5.0 of the guidelines where chains of backchannels were tagged as one SU and separated upon post-processing.

Backchannel	yes	yes	yes
Discourse response	no	yes	yes
Discourse marker	no	no	no

Sometimes it is clear from the discourse that while the non-dominant speaker has uttered a backchannel and has not intended to try to take control of the floor, the dominant speaker misinterprets the backchannel as a discourse marker or discourse response signaling an attempted interruption:

```
A: It's one of the things the kids are complaining about /.
B: Yeah /@
A: There's a lot of police presence /.
B: Yeah /@  <pause>
B: So Ø that must cut down on crime at least /.
```

In this example, Speaker B pauses after the second *yeah,* presumably leaving room for Speaker A to continue speaking. Although it seems that Speaker A has misinterpreted the *yeah* as an attempted interruption rather than as a passive backchannel, annotators should still treat such examples as backchannels when it is clear that the speaker intended it to function as a backchannel.

Note too that words that can act as backchannels can sometimes play other roles as well. They can act as questions themselves (e.g., *huh?*) or as direct responses to questions, or as content words elsewhere in the discourse. Annotators should label only those cases in which these words are functioning in a way that is clearly recognizable as a backchannel.

4.4.4 Incomplete SUs

When an utterance does not constitute a grammatically complete sentence, phrase or continuer, and does not express a complete thought, it is labeled as an incomplete SU. Incomplete SUs occur in two situations. When a speaker trails off at the end of their turn and abandons it completely, without then restructuring it and continuing along the same lines, an incomplete SU exists. For instance:

```
A: If you put enough patience into <pause> /-
B: <pause> Yeah just be consistent and diligent /.
```

In standard writing, this kind of incomplete SU might be followed by ellipses (…).

The other type of incomplete SU occurs when one speaker's turn is cut short by an interruption from the other speaker:

```
A: Yeah but the thing about /-
B: No see you have to take inflation into account /.
A: Oh okay I get it now /.
```

In standard writing we might see a double dash following such incomplete SUs. Note that not all interruptions create an incomplete SU. In some cases, the original speaker maintains control of the floor and completes the utterance after (or during) the interruption, resulting in a complete SU by the dominant speaker:

```
A: We did get the wedge cut out by building some kind of
B: a cradle for it /.
A: cradle for it /.
```

4.4.4.1 Incomplete SUs and Restarts

Incomplete SUs can be confused with restart disfluencies (see Section 3.4.3). For purposes of SimpleMDE annotation, restarts occur only when a speaker interrupts him/herself and then restructures the utterance and continues speaking on the same topic, e.g.:

```
A: [Do you know] * How would he like it to be cooked /?
```

Restarts are labeled as deletable regions, and are contained within the larger SU. Restarts do not induce a separate SU break of their own. They always occur within a single speaker turn.

Incomplete SUs, on the other hand, occur when a speaker is interrupted or when the speaker trails off, failing to complete the utterance within a turn:

```
A: Yeah but the thing about /-
B: No see you have to take inflation into account /.
A: Oh okay I get it now /.
```

Incomplete SUs generate a sentence-level SU break, and can only occur at the end of a speaker's turn. It can be followed by another turn from the same speaker, or by a new speaker's turn.

Simply put, annotators can decide between restarts and Incomplete SUs by deciding whether a turn boundary exists[13]. According to this rule, even cases like the following, where one could argue that the initial attempt is abandoned and not restructured, should be treated as a restart because it occurs within a single speaker turn:

```
A: But when I saw the way they treated the animals /, [ I was you
know just ] * it was really awful /.
```

---

[13] Some examples provide evidence that both I-SUs and restarts can actually occur anywhere, and should not be restricted by their placement in a turn. An argument can be made that deciding between I-SUs and restarts should be a matter of determining whether the abandoned utterance is restructured and reattempted, or whether it is simply abandoned. This approach works for some cases, but many examples present a real challenge because annotators cannot agree on whether the material has been restructured or not (or whether the reworked material is in fact indicative of a revision edit disfluency, rather than a restart). In keeping with the streamlined approach adopted for SimpleMDE, we defer treatment of this complex problem and instead adopt the simple decision rule outlined within this section.

```
     B: Mhm /@
```

4.5   Recognizing SU breaks

Although deciding between SU types is relatively straightforward, it can be
difficult to determine when to insert a new SU boundary and when to place two
units within the same SU. The following sections provide a number of guidelines
for recognizing SU boundaries for all SU types.

4.5.1   Phrases

Short phrases that do not constitute grammatically complete sentences but
nevertheless function as a complete utterance (for instance, in a direct response
to a question) receive the statement SU break:

```
     A: What did you get for your birthday /?
     B: A bunch of CDs /.
```

4.5.2   Simple sentences

An SU can comprise a simple sentence with a single main clause:

```
     The last great explorer Jacques Cousteau has died in Paris at age
     eighty-seven /.
```

4.5.3   Compound and complex sentences

Compound sentences consist of two or more clauses joined by coordination,
while complex sentences contain some kind of subordination. Compound and
complex sentences present a significant challenge to the annotation process, but
they are very common in discourse. The number of SUs present in a given
complex or compound sentence will vary depending upon both the number and
type of clauses contained within the sentence.

4.5.3.1   Coordination

Compound sentences involve some kind of coordination; that is, they contain
multiple phrases or clauses joined by a coordinating conjunction (CC). We will
consider correlative conjunctions (conjunctions used in pairs) as part of this set.

The class of CCs is semi-closed: there may be new elements, but these are
limited in number and frequency and typically amount to synonyms of current
members. Moreover, coordination can occur even without the presence of an
explicit coordinating conjunction. A list of some of the commonest CCs follows.
Note, once again, that this list should be considered illustrative, rather than
exhaustive.

```
     --NULL--                              or
     and, and so, and yet                  so
     but, but also, but only               yet
     for                                   both… and…
     nor                                   not only… but also…
```

```
not because… but because…                neither… nor…
either… or…
```

4.5.3.1.1 Coordination of main clauses

Coordination can occur between two or more main clauses. This is the simplest type of compound sentence. A sentence containing two or more main clauses joined by a coordinating conjunction should be divided with a sentence-level SU break before the conjunction:

```
Most folks on this planet spend virtually all of their lives on
dry land /. but Cousteau was rarely found out of the water /.
```

Each main clause element on either side of the coordinating conjunction is part of a separate SU, with the division point coming before the coordinating conjunction. Two separate compound sentences joined with a coordinating conjunction are treated in the same way:

```
We went to the ski show /& and looked at all of the resort
information /.  And we picked up a lot of brochures /.
```

4.5.3.1.1.1 Main clauses with semantic dependency

When the coordinating or correlating conjunction joins two main clauses that depend on one another for completion of an idea, a sentence-internal coordination SU break is imposed.

```
Not only do we methodically destroy the coastal fringe /& but we
also throw back our toxic effluence directly in the sea /.
```

The clauses on either side of the conjunction cannot stand alone as separate sentences, but must remain linked to one another in order to express a complete idea: In the first example, the SU is structured with the non-continuous expression *not only … but*; in the second, the fragment *but he did* completes the idea expressed as the complement of the negation *did not*.

In cases like this, we separate the clauses with a coordination SU and preserve a single sentence.

4.5.3.1.1.2 Main clauses without a conjunction

When a sentence contains multiple main clauses that are coordinated without conjunctions, they should be separated into individual sentence-level SUs of the appropriate type. Each unit expresses a complete idea.

```
NATO defense ministers in Brussels yesterday issued the toughest
warning yet to Yugoslav president Slobodan Milosevic /. Stop the
killing of ethnic Albanians who live in Kosovo /. Withdraw your
heavy forces /. And begin peace talks /. or face the possibility
of military action by the west /.
```

These clauses might be separated by additional punctuation (dashes, commas, semicolons, colons) in standard writing, but punctuation is highly variable and should not be the deciding factor in determining where to place SU boundaries.

4.5.3.1.2 Coordination of non-main clauses

Coordinating conjunctions can also join two or more non-main clauses. Subject and predicate clauses are not main clauses; that is, they cannot stand alone as a complete sentence. Coordination between non-main clauses imposes a sentence-internal coordination SU break.

For instance, consider this case of predicate coordination:

```
You can't stand in front of a counter /& and do that kind of
stuff all day /& and feel like you're contributing something to
scientific knowledge /.
```

or this case of subject coordination:

```
Also vacation time too is a big one for me /& and support for
whatever other kinds of professional activities you have at work
/.
```

or this case of prepositional phrase coordination:

```
And then I've worked with regular industry type places /& and
then with others that are kind of half way between /.
```

For non-main clause coordination, the clauses remain within the same sentence and receive only one end-of-sentence SU boundary, but the clauses are also separated by a coordination SU break. The following examples illustrate the difference between coordination of main clauses, resulting in multiple sentence-level SU breaks:

```
A: I had to cut it into slabs /. And I had to build up this big
square /. And then I turned it on the lathe /.
```

versus coordination between non-main clauses, resulting in multiple coordination SU breaks but only one sentence-level break:

```
A: I had to cut it into slabs /& and build up this big square /&
and turn it on the lathe /.
```

4.5.3.1.3 Conjoined simple subjects

When multiple subjects or other NPs are conjoined with a single predicate, they do not receive a SU break:

```
If American young men Ø and women are in harm's way
```

```
John Ø and Mary Smith …
```

```
        I now have a letter from Trent Lott Ø and other Senate leaders /.
```

4.5.3.1.4 Lists

Strings of NPs or predicates that function as a list should not be separated by SU breaks:

```
        And I take Discover Magazine Ø and also CD and Stereo Review Ø
        and National Geographic I think Ø and Boys Life /.
```

4.5.3.1.5 Idiomatic expressions

Idiomatic or frozen expressions should not be separated by a coordinating SU break. For instance (idioms italicized for clarity):

```
        I'm going to try Ø and see if she's there /.

        she went Ø and learned all about boxing /.
```

4.5.3.2 Subordination

In cases of subordination a clause depends on the rest of the sentence for its meaning. Subordinate clauses cannot stand on their own as complete sentences without changing the meaning of the construction as a whole; therefore, they cannot by themselves constitute a sentence-level SU.

When the clause begins with a subordinating word, it is no longer an independent clause; it is a dependent or subordinate clause because it depends on something else (the independent clause) for its meaning. Consider the following sentence:

```
        If the conflict cannot be resolved /, European Champion Germany
        could replace France.
```

In this sentence, *if the conflict cannot be resolved* is the dependent or subordinate clause; *European Champion Germany could replace France* is the independent clause. The two clauses are semantically linked: they depend on one another for meaning.   More specifically, we can say that the subordinate clause is semantically dependent upon the independent clause.

Subordination is an important concept for SU annotation in that two clauses that have a subordination relationship cannot be separated into two distinct sentences, because they depend on one another for meaning. Instead of creating a sentence-level SU break, subordination creates a sentence-internal clausal SU break.

4.5.3.2.1 Types of subordination

There are a number of ways that clauses can be subordinated to larger structure. The following sub-sections attempt to delineate the cases.

4.5.3.2.1.1 Subordinating conjunctions

Most subordination relationships are created by subordinating conjunctions. A large number of words and phrases can act as subordinating conjunctions. A non-exhaustive list of some of the most common subordinating conjunctions follows:

```
after              even though          unless
although           if                   until
as                 rather than          when
as if              since                while
because            so that
before             though
```

The following sentences contain clauses joined by subordinating conjunctions. For clarity's sake, the conjunctions appear in *italics.*

It's an opportunity to stay alive really **/,** *instead of* trying to exist on the good health of my youth **/.**

And *now that* El Nino is virtually gone **/,** there is La Nina to worry about **/.**

While mourners gather for his funeral **/,** colleagues and friends remember his contributions to our understanding of the last frontier **/.**

because there is not enough to go around for everybody **/,** you don't expect anything **/.**

He'd probably look funny **/,** *if* we did that **/.**

Legally speaking, the sanctions cannot be lifted **/,** *until* the special commission issues a positive report to the security council **/.**

And they will not have Reeves on the sidelines **/,** *when* they meet Detroit on Sunday **/.**

As we continue with this special edition of NBC Nightly News **/,** we want to go to Charleston, West Virginia, where former Secretary of State Henry Kissinger has been making an appearance this evening **/.**

Note that the order of the clauses may be inverted so that the conjunction comes at the beginning of the sentence rather than between the two clauses. In any case of subordination due to a subordinating conjunction, annotators should tag the boundary between the clauses with a clausal SU.

4.5.3.2.1.2 Subordinating verbs

Subordination can also be caused by structures that do not include a subordinating conjunction, as in the case of certain verbs. These verbs often express thinking, feeling, saying and the like, and can take a complementizer like

that or *whether* (although the complementizer may not be realized). A non-exhaustive list of verbs that trigger subordination follows:

> *verbs of perception and discovery*
> > find, hear, notice, realize, recognize, see, detect, sense

> *verbs of thinking and feeling*
> > believe, decide, feel, figure, hope, know, prefer, remember, seem, suppose, suspect, think, wish, assume, appreciate

> *verbs of declaration*
> > deny, discuss, explain, say, mean, tell, claim, mention, articulate

> *verbs of questioning*
> > ask, guess, wonder, question

In these cases, a clausal SU break should be placed after the verb and before the complementizer, if present. For instance,

> ```
> A friend of mine at work here said /, that he tried it with his
> dog /.
> ```

When such verbs occur medially within a clause, they do not invoke subordination nor do they take a complementizer; therefore they do not trigger a clausal SU break:

> ```
> That would I think Ø be a problem.
> ```

Note, however, that this rule only holds for subordinating verbs that occur **clause** medially. For the conventions describing the annotation of more complex quotes see Section 4.6.4.5.

4.5.3.2.1.3 Conditionals

Another common cause of subordination is a conditional construction. If-then statements are the most common type of conditional, e.g.

> ```
> If someone were studying economics /, then the State Department
> could offer to send them to a foreign country for two years /&
> and teach them how to run the country /.
>
> The Jags will clinch the AFC Central title /, if the Tennessee
> Oilers lose /& or tie the Green Bay Packers /.
> ```

Note that not all conditionals follow this exact structure. The order of the clauses is not fixed, the "then" may be missing, or other words (like "given (that)") may stand in for the "if". In all of these cases, the boundary between clauses that are joined by a conditional construction should be annotated with a clausal SU.

Also notice that this type of subordination is quite similar to the subordinating conjunction case described above. We can almost think of this type of construction as a 'special case' of the constructions triggered by the use of subordinating conjunctions.

4.5.3.2.1.4 Subordination in questions

Subordination can also occur within questions, as in the following example:

```
And if there's no inspectors /, what is the recourse /?
```

As with statement SUs involving subordination, these cases also invoke a clausal SU within the sentence.

4.5.3.2.1.5 Embedded or multiple subordinations

While the examples above are relatively straightforward cases of subordination, more complex cases can lead to confusion. One such difficult case is that of additional subordination within the sentence. This is the cleanest of the difficult cases because we simply apply the rule for subordination anew and continue. In other words, if two or more clauses within a sentence stand in subordination, and one of those clauses (either main or dependent) itself contains another subordination, then this internal subordination is also signaled with a clausal SU. For instance, consider the following:

```
The administration had been on the record saying /, when all
of this came up in early November /, that if he didn't cooperate
one more time /, they would pull the trigger /.
```

The application of this rule can continue ad infinitum, as long as the embedded constructions are themselves subordinations. For instance:

```
World Cup Football Champion France has said /, that it will not
take part in next year's Confederation cup in Mexico /, unless
the tournament date is changed /.

Today we took some important steps to ensure /, that Mr.
Milosevic knows /, that his indiscriminate use of force is
unacceptable /.

Now the Falcons coach says /, he will modify his behavior /,
unless the game gets tight /.

They know /, if it goes to a vote in the Senate /, they can't
count on not having sixty-seven votes against the president /.
```

4.5.3.3 Interactions between coordination and subordination

In many cases, subordination and coordination interact within a sentence, thereby complicating rules for SU annotation.

4.5.3.3.1 Coordination in the dependent clause

In many cases of subordination, there not only exists subordination between the independent and dependent clauses, but also coordination within the dependent clause.

For instance, in this sentence:

```
If you write papers /& or you do design studies /, then it's real
hard to rate something like that /.
```

the dependent clause contains the coordination:

you write papers *or* you do design studies.

Under normal rules of coordination, we would expect to apply a sentence-level SU break between the two coordinated clauses, since they contain both subject and predicate. However, in this case, these clauses are part of the larger dependent clause caused by the if-then subordination.   Rather than applying a sentence-level SU break, the two coordinated dependent clauses are separated by a coordinating SU break.

4.5.3.3.2 Coordination in the independent clause

Likewise, in many cases of subordination the independent clause is itself a coordinated construction. For example, in:

```
The settlement requires /, that the company clearly state in its
mailings /, that no purchase is necessary to enter the
sweepstakes /& and it cannot tell consumers that placing an order
will improve their chances of winning /.
```

the independent clause is the coordination:

The company clearly state in its mailings that no purchase is necessary to enter the sweepstakes *and* it cannot tell consumers that placing an order will improve their chances of winning.

Similarly, the normal rules for coordination would impose a sentence-level SU break between the two clauses. Once again, the rule for coordination is overridden and we apply a sentence-internal coordinating SU. The result is that the entire subordination (the dependent clause and both of the conjoined clauses in the independent clause) form a single sentence with one sentence-final (sentence-level) SU break.

Another example follows:

```
If they've given up on the inspection regime /, then we're
talking about either containment and perpetuity which I think is
unreasonable /& or we're talking about some sort of action which
will lead to the removal of Saddam Hussein /.
```

4.5.4 Additional SU considerations

4.5.4.1 Conversational introductions

It is common in the MDE telephone data for conversations to begin with an utterance that establishes common ground between the two speakers with respect to the assigned topic of conversation, e.g.,

> **So** air quality **/.**

Such utterances should be classified as a statement SU. Note that these introducers are often prefaced by a discourse marker. (See Section 4.5.3.1 for further discussion.)

4.5.4.2 Progressives, participles and infinitive verb phrases

Sometimes speakers use a clause beginning with the progressive or participial form of a verb as a kind of modifier. This happens most often in Broadcast News but can also happen in telephone speech files. In these cases, a clausal SU break is applied before the progressive form of the verb. For example:

> In Washington it'll be back to work for the president **/,**
> preparing for the State of the Union address later this month **/.**
>
> He's up there moving around **/,** smiling like he's happy **/.**
>
> Thousands of frenzied Beenie Baby collectors gathered outside
> this McDonalds early today **/,** causing a traffic jam **/.**
>
> It has been quite a day for plush toy enthusiasts **/,** going down
> in our memories as the cutest morning in history **/.**
>
> Cousteau set off in April **/,** accompanied by his longtime crew **/.**

Similarly, when a clause begins with an infinitive, a clausal SU break is introduced before the "to":

> Maurice Evans scores twenty-eight **/,** to lead the Longhorns to a
> ninety-four to seventy-eight victory **/.**
>
> Chris Duhan connects from fifteen feet at the buzzer **/,** to lift
> the Blue Devils to an eight-two to eight win **/.**

As you can see from the examples, this type of construction is especially common in sports reporting.

(Note the contrast with the treatment of relative clauses, Section 4.6.4.6 below.)

4.5.4.3 Prefacing statements

Speakers often preface opinions and other statements with a kind of term, ending in a finite form of *to be* and optionally followed by *that*. These are difficult to characterize precisely, yet they occur with some frequency in conversational

speech. For instance:

```
The first thing for me is /, I see a couple of different ways of
talking about what privacy is /.

The thing is /, is that /, they're also expecting us to work
overtime.

Another thing that occurred to me is /, there's not so much
invasion of my privacy /.

The question is /, did he really know it was going to be good /&
Or did he just do it /.

What we run into is /, we have the Texas Air Control Board TACB
that send out jurisdictions under which we have to reply to /.
```

These prefacing statements will receive a clausal SU break.

4.5.4.4 Refinements and generalizers

Refinements follow an original statement with some additional or revised
information. They should not be confused with edit disfluencies that contain a
deletable region. Refinements are tagged with a clausal SU break, as in these
examples:

```
Right now I believe /, there are two policies that the government
uses to modulate immigration /, three actually /.

We went camping this past weekend with some friends that had a
minivan /& and pulled the little trailer behind /, the pop up
trailer /.
```

Generalizers are phrases like "stuff like that" or "and so on" that are used by a
speaker to extend their specific examples to a more general observation. They
often follow lists. Generalizers invoke a clausal SU break, for instance:

```
Asleep At The Wheel, Sons of the Pioneers /, stuff like that /.

In the summer I like one piece dresses, short sleeves /, things
like that /.
```

Generalizers often occur alongside refinements, as in this example:

```
And he wanted to drive that van /& and travel /& and wanted
everything in there /, T V /, the whole bit /.
```

(See Section 2.2.2.1 for treatment of discourse marker "you know" behaving as a
generalizer.)

4.5.4.5 Quotes

Quotations are a very frequent occurrence, particularly within broadcast news data. Whether direct or indirect, quotations impose a clausal SU break. For instance, in the following sentences:

```
But the FDA says /, the drug can have a dramatic impact on people
who suffer from Crohn's disease /.

Buccaneers coach Tony Dungy says /, that even if Tampa Bay does
not make the post season /, the team is playing the kind of
football it likes /.
```

the indirect quote causes a clausal SU break after the quotative verb (say). The clauses remain joined in a single sentence with one sentence-level SU boundary.

Direct quotes follow the same basic guidelines as indirect quotes. A direct quotation imposes a clausal SU break after the verb of quotation:

```
When it was over /, Samuel Haile, a fifth-grader, said /, "I want
to learn more about John Glenn." /.
```

When the attribution follows, rather than precedes, the initial quote, the clausal SU break occurs directly after the quote:

```
"There was a President of the U.S. who really made a commitment
to the country /, that before the end of the 20th century we were
going to have a man on the moon," /, Principal Plaut said /.
```

Note that the sentence-level SU type should be assigned based on the sentence as a whole rather than just the quoted material, so that a sentence like the following

```
The principal then asked /, "Does anybody know who that president
was?" /.
```

receives an sentence-level SU label of statement rather than question.

All of the standard rules for annotation of complex SUs apply within quotations, whether indirect or direct, with the following exception. When the quoted material continues for several sentences and the quotative verb occurs **inside the quote**, additional SUs may be necessary. Consider the following example:

```
"Gerry has got that revolutionary aura, that whiff of cordite,"
/, says one friend /. "Bianca Jagger calls him /. All the rich
American society hostesses throw themselves at him /. He's amused
by it." /.
```

The first complete sentence containing the quote and its quotative verb is tagged as a single statement SU, with an additional clausal SU break preceding the quotative. The sentences that follow are all treated as independent statement SUs.

4.5.4.6 Tag questions

A tag question is a phrase added to the end of an utterance that invites the listener to give feedback. Tag questions should be labeled as questions**,** and should be separated from the previous SU using a clausal break:

```
It seems like winter will never end /, doesn't it /?

You've been working there for years /, haven't you /?
```

4.5.4.7 Relative Clauses and *where* modifiers

Relative clauses modify a Noun Phrase (NP), providing more information about the NP. They are often signaled by the words "who", "which" or "that". Relative clauses **do not** motivate an SU boundary of any kind:

```
He usually goes to the store Ø that has the good ice cream /.

I now have a letter from Trent Lott and other Senate leaders,
including Jesse Helms, Ø who is the chairman of the Senate
Foreign Relations Committee /.

Passengers Ø leaving on Flight738 reported strong turbulence /,
before the plane went down /.

And we are presently in receipt of a site permit Ø which will
allow us to have certain emissions up to a certain tonnage /.
```

Similarly, "where" can begin a phrase that modifies a previous NP. These "where" clauses should be treated like relative clauses, receiving no SU break:

```
But in reality I think /, what you would wind up with is a
political football Ø where they would see all these body counts Ø
that they can use for their own will /.

We both were raised that way Ø where our parents sort of stuck us
in front of a TV to be the babysitter /.
```

4.5.4.7.1 A special case of *which*

While relative clauses do not motivate an SU boundary, the word *which* has a special usage which receives a different treatment. On occasion, *which* is used to comment on the **entire preceding clause** (or perhaps even the entire preceding SU). In these cases, *which* is preceded by a lengthy pause, and the preceding clause has sentence-final intonation.

For example,

```
They said /, that TI was looking into it to purchase extra
vacation days /. <pause>  Which I thought sounded like a good
idea /.
```

In cases where there is strong motivation to consider the preceding clause a complete sentence-level SU, annotators should treat *which* as starting a new SU.

4.5.4.8 Broadcast news conventions

Stylistic conventions in the reporting of broadcast news lead to challenges for SU labeling. Broadcasters often employ strings of sub-sentence units rather than longer complex sentences. For the most part, these shorter units will be treated as separate SUs. Keep in mind that punctuation in these cases is highly variable and should not be relied upon to provide meaningful insight into the proper identification of SUs.

4.5.4.8.1 Introductory headlines

These often take the form of a series of brief phrases:

```
Now that El Nino is virtually gone there is La Nina to worry
about /. One hot and one cold /. We'll take a closer look /.
```

Sometimes a statement (like a story title) is followed by a question:

```
Airline passengers and outrageous behavior at 30000 feet  /. what
can an airline do /?
```

These brief bullet-point phrases or sentences should generally be treated as individual SUs because they can stand alone and express a complete idea.

4.5.4.8.2 Formulaic constructions

Broadcast news reports often include short "bylines", or formulaic introductory announcements consisting of e.g., the date, program title, announcer's name, listing of today's top headlines and the like. These short statements should be considered separate statement SUs, as long as they do not constitute a complete clause.

```
Richard Harris /. NPR News /. Washington /.

Angela Astore /. CNN /. Reporting /.
```

When the reporter signoff consists of a well-formed sentence (a complete main clause) it should be tagged as a single (statement) SU:

```
I'm Angela Astore /, reporting from Dallas for CNN /.

And this is VOA News Now /. I'm Neil Curry with Theresa Erikson at
22:53 Universal Time /.
```

A similar construction occurs when one broadcast reporter turns the floor over to someone else by saying the person's name. Anchors and reporters often finish their story introductions and conclusions by adopting this technique:

```
Finally, the governor's stance on rising taxes /. We go to Angela
Astore in Dallas /. Angela /.
```

```
The governor said /, he would not comment on future tax cuts /.
Peter /.
```

These cases merit a statement SU break, even if the speaker's intonation rises on the name. Since this construction is merely a way of passing the turn to the next person and is not an actual question, the annotator should not tag the name with a question SU break.

Sometimes, the return of the floor will be made more explicit by adding a short statement after the name of the anchor. In these cases, annotators should include the anchor's name and the short statement in a single Statement SU:

```
Peter Ø back to you /.
```

Similarly, a reporter will often preface part of a report with the anchor's name. In this case, the name does **not** merit a separate statement SU break, and will be considered a part of the larger complete SU:

```
Peter Ø the situation here is hectic /.
```

4.5.4.8.3 Apposition

Appositive constructions are very common in broadcast news speech. As a rule, annotators should not create an SU break around appositives.

```
The last great explorer Ø Jacques Cousteau has died in Paris at
age eighty-seven /.
```

4.5.4.8.4 Multiple overlapping speakers

Because broadcast data includes many speakers on a single channel, there are occasionally periods of overlapping speech that involve multiple speakers. In these cases, SU assignment can be quite difficult due to multiple interruptions, sequential or nested overlaps and the like. Annotators should use the "NoRT Annotation" label (fully described in Section 9 below) whenever a span of speech/transcript is too difficult to annotate accurately.

4.5.5 SUs across segment boundaries

Complete SUs can continue across segment boundaries in the transcription file. This happens most frequently in the case of failed interruptions, as in this example:

```
A: We did get the wedge cut out by building some kind of
B: a cradle for it /.
A: cradle for it /.
```

4.6 Deciding among SU types

In some cases it may be difficult to decide among SU types, depending on the function of the SU within the discourse. Annotators should adhere to the following hierarchy:

Sentence-external
   Incomplete SU ⇨ Question ⇨ Backchannel ⇨ Statement

↓

Sentence-internal
   Coordination ⇨ Clausal

Sentence-external SU breaks are the default; that is, if there is motivation for a sentence-external SU break it should be assigned over and above a sentence-internal break. For instance, in the following utterance a question SU "wins out" over a coordination SU in the first clause:

```
Who is he /? and what is he doing /?
```

Within external and internal SU breaks, different break types are also ordered. Given this hierarchy, a backchannel that is spoken with a question intonation (rising final intonation) and functions as an interrogative should be labeled as a question, not a backchannel.

```
A: I almost got hit by a bus once on my way to work /.
B: Yeah /?
```

4.7 SU conclusions

Most SUs will be easy to identify and classify, but there will be numerous cases of ambiguity, particularly in spontaneous speech. Annotators should keep in mind the ultimate goal of metadata annotation: to enhance transcript readability. When in doubt about where to place an SU boundary, annotators should rely primarily on the semantic information conveyed by the utterance and should apply the SU break in accordance with the primarily syntactic rules detailed above. Prosodic features, while interacting with SU boundaries, play only a limited role in SU assignment.

5 Interactions among MDE types

5.1 At the start of a file

At the beginning of a new file it can be difficult to know exactly how to tag items, since the context of the preceding utterances may be missing. If the annotator clearly understands the function of such an element in the discourse, it should be tagged according to the normal rules (as a backchannel, filled pause, etc.). When in doubt about the status of the item, annotators should tag it with NoRT (See Section 9 for an explanation of NoRT).

5.2 Disfluencies and SUs

Edit disfluencies are always contained within a larger SU rather than standing alone as a single unit:

```
[Her] * all her kids were sick /.
```

Furthermore, annotators should not indicate SU breaks of any kind within the delregs they identify. (See Section 4.5.4.1 for a related discussion of restarts and Incomplete SUs.)

5.3 Asides/Parentheticals vs. complete SUs

One particularly difficult decision point results from the current treatment of asides and parentheticals (see Section 2.4 for discussion). Although A/Ps are usually semantically "off-topic", decisions about whether to treat an utterance as an aside should be informed by its syntactic status as well. A/Ps interrupt a larger SU and are nested within it. A phrase that can stand alone as a statement SU and does not interrupt a sentence-level SU should be tagged as a separate SU, not as an aside. For instance, the following are examples of an aside:

```
He has now for about {oh gosh how long has it been} ten years /,
I guess

It's really {I don't know} incredible the way it all happened /.
```

while these are not:

```
He has now for a while /. Oh gosh how long has it been /? Ten
years I guess /.

I'll ask my husband /.  Honey do we still have the death penalty
/? He says /, we do /.

It just all seems so crazy /. I don't know /. I guess /, it's
going to be fine /.
```

5.4 Fillers, turns and SUs

Fillers often occur at the edges of SUs and annotators must decide where to place the SU boundary relative to the filler. When there is no turn boundary preceding or following the filler, the default rule is to treat the filler as outside of the extent of the preceding SU[14]. For instance:

```
add plain old filler example
```

---

[14] Some have argued that we should allow SU-final fillers (specifically, DMs) when they have a post-nuclear intonation (that is, when they occur after the most prominent accent in the SU) rather than having an SU break before the filler. However, this decision rule appeals exclusively to prosody, which is beyond the scope of SimpleMDE.

This is the case even when a segment boundary (but not a turn boundary) appears within the region in question:

```
A: I'm usually really insulted by them /. so
A: I just hang up /, as soon as I recognize what they are /.
```

More difficult are cases of "trailing fillers". Trailing fillers are followed by a noticeable pause (and thus, a turn boundary). In such cases, the default treatment is to exclude the filler from the span of a preceding SU, and tag it as an incomplete SU of its own:

```
A: Yeah that was a lot to go through /. so /-
B: You must be relieved to be over that /.
```

Sometimes a speaker will conclude an incomplete SU with a trailing filler. When a filler occurs at the end of a larger incomplete SU, it should be included within the span of the preceding SU, as in the following:

```
B:  So maybe it would have a bad effect on society and make them
a bunch of uh /-
A:  Airheads /.
B:  Yeah /@
A:  It would chill everybody out a little too much /.
```

5.4.1 Fillers at the start of a turn

Fillers (including asides/parentheticals) that come at the beginning of a speaker turn are typically captured within the following SU; no separate SU should be introduced after the filler.

```
See Ø my company has a much stricter policy than yours it sounds
like /.
```

When a filler directly precedes an edit disfluency but is not part of that edit, it is contained within the same SU as the edit. It should **not**, however, be labeled as part of the deletable region:

```
Um Ø [the, th-] the one thing I'm thinking is /, that it might be
hard to see the stage from way back there /.
```

Annotators should **not** look inside fillers for additional SU breaks. This is particularly relevant for lengthy asides/parentheticals. Like all fillers, these should be wholly contained within the larger SU and should not contain any internal SU breaks:

```
I read a thing {I don't even remember if it was in the Dallas
Site or the Inside one} about companies allowing you to purchase
extra vacation days/.
```

5.4.2 Fillers as complete turns

If a turn consists solely of a filler, it should be tagged as an incomplete SU:

```
A: But it's just really bizarre /, if you ask me /.
B: Um /-
A: The whole criminal justice system /.
B: Um Ø but I don't think /, the police are the biggest problem
   /.
```

Typically such examples can be interpreted as failed interruptions. With stand-alone items like these, annotators must be careful to distinguish fillers from backchannels, since some words can behave as either fillers or backchannels.

5.5   *Yeah:* a case study

The word *yeah* (and variations like *yup*) can play many different roles in the discourse. As such *yeah* serves as a particularly useful illustration of how to decide among SU types, how to recognize fillers (in this case DMs) that are included within larger SUs rather than acting as SUs of their own, and how to decide between regular DMs and discourse responses.

Yeah sometimes serves as a direct response to a question, in which case it should be tagged as a statement SU:

```
A: You said /, you had ten cats /?
B: Yeah /. When I moved /, I gave them to my mother /.

A:  Did you say /, you're calling from Dallas /?
B:  Yeah /.
```

On occasion, *yeah* functions as a question itself, displaying question intonation:

```
A: I almost got hit by a bus once on my way to work /.
B: Yeah /?
```

In such cases annotators should tag *yeah* as a question SU.

More frequently, *yeah* clearly functions as a backchannel and can be interpreted as roughly equivalent to "I'm listening – keep talking". For instance,

```
A:  I'm sure /, everybody knows /, everybody is getting high /.
You might as well just legalize it then /.
B:  Yeah /@
A:  It would probably bring down the price a lot /, if it was
legalized /.
B:  Yeah /@
A:  So the drug dealers wouldn't like it being legalized too much
   /.
```

Backchannel *yeah* typically occurs in a separate turn, or in a turn with other backchannels. As a backchannel, it cannot preface or be imbedded within a longer SU, an interruption, or any kind of new turn.

As a rule, when *yeah* occurs at the beginning of a turn or within a larger turn, but is not acting as a direct response to a question, it will be tagged as a discourse marker. In such cases *yeah* does not invoke a separate SU break. Discourse marker *yeah* can provide structure to the discourse:

```
Yeah Ø it's like /, I'm young /& and I should get started getting
into shape /.
```

Discourse marker *yeah* can also be invoked to provide a (positive) reaction, affirmation or endorsement of something the other speaker has said, in which case it is tagged as a discourse response (indicated in italics):

```
A:  It's supposed to alleviate some of their pain /.
B:  Yeah Ø why not /? Legalize a hundred percent /.
A:  Yeah Ø I don't see what the big deal is /.

A:  Where with teenagers around here /, you gotta drink to be
cool /.  So /-
B:  Yeah Ø you got it /.
A:  Who knows /?  Where are you from /?
```

Discourse marker *yeah* can also serve to affirm to the speaker's own statement, in which case it should be tagged as a plain DM and not a DR (since DRs are restricted to turn-initial position):

```
Of course my family has always had cars like that /, in that yeah
Ø my father has always had a VW bus /.
```

Several complex examples containing many instances of *yeah* will illustrate the various interpretations in context:

```
B:  Right now I'm in New Jersey /.
A:  Yeah /?
B:  Yeah /.  In a couple of days I'm Aloha bound for /-
A:  Yeah Ø right on /.
B:  Yeah Ø for the whole winter /.  So I'm stoked /.


A:  ...because there's much more kiddie stuff /.
B:  Yeah Ø the board walk /?
A:  Yeah /.  I per- /-
B:  Board walk's great /.  Board walk is so great /.
A:  Yeah Ø it is /.  Do you surf /, or something /?
B:  Yeah /.
A:  Yeah Ø is that why you're heading out to Hawaii /?
B:  Yeah /.
```

While the examples in this section all involve "yeah", the rules discussed can be extended for other words, for instance "no":

```
A:  I live in Northeast Philly /. but I used to go to Bringatine
/, when I was a kid /.
B:  Not a bad spot /.
A:  No ∅ it's not /.  I like it /.
```

6   Noises and other non-speech phenomena

Speaker noises like **breath**, **cough** and **laugh** as well as background **noise** tags
appear with some frequency in the transcripts. These tags appear in **light green**
font in the annotation tool. Given a more robust annotation tool, noises would be
masked from the transcript file because their presence has no bearing at all on
the metadata task as currently defined. As this is not currently possible, as a rule
of thumb annotators should exclude noises at the edges of deletable regions and
SUs (on analogy with fillers at the edges of delregs and SUs).

```
..when I go out to eat on Thanksgiving /, I still have a turkey
dinner /. <laughter>


I agree /. <noise> That's the kind of direction you wanna go with
food anyway /.
```

Note that isolated noise, whether background or speaker noise, does not count
as a turn, and that noises at the edges of turns are not considered part of that
turn.

On occasion, transcripts contain punctuation including commas, periods, dashes
and the like. Punctuation characters should be treated like noises – that is, they
should not influence assignment of SUs or other MDE types, but as a convention
they will be placed after the SU break.

7   "Difficult Decision" label

Some annotation decisions will be harder than others, and annotators might end
up spending an inordinate amount of time deciding how to handle a small
number of very difficult cases. In order to avoid spending too long on any one
annotation decision, annotators should use the "difficult decision" label to indicate
that a given decision was very hard to make. This label can apply to any type of
annotation decision. Because difficult decisions are also likely to have lower
annotator confidence, all "difficult decision" annotations will be revisited during a
separate quality control pass over the data. Annotators should use the "difficult
decision" label liberally.

8   "Questionable Transcription" label

As annotators label metadata in the existing speech and text files, they are likely
to encounter transcription errors in the original transcript files. Annotators are not
permitted to edit or repair existing transcripts[15], but should label the extent of a

---

[15] Transcript correction is not possible under the current task definition. This issue may be
revisited in future MDE efforts.

questionable transcription as such so that these errors can be readily located and corrected later.

It is likely that many of these questionable transcriptions will occur in areas of disfluency, since these are notoriously difficult to transcribe. **In such cases, the annotator should add metadata annotation if possible, taking the transcript file (not the audio) as accurate.** The entire extent of the region of questionable transcription should also be identified as such using the "questionable transcription" tag. If the mismatch between the transcript and audio interferes with the annotator's ability to accurately label the region for MDE, the annotator should select and tag the whole region as "NoRT Annotation" (Section 9).

9 "NoRT Annotation" label

While the "difficult decision" label is used when the annotator is unsure about whether the given annotation is correct, in some cases it may be impossible to make any annotation decision at all. When the audio signal is distorted, when there is a great deal of overlapping speech that makes it impossible to sort out who is saying what or where interruptions vs. incomplete SUs occur, when the transcript presents a commercial or other material that should not be annotated, or when the transcript is so inaccurate as to inhibit understanding, the annotator should select the entire span of the problematic region and label it with the tag "NoRT Annotation". This label should also be used at the beginning or end of a file when a particular turn or SU is artificially truncated. Regions labeled " NoRT Annotation " should receive no annotation of any kind[16].

---

[16] NoRT annotation can impact the assignment of SUs. When an SU contains words that are marked NoRT, the entire span of that SU will be labeled as unannotated SU on post-processing and the SU type will be changed to unannotated. Otherwise, the words marked NoRT would be deleted from the SU, leaving an non-contiguous stretch of text erroneously tagged as a full SU.