

Comparison of Transcription Symbols

	MEANING	DT1	DT2	CA
Unit				
1.	word	<u>SPACE</u>	<u>SPACE</u>	<u>SPACE</u>
2.	intonation unit	<u>LINE</u>	<u>LINE</u>	
Pause				
3.	pause, timed	...(1.2)	(1.2)	(1.2)
4.	pause, short	(.)
5.	pause, long (untimed)	
6.	latching	(0)	=	=
7.	lag (prosodic length)	=	:	:
Sequence				
8.	overlap (single)	[]	[]	[]
9.	overlap (2nd)	[2 2]	[₂]	[]
Disfluency				
10.	truncated/cut-off word	wor-	wor-	wor-
Vocalism				
11.	breathe (in)	(H)	(H)	.hhh
12.	exhale	(Hx)	(Hx)	hhh
13.	vocalism	(SNIFF)	(SNIFF)	(sniffle)
14.	click	(TSK)	(TSK)	.t
15.	laugh pulse	@	@	heh
16.	laughing word	wo@rd	wo@rd	wo(h)rd
17.	glottal stop, creak	%	(%)	
18.	glottalized word	w%ord	w%ord	wghord
Manner/Quality				
19.	manner/long feature	<A A>	<A> 	
20.	vox	<Q Q>	<VOX> </VOX>	
21.	piano, attenuated speech	<P P>	°words°	°words°
22.	smile quality	<SM SM>	<☺> </☺>	£
Metatranscription				
23.	unintelligible	XXX	###	()
24.	uncertain	<X word X>	#word	(word)
25.	comment	((WORDS))	((WORDS))	((words))
26.	pseudograph		~Jill	
Participation				
27.	speaker/turn attribution	JILL:	JILL;	Jill:
28.	unidentified speaker	X:	#;	():
29.	uncertain speaker		#JILL;	(Jill):
Boundary/Closure				
30.	terminative	.	.	.
31.	continuative	,	,	,
32.	truncated intonation unit	--	—	
33.	appeal (final)	?	?.	?
34.	appeal (continuing)	?	?,	?
Prosody				
35.	primary accent	^	^	<u>word</u>
36.	secondary accent	`	`	<u>word</u>
37.	forte	<F word F>	<F> word </F>	WORD
38.	high pitch (top)		↑	↑
39.	high pitch		↑	<u>word</u>
40.	low pitch		↓	
41.	low pitch (bottom)		↓	↓

Commentary

This Table presents a comparison between three sets of conventions for symbols used to represent basic transcription categories. The three sets are:

- ◆ **DT1** *Discourse Transcription* conventions as in the original version
(Du Bois 1991, Du Bois et al. 1992, Du Bois et al. 1993)
- ◆ **DT2** *Discourse Transcription* conventions as currently revised
- ◆ **CA** *Conversation Analysis* conventions as currently revised (Jefferson 2002)

A careful comparison of these transcription systems reveals both similarities and differences. A meaningful evaluation of such similarities and differences must penetrate more than skin deep, however. The comparison must move beyond mere surface equivalences or differences in symbols to consider the symbol's function and meaning in terms of transcription categories and transcription practices. Consider the Discourse Transcription systems, on the hand, as compared and contrasted with Conversation Analysis on the other. Here we find that a difference in symbols in some cases disguises an underlying similarity in the transcription categories, or in de facto transcriber practice. Conversely, the use of the same or a similar symbol does not guarantee an identical transcription category, nor equivalent analytical practice. Even similar names or glosses for the transcription categories in question should be taken only as a rough guide to the actual analytical practice. Where transcription practices differ, the effective meaning of the categories and symbols used is likely to differ as well. Thus, any full comparison of these transcription systems must ultimately make reference to a detailed consideration of the actual transcription practices as described in the relevant literature for each system, and as actually carried out in the body of published research.

Still, it can be stated that in many cases, the similarity suggested in a given row of compared symbols is real, even if the notations may differ. In the end, the choice of symbol matters less than the conceptualization of the category being represented, and the analytical practice it indexes.

Some observations on the comparison of individual transcription notations, categories, and practices is given below.

Speaker label. DT2 places a semi-colon after the speaker label, where DT1 and CA write a colon. While the colon is admittedly more aesthetic, it can be confused with another important function, that of marking prosodic lengthening or "lag". The use of semicolon in DT2 is designed to avoid this confusion. This is especially important for any computer processing of transcription files, as the boundary between the "text" and "pre-text" information in any transcribed line is demarcated precisely by the speaker label. There are few symbols which need to be as reliably distinguished from all others as this particular division of transcription information.

Truncated/Cut-Off Word. DT2 uses the en dash, where DT1 and CA use a simple hyphen. The reason for preferring en dash in DT2 is to avoid confusion with the use of hyphen in compound (hyphenated) words and other functions (e.g. morpheme boundaries, if these are marked). However, the plain hyphen can be used in DT2 as an interim measure, for convenience.

Truncated Intonation Unit. DT1 uses two hyphens, while DT2 uses an em dash. The use of em dash in DT2 is designed to distinguish intonation unit truncation from word truncation, as well as from

Du Bois: Comparison of Transcription Symbols

compound (hyphenated) words and morpheme boundaries. However, it is possible to use two hyphens in DT2 as an interim measure, for convenience.

Piano. DT2 provides two alternate conventions for piano, attenuated, or quiet speech. The standard DT2 notation for manner can be used, in which the words spoken quietly are enclosed within angle brackets marked with a “P”: <P> words </P>. Alternatively, the Conversation Analysis convention of enclosing the words within degree signs can be used: °words°.

Pause. Standard practice in DT2 is to measure all pauses to the tenth of a second (except micropauses, i.e. pauses of less than 150 millisecond in duration). Thus, the symbols for (unmeasured) medium and long pauses are to be used only in the broadest levels of transcription delicacy.

Pseudograph. This term refers to a notation in which a person’s name (or other sensitive indicator of identity, such as an address or cell phone number) has been changed to preserve anonymity. It is not used for speaker labels, but only where the word was actually uttered.

References

- Du Bois, John W. 1991. Transcription design principles for spoken discourse research. *Pragmatics* 1:71-106.
- Du Bois, John W., Cumming, Susanna, Schuetze-Coburn, Stephan, and Paolino, Danae. 1992. Discourse transcription. *Santa Barbara Papers in Linguistics* 4:1-225.
- Du Bois, John W., Schuetze-Coburn, Stephan, Cumming, Susanna, and Paolino, Danae. 1993. Outline of discourse transcription. In *Talking data: Transcription and coding in discourse research*, eds. Jane A. Edwards and Martin D. Lampert, 45-89. Hillsdale, NJ: Erlbaum.
- Jefferson, Gail. 2002. Is “no” an acknowledgement token? Comparing American and British uses of (+)/(-) tokens. *Journal of Pragmatics* 34:1345-1383.

[Rev. 23-Jun-2006]