



Woefzela 8.0 User Guide

Author: Nic de Vries

Updated: June 2011

If you find this application useful, please cite the associated paper:

N.J. de Vries, J. Badenhurst, M.H. Davel, E. Barnard and A. de Waal, "Woefzela - An open-source platform for ASR data collection in the developing world," in *Proc. INTERSPEECH*, Florence, Italy, August 2011, pp. 3177-3180.

BibTeX-style reference:

```
@inproceedings{DeVries2011Woefzela,  
  author = {de Vries, Nic J. AND Badenhurst, Jaco AND Davel, Marelie H. AND Barnard, Etienne AND de  
    Waal, Alta},  
  title = {{Woefzela - An open-source platform for ASR data collection in the developing world}},  
  booktitle = {Proc. INTERSPEECH},  
  address = {Florence, Italy},  
  month = {August},  
  year = {2011},  
  pages = {3177--3180},  
}
```

Table of Contents

[Introduction](#)

[Purpose](#)

[Requirements](#)

[Software dependencies](#)

[Important Notices](#)

[Getting Started](#)

[Installation](#)

[Application download and installation](#)

[Installing application dependencies](#)

[Download and installation of test corpora](#)

[A number of ways exist to 'install' the test corpora for Woefzela.](#)

[Preparing prompt lists](#)

[Input corpus file names and file format](#)

[Prompt list file name](#)

[Prompt list file location](#)

[Prompt list format](#)

- [Prompt list length](#)
- [Recording a session](#)
 - [Fieldworker Information](#)
 - [Fields](#)
 - [Action buttons](#)
 - [Respondent Information](#)
 - [Fields](#)
 - [Drop-down lists \(Spinners\)](#)
 - [Action buttons](#)
 - [Session Information](#)
 - [Information \(only\) fields](#)
 - [Drop-down lists \(Spinners\)](#)
 - [Fields](#)
 - [General input comments](#)
- [Main recording screen](#)
 - [Prompt area](#)
 - [Action buttons](#)
 - [Statistics area](#)
 - [Progress line](#)
 - [Quality line](#)
- [General application navigation](#)
- [Retrieving data](#)
 - [Data retrieval process](#)
 - [Folder structure of retrieved data](#)
 - [Top level tree:](#)
 - [Next level tree:](#)
 - [Output data tree:](#)
 - [Profiles tree:](#)
 - [Basic file types](#)
 - [Folder: /sdcard/Woefzela/OutputData](#)
 - [Folder: /sdcard/Woefzela/Profiles](#)
 - [Folder: /sdcard/Woefzela/Tracking](#)
 - [XML Metadata](#)
- [Clearing the SD card for further recordings](#)
- [Additional settings](#)
 - [Recording target](#)
- [Troubleshooting](#)

Introduction

Purpose

The primary purpose of the *Woefzela* software application is to record a list of prompts by a number of different speakers. The resultant output is then intended to be used as training and/or testing data (audio and associated transcriptions) for building Automatic Speech Recognition (ASR) Systems in the selected language.

Requirements

Woefzela requires Android version 2.1-update1 or higher.

The application has been tested on:

- HTC Magic running Android 2.1-update1 (CyanogenMod-5.0.8-DS)
- HTC Dream running on Android 2.2 (CyanogenMod-6.0.0-DS-RC1)
- Motorola DEFY/Motorola MB 525 (in South Africa) running Android 2.3.4 (CyanogenMod; MB525 Build/GRJ22)

Software dependencies

The application FileManager-1.1.3 is currently a software dependency used via an intent from within Woefzela. This application may be obtained from

<http://code.google.com/p/openintents/downloads/detail?name=FileManager-1.1.3.apk>.

Important Notices

Please note that due to the sensitive information contained in both Fieldworker and Respondent profiles on the SD cards used for recordings, access to this information should be carefully controlled. It is suggested that these profiles be often removed from the SD cards and archived in an access-controlled location.

Getting Started

Installation

Application download and installation

Woefzela can be installed on any compatible Android device in a number of ways (see the “Requirements” section for Android version minimum requirements):

Pre-requisite for all methods: Enable applications installed from Unknown sources by selecting: Menu > Settings > Applications, and checking the ‘Unknown sources’ checkbox. Note that this may pose certain risks when installing other applications, but if the actual source of the application package (apk) is trusted, it should pose no additional risks. After installation this option may be unchecked if desired.

1. Via QR code
 - a. Scan the QR-code displayed on the Woefzela site home page with any QR-enabled bar code reader on the Android device e.g. ixMAT Barcode Scanner, and open the link in a browser. This should download the application (apk) onto the Android device automatically.
 - b. Select the downloaded apk file to install the application.
2. Via WiFi
 - a. Enable WiFi on the Android device and connect to the Internet via WiFi.
 - b. Using the device’s browser: browse to the package (apk) on the download page on the Woefzela site i.e.
<https://sites.google.com/site/woefzela/home/download/Woefzela.apk> (note that the package name should be included in the URL). This will cause the package to be automatically downloaded and subsequently installed on the device.
3. Via SD card with ‘normal’ PC
 - a. Download the apk onto a PC from the Woefzela site’s Download page.
 - b. Remove the SD card in the Android device by unmounting it by selecting: Menu > Settings > SD card & phone storage > Unmount SD card on the device. Physically remove the SD card and place it in a SD card reader connected to the PC.
 - c. On the PC: copy the apk file onto the SD card in any location and safely remove the card from the PC.
 - d. Place the SD card back into the Android device and wait a couple of seconds for the SD card to mount automatically.
 - e. Using any file manager on the Android device e.g. OI File Manager or ASTRO

(available on Android Market), browse to the apk on the SD card.

- f. By selecting the apk on the SD card, Android will automatically attempt to install the application on the device.
4. Via SD card with PC loaded with the Android Debug Bridge (adb)
See: <http://developer.android.com/guide/developing/tools/adb.html> for information on adb.
 - a. Download the apk onto a PC from the Woefzela site's Download page.
 - b. Connect the Android device to the PC with a USB cable.
 - c. On the command line of the PC, browse to the folder where the apk was downloaded.
 - d. On the command line of the PC, invoke:
 - i. `adb devices` to ensure that the device is recognised.
 - ii. `adb install Woefzela.apk` to install the package. The displayed message should indicate success.
(If a different version of Woefzela is already installed on the device, it could be indicated by getting a "Failure [INSTALL_FAILED_ALREADY_EXISTS]" message. Invoke `adb uninstall org.meraka.nchlt.woefzela` on the command line, and then the previous install command.)
5. Via the Android Market
The application is currently not available on the Android Market.

Installing application dependencies

Any required application dependencies, such as OI File Manager, can be installed from the Android Market by searching for that dependency. e.g. Search for "OI File Manager"

Download and installation of test corpora

A number of ways exist to 'install' the test corpora for Woefzela.

Pre-requisite for all methods:

1. Download and install Woefzela as explained in the relevant sections above.
2. Invoke the Woefzela application until the first 'form', called "Fieldworker Information", is displayed.
3. Terminate the application by pressing the BACK key at this point and selecting the 'Quit' option.
4. At this point all the necessary folders have been created by the Woefzela application, ready for uploading data (e.g. corpora) to the relevant folders.
5. Follow any of the options listed below.

1. Via WiFi

- a. Enable WiFi on the Android device and connect to the Internet via WiFi.

- a. Using the device's browser: browse to a corpus file on the download page on the Woefzela site e.g.
https://sites.google.com/site/woefzela/home/download/xx_ZZ_train.txt (note that the corpus file name should be included in the URL). This will cause the corpus file to be automatically downloaded onto the device's SD card in a default download location.
- b. Repeat the previous step for all the corpus files.
- c. By using any file manager installed on the device, move the desired corpus files from the default download location to the following folder on the SD card:
Woefzela/CorpusInput

2. Via SD card with 'normal' PC

- d. Download the corpus files onto a PC from the Woefzela site's Download page. e.g. xx_ZZ_train.txt and xx_ZZ.txt.
- e. Remove the SD card in the Android device by Unmounting it by selecting: Menu > Settings > SD card & phone storage > Unmount SD card on the device. Physically remove the SD card and place it in a SD card reader connected to the PC.
- f. Copy the corpus files onto the SD card into the folder `Woefzela/CorpusInput` (which should already exist if above pre-requisite section was successful) and safely remove it from the PC.
- g. Place the SD card back into the Android device and wait a couple of seconds for the SD card to mount automatically.

3. Via SD card with PC loaded with the Android Debug Bridge (adb)

See: <http://developer.android.com/guide/developing/tools/adb.html> for information on adb.

- h. Download the corpus files onto a PC from the Woefzela site's Download page.
- i. Connect the Android device to the PC with a USB cable.
- j. On the command line of the PC, browse to the folder where the corpus files was downloaded.
- k. On the command line of the PC, invoke:
 - i. `adb devices` to ensure that the device is recognised.
 - ii. `adb push xx_ZZ_train.txt /sdcard/Woefzela/CorpusInput/xx_ZZ_train.txt` to copy the file to the correct location on the SD card. The displayed message should indicate success.
 - iii. Repeat the above step for all required corpus files.

Preparing to record

Preparing prompt lists

In order to present a list of prompts for the user to read, two (2) similar lists must be prepared and loaded onto the SD card of the target device for each desired language, namely a 'training corpus' and a 'recording corpus'.

Input corpus file names and file format

Prompt list file name

The training input corpus file name has the format xx_YY_train.txt and the recording input corpus file name has the format xx_YY.txt, where xx indicates the two-letter lowercase ISO language code (such as "en") as defined by ISO 639-1, and YY indicates the two-letter uppercase ISO country code (such as "ZA") as defined by ISO 3166-1. (For more information, see: <http://developer.android.com/reference/java/util/Locale.html>).

e.g. For the Zulu language in South Africa, the following files must exist: zu_ZA_train.txt (for the training corpus) and the associated zu_ZA.txt (for the recording corpus).

Note, that both the training and recording corpora must be present by the time the Respondent's accent is selected on the Respondent Information screen. The software will warn the user if the relevant corpus, based on the selection of the speaker's accent, is not present. e.g. If the respondent has indicated as having an af_ZA (Afrikaans - South Africa) accent, the files that must be present are af_ZA_train.txt and af_ZA.txt.

Prompt list file location

Both the training and recording prompt list files, with names as explained above, must be present in the folder /sdcard/Woefzela/CorpusInput.

These files may be directly loaded onto the SD card into the correct location or may be loaded using the Android Debug Bridge (<http://developer.android.com/guide/developing/tools/adb.html>) if available.

Note: The first time that the application is run on the device, all the necessary folders will be created on the SD card. If the user thus chooses to, the application may be started and then terminated at the first input screen (Fieldworker Information). Recording corpora may then be loaded into the required folders and the application re-launched.

Prompt list format

The general format of each line of both the training and recording corpora are as follows:
<utterance length expectation>;<utterance category>;<utterance text><End Of Line>

e.g. 1.173;normal;salt lake city

Notes:

- Utterance length expectation (duration in seconds) is intended for quality control purposes - this field must be present, but is currently not used. i.e. "1.000" would be an acceptable default value.
- Utterance category was also planned for quality control purposes in conjunction with the utterance length expectation - this field must also be present, but is currently also not used. i.e. "normal" would be an acceptable default value.
- No quotation marks must be present unless it forms part of the actual prompt string.
- No empty lines must be present (except for the last line in the corpus) as this will signify the End-Of-File for that corpus.
- No leading spaces must be present at the start of a line.
- No leading or trailing spaces must be present before or after semicolons (";").
- Each entry (i.e. prompt text with associated utterance length expectation and category) must be on a separate line in the file.
- File encoding to be used must be UTF-8 if any diacritics or special characters are required.

Prompt list length

Although there is no absolute minimum or maximum size for the training and test prompt corpora, it is **imperative** that both the corpora have a significantly larger number of prompts (lines in this case) in the files than the target number of prompts set for the training and test sessions (pre-compiled).

As a good example: A typical training corpus has around 30 lines and the training target is set to 15 prompts, while the recording corpus typically has 10 000 lines and the target is set for 600 prompts per recording session.

Note: Due to some constraints in the current version of the software, it is required that the prompt list/corpus be much larger than the prompt target i.e. to avoid having to wrap around to the top of the corpus multiple times during prompt loading, but this constraint may be alleviated in the future versions of the software.

Recording a session

Once the required prompt lists (input corpora) have been loaded onto the device's SD card for the desired language, the program may be launched.

Fieldworker Information

Fieldworker Information screen

Fields

First name(s) Enter name or names of person. e.g. Xolani Mandla (Compulsory)

Surname(s) Enter surname/family name/last name of person. e.g. Nelisiwe
(Compulsory)

ID No. Enter a 13-digit identity number of person with the format YYMMDDxxxxxxx. This field is used for identification purposes in certain countries. Any number may be substituted here as long as the first 6 digits (YYMMDD) contains the date of birth of the person in order for the age of the person to be easily derived. e.g. 8012311234567 for someone born on 31 December 1980. (Compulsory)

Mobile number Enter a 10-digit mobile phone number of the format dddddddddd. This number may be substituted for any number of digits less than 10 digits as long as at least one digit is present. (Optional)

Email address Enter the email address of the person. (Optional)

Action buttons

“Load existing profile”

This button will enable you to load any Fieldworker profile that already exists on the SD card. Please, note that Fieldworker profiles (stored under /sdcard/Woefzela/Profiles/Fieldworkers) can not be mixed with Respondent profiles and vice versa.

In order to use an existing profile, select the desired profile from the list and press “Go”. If the desired profile does not exist, press the BACK key on the keyboard and enter the required information in the fields. This profile will then automatically be saved when moving to the next screen.

“Reset fields!”

This button is intended to clear all fields of the current input screen. This functionality is currently not implemented.

“Next...”

This button allows the user to proceed to the the next input screen (Respondent Information).

Respondent Information

Respondent Information screen (overlap shown)

Fields

- First name(s) Enter name or names of person. e.g. Jack (Compulsory)
- Surname(s) Enter surname/family name/last name of person. e.g. Blake (Compulsory)
- Age Enter a 2-digit of person. This field is used for age categorization of the speech data and is thus required to be accurate. (Note: ID No. in older versions) (Compulsory)
- Mobile number Enter a 10-digit mobile phone number of the format dddddddddd. This number may be substituted for any number of digits less than 10 digits as long as at least one digit is present. (Optional)
- Email address Enter the email address of the person. (Optional)

Drop-down lists (Spinners)

- Spoken accent Select the accent of the current speaker from a pre-compiled list.
Note that the input corpus (prompts) for this accent must already be available on the SD card at this point.
- Gender Select the person's gender. Generally used for training gender-specific ASR models.

Terms and Conditions This section consists of a button (to read the terms and conditions) and a check-box to acknowledge acceptance of the terms. Terms can only be accepted if the Terms has at least been viewed once (assumed from someone pressing the “Read the Terms and Conditions” button).

Action buttons

“Load existing profile”

This button will enable you to load any Respondent profile that already exists on the SD card. Please, note that Respondent profiles (stored under /sdcard/Woefzela/Profiles/Respondents) can not be mixed with Fieldworker profiles and vice versa.

In order to use an existing profile, select the desired profile from the list and press “Go”. If the desired profile does not exist, press the BACK key on the keyboard and enter the required information in the fields. This profile will then automatically be saved when moving to the next screen.

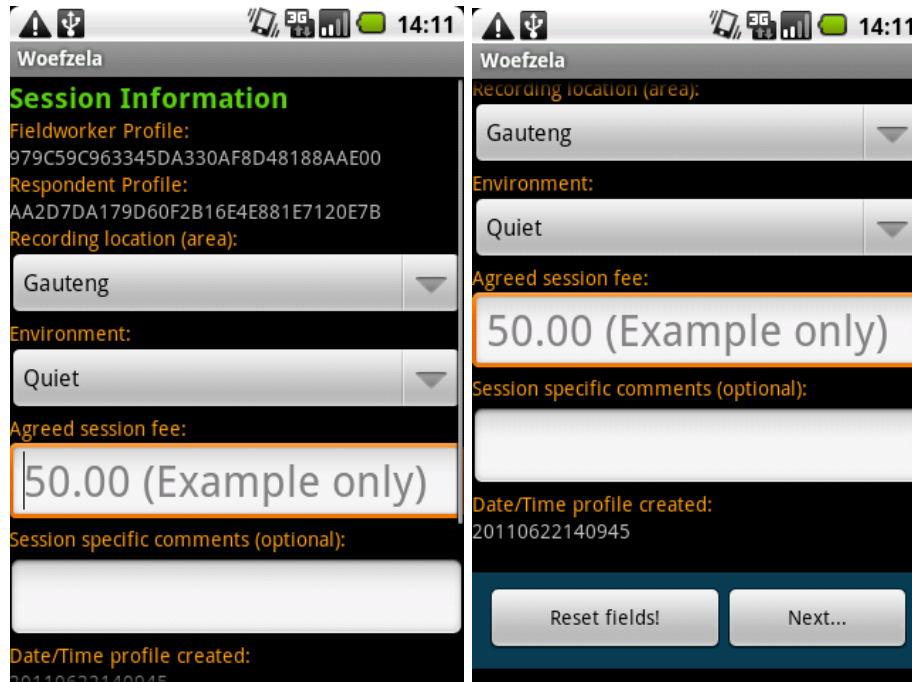
“Reset fields!”

This button is intended to clear all fields of the current input screen. This functionality is currently not implemented.

“Quit!/Next...”

This button changes it’s text based on the acceptance/non-acceptance of the Terms and Conditions as indicated by the check box. A warning message will be shown should this button be accidentally pressed in the “Quit!” state. Upon pressing “Next...”, the user proceeds to the the next input screen (Session Information).

Session Information



Session Information screen (overlap shown)

Note: The purpose of collecting information associated with each session is to annotate the database/information with meta-data relevant for each session.

Information (only) fields

Fieldworker Profile	This is a unique number associated with the current Fieldworker profile.
Respondent Profile	This is a unique number associated with the current Respondent profile.

Drop-down lists (Spinners)

Recording location (area) A pre-compiled list of geographic locations where recording of the current session is taking place.

Environment A pre-compiled list of environment types where recording of the current session is taking place.

Fields

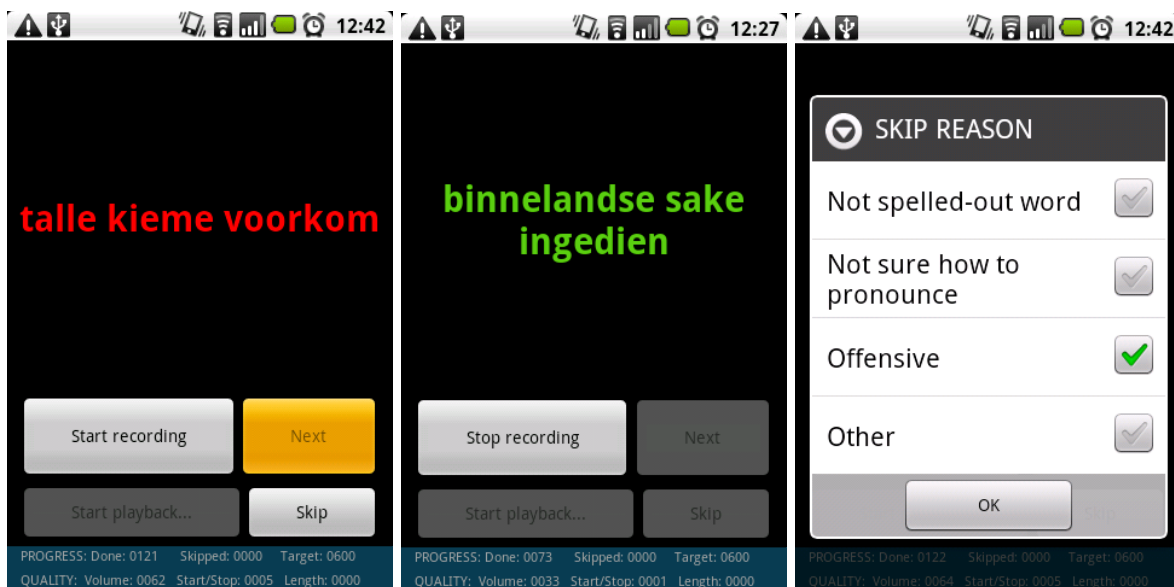
Agreed session fee If a Respondent will be refunded for his/her efforts, this field could be used to track such an agreement. This field may be left blank.

Session specific comments (optional) Any comments (such as testing scenarios) may be entered here. This field may also be left blank. This field will form part of the XML data associated with each recorded file.

General input comments

- If a specific input field does not meet the necessary criteria, an appropriate warning message will be displayed, starting from the first field not meeting its criteria.
- Long-press on a key to get extended characters.
- As a general rule, Fieldworker information has more stringent requirements (less optional fields) as they are often tied into some basic contract.
- If not all fields on the screen are immediately visible, swiping the screen upwards will scroll the display.

Main recording screen



Main recording activity screen

Prompt area

This is the area where the prompt string is displayed. Depending on the length of the string, the text may wrap to the next line.

Color significance: Red and Green

It is important to note that the text is normally red and only when the text is green, may recording proceed. This is to allow time for the recording equipment to prepare between each prompt.

Action buttons

“Start recording/Stop recording/Start re-recording”

Use this button to initiate or terminate recording or to re-record a prompt if the user is dissatisfied with the current recording.

“Start playback.../Stop playing”

This button is used to listen to the current recording prior to moving to the next prompt or opting to re-record the current prompt. Playback may be stopped while in progress.

“Next”

The “Next” button advances to the following prompt in the queue, *implying that the user was satisfied with the recording of the current prompt.*

“Skip”

This button will cause a prompt to be skipped either prior to recording the audio or after recording the audio. The meta-data associated with this prompt is then annotated with the reason provided from the pre-compiled list.

Statistics area

The bottom part of the recording screen (in blue) indicates some basic statistics associated with the current session. A brief description of each item follows:

Progress line

e.g. PROGRESS: Done: 0010 Skipped: 00001 Target: 0015

Done: This value indicates the number of prompts that has been recorded and that has passed the on-line quality checks. i.e. it indicates only the number of “good” recordings and thus this value is mutually exclusive from all other values in the statistics area (as they are ‘negative’ values). In the example, 10 audio files are deemed of sufficient quality. NOTE that the results of the quality checks generally lags the recordings by a number of prompts depending on the processing power available on the device.

Skipped: The number of prompts that has been skipped by the user. This value is incremented if the skip button was selected regardless of whether any audio of the associated prompt has been recorded or whether an actual skip reason has been selected or not. In the example, one prompt has been skipped.

Target: This is the target number of (good) prompts to be reached for this session. Note that this value may be (and usually is) different for the training and recording phases of a session. Typical (pre-compiled) values for training sessions are 15 and for recording sessions are 600. In the example, a target of 15 recordings of sufficient quality (refer to the “Done” field) is being pursued.

Quality line

e.g. QUALITY: Volume: 0002 Start/Stop: 0003 Length: 0000

Volume: This value indicates the number of audio files that either was not of sufficient volume or the volume was too loud. i.e. the “number of volume errors” per session. In the

example, two recordings were of insufficient volume by being either too loud or too soft.

Start/Stop: This value shows the number of recordings (audio files) which were either started too late or stopped too early. This decision is derived from the amount of energy present at the start or end of the audio file inferring from too much energy present that recording started too late or ended too soon. Please note that a large amount of energy present in background noise could cause this threshold to be reached since high levels of background noise is generally not desired in resources required to develop ASR systems. i.e. This value could be seen as the “number of times a recording was started too late or stopped too early.” In the example, three recordings were deemed to either have been started too early or stopped too late due to the amount of energy present at the start and/or end of the audio files.

Please view this error in conjunction with the text changing from red to green and green to red to indicate that the actual recording hardware is ready and not that the user is ready to make a recording.

Length: This field should indicate the number of prompts that fell outside the tolerance limits of the expected utterance duration. The tolerance will be based on the utterance category. This functionality is currently not implemented and thus this value should remain at zero. In the example, no files fell outside the tolerance values of the utterance length (as this metric has not been fully implemented).

General application navigation

A few general tips for navigating the application:

- The 'BACK' key's function is set to provide a means of quitting the application. A warning message will be raised to ensure that the user does not accidentally press this button. Since specific information is required on each screen, the user will not be allowed to navigate back to the previous screen (the *normal* functioning of this button), but will have to terminate and re-initiate the program should any corrections be necessary. Also for this reason, a user's previously entered profile may be loaded from the SD card and does not have to be re-entered.
- The 'HOME' key maintains its usual Android functionality allowing the user to switch directly to the home screen. TIP: If the home key is accidentally pressed during the running of the application, it can simply be LONG-PRESSED (kept in for about one second) to reveal a menu from which the currently running instance can be recalled. However, pressing the HOME key should in general be avoided during the running of this application.

Retrieving data

Data retrieval process

All data, meta-data and profiles are stored on the device's SD card.

Data may be retrieved in one of two ways:

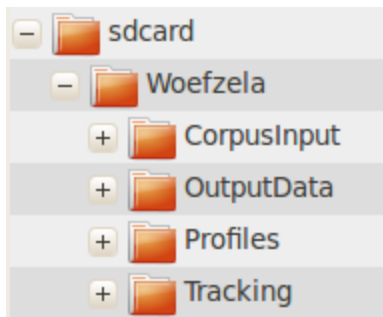
1. Unmount and remove the SD card from the device, place it in a SD card reader, and copy the Woefzela folder from the SD card using standard software on the PC.
2. Connect the device (with the SD card still mounted inside it) to a PC via a USB cable. Obtain the Android Debug Bridge program (<http://developer.android.com/guide/developing/tools/adb.html>) and 'pull' the data from the SD card with a command similar to:

a. `adb pull /sdcard/Woefzela Woefzela`

Folder structure of retrieved data

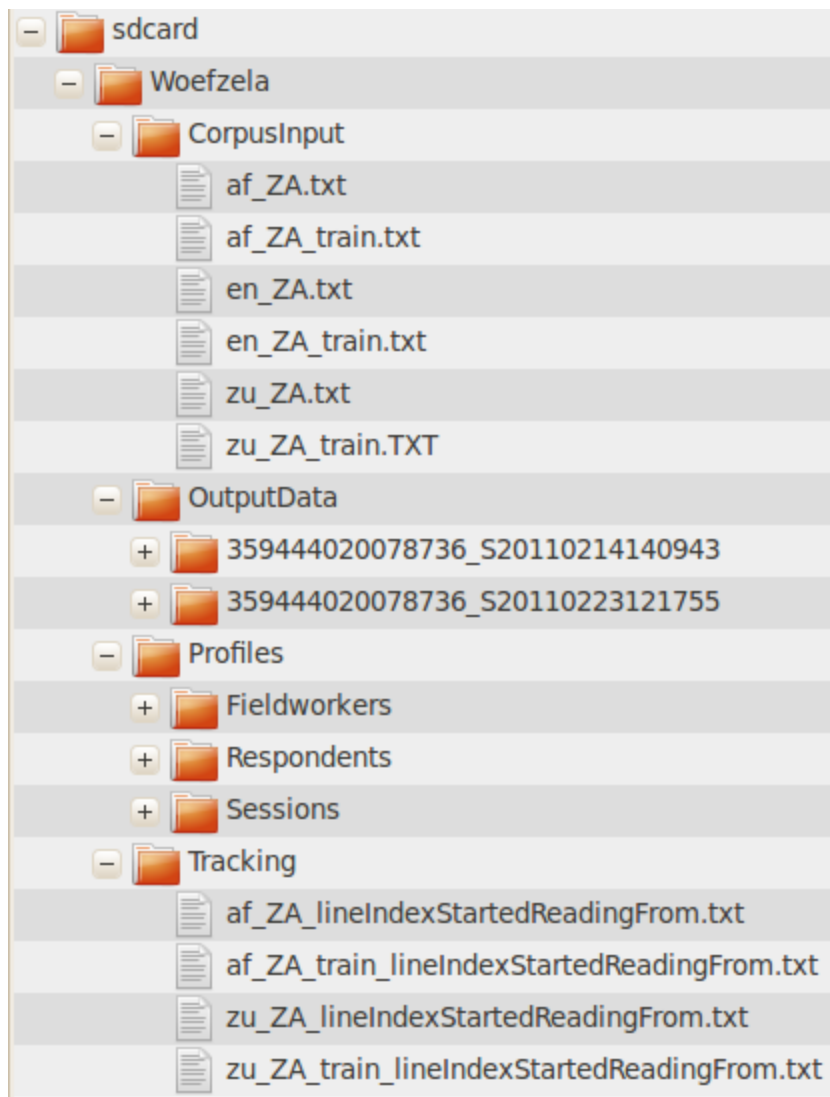
An example directory structure is shown below:

Top level tree:

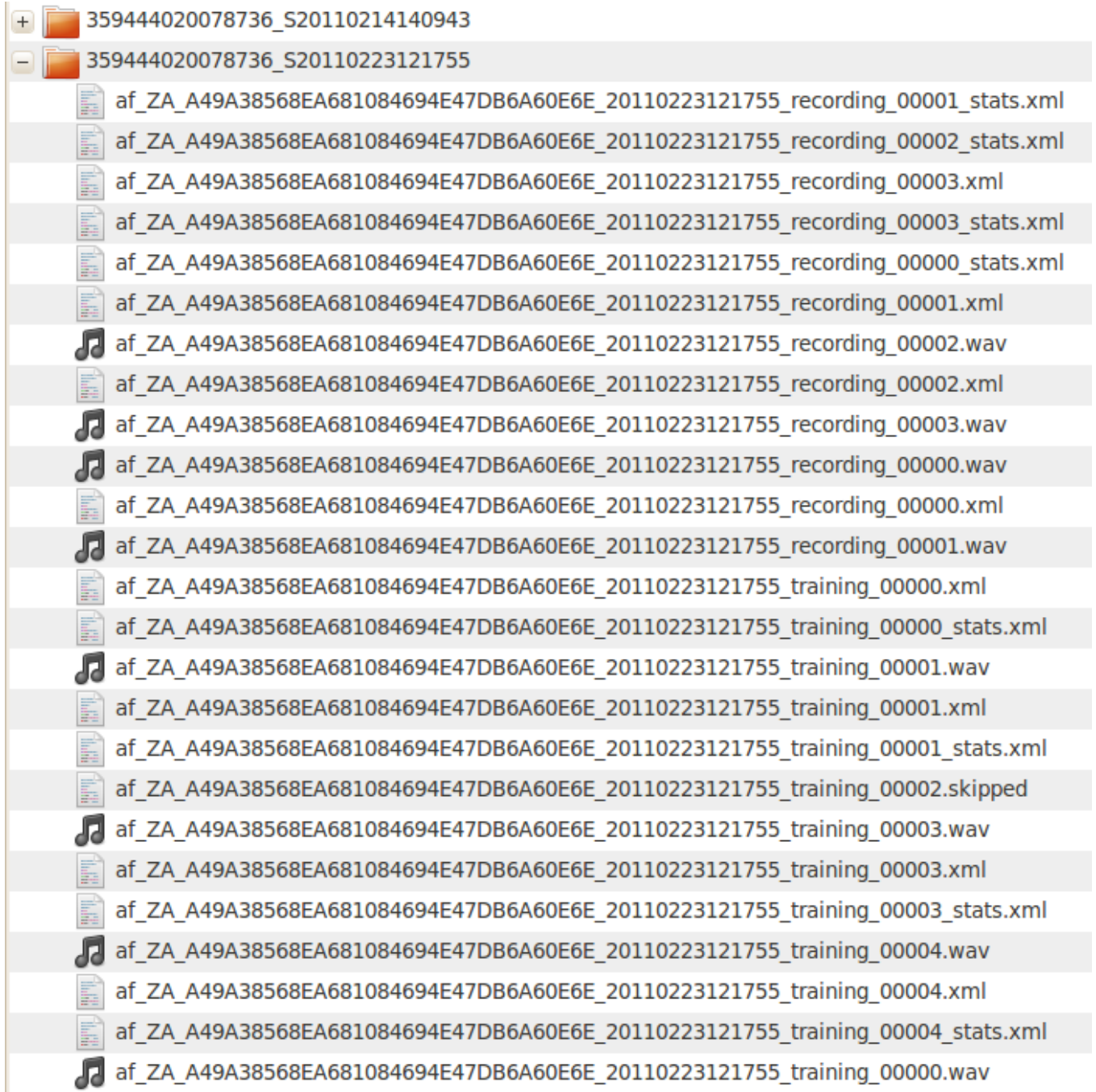


Next level tree:

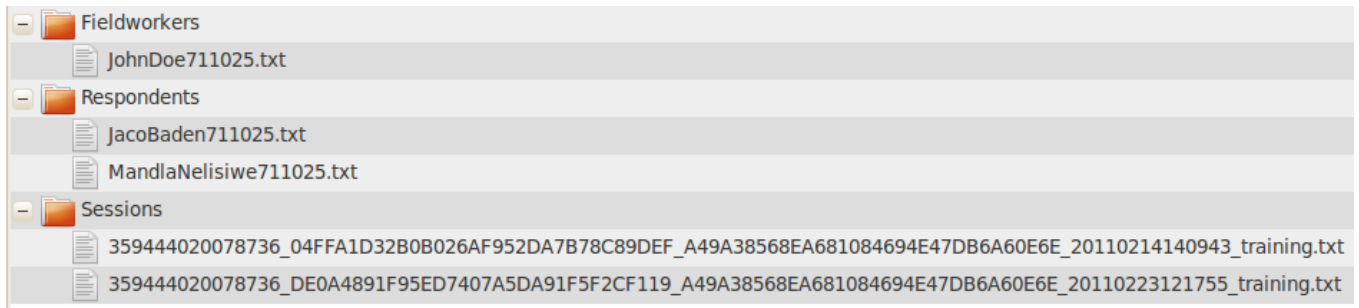
(one-level expanded top level tree)



Output data tree:



Profiles tree:



Basic file types

Folder: /sdcard/Woefzela/OutputData

Filename	Contents
X.wav	Audio file (16-bit, mono, PCM)
X.xml	Transcription/prompt text of audio file
X_stats.xml	Results of statistical tests on audio file
X.skipped	Prompt text and reasons for skipped file

Folder: /sdcard/Woefzela/Profiles

Fieldworkers: Personal information associated with each Fieldworker.
 Respondents: Personal information associated with each Respondent.
 Sessions: Key information associated with each recording session linked to both the Fieldworker and the Respondent profiles via their unique profile hash keys.

Folder: /sdcard/Woefzela/Tracking

This folder contains certain tracking information regarding each corpus.

Note that this information may be aggregated across sessions and devices, but does not have to be maintained/transferred between SD cards.

XML Metadata

A typical metadata file associated with each recording is shown in the image below.

```
<?xml version="1.0" encoding="UTF-8"?>
<Woefzela>
  <prompt>local nomadic tribes</prompt>
  <accent>en_ZA</accent>
  <age>28</age>
  <gender>Female</gender>
  <environment>Quiet</environment>
  <comments>Stellenbosch</comments>
</Woefzela>
```

XML metadata file for each recording

Clearing the SD card for further recordings

In order for the SD card not to become too full, follow this method:

1. Retrieve all the data from /sdcard/Woefzela and ascertain success of retrieval i.e. no data loss.
2. Delete the complete Woefzela folder on the SD card.
3. Create a new folder structure Woefzela/CorpusInput on the card.
4. Place the necessary corpora (training and recording) for the desired language in the CorpusInput folder on the SD card.
5. Launch the application on the mobile device and start using it.

Additional settings

Recording target

To change the recording target (only for recording sessions) without code recompilation, create a plain text file called recordingTarget.txt and place it on the device in the folder /sdcard/Woefzela/Tracking. The number inside this file will be used as the recording target.

Semi-real-time QC

To turn all QC functionality off:

1. Create a file called advancedQCOFF.txt in the folder /sdcard/Woefzela/Tracking
2. Place the numeric value of 1 on the first line of this file

Notes:

- If no such file exists: QC is ON (wanted default behaviour to have QC ON)
- A value of 1 in file means QC is OFF (i.e. 1=TRUE)
- For any other value in this file (e.g. 0, -1, 2, or even the word OFF, etc), QC is ON

Thus to force QC to be OFF, create such a file in the correct folder, AND place the value of 1 on it's first line.

Troubleshooting

Problem: Program crashes without any reason.

Solution: Navigate to the following file and view the error log information with the appropriate date/time stamp: /sdcard/Woefzela/Tracking/WoefzelaErrorLOG.txt

Problem: Playback volume too soft

Solution: Push up Media Volume using appropriate buttons *during* playback.

Appendices

ISO language codes

afr	af_ZA	Afrikaans
eng	en_ZA	English (South African English)
nbl	nd_ZA	isiNdebele
nso	ns_ZA	Northern Sotho = Sepedi
ssw	ss_ZA	SiSwati
sot	st_ZA	Southern Sotho = Sesotho
tsn	tn_ZA	Setswana
tso	ts_ZA	Xitsonga
ven	ve_ZA	Tshivenda
xho	xh_ZA	Xhosa
zul	zu_ZA	Zulu