COLLABORATIVE TRANSCRIPTION IN AUSTRALIAN ABORIGINAL COMMUNITIES

Éric Le Ferrand*†, Mat Bettinson*, Steven Bird*, Laurent Besacier†

*Charles Darwin University †Université Grenoble Alpes

Abstract

We present the deployment of an interactive transcription pipeline in two Australian Aboriginal communities using Spoken Term Detection and human-in-the-loop. We have designed an activity and a mobile app meant to be used by the people to confirm the sparse transcription [2] automatically generated.

Word spotting for transcription

In order to support speech transcription in aboriginal communities, we set up a semi-automatic workflow in which the transcription is sparsely done by a spoken term detection system and verified by native speakers in an iterative process. In [3] we have shown that we can transcribe up to 45% of the words from a small lexicon on spontaneous speech with an average precision of 22%
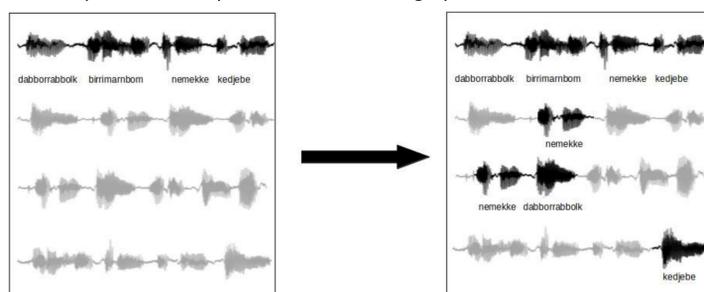

Fig. 1: Word spotting for sparse transcription

Application for word confirmation

Based on [1], we have designed a word confirmation app which takes as input the words spotted by the spoken term detection system and display them to the speakers. They would have the possibility to listen to the query term, the utterance and confirm if the words have been correctly retrieved
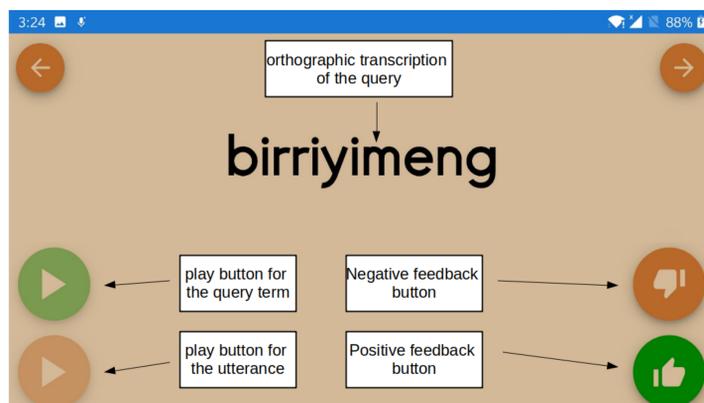

Fig. 2: Word confirmation app

Deployment

Two Aboriginal communities:

- Gunbalanya a town of $\sim$ 1100 people
- Manmoyi a remote community of $\sim$ 50 people

Two Deployments:

- First: Words extracted from a speech collection A and spotted in the same speech collection
- Second: Words extracted from a speech collection B and spotted in the speech collection C, D and E

Introduction of the task:

- Explanation: "Can you tell us if the first word you hear is in the sentence"
- Demonstration: Only for some participant, we show how to use the app for a few examples

Results

Speaker	Age	Location	App	Activity	Demo	Behavior
1	27	Gunbalanya	1	1	yes	T
2	47	Gunbalanya	0	1	no	T
3	35	Gunbalanya	3	1	no	T, R, C
4	$\sim$50	Manmoyi	0	2	no	C
5	$\sim$50	Manmoyi	2	2	yes	T, R, O, E
6	$\sim$60	Manmoyi	0	2	yes	E

Fig. 3: Deployment description

Fig 3 describes the results of the deployment.
App refers to the user confidence with the app:

- 0 : the user didn't use the app
- 1 : the user used the app but was not confident
- 2 : the user used the app but was only able to use the play buttons
- 3 : the user used all the functionality of the app

Activity refers to the configuration of the of the deployment described in the previous section: First (1) or Second (2).
Behavior refers to the feedback that the people were providing:

- Translating (T): Translation or explanation in English of the query/utterance
- Respeaking (R): Careful respeaking of the query/utterance
- Confirming (C): Completion of the word confirmation task
- Explaining (E): Clarification in language of the meaning of the query/utterance
- Other (O) : Not classified behavior

Discussion

Fig. 4: Deployment of the word verification app

Activity design:

- The feedback required was not natural for them to provide
- The activity lacked concreteness
- The task was disconnected to the idea of transcription

Application design:

- The purpose of the feedback buttons was unclear
- The query and utterance play buttons were hard to differentiate

Engagement design:

- The application was redundant
- The activity didn't bring anything useful to the community
- The activity put speakers in an unfamiliar environment

Acknowledgements

References

[1] Mat Bettinson and Steven Bird. "Developing a suite of mobile applications for collaborative language documentation". In: *Proceedings of the 2nd Workshop on the Use of Computational Methods in the Study of Endangered Languages*. 2017, pp. 156–164.

[2] Steven Bird. "Sparse transcription". In: *Computational Linguistics* 46 (2020), pp. 1–32.

[3] Eric Le Ferrand, Steven Bird, and Laurent Besacier. "Enabling Interactive Transcription in an Indigenous Communit". In: *COLING 2020 (short paper)*. 2020.