

Lexicon for the Transliteration of Spontaneous Speech Verbmobil II Subset for the transcription of the ISL Meeting Corpus

Susanne Burger
Interactive Systems Laboratories, Carnegie Mellon University, Pittsburgh, PA, USA
sburger@cs.cmu.edu

Abstract.....	2
1 Introduction	2
1.1 Entire Symbol Inventory of the VERBMOBIL_II Conventions	2
1.2 TRLparse Grammar	3
1.3 Objects to be symbolized in meeting transcription	3
2 Lexicon for the Transcription of the ISL Meeting Corpus	4
2.1 Part I: Structure of a transcription text.....	4
2.1.1 Global text file format	4
2.2 Part II: Transcription Symbols.....	5
2.2.1 Aborted Articulation	5
2.2.2 Acronyms and Letter Spelling.....	5
2.2.3 Breathing	6
2.2.4 Capitalization	6
2.2.5 Empty Pause.....	7
2.2.6 False Start.....	7
2.2.7 Filled Pauses.....	8
2.2.8 Foreign Word	9
2.2.9 Global Comment	11
2.2.10 Human Noise.....	11
2.2.11 Interjection	12
2.2.12 Interruption of Lexical Items.....	13
2.2.13 Proper Names	14
2.2.14 Neologisms.....	15
2.2.15 Numbers	15
2.2.16 Contractions	15
2.2.17 Punctuation.....	16
2.2.18 Repetition or Correction.....	19
2.2.19 Technical Interruption	19
2.2.20 Turn Break	20
2.2.21 Turn Identification.....	21
2.2.22 Unidentifiable - Hard to Identify	21
3 List of Symbols	23
4 Transcription Sample	24

Abstract

The following lexicon describes the annotation system used for the transliteration of the spontaneous speech of meetings according to the transcription system developed for the second part of the VERBMOBIL project (see <http://verbmobil.dfki.de/>). The transcription of the ISL meeting corpus uses a subset of this convention system.

The original system contains a series of rules in addition to the orthographic rules of the languages involved in VERBMOBIL (German, English, Japanese), labels for the annotation of spontaneous events, and formatting details for text files containing transliteration of spontaneous dialogues. The system is aimed at making transcription of spontaneous speech automatically processable and parseable, but also readable.

The subset does not use symbols for technical noises, interferences, proper names, digits and spelling, or rules for pronunciation comments.

1 Introduction

1.1 Entire Symbol Inventory of the VERBMOBIL_II Conventions

The major challenges for the development of a set of labeling conventions for the transliteration of spontaneous speech for the VERBMOBIL project, but also for similar projects working on automatic speech recognition, translation and synthesis primarily relate to the fact that the different parties involved require the transliteration of different features on different levels of annotation, including both acoustic (speech and non-speech sounds) and non-acoustic (grammatical and categorical) tags.

The Verbmobil Conventions, therefore, are divided into Time Consuming Events and Non Time Consuming Events (see figure 1):

Time Consuming Events are acoustic events filling a space of time. In other words, the event can be heard and generally also seen on a specific audio signal. We separate Time Consuming Events into *Lexical Items* and *Non-Lexical Events*. *Lexical Items* are words spoken in a defined language that are found in the dictionary for that language. The spelling of these items follows the normal orthographical rules. Special cases are defined by a vocabulary list, providing consistent spelling in cases of ambiguity. Spontaneous speech is characterized by the occurrence of additional phenomena such as breaks, repetitions or filled pauses. Words, in addition, may be aborted, interrupted, difficult to identify, or pronounced in non-standard ways. Cases of spontaneous phenomena on the word level are categorized as *Word Irregularities*. *Non-Lexical Events* include all acoustic events, excluding words. These events are separated into *Events of Human Origin* and *Events of Non-Human Origin*. The *Events of Human Origin* have a partly prosodic function, such as pauses in speech, breathing, sound lengthening, and filled pauses. Additionally, *Events of Human Origin* include human noises, such as laughing and non-identifiable utterances. *Events of Non-Human Origin* are all other noises, like background hum, microphone crackling, or knocking sounds.

Non-Time Consuming Events are tags added to provide categorical information about parts of speech. There are word tags providing information about special word categories, e.g. foreign words, spelled letters, proper names and neologisms. To offer structural information within spontaneous speech, we extend the syntactic structure of sentence punctuation by adding break symbols and brackets for non-grammatical phrases.

<i>Time consuming events</i>	
Lexical items: Orthography VMB vocabulary list Word irregularities: Abortion ..= Interruption .._ Hard to identify ..% Pronunciation deviation <!n ..>	Non-lexical events: Events of human origin: Speech Pause <P> Breathing Sound lengthening ..<L>.. Filled pause <uh> <uhm> <hm> <hes> Non-identifiable: <%o> Human noises: <Snack> <Swallow> <Throat> <Cough> <Laugh> <Noise> Non-Human <#Knock> <#Rustle> <Squeak> <#> Interferences: Speaker interferences ..n@..@n.. Noise interferences <.<> ..>
<i>Non-time consuming events</i>	
Category tags: Foreign word < *FOR>.. Spelled letter \$. Proper name ~.. Digit #.. Neologism *..	Structural marking: Punctuation . , ? Structure break < *T> < *T>t Non-grammatical phrases: Repetition/correction +/..../+ False start -/..../-

Figure 1. Inventory of the Verbmobil II conventions for the transliteration of spontaneous speech.

1.2 TRLparse Grammar

The TRL convention system employs a parser based on a GNU-Bison grammar, developed for the entire system. The grammar defines elements from the highest level to the lowest level in Backus-Naur format. For example, it first defines a turn, then the elements a turn contains, and ends by defining what characters constitute an element.

The parser filters out non-required information. The filter functions allow combinations for a special output of label groups—for example, only words and word tags plus punctuation, or only words plus technical noises. It is also used as the final consistency checking device before data is published; if edited accordingly, it can also be used as conversion tool from TRL to other finite annotation formats.

1.3 Objects to be symbolized in meeting transcription

The following broad categories are the elements to be transcribed in a meeting contribution:

- lexical units
- syntactic-semantic segmentation
- non-verbal articulated productions
- pauses in speech
- comments
- special comments

2 Lexicon for the Transcription of the ISL Meeting Corpus

2.1 Part I: Structure of a transcription text

2.1.1 Global text file format

A transliteration file contains of

- a header with information on name, date, and transcriber of the recording
- the transliteration of the conversation contributions (turns)

Some General Rules for creating headers:

- Every line of the header starts with a SEMICOLON followed by ONE white space.
- Names of fields begin with a capital letter.
- Names of fields end with a colon, followed again by ONE white space.

The VMB transliteration header contains the following field names:

In the first line of the header, CDR indicates the CD-ROM version of the recorded speech data; TRV is the actual version of the transliteration; File is the file name. The next three or more lines show the last-updated date for the transcription file, the initials of the transcriber, and global comments on the recorded meeting (quality of speaking, technical details, warnings about disturbances in the signal, etc.).

```
; CDR: 00.00
; TRV: 00.00
; File: m100
; Last changes made on 01/15/2003
; Transcriber: dnh
; Comments: Speaker ABC stutters.
; Comments: disturbances between turn 004 and 020.
;
```

2.2 Part II: Transcription Symbols

2.2.1 Aborted Articulation

Symbol

=

Definition

A word is aborted when it is cut off or terminated at a point where some of the component sounds have been voiced, but the rest of the word has been omitted.

Description

Aborted articulation is represented by the equal sign [=]. Examples include stutters and interruptions. Even if it seems clear how the word would end, only the voiced part is transcribed, followed by the equal sign.

Example cases are:

1. When a speaker has begun to say a word, but does not finish enunciating it
2. During a series of stutters where the beginning of a word is repeated but not completely pronounced

Rules for Transcribing the [=]

The voiced part of the word is transliterated; the equal sign [=] marks the position of abortion.

Examples

1. uh-huh , you are wa= <*T>t
 2. ah, so . -/this was the prob=/- mhm .
 3. mhm . no , +/it shoul=/+ it should not be incredibly detailed .
 4. +/cou=/+ could that be ?
 5. of course , they are expecting us to have it on the n= <*T>t
 6. +/sh=/+ +/sh=/+ she came back from it a year ago , so .
-

2.2.2 Acronyms and Letter Spelling

Definition

Acronyms are words using only letters as abbreviations for longer expressions.

When the letters are still pronounced as single letters, acronyms are written entirely in capital letters:

CD “cee dee” (*Compact Disc*)

ASR “ey ess ar” (*Automatic Speech Recognition*)

In cases where the acronym is pronounced as a word, it is written as if it were a complete word, in lower case. Since most acronyms are also proper names, the first letter is usually capitalized:

“Icasp” for ICASSP, because it is pronounced “Icasp” not “aye cee ey ess ess pee”

“Faz” for FAZ (pronounced “fats,” not “eff ah zet”)

“asap” for ASAP when the acronym is pronounced “ey-sap”

If words are spelled out for better understanding, each single letter is written between spaces:

“My name is Huber , H U B E R .”

Examples:

1. we are working on ASR
 2. my name is Huber , H U B E R
 3. I read the Faz , also known as FAZ or Frankfurter_Allgemeine-Zeitung.
-

2.2.3 Breathing

Symbol

Definition

The inhalation or exhalation that occurs during spontaneous speech that often happens at prosodic or syntactic boundaries. In a transcript, only breathing that can be heard on the recording is transcribed. Breathing often occurs in combination with <Smack> (lip smack, tongue smack), which is described in the Human Noises section, or in combination with an empty (silent) pause. An exhalation that follows a word-final plosive consonant (like the *t* in ‘wet’) should be interpreted as a lengthening of the plosive consonant, rather than as breathing.

*Rules for Transcribing the *

1. If punctuation and breathing collide, punctuation is always transliterated first followed by the breathing element
2. When an inhalation is followed by an audible exhalation—e.g. the breath is in and out—only one is transcribed.

Examples

1. when do you <uhm> want to go ?
 2. why ? do you know her ?
 3. oh , <P> <uh> okay .
-

2.2.4 Capitalization

Definition

In transcription, we do not exactly follow the capitalization rules of a language. Only proper nouns and spoken letters, or those part of an abbreviation, are capitalized. A new sentence begins with a lower case letter, rather than a capital, unless a proper noun starts the sentence.

Examples

1. she wants to take two of them to Japan .
2. which reminds me , I forgot to go see Kim .

3. Susi generates an SPR file or something +/-for/+ <P> with data
4. the Opec one is even bigger
5. let us think about it . that may be good .

2.2.5 Empty Pause

Symbol

<P>

Definition

A pause, in general, is a temporary absence of any acoustic signal. When referring to spontaneous speech, it can be defined as a temporary silent hesitation or stop in speech. When evaluating an empty pause, it is the general rule to use the length of time it takes for the speaker to articulate a 2-syllable word at his normal rate of speech. If this length of time fits within the pause in question, then it is a pause. If not, no pause is transcribed.

Someone people who talks faster will not have the same criteria for evaluating length of time between two utterances as a slow talker.

Rules for Transcribing the <P>

1. Pause is set when a speaker has stopped or halted his speech.
2. The pause label <P> is always separated by spaces from the rest of the text.
3. The minimum length of a pause is the time it would take the speaker to say a two-syllable word at his normal speech rate.
4. The pause is UNFILLED (or there are no vocalizations or breathing)
5. The pause or halt appears INSIDE an utterance or sentence; that is, there is no pause at the beginning or end of a turn.
6. Pauses can appear after a repetition or correction

When NOT to use the <P>

1. Pauses do not automatically delinate sentences. A <P> may follow a period, but only if there is a two-syllable or longer pause before the next sentence.
2. If the pause is FILLED—there are vocalizations, human or non-human noises, hesitations or breathing present during the break—do not use <P>.
3. <P> cannot be used at the beginning or end of a turn.

Examples

1. so <uh> <P> what time did you say the meeting was ?
2. well , -/I was trying/- <P> I managed to get here just in time
3. yes , well <P> he did tell us that , <P> but I wasn't so sure .
4. yeah , next Monday , this will be <P> ready for processing .

2.2.6 False Start

Symbol

-/ . . ./-

Definition

A false start is the act of beginning an utterance but aborting it prior to completion – a new idea or train of thought interrupts the old stream of thought.

Description

False starts occur most often when a conversation becomes intense, with many speakers speaking at once, or after a speaker is interrupted. They can also occur when a speaker becomes excited and begins to skip from one topic to another. An interruption that ends a turn is not handled as a false start; a turn break symbol <*T>t is used instead.

When to use [-/. . ./-]:

When evaluating false starts, it is helpful to look at the sentence structure. A speaker may begin a question, abruptly end the question, and then start all over with a statement. Although the context of both the question and statement are similar, this occurrence would be considered a false start because of the change in structure. For example:

1. -/do you think we should/- I think we should go tomorrow .
2. +/do you think/+ you think we should go tomorrow ?

Notice the differences in structures. The first utterance contains a false start, but the second does not.

When not to use [-/. . ./-]:

False starts cannot be transliterated at the end of a turn because no new phrase is started. In this case, a turn break [<*T>t] is transliterated. When the verbal disruption is followed by a continuation of the same thought or idea it becomes a repetition/correction, rather than a false start.

Rules for Transcribing the False Start [-/. . ./-]:

1. There is no punctuation between the false start and the new start, nor after it.
2. The entire false start is marked, beginning with the first element of the aborted phrase and ending after the last element of the aborted phrase.
3. There is no space between the left marker [-/] and the first word of the aborted phrase. Similarly, there is no space between the last item in the phrase and the right marker [/-].

Examples

1. -/I'm trying to figure out/- I guess we can talk on the plane .
2. -/it really is in this/- they want to rapidly +/settle on/+ get to one
3. -/and ideally we would have a full event log where/- <P> the fact that I don't have an observation doesn't mean nobody has seen them .
4. -/and I/- that was the space issue that I was talking about yesterday .
5. -/what do you/- yeah , I mean the most vulnerable time to hit those guys is when they are in a straight line

2.2.7 Filled Pauses

Symbol

<uh> <uhm> <hm> <hes>

Definition

The sound produced during spontaneous speech that represents a pause filled by a vocalization.

Description

A filled pause, or hesitation, is an articulation by the speaker that may be encountered between utterances but is not to be mistaken for a lengthened sound within a word.

A filled pause occurs most often when a speaker is thinking. It is a filled pause in that the speaker actually breaks off speech while continuing to articulate. However, the articulation is neither a word nor part of a word.

Filled pauses include:

<uh> : vowel such as /uu/ /aa/ /ooo/ /eee/ /er/ /ih/ etc.

<uhm> : vowel + nasal such as /eem/, /aam/, /uum/

<hm>: nasal such as /mmm/, /nnn/

<hes>: trash such as /pfff/, /tsss/

A filled pause is not a word, and therefore should not be treated as such. Punctuation cannot follow a filled pause; it always comes before. A filled pause, however, may stand alone as a turn of its own.

Examples

1. yes . <hes> yes .
2. noise here, <uh> a cough , it seems <uhm> a cold .
3. but <hm> I don't know if you want to ask to +/ea=/+ <uh> each person , each time
4. how far is the <uhm> next gas station ?

Note

If a sound can be substituted with something meaningful, such as an interjection - 'well' or 'ah' among others - then the sound is NOT a filled pause.

For example:

“huh , what did you say ?”

2.2.8 Foreign Word

Symbol

<*FOR>word

Definition

In transcription, the Foreign Word refers to words that appear in a transcript that are not of the same language as the primary language of the transcript.

Description

The Foreign Word label can be used “as is” instead of a word of an unknown language:

“<*FOR>word is a foreign word and we don’t know what it means how it is spelled or what language it is .
“

Or, if the language was identified, a language code could be inserted instead of “FOR”:

“<*GER>word is a German word , but we don’t know how it is spelled .”

Or, if the language is known and the transcribers knows the foreign spelling, too:

“<*GER>Badewanne is a German word . “

When to use <*FOR>word

1. When a word appears in a transcript that is non-native to the primary language.
2. When the foreign word has not become integrated into the primary language.
3. When the foreign word may be familiar to speakers of a given language, but has not been assimilated into the vocabulary of that language.

When not to use <*FOR>word

1. When the word has been fully integrated into the primary language, and is a part of its standard vocabulary.
2. When the foreign word can be found in the dictionary of the primary language
3. Foreign names like Adamello-Brenta, Val-di-Fiemme.

Note

Many foreign words and phrases exist in our colloquial language. Some, like "Hasta la vista," have been made famous by the movies. Others, like "c'est la vie" and "ciao," are foreign, but their meaning is clearly understood. All three of these would be transcribed using the <*FOR>word convention. For example, "Hasta la vista" would be written as: <*SPA>hasta <*SPA>la <*SPA>vista. Even though the words are recognized in English, they are not yet component members of English, and so would be transcribed using the foreign word convention.

There are many words, however, such as "kindergarten" or "fait accompli", that have foreign roots, or may indeed still be a part of their language of origin, but which have been fully integrated into the English language. An indication is when a word can be found in the dictionary. These words should be considered as part of the vocabulary of the English language and should NOT be transcribed according to the foreign word convention.

Rules for Transcribing <*FOR>word

FOR stands for a three letter symbol to represent the foreign language. When using the convention, the tag should come first, or before the foreign word that the convention is representing (see examples below).

ENG = English
FRA = French

GER = German
ITA = Italian
JAP = Japanese
SPA = Spanish

If the foreign language cannot be identified, but it is clearly foreign, the convention <*FOR>word is used.

Examples

1. <*JAP>sayonara , Mister ~Fujisaki .
2. yeah . well , I think that is pretty <*GER>komisch .
3. si , y quiero discutir la <*ENG>interface contigo .
4. ja , aber das ist nicht besonders <*ENG>nice .
5. I am not sure if that is <*FOR>word or not .

Notes

There is NO space between the last convention bracket and the beginning of the foreign word, i.e:
<*FRA>apprivoiser

--> No space here between the ">" and the "a"

2.2.9 Global Comment

Symbol

;

Definition

In the event that a comment needs to be made regarding a specific feature of a turn or the entire turn, a global comment may be used. This makes comments more concise. Rather than placing a comment beside each word of the turn, one larger comment is made directly following the turn.

Description

A global comment may involve a technical occurrence, irregularities of the articulated phenomena of the speaker, or anything the transcriber may wish to state.

The global comment follows the turn. It begins with a semicolon [;]. Each additional line of the comment also begins with a semicolon.

Examples

e123ar2_001_BBP_120002: hello ~Adam . we have to arrange an appointment .
; BBP speaks with hoarse voice

Notes

For comments that are necessary throughout the entire transcription, global comments are not needed. Rather, comments like these may be placed in the header comments.

2.2.10 Human Noise

Symbol

<Laugh> <Throat> <Cough> <Smack> <Swallow> <Noise>

Definition

There are many sounds a speaker may produce that have no real meaning. Examples include laughing <Laugh>, clearing one's throat <Throat>, coughing <Cough>, smacking of lips <Smack>, and swallowing <Swallow>.

A variety of others, such as hiccupping and a glottal sound, are included in the 'trash' category <Noise>. For the most part, these are self-explanatory and are not difficult to become accustomed to.

Description

- <Smack> sound made during spontaneous speech by pressing the lips together and pulling them apart quickly. Often followed by breathing .
- <Laugh> is accompanied by a chuckle or an explosive sound together with heavy breathing - the sound produced when something is funny.
- <Swallow> sound produced when saliva has to be removed; often in combination with beathing
- <Cough> sound produced for a variety of reasons, such as when there is irritation in the windpipe or to indictae embarrassment or outrage
- <Throat> Sound produced, for example, when phlegm has to be dissolved. Also produced because of irritation, dryness, or some other physiological disturbance.
- <Noise> Spontaneous speech can produces noises that are unique to spontaneous spoken situations. Defined as the human noise "trash-class," it is used to represent human noises for which no other conventions exist to describe them. Sometimes a grunt or a hum or a burp, the <Noise> convention

represents any human sound that has no "proper" term or convention to describe it.

Rules for Transcribing Non-Verbal Articulations:

1. In the case that one articulation occurs for a prolonged period of time without interruption - ie. a speaker laughs for a period of time - one mark will suffice.
2. These symbols are not to be treated as normal words. Punctuation may not follow them, it must go first.
3. In the case that one speaker interferes with another by making one of these noises (laughing, coughing, etc is heard in the background), it is transcribed as <Noise>.

Examples

1. sorry . <Swallow> my mistake .
 2. +/well/+ <Throat> well , that is okay .
 3. yes . <Cough> sorry . now we should find a second date
 4. we could <Noise> go there tomorrow .
 5. <Smack> -/it really is in this/- they want to +/settle on/+ get to one .
-

2.2.11 Interjection

Symbol

uh-huh mhm mm uh-uh

Definition

Expressions of surprise ('oh'), affirmation, negation and discourse particles (such as 'well', 'anyway') are examples of interjections. Four of them utilized by speakers engaging in spontaneous speech will now be described in greater detail: two of them fall into the category "affirmation," and the other two belong to the category "negation".

- **Affirmation**

Symbol

uh-huh mhm

Definition

These two variations of the affirmation are used by speakers engaging in spontaneous speech to indicate agreement with, certainty about, or understanding of something that has just been said.

In transcription, the two are treated in the same manner as a word and may serve as turns by themselves. They are written as they sound.

- **Negation**

Symbol

mm uh-uh

Definition

These two sounds have the opposite meaning when utilized in speech as that of the affirmation. The 'mm' and 'uh-uh' are used to represent disagreement, and are virtually synonymous in meaning with the word 'no.' The sound 'mm' is produced nasally with the lips closed. It is a two-syllabic vocalization with a glottal stop in the middle, and is generally accompanied with a shake of the head carrying the meaning 'no.'

The sound 'uh-uh' is also a two-syllabic vocalization produced at the back of the throat characterized by a

vowel sound and glottal stop in the middle. It also carries the meaning 'no.'
Negations are responses or reactions, NOT filled pauses.

Notes

There are two functions for 'hm,' the filled pause and the articulated 'hm.' The articulated version carries meaning, either as an inquiry, 'hm?' or as surprise.

Examples Affirmation

uh-huh:

1. Speaker_1: channel number six is this microphone here
....Speaker_2: uh-huh .
2. Speaker_1: primary school , where this person went to grammar school .
....Speaker_2: uh-huh
3. uh-huh . <uh> once we hit the record button we just let it go , no matter what

mhm:

1. mhm .
2. mhm . his information on the web page is <uh> like a year out of date or something .

Examples Negation

uh-uh:

1. uh-uh .
2. uh-uh . at least , I don't think so .

mm:

1. mm .
2. mm . I don't think he is going to like that .
3. mm , mm , mm , that is not what I meant .

2.2.12 Interruption of Lexical Items

Symbol

—

Definition

Lexical items can be interrupted by various phenomena. These can include pauses, filled pauses, breathings, mispronunciations, and certain combinations of these. After the interruption the lexical item is continued.

Rules for Transcribing the Interruption [_]

1. At the point of interruption an underscore [_] is added to the word fragment.
2. This is followed by the interrupting element or elements (a lexical item can of course be interrupted

more than once).

3. Finally the remaining part of the interrupted lexical item is transcribed with another underscore at the beginning.

Note

White spaces have to be placed between the underscores and the interrupting elements. If there is more than one interruption element, then white spaces need to be placed between these as well. No white spaces occur between the lexical item fragments and the underscores.

Examples

1. meaning_ _ful .
2. it is un_ +/b=/+ <P> _believable .
3. I need more con_ <uh> <uh> _cen_ +/tri=/+ _tration .

Special Cases

The way to transcribe interrupted compositions like "full-form" that have a hyphen separating the two parts is as follows:

“full_ <uhm> _-form”

If the interruption takes place here, between the 'full' and the 'form,' the hyphen is to the right of the second underscore.

2.2.13 Proper Names

Definition

Proper names are capitalized. Names consisting of more than one lexical unit are hyphenated. This keeps the names together so they are treated as one element

General Rules

surnames and first names of people:

Henry Smith

names of streets:

Fifths-Avenue

names of hotels and restaurants:

the Waldorf-Astoria

company names and names of institutions:

Carnegie-Mellon-University

IBM

names of geographic places:

Montana

New-York

names of public holidays:

Dr-Martin-Luther-King-Day

Book or movie titles:

The-Name-of-the-Rose

Keeping-up-Appearances

Note

Also names of Days and Months are capitalized.

Monday

January

2.2.14 Neologisms

Symbol

*

Definition

A neologism is the term used to describe a word that has been made up or invented by a speaker, which appears in a transcript of spontaneous speech dialogue. It can also be described as a word which does not appear in the dictionary of the primary spoken language, but which is also not a foreign word.

Description

A neologism may be a slang word that has yet to find its way into mainstream conversation, or it may be the creation of a non-native speaker who has made for example a grammatical error. The so-called slip of the tongue may also be seen as neologisms.

It is marked with the asterisk [*] at the beginning of the "new" or neologism word.

Examples

1. that candy has a *tutti-fruity taste
2. the new LXD-forty is a *narly model .
3. she cared for the new puppy in a *perspecial way
4. it was *percept that we all wanted more bagels .

2.2.15 Numbers

Description

All numbers are written out alphabetically, NOT numerically: twenty, nine, fifteen, etc. As is standard in regular typography, numbers, which are written as two words with a hyphen between them, such as twenty-two and thirty-three, are still hyphenated. However, this applies only for two digit numbers. Three digit numbers are not hyphenated.

2.2.16 Contractions

Definition:

Although contractions are appropriate in both spoken and written language, in transcription we use full-form in order to clarify meaning in case of ambiguities. For example, the contraction 'I'd' could mean either 'I would,' 'I could,' or 'I had.' "It's" could refer to the possessive or express the reduction of it is.

Reductions which are unique and already used in literature and dictionaries may be used as such:

Negations (can't, won't ..)

gonna, wanna, gotta

I've, I'm, we're ...

2.2.17 Punctuation

Symbol

, ? .

Definition

The inserting of standardized marks or signs in written documents clarifies the meaning and the separate structural units of sentences and utterances.

Description

Spontaneous speech is difficult to punctuate correctly. The path of the conversation is generally unplanned, and spoken utterances are often based upon impulsive decisions. For this reason, punctuation is a difficult sentence structure convention.

Below, we describe in detail the use of the transcription punctuation set: the comma [,], the period [.] and the question mark [?]. These are the only possibilities for punctuation within a turn. Some general rules, however, apply to the entire punctuation set (the [,] [.] and [?]).

General Rules

1. A white space [blank space] is always set before and after the punctuation
Example: I would like to go home now .
notice: the space between the "now" and the "."
2. After a period or question mark, the transliteration is continued in lower case unless the word category requires capitalization.
Example: I would like to go home now . would that be all right ?
notice: the "would" is not capitalized after period
3. The setting of punctuation generally follows the punctuation rules of the specific language.
4. Elements inside of a set of <> always follow a punctuation mark.
5. There is no punctuation after <%> (symbol for "not-understandable").
6. If a turn consists only of a filled pause, breathing, or a human noise, no punctuation, nor a turn break, follows.
Example: m021_8_0599_JOHN_00: <uh>
Otherwise,
7. The end of a turn must have either a period [.] , a question mark [?], or a turn break <*T>t in the event that the turn was broken off or aborted.

• Comma

Symbol

,

Definition

The punctuation mark is used in writing to indicate separation of ideas or elements within the structure of a sentence.

Description

In transcription, the comma is used almost analogously to the manner in which it would be used in writing. This is difficult because English has few ground rules for the usage of the comma.

Rules for Transcribing the [,]

1. When it is necessary to make a transcript more readable
2. When the conventions dictated by regular written English require the use of [,]
3. During confusing situations in the transcription. Since dashes are not a permitted form of punctuation, use the comma in a similar manner as the dash to enhance the clarity of the sentence.

General Rules

The following rules should be viewed as general guidelines to the usage of the [,]. These rules may not always apply, but are for the most part good indicators of when a comma will be required. Therefore, [,] should be used

1. After an interjection, such as:
oh , well , yeah .
okay , this sounds good .
yeah , this is good .
2. Before or after an address , such as:
I say Bill , <uhm> you are in charge of these units .
I didn't mean to give you such a long winded answer , Steve .
Bill , we have you until eleven thirty ?
3. After "good morning" or similar salutation or phrase:
good morning , Mister Smith .
how are you , Susi .
4. Between two phrases where an interjection is a distinct or separate idea:
that hotel , like the one we stayed at before , would probably be a good idea .
5. To connect thoughts, ideas, or items in a list
I bought two oranges , two mangos , and a bunch of concord grapes .
6. When it serves as a syntactical pause within a statement:
I think so , but I'm not sure .
7. To separate lists:
well , you know how to increase group cohesiveness , you have to <P> interact frequently , you have to have a small group , and you have to reward them .
8. When a comma helps to clarify the meaning of a statement, or to indicate the separation between clauses or interjections:
I would like to , but I'm afraid I have class at that time .
she said she would go , you know , but she had too much to do.

• Question Mark

Symbol

?

Definition

A mark that is used in writing and printing at the conclusion of a sentence or utterance to indicate a direct question.

Rules for Transcribing the [?]

1. Use [?] to identify questions, sentences, or phrases that are used to interrogate or gather information
2. When the utterance is structurally a question, but has falling tone at the end
3. When the utterance is a statement with rising tone at the end - in this case, the statement would become a question. Questions are transliterated on the basis of:
 - Interrogative statement
 - Construction of the sentence
 - Intonation (speakers tone rises at the end of the speech)
 - Context

Example

1. what time does that flight leave on Sunday ?
2. perhaps she does it too detailed ? +/cou=/+ could that be ? did we see something ?
3. really ?
4. don't we have a lapel mic for him ?
5. hi , John . how are you ?

• **Period**

Symbol

.

Definition

The punctuation mark indicating a full stop placed especially at the end of declarative sentences.

Description

The period is used in transcription as it would be used in normal text. It represents the end of a complete thought or idea.

Difficulty arises in transcription between when the use of a period or the use of a comma should be employed, especially when the speaker talks fast and has a tendency to babble. In such cases, it may be difficult to decide whether to use a period or a comma. The transcriber will have to decide if the ideas are similar or connected, or if there would be a natural sentence break at that particular point in writing. This decision will be made based upon transcriber's judgment, the context of the speech, and intonation.

Spontaneous speech is rarely grammatically correct. Therefore, the decision to place a period should be based on the following:

Rules for Transcribing the [.]

The decision to use a [.] is based on:

- Grammar
- Intonation
- Pauses, respiration
- Beginning of a new thought or idea

Example

1. she was in New-Guinea or New-Zealand or someplace on vacation .
2. oh , yeah .
3. ah , I wasn't around .

4. he has a name for it , too , but I can't remember the name of +/-the heli=/+ the landing pad .
-

2.2.18 Repetition or Correction

Symbol

+/. . ./+

Definition

The tendency of speakers engaging in spontaneous discourse to stutter or re-express things during moments of excitement, uncertainty, or temporary verbal incapacitation, and to subsequently correct themselves after such errors of speech.

Description

The +/- . . ./+ convention is used to symbolize when this phenomenon actually occurs in a transcript when a speaker either repeats or corrects himself.

When to use [+/. . ./+]

1. When a word or phrase is repeated by the same speaker in succession
2. When a speaker corrects or substitutes a new word for one he previously started, but continues with the same idea or train of thought

Rules for Transcribing [+/. . ./+]

The phrase that is corrected (or repeated) is set in slash marks in such a way that the removal of the marks would leave a more or less grammatically correct sentence.

Example: we saw +/-the gor=/+ the two gorillas at the zoo

Examples

1. yeah , so it is good for +/-mor=/+ +/-high/+ help morale and welfare .
 2. and always ask for help , if I +/-can't find/+ don't know the answer myself , I will find it .
 3. +/-if/+ if
-

2.2.19 Technical Interruption

Symbol

<*T> (technical interruption within a turn)

<*T>t (turn break)

<T_>rd (beginning of turn/word missing)

wo<_T> (end of word missing)

Definition

A temporarily broken or missing section of the audio signal caused by technical disruptions, distortions, or disturbances in the recording equipment, or as a result of recording mistakes.

Description

Technical interruptions can occur for a variety of reasons, and may be found within and at the ends of a turn.

1. <T_> is used when the beginning of the first word of the turn is missing. In this case it is appended to the beginning of the first occurring lexical item without a white space, regardless whether the lexical item seems to be complete or just a fragment.
2. <*T> covers those cases in which larger parts of an utterance are cut off within a turn. It is transcribed instead of the missing turn elements and is preceded and followed by a white space.
3. <*T>t is used if the end of a turn is missing, preceded by a white-space. This tag is used not only for interruptions resulting from technical artifacts, but **also when the speaker himself stops speaking before finishing his sentence.**
4. <T_> or <_T> may be used if parts of lexical items are cut off due to technical interruptions. <T_> is again used when the initial part of the lexical item is cut off. <_T> is used if the last part is cut off. They are appended to the lexical item without a white space, <T_> in front, <_T> behind.

Note

<_T> occurs only in combination with lexical items, excluding unidentifiable ones (<%>). Unidentifiable elements are excluded because it is impossible to decide whether these are cut off or not.

For the same reason, <T_> or <_T> and the "hard-to-identify"-tag (%) cannot be applied to the same lexical item.

No punctuation is allowed immediately before or after any of these tags.

If necessary, these tags can be combined quite freely, with the exception of <*T>, which can only be combined with <_T>.

Examples

1. I went swi<_T> yesterday .
2. no <T_>se shoes I have never seen before .
3. <T_>ect answer you gave .
4. <T_>cheap , that is for sure
5. wow <*T> better so far ?
6. does an<_T> <*T> <T_>orite book ?
7. that is really <*T>t
8. I think , I have to wr<_T> <*T>t

Note the difference between [=] and <_T> in

no prob<_T> <*T>t

no prob= <*T>t

<_T> means that the speaker is no longer being recorded, while [=] means that the speaker himself has stopped speaking.

2.2.20 Turn Break

Symbol

<*T>t

Definition

If a speaker contribution is aborted prematurely, it is not possible to use [...] or [...] at the end. The <*T>t symbol is thus used instead of punctuation.

Description

A turn break is the result of either a) technical disturbance affecting the recording, or b) the action of the speaker himself, who may be interrupted by his dialogue partner or for some reason, ceases to speak in the middle of a turn. Furthermore, a speaker contribution, that ends with or consists only of a non-identifiable utterance (<%>) is considered incomplete and requires the <*T>t tag.

Examples

1. thanks , we <uh> <*T>t
 2. Speaker 1: no , it is <*T>t
....Speaker 2: of course it is .
 3. I can <%> <*T>t
 4. <%> <*T>t
 5. do you want t= <*T>t
-

2.2.21 Turn Identification

Definition

In order to make transcripts easier to read, and to facilitate referencing names within a transcript, turn identification is used. The turn identification appears before each turn and identifies the speaker, the speaker channel, the name of the transcription, and the number of the turn.

Description

A turn identification looks like the following:

m100_4_0000_COU_00

m100 refers to the file name

4 - is the channel number.

0000 - this represents the turn number. In this case it happens to be the first. Turn numbers begin with zero '0'. The next turn number would be '0001' and so on.

COU - is the speaker identification. Using the speaker's initials usually creates the speaker identification.

00 – space for extra coding or version numbers.

Note that each piece of the identification is separated by an underscore. The identification ends with a colon [:] and is then followed by one white space before the transcription of the turn starts.

2.2.22 Unidentifiable - Hard to Identify

Symbol

<%> , %

Definition

The inability to understand what someone is saying. The unidentifiable word or words may be either completely incomprehensible <%> or difficult to identify (.%). There may be partial, but not complete, certainty that something perceived is actually what was said by the speaker.

Description

One of the most difficult tasks of transcription is trying to make sense of what another person says. There are instances when a person cannot be understood for a variety of reasons. Speaker interferences, rapid or soft speech, and articulation and enunciation can make hearing and understanding difficult. Dialectal or

articulated variants may also make equivalents in the standard language hard to distinguish. Finally, comprehensibility can be affected by technical factors (e.g. amplitude, clipping, cutoffs).

Use #1: In order to compensate for these instances, the percentage sign <%> mark is used. There are two related uses of this mark.

The first is for a completely unidentifiable portion of speech. Which sounds like speech, but it is not possible to make sense of it. In this case, the percentage sign within carat marks is used. If this appears at the end of a turn, it is followed by a <*T>t (the turn break symbol), not by punctuation. This symbol is not to be treated as a lexical item and may not be followed by punctuation; it is always written after punctuation.

This mark may represent a whole turn, a word, or a collection of words within a turn. It may also be used more than once in a row to represent a series of unidentifiable words.

Use #2: Words within a turn that are somewhat understandable are marked by a final % . The transcriber might be able to make out or have a faint idea of what is being said, but is not entirely certain.

In this case, the percentage sign by itself follows each word or words, which are only somewhat clear. The words followed by a percent sign are treated as normal words and may be followed by punctuation.

Examples

1. <%> database this Thursday
 2. the% speaker% database today .
 3. <%> <*T>t
 4. oh% , I% see%
 5. yes , I understand that but one must also <%> <%> <%> understand that ?
 6. yes, I understand that but one must also think% about% the% sacrifice% . can%
you% understand that ?
-

3 List of Symbols

Symbol	Description
abort=	Aborted articulation
(CMU, asap)	Acronyms and letter spelling
	Breathing
(Proper names)	Capitalization
<P>	Empty pause
-/false start/-	False start
<uh> <uhm> <hm> <hes>	Filled pauses: vowel category vowel plus nasal category nasal category other
*<FOR>word	Foreign word
;global	Global Comment
<Laugh> <Throat> <Cough> <Smack> <Swallow> <Noise>	Human Noise laughing throat clearing coughing lip smack swallowing other
(mhm, uh-uh)	Interjection
inter_ .. _ruption	Interruption of lexical items
(New-York)	Proper names
neologism*	Neologism
(twenty-one)	Numbers
(gonna)	Contractions
, ? .	Punctuation comma question mark period
+/repetition correction/+	Repetition and correction
<*T> <*T>t <T_>reak brea<_T>	Technical interruption interruption turn break initial word break final word break
<*T>t	Turn break specials
<%> mumble%	Unidentifiable Hard to identify

4 Transcription Sample

; CDR: 00.00
; TRV: 00.00
; File: m100
; Last changes made on 01/15/2003
; Transcriber: dnh
; Comments: Speaker TBVSTT speaks very fast
;
m100_4_0000_COU_00: it was unfortunate . <P> are we good ? let's roll .

m100_3_0001_YWTKTM_00: ready ?

m100_2_0002_TBVSTT_00: yup .

m100_4_0003_COU_00: Hum-Vs .

m100_3_0004_YWTKTM_00: it is not cut , <*GER>nein .

m100_2_0005_TBVSTT_00: yes , so +/I was/+ +/I was/+ as I was +/driv=/+
driving over here today , there was +/thi=/+ -/this/- <P> +/some/+
some brand new shiny <*T>t

m100_4_0006_COU_00: what kind of SUV , you don't know ? was it like
a big <*T>t

m100_2_0007_TBVSTT_00: +/it was a/+ +/it was/+ I think it was a Honda
thing , it was a brand new black shiny one .

m100_4_0008_COU_00: Honda makes an SUV ?

m100_3_0009_YWTKTM_00: oh yeah . they make like at least two .
Toyota makes like four different SUVs .

m100_1_0010_NWXRPL_00: oh , everybody does .

m100_4_0011_COU_00: Honda ? Honda the <*T>t

m100_1_0012_NWXRPL_00: Volkswagen .

m100_3_0013_YWTKTM_00: they make the one that is like a car size .

m100_4_0014_COU_00: <%> I wanted them to be like the speakers for
the environment .

m100_2_0015_TBVSTT_00: yeah .

m100_2_0016_TBVSTT_00: well <*T>t

m100_4_0017_COU_00: and to know better than that .

m100_2_0018_TBVSTT_00: I= <*T>t

m100_2_0019_TBVSTT_00: so +/people are/+ +/pe=/+ people are buying them , mean -/you can/- you know <Laugh> +/there/+ +/there is/+ +/there is very little/+ +/there is/+ there is very little +/you can/+ +/you/+ you can do +/about/+ about that directly but <%> this guy is driving and the road is +/ki=/+ getting kind of narrow and -/there is li=/- as we're getting <%> towards this bridge he is got his left turn signal on , and he is in the right lane with two lanes and a lot of cars +/and/+ and I'm trying to figure out +/w=/+ why he is got +/his/+ his turn signal on -/and then as/- and +/he/+ he goes very close and he is got a side mirror and it just hits against the side +/of the/+ <Laugh> of the bridge , you know he <*T>t

m100_3_0020_YWTKTM_00: <Laugh>

m100_4_0021_COU_00: good .

m100_3_0022_YWTKTM_00: <uh> you have to kind of laugh at this . <Laugh>

m100_2_0023_TBVSTT_00: well , I% mean% <*T>t

m100_2_0024_TBVSTT_00: well +/y=/+ +/you/+ +/you/+ +/you get/+ you get a car that is too big , I mean +/that is/+ +/that is/+ <P> -/that is what you/- that is a disadvantage of them <P> among other things you got . <P> they're just +/too/+ too wide , and obviously the road was designed for +/much/+ much smaller cars .

m100_3_0025_YWTKTM_00: yeah .

m100_3_0026_YWTKTM_00: +/I think/+ I think the problem with SUVs is +/that/+ that <P> +/they're/+ they're really good for people who know how to drive them and have an actual use for them but they have to much appeal <P> to everybody , and everybody wants one .

m100_2_0027_TBVSTT_00: mhm .

m100_3_0028_YWTKTM_00: and <uh> +/they/+ +/they/+ they're like a family vehicle now , which is kind of ridiculous I think .

m100_4_0029_COU_00: they're not safe .

m100_4_0030_COU_00: yeah .

;

;EOF