
# TRANSCRIPTION CONVENTIONS [2.1]

## Mark-up conventions

VOICE
Vienna Oxford International Corpus of English

# Mark-up conventions

Version 2.1 June 2007

| 1. SPEAKER IDS | |
|---|---|
| S1:<br>S2:<br>… | Speakers are generally numbered in the order they first speak. The speaker ID is given at the beginning of each turn. |
| SS: | Utterances assigned to more than one speaker (e.g. an audience), spoken either in unison or staggered, are marked with a collective speaker ID **SS**. |
| SX: | Utterances that cannot be assigned to a particular speaker are marked **SX**. |
| SX-f:<br>SX-m: | Utterances that cannot be assigned to a particular speaker, but where the gender can be identified, are marked **SX-f** or **SX-m**. |
| SX-1:<br>SX-2:<br>… | If it is likely but not certain that a particular speaker produced the utterance in question, this is marked **SX-1**, **SX-2**, etc. |
| **2. INTONATION** | |
| Example:<br>S1: that's what my next er slide**?** does | Words spoken with rising intonation are followed by a question mark "**?**" . |
| Example:<br>S7: that's point two**.** absolutely yes**.** | Words spoken with falling intonation are followed by a full stop "**.**" . |
| **3. EMPHASIS** | |
| Example:<br>S7: er internationalization is a very IMPORTANT issue<br><br>Example:<br>S3: toMORrow we have to work on the presentation already | If a speaker gives a syllable, word or phrase particular prominence, this is written in capital letters. |
| **4. PAUSES** | |
| Example:<br>SX-f: because they all give me different **(.)** different **(.)** points of view | Every brief pause in speech (up to a good half second) is marked with a full stop in parentheses. |
| Example:<br>S1: aha **(2)** so finally arrival on monday evening is still valid | Longer pauses are timed to the nearest second and marked with the number of seconds in parentheses, e.g. (1) = 1 second, (3) = 3 seconds. |

| **5. OVERLAPS** | |
|---|---|
| Example:<br>S1: it is your best **<1> case </1>** scenario (.)<br>S2: **<1> yeah </1>**<br>S1: okay | Whenever two or more utterances happen at the same time, the overlaps are marked with numbered tags: **<1> </1>**, **<2> </2>**,…<br>Everything that is simultaneous gets the same number. All overlaps are marked in blue. |
| Example:<br>S9: it it is (.) to identify **some<1>thing** </1> where (.)<br>S3: <1> mhm </1> | All overlaps are approximate and words may be split up if appropriate. In this case, the tag is placed within the split-up word. |
| **6. OTHER-CONTINUATION** | |
| Example:<br>S1: what up till (.) till twelve?<br>S2: **yes=**<br>S1: **=really**. so it's it's quite a lot of time. | Whenever a speaker continues, completes or supports another speaker's turn immediately (i.e. without a pause), this is marked by "=". |
| **7. LENGTHENING** | |
| Example:<br>S1: you can run faster but they have much **mo:re** technique with the ball | Lengthened sounds are marked with a colon "**:**". |
| Example:<br>S5: personally that's my opinion the: **er::m** | Exceptionally long sounds (i.e. approximating 2 seconds or more) are marked with a double colon "**::**". |
| **8. REPETITION** | |
| Example:<br>S11: e:r i'd like to go t- t- **to to** this type of course | All repetitions of words and phrases (including self-interruptions and false starts) are transcribed. |
| **9. WORD FRAGMENTS** | |
| Example:<br>S6: with a minimum of (.) of **participa-**<br>S1: mhm<br>S6: **-pation** from french universities to say we have er (.) a joint doctorate or a **joi-** joint master | With word fragments, a hyphen marks where a part of the word is missing. |
| **10. LAUGHTER** | |
| Example:<br>S1: in denmark well who knows. **@@**<br>S2: **<@>** yeah **</@>** **@@** that's right | All laughter and laughter-like sounds are transcribed with the **@** symbol, approximating syllable number (e.g. ha ha ha = @@@). Utterances spoken laughingly are put between **<@>** **</@>** tags. |

| **11. UNCERTAIN TRANSCRIPTION** | |
|---|---|
| Example:<br>S3: i've a lot of very **(generous)** friends<br><br>Example:<br>SX-4: they will do whatever they want because they are a compan**(ies)** | Word fragments, words or phrases which cannot be reliably identified are put in parentheses **( )**. |
| **12. PRONUNCIATION VARIATIONS & COINAGES** | |
| Example:<br>S4: i also: (.) e:r played (.) tennis e:r **<pvc> bices </pvc>** e:r we rent? went? | Striking variations on the levels of phonology, morphology and lexis as well as 'invented' words are marked **<pvc> </pvc>**. |
| Example:<br>S9: how you were controlling such a thing and how you **<pvc> (avrivate) </pvc>** (it) | What you hear is represented in spelling according to general principles of English orthography. Uncertain transcription is put in parentheses **( )** . |
| Example:<br>S6: what we try to explain here is the foreign direct investment growth (2) in a certain industry (.) and a certain **<pvc>** compy **{company} </pvc>** | If a corresponding existing word can be identified, this existing word is added between curly brackets **{ }**. |
| Example:<br>S2: anyway i make you an a total (.) **<pvc>** summamary {summary} **<ipa>** sʌməˈmærɪ **</ipa> </pvc>** of destinations | Particularly when it comes to salient variations on the level of phonology, e.g. sound substitution or addition, a phonetic representation should be added between **<ipa> </ipa>** tags. |
| **13. ONOMATOPOEIC NOISES** | |
| Example:<br>S1: it may be quite HARMLESS and at the end of the day you (.) **<ono>** dəʃ dəʃ dəʃ **</ono>** (.) somebody | When speakers produce noises in order to imitate something instead of using words, these onomatopoeic noises are rendered in IPA symbols between **<ono> </ono>** tags. |
| **14. NON-ENGLISH SPEECH** | |
| Example:<br>S5: **<L1de> bei firmen </L1de>** or wherever | Utterances in a participant's first language (L1) are put between tags indicating the speaker's **L1**. |
| Example:<br>S7: er this is **<LNde> die seite? (welche) </LNde>** is | Utterances in languages which are neither English nor the speaker's first language are marked **LN** with the language indicated. |
| Example:<br>S4: it depends in in in **<LQit> roma </LQit>** | Non-English utterances where it cannot be ascertained whether the language is the speaker's first language or a foreign language are marked **LQ** with the language indicated. |
| Example:<br>S2: erm we want to go t- to **<LNvi> xx xxx </LNvi>** island first of all | Unintelligible utterances in a participant's L1, LN or in an LQ are represented by **x's** approximating syllable number. |
| Example:<br>S4: and now we do the boat trip (1) **<L1xx> xxxxx </L1xx>**<br>S3: mhm | Utterances in a language one cannot recognize are marked **L1xx**, **LNxx** or **LQxx**. |

| | |
|---|---|
| Example:<br>S3: <L1fr> oui un grand carre **{yes like a big square}** </L1fr> (.) i <fast> think it would </fast> be better if we put the tables a <soft> different way </soft> | If possible, translations into English are provided between curly brackets **{ }** immediately after the non-English speech. |
| **15. SPELLING OUT** | |
| Example:<br>S1: and they (3) created some (1) some er (2) JARGON. do you know? the word JARGON? (.) **<spel> j a r- </spel> <spel> j a r g o n? </spel>** jargon | The **<spel> </spel>** tag is used to mark words or abbreviations which are spelled out by the speaker, i.e. words whose constituents are pronounced as individual letters. |
| **16. SPEAKING MODES** | |
| Example:<br>S2: because as i explained before is that we have in the **<fast>** universities of cyprus we have **</fast>** a specific e:rm procedure<br><br><fast> </fast><br><slow> </slow><br><loud> </loud><br><soft> </soft><br><whispering> </whispering><br><sighing> </sighing><br><reading> </reading><br><reading aloud> </reading aloud><br><on phone> </on phone><br><imitating> </imitating><br><singing> </singing><br><yawning> </yawning> | Utterances which are spoken in a particular mode (fast, soft, whispered, read, etc.) and are notably different from the speaker's normal speaking style are marked accordingly.<br><br>The list of speaking modes is an open one. |
| **17. BREATH** | |
| Example:<br>S1: so it's always **hh** (.) going around (2) yeah | Noticeable breathing in or out is represented by two or three **h's** (hh = relatively short; hhh= relatively long). |
| **18. SPEAKER NOISES** | |
| <coughs><br><clears throat><br><sniffs><br><sneezes><br><snorts><br><applauds><br><smacks lips><br><yawns><br><whistles><br><swallows> | Noises produced by the current speaker are always transcribed. Noises produced by other speakers are only transcribed if they seem relevant (e.g. because they make speech unintelligible or influence the interaction).<br>The list of speaker noises is an open one. |
| Example:<br>S1: yeah <1> what </1> i think in in doctor levels | These noises are transcribed as part of the running text and put between pointed brackets **< >.** |

| | |
|---|---|
| S7: <1> **<clears throat>** </1> | |
| Example:<br>SX-m: but you NEVER KNOW when it's popping up you never kno:w<br>S3: **<coughs (6)>** | If it is deemed important to indicate the length of the noise (e.g. if a coughing fit disrupts the interaction), this is done by adding the number of seconds in parentheses after the descriptor. |
| **19. NON-VERBAL FEEDBACK** | |
| <nods><br><shakes head> | Whenever information about it is available, non-verbal feedback is transcribed as part of the running text and put between pointed brackets **< >**. |
| Example:<br>S3: but i think if you structure corporate governance appropriately you can have everything (1)<br>S7: <soft> mhm </soft> **<nods (2)>** | If it is deemed important to indicate the length of the non-verbal feedback, this is done by adding the number of seconds in parentheses. |
| **20. ANONYMIZATION** | |
| | A guiding principle of VOICE is sensitivity to the appropriate extent of anonymization.<br>As a general rule, names of people, companies, organizations, institutions, locations, etc. are replaced by aliases and these aliases are put into square brackets **[ ]**. The aliases are numbered consecutively, starting with 1. |
| Example:<br>S9: that's one of the things (.) that i (1) just wanted to clear out. (2) **[S13]**?<br><br>Example:<br>S6: so: (1) ei:ther MYself or mister **[S2/last]** or even boss (.) should be there every year<br><br>Example:<br>S8: so my name is **[S8]** **[S8/last]** from vienna | Whenever speakers who are involved in the interaction are addressed or referred to, their names are replaced by their respective speaker IDs.<br><br>A speaker's first name is represented by the plain speaker ID in square brackets **[S1]**, etc.<br><br>A speaker's last name is marked **[S1/last]**, etc.<br><br>If a speaker's full name is pronounced, the two tags are combined to **[S1]** **[S1/last]**, etc. |
| Example:<br>S2: that division is headed by (1) **[first name3]** **[last name3]** (1) | Names of people who are not part of the ongoing interaction are substituted by **[first name1]**, etc. or **[last name1]**, etc. or a combination of both. |
| Example:<br>S5: erm she is currently head of marketing (and) with the **[org2]** (1) | Companies and other organizations need to be anonymized as well. Their names are replaced by **[org1]**, etc. |
| Example:<br>S1: i: i really don't wanna have a: a joint degree e:r with the university of **[place12]** (.) | Names of places, cities, countries, etc. are anonymized when this is deemed relevant in order to protect the speakers' identities and their environment. They are replaced by **[place1]**, etc. |

| | |
|---|---|
| Example:<br>S8: he get the <L1cs> diplom {diploma} </L1cs> of **[name1]** university (.) and french university can give him also the <L1cs> diplom {diploma} </L1cs> | Other names or descriptors may be anonymized by **[name1]**, etc., as in e.g. Charles University. |
| Example:<br>S3: erm i- in the **[thing1]** is very well explained. so <2> i can </2> pa- <3> er pass you this </3> th- the definitions.<br>S4: <2> aha </2><br>S4: <3> okay <@> okay </@> </3> | Products or other objects may be anonymized by **[thing1]**, etc. |
| **21. CONTEXTUAL EVENTS** | |
| {mobile rings}<br>{S7 enters room}<br>{S2 points at S5}<br>{S4 starts writing on blackboard}<br>{S4 stops writing on blackboard}<br>{S2 gets up and walks to blackboard (7)}<br>{S3 pours coffee (3)}<br>{SS reading quietly (30)}<br>…<br><br>Example:<br>S3: one dollar you get (.) (at) one euro you get one dollar twenty-seven. (.)<br>S4: right. **{S5 gets up to pour some drinks}**<br>S3: right now at this time (3)<br>S1: er page five is the er (4) **{S5 places some cups and glasses on the desk (4)}**<br>S1: i think is the descritip- e:r part of what i have just explained (.) | Contextual information is added between curly brackets **{ }** only if it is relevant to the understanding of the interaction or to the interaction as such. If it is deemed important to indicate the length of the event, this can be done by adding the number of seconds in parentheses.<br><br>Explanation:<br>The pause in the conversation occurs because of the contextual event. |
| **22. PARALLEL CONVERSATIONS** | |
| Example:<br>S1: four billion <spel> u s </spel> dollars. (.)<br>S4: quite impressive (.)<br>S1: er **<to S2>** not quite isn't it **</to S2>** (.) i understand some other countries we handle | To indicate that a speaker is addressing not the whole group but one speaker in particular, the stretch of speech is marked with (e.g.) **<to S1> </to S1>**, choosing the speaker ID of the addressee. |
| Example:<br>S7: i've i've found the people very stressed<br>SS: @@@<br>S7: that's (.) i don't know how many of you study here but it's VERY important to push the close the door button in that elevator. this is something i've never <3> seen in sweden </3> **{parallel conversation between S1 and S3 starts}** or anywhere else <4> but it's very | Wherever two or more conversational threads emerge which are too difficult to transcribe, as a general rule only the main thread of conversation is transcribed. The threads which are not transcribed are treated like a contextual event and indicated between curly brackets **{ }**. |

| | |
|---|---|
| important to push this button </4><br>SS: <3> @@@@ </3><br>SS: <4> @@@@@@@@ </4> @@<br>S7: <5> i never even saw this button in another el- elevator </5><br>SS: <5> @@@@@@@@@@ </5><br>**{parallel conversation between S1 and S3 ends}** @@@ | |
| **23. UNINTELLIGIBLE SPEECH** | |
| Example:<br>S4: we **<un> xxx </un>** for the <7> supreme (.) three </7> possibilities<br>S1: <7> next yeah </7> | Unintelligible speech is represented by **x's** approximating syllable number and placed between **<un> </un>** tags. |
| Example:<br>S7: obviously the the PROCESS will **<un>** x **<ipa>** θeɪŋ **</ipa> </un>** (.) w- w- will (.) will take (.) at least de- decade | If it is possible to make out some of the sounds uttered, a phonetic transcription of the x's is added between **<ipa> </ipa>** tags. |
| **24. TRANSCRIPTION BORDERS** | |
| <**beg** CD1_4_00:35> | The **beg**inning of the transcript is noted by indicating the CD number, the track number and the exact position of the respective track in minutes and seconds. |
| <**end** CD1_21_01:27> | The **end** of the transcript is noted in the same way. |
| <end CD1_19_01:27><br>**(gap 00:06:36) {multiple parallel conversations, hardly intelligible}**<br><beg CD1_21_02:03> | A **gap** in the **transcription** is indicated in parentheses, including its length in hh:mm:ss. Curly brackets **{ }** are used in order to specify the reasons for or the circumstances of the gap. |
| <end CD1_24_3:02><br>**(nrec 00:00:45) {change of minidisk}**<br><beg CD2_1_00:00> | An **interruption** in the **recording** is indicated in the same way, but abbreviated as "**nrec**" (i.e. non-recorded). The length you indicate will normally be a guess. |

In addition to the regular mark-up, transcribers supplement the transcripts with Transcriber's Notes in which they provide additional contextual information and observations about other features of the interaction not accounted for in the transcript.

For a detailed discussion of specific aspects of the transcription conventions cf. Breiteneder, Pitzl, Majewski, Klimpfinger. (2006). "VOICE recording – Methodological challenges in the compilation of a corpus of spoken ELF". *Nordic Journal of English Studies*, 5/2, 161-188.