



TRANSCRIPTION CONVENTIONS [2.1]

Spelling conventions

The **VOICE Transcription Conventions** are protected by copyright. Duplication or distribution to any third party of all or any part of the material is not permitted, except that material may be duplicated by you for your personal research use in electronic or print form. Permission for any other use must be obtained from VOICE. Authorship must be acknowledged in all cases.

Spelling conventions

Version 2.1 June 2007

1. CHARACTERS	
a b c d e f g h i j k l m n o p q r s t u v w x y z	Only alphabetic Roman characters are used in the transcript. No diacritics, umlauts or non-roman characters are permitted in the running text.
2. DECAPITALIZATION	
<u>Example:</u> S8: so you really can <@> control my english </@>	No capital letters are used except for marking emphasis (cf. mark-up conventions).
3. BRITISH SPELLING	
British spelling	British English spelling is used to represent naturally occurring ELF speech. The Oxford Advanced Learner's Dictionary (OALD), 7th edition , is used as the primary source of reference. If an entry gives more than one spelling variant of a word, the first variant is chosen. If there are two separate entries for British and American spelling, the British entry is selected.
4. SPELLING EXCEPTIONS	
center, theater behavior, color, favor, labor, neighbor defense, offense disk program travel (-l-: traveled, traveler, traveling)	The 12 words listed on the left and all their derivatives are spelled according to American English conventions (e.g. colors, colorful, colored, to color, favorite, favorable, to favor, in favor of, etc.).
<u>Example:</u> S2: we are NOT quite sure if it will REALLY be (.) privatized next year	In addition, all words which can be spelled using either an <i>-is</i> or an <i>-iz</i> morpheme are spelled with -iz (e.g. to emphasize, organizations, realization, recognized, etc.).
5. NON-ENGLISH WORDS	
<u>Example:</u> S1: <L1de> wieso o esterreich? {why austria} </L1de> <u>Example:</u> S3: <LNfr> c'est ferme? {is it closed} </LNfr>	Non-English words are rendered in the standard variant of the original language (i.e. no non-standard dialect). The roman alphabet is always used, also in the case of languages like Arabic or Japanese. No umlauts (e.g. NOT österreich), no diacritics (e.g. NOT fermé) and no non-roman characters are permitted.

<p align="center">6. FULL REPRESENTATION OF WORDS</p>	
<p><u>Example:</u> S7: the students that (.) decide freely to enter (.) this kind of master knows (.) for example that he can (.) at the end achieve (.) sixty credits</p>	<p>Although words may not be fully pronounced or may be pronounced with a foreign accent, they are generally represented in standard orthographic form.</p> <p><u>Explanation:</u> S7 is Italian and pronounces the <i>he</i> in <i>he can</i> as /ɪ/, swallowing the initial h. Nevertheless, this is regarded as a minor instance of L1 accent and therefore represented in standard orthography (<i>he</i>).</p>
<p align="center">7. FULL REPRESENTATION OF NUMBERS, TITLES & ABBREVIATIONS</p>	
<p>oh/zero, two, three, ... one hundred, nineteen ten, eighteen twenty-seven, ...</p>	<p>Numbers are fully spelled out as whole words. British English hyphenation rules apply.</p>
<p>missis (for <i>Mrs</i>), mister, miss, mis (for <i>Ms</i>), doctor, professor, ...</p>	<p>Titles and terms of address are fully spelled out.</p>
<p>et cetera, saint thomas, okay,...</p>	<p>Forms that are usually abbreviated in writing, but spoken as complete words are fully spelled out.</p>
<p align="center">8. LEXICALIZED REDUCED FORMS</p>	
<p>cos gonna, gotta, wanna</p>	<p>Lexicalized phonological reductions are limited to the four on the left. All other non-standard forms are fully spelled out (e.g. /hæftə/ = <i>have to</i>).</p>
<p align="center">9. CONTRACTIONS</p>	
<p>i'm, there're, how's peter, running's fun, ... i've, they've, it's got, we'd been, ... tom'll be there, he'd go for the first, ... we aren't, i won't, he doesn't, ...</p> <p>what's it mean, where's she live, how's that sound ... let's</p>	<p>Whenever they are uttered, all standard contractions are rendered. This refers to verb contractions with <i>be</i> (<i>am, is are</i>), <i>have</i> (<i>have, has, had</i>), <i>will</i> and <i>would</i> as well as <i>not</i>-contractions.</p> <p>Additionally, 's is used to represent <i>does</i> when reduced and attached to a <i>wh</i>-word. It is also used to represent the pronoun <i>us</i> in the contracted form <i>let's</i>.</p>
<p align="center">10. HYPHENS</p>	
<p><u>Example:</u> S3: more than thirteen years of experience er working in (.) er (.) design and development (.) er of (1) real-time software (.) er for industrial (.) implications</p>	<p>Hyphens are used according to British English hyphenation rules. The OALD, 7th edition, is used as the primary source of reference.</p>

<u>Example:</u> S2: we would allow that within er an international cooperation (.)	If an entry gives more than one spelling variant of a word, the first variant is chosen.
11. ACRONYMS	
<u>Example:</u> S10: for the development of joint programmes within the unica networks.	Acronyms (i.e. abbreviations spoken as one word) are transcribed like words. They are not highlighted in any way.
12. DISCOURSE MARKERS	
	All discourse markers are represented in orthography as shown below. The lists provided are closed lists. The items in the lists are standardized and may not represent the exact sound patterns of the actual discourse markers uttered.
yes, yeah, yah okay, okey-dokey	Backchannels and positive minimal feedback
mhm, hm aha, uhu	(closed sound-acknowledgement token) (open sound-acknowledgement token)
no n-n, uh-uh	Negative minimal feedback
er, erm	Hesitation/filler
huh	tag-question
	Exclamations
yay, yipee, whoohoo, mm:	joy/enthusiasm
haeh	questioning/doubt/disbelief
a:h, o:h, wow, poah	astonishment/surprise
oops	apology
ooph	exhaustion
ts, pf	disregard/dismissal/contempt
ouch, ow	pain
sh, psh	requesting silence
oh-oh:, u:h	anticipating trouble
ur	disapproval/disgust
oow	pity, disappointment
<u>Example:</u> S3: <L1ja> he: </L1ja>	What are clearly L1-specific discourse markers are marked as foreign words. Due to the wide range of these phenomena in different languages, the L1-list is open-ended. A translation is added whenever this is possible.
<u>Example:</u> SX-m: <L1de> ach ja {oh yes} </L1de>	

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.