THE TRANSCRIPTS

ACKNOWLEDGEMENT

The team carrying out the current research project is deeply indebted to the team of researchers who have developed the Vienna-Oxford International Corpus of English (VOICE), particularly as regards the system of transcription and annotation they have developed for their corpus of English as lingua franca conversations. The system of transcription used in the current project is based on the following systems developed by the VOICE team:

VOICE Project. 2007. "Mark-up conventions". VOICE Transcription Conventions [2.1].

http://www.univie.ac.at/voice/documents/VOICE_mark-up_conventions_v2-1.pdf

(date of last access 04/06/2013).

VOICE Project. 2007. "Spelling conventions". VOICE Transcription Conventions [2.1].

http://www.univie.ac.at/voice/documents/VOICE_spelling_conventions_v2-1.pdf

(date of last access 04/06/2013).

MAJOR DIFFERENCES BETWEEN THE TRANSCRIPTION AND ANNOTATION SYSTEMS USED IN VOICE AND THE CORPUS OF OFFICE HOURS' CONSULTATIONS

Towards a multi-modal transcription of the office hours' consultations

The nature and purpose of the recording of these office hours' consultations contribute to some of the major differences between these transcripts and those developed for VOICE. The conversations were video-recorded, which allowed the transcribers to observe many non-verbal aspects of the interaction (gesture or eye gaze, for example) that contributed in meaningful ways to the conversations that took place. The information provided in the 'Contextual Information' document that accompanies each transcript, as well as the annotation of non-verbalfeedback, aims to allow readers of these transcripts to gain insights into the role played by gesture, facial expressions, and other types of non-verbal behaviour in each interaction.

- **LAYOUT**
- The transcripts of the office hours' consultations are organized in six columns, which record (from left to right) (1) the speech entry number (which does not necessarily correspond to a turn at talk, as in VOICE); (2) speaker identification (either the initials of the participants' pseudonyms or R for 'researcher'); (3) transcribed speech (4) time stamp (every 30 seconds) and word/fragment of speech at which it occurs; (5) transcription of lecturer's non-verbal behaviour; (6) transcription of student's non-verbal behaviour.

ANNOTATION OF THE NON-VERBAL BEHAVIOUR

The positioning of participants has on occasion made it impossible to annotate all non-verbal behaviour. Where this has been possible, the non-verbal behaviour annotated in the 5th and 6th columns of each transcript record gesture, eye gaze, head movement, and facial expression. These are distinguished by their colour coding.

- **Gesture is marked in red**. The annotation of gestures in the transcript at present only record when they occur, but do not describe in detail what they consist of or offer any interpretation of their meaning. Two types of annotation have been used. One records a clearly delimited movement of one or both of a participant's hands (i.e. consisting of one stroke) and the other what we have termed a gesture phase (i.e. movements of one or both hands which succeed each other and in which there is no clear retraction – or return of the hands to a position of rest –between each stroke). The former type are annotated as **Gesture** (left hand / right hand / both hands) **occurs at** (time/word) . The second are annotated as **Gesture phase** (left hand / right hand / both hands (alternately)) **begins at** (time/word) and **ends at** (time/word).
- **Eye gaze is marked in green.**

  The position of the participants, whether they were wearing glasses, whether their hair covered part of their faces, and so on, made it impossible to follow the participants' eye gaze with any degree of reliability

in many cases. However, where it was possible to follow what participants were looking at, this is annotated in the following ways. The default eye gaze position is taken to be that directed at the conversational partner, and so this position is not annotated. Shifts of gaze from this position are annotated as X looks away at (word); X looks at (eg monitor/notebook s/he is holding /participant) at (word); and X looks at ceiling at (word). When the shift in eye gaze is accompanied by a movement of the body, this is annotated as X turns away at (word) or X turns to Y at (word). Likewise, eye gaze shifts when a participant is writing (shifting to the piece of paper). This is annotated as X writes on Y at (word); X writes on Y from (word) to (word); or X begins to write at (word) Stops writing at (word).

- **Head movement providing non-verbal feedback is marked in blue.**

  Two types of head movement (nods ans shakes of the head) are annotated as follows: X nods (word); X begins to nod at (word); X stops nodding at (word); and X shakes head (word). X begins to shake head at (word). X stops shaking head at (word).

- **Facial expression is marked in purple.**

  The annotation of facial expressions must be taken against the background of the participants' demeanour generally, as noted in the 'Contextual Information'. The following types of facial expression are noted: X frowns (word) [only the first word]; X smiles (word) [only the first word]; X looks surprised (word) [only the first word]; X looks puzzled (word) (where facial expression is interpreted as signalling incomprehension) [only the first word]; and X rolls eyes (word) [only the first word].

MINOR DIFFERENCES BETWEEN THE TRANSCRIPTION AND ANNOTATION SYSTEMS USED IN VOICE AND THE CORPUS OF OFFICE HOURS' CONSULTATIONS

1. MODIFICATION OF VOICE MARK-UP CONVENTIONS (VERSION 2.1 JUNE 2007)

- **1. Speaker IDs** are initials, corresponding to the pseudonyms assigned to each participant. 'R' stands for an unidentified member of the research team who made the recording in question.

- **2. Intonation.** No attempt has been made to distinguish fall from non-fall tones in these conversations and render these through punctuation marks (i.e., . or ?). The use of ? is used to disambiguate an utterance intended a

question (e,g, 'that's clear?') rather than a statement (e.g. 'that's clear') in the context in which it is used. ? is not used where the form of the utterance (e.g. inversion of operator and subject as in 'do you understand') makes it clear that the speaker is asking a question.

- **5. Overlaps.** For reasons of readability, those overlaps that occurred in part of a word or in certain syllables were not split up but instead highlighted in blue within the tag. That is, the entire word goes into the tag but only the part of the word that overlaps is marked in blue. For example: JW: that are relevant to all <19> students regardless </19> DI; <19> yeah okay </19>

- **12. Pronunciation variations and coinages.** In addition to the mark-up conventions described in the VOICE Transcription Conventions, wordings and expressions which have been identified as being the result of influence of the L1 or another language which is not the speaker's mother tongue are marked up as <L1[language]inf> </L1[language]inf> or <LN[language]inf> </LN[language]inf>, respectively. For example, HV: the <L1nlinf> red line (.) of their lecture </L1nlinf>

- **16. Speaking modes.** An important addition to the list of speaking modes described in the VOICE Transcription Conventions is reading aloud from a document or other visual display. This marked <reading aloud> </reading aloud>.

- **19. Non-verbal feedback**. All non-verbal feedback is recorded in the 5th and 6th columns and is therefore not marked up in the transcript of the speech.

- **20. Anonymization.** All the names of people and places mentioned in these conversations that might lead to identification of the participants or others they allude to have been assigned a pseudonym. These are marked up as <pseud> </pseud>. For example, JW: i'll make a note of that for <pseud> betty </pseud> because (.) erm <pseud> betty </pseud> does the registration.

- **21. Contextual events.** These are not marked up in the speech transcripts, but in the 5th and 6th column or in the 'Contextual Information' document accompanying the transcripts.

- **22. Parallel conversations.** Due to the nature of the conversations recorded, no parallel conversations needed to be marked up.

- **23. Unintelligible speech.** All unintelligible speech is represented by x's between <un> </un> tags. The use of phonetic transcription has been avoided.
- **24. Transcription borders.** These are noted using the name assigned to each recording. For example, <beg UE1 00:35> <end UE1 10:35>

2. <u>**MODIFICATION OF VOICE SPELLING CONVENTIONS (VERSION 2.1 JUNE 2007)**</u>
- **3. British spelling.** The dictionary used for spelling conventions was the Oxford English Dictionary (OED).
- **4. Spelling exceptions.** No spelling exceptions are made to those found in the OED.
- **8. Lexicalized reduced forms.** All phonologically reduced forms are represented in their standard orthographic form (for example, 'want to' for [wɒnə] rather than 'wanna'). The only exceptions to this are " 'cause' (rather than "cos") and "gonna" when it is followed by "to VERB". For example,