# ASR in Classroom Today: Automatic Visualization of Conceptual Network in Science Classrooms

Daniela Caballero[1(✉)], Roberto Araya[1], Hanna Kronholm[2], Jouni Viiri[2], André Mansikkaniemi[3], Sami Lehesvuori[2], Tuomas Virtanen[4], and Mikko Kurimo[3]

[1] CIAE, Universidad de Chile, Periodista José Carrasco Tapia 75, Santiago, Chile
daniela.caballero@ciae.uchile.cl
[2] University of Jyväskyla, Alvar Aallon katu 9, 40014 Jyväskylä, Finland
[3] Aalto University, 00076 Aalto, Finland
[4] Tampere University of Technology, Korkeakoulunkatu 10, 33101 Tampere, Finland

**Abstract.** Automatic Speech Recognition (ASR) field has improved substantially in the last years. We are in a point never saw before, where we can apply such algorithms in non-ideal conditions such as real classrooms. In these scenarios it is still not possible to reach perfect recognition rates, however we can already take advantage of these improvements. This paper shows preliminary results using ASR in Chilean and Finnish middle and high school to automatically provide teachers a visualization of the structure of concepts present in their discourse in science classrooms. These visualizations are conceptual networks that relate key concepts used by the teacher. This is an interesting tool that gives feedback to the teacher about his/her pedagogical practice in classes. The result of initial comparisons shows great similarity between conceptual networks generated in a manual way with those generated automatically.

**Keywords:** Automatic Speech Recognition · Conceptual network · Classroom dialogue · Teacher discourse

## 1 Introduction

A single teacher can teach hundreds of hours of classes per year. In most of the times he/she doesn't get automatic and quick feedback regarding his/her class. Moreover, the ubiquity of smartphones makes them an easy to find and economic tool to collect data, like teachers' class audio. In previous work the structure of the classroom speech was analyzed [1], considering that the content of the lesson has an impact on the conceptual structure of students, and specifically, the connections of the lesson's content relate to students' learning and in the quality of the lesson [2].

Automatic Speech Recognition (ASR) has many applications such as automatically detection of teachers' questions [3], identify keywords to captioning systems [4], indicate difficulties in second language learners [5], and several others [6]. This paper aims to provide another application of ASR in learning by showing preliminary results using this technology in Chilean and Finnish middle school to automatically provide teachers a visualization of the structure of concepts present in their speech in physics classrooms,


É. Lavoué et al. (Eds.): EC-TEL 2017, LNCS 10474, pp. 541–544, 2017.
DOI: 10.1007/978-3-319-66610-5_58

which, as shown in previous work [1] could be related with students' learning gains, specifically measuring different concepts and the number of pairs of related concepts. The research question presented here is: what is the correspondence between automatic and manual concept network, in other words whether the ASR technology allows us to present a suitable feedback for teachers.

## 2 Method

### 2.1 Participants

Two teachers participated in our pilot study: one teacher from Chile and other from Finland. Both recorded themselves in a regular physics class. The classes were about cinematic in Chile ($11^{th}$ grade) and electricity for the Finnish classroom ($9^{th}$ grade).

### 2.2 Procedure and Equipment

During their classes the Chilean teacher wore a SmartLav+ microphone and the Finnish teacher wore two microphones, AKG C520 and DPA IMK-SC4060, and a dictator ZOOM H4N. Both teachers recorded a single lesson of 43 and 36 min for the Chilean and Finnish class respectively.

### 2.3 Automatic Speech Recognition Systems

To run the ASR experiments on Spanish, the Google Cloud Speech API [6] was used. It supports over 80 languages including Spanish from Latin America. It showed higher levels of recognition in Spanish than in Finnish in our early experiments. For that reason, ASR experiments on Finnish were run using the Kaldi toolkit [7]. The recognition system was based on time-delay neural networks (TDNNs) combined with long short-term memory (LSTM) layers.

## 3 Data Analysis

Each of the audio file is split into small slices of audio (5–10 s). A person transcribed the audio and also did, automatically, the ASR systems. With both transcriptions (manual and ASR) and a keyword list based on curricular directions, we build a connectivity matrix which relates the frequency of two pairs of concepts that appear together within a 10 s window (it is the same process shown in previous work [1]). Finally, with this matrix the conceptual network is automatically generated with R's library *igraph*.

## 4 Preliminary Results

For both classrooms, the ASR has shown to be quite trustful. The resume of the results are shown in Table 1, where the recognition rates for the keywords are calculated.

**Table 1.** Recognition rates for Chilean and Finnish classrooms.

| Classroom | Number of keywords | Number of total keywords appearance | Number of keywords recognized | Recognition rate |
|---|---|---|---|---|
| Chilean | 12 | 146 | 109 | 74.7% |
| Finnish | 17 | 324 | 212 | 65.4% |

The Figs. 1 and 2 shows the conceptual network for the Chilean and Finnish class. Each of the keyword is located in one vertex of a polygon. Two pairs of concepts are related (i.e. there is a line connecting them) if they were mentioned in a 10 s slot. The width of the line is related to the amount of time those two concepts were mentioned in the whole class. For instance, in both concept map of Fig. 2 *"distancia"* (distance) and *"tiempo"* (time) are the two concepts which were highly mentioned together.

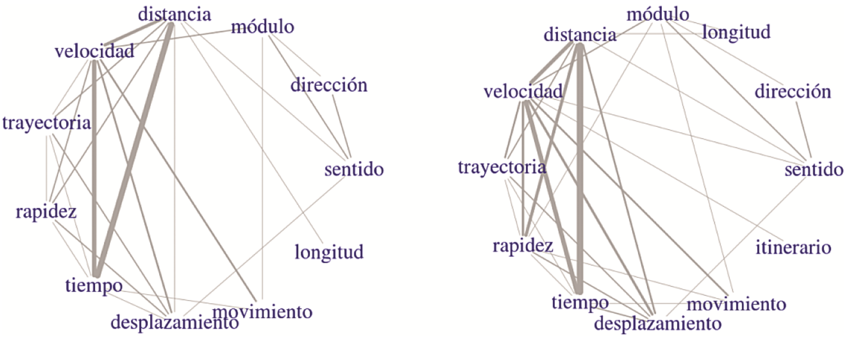


**Fig. 1.** Conceptual network from Automatic (left) and Manual (right) Transcription for Chilean class.

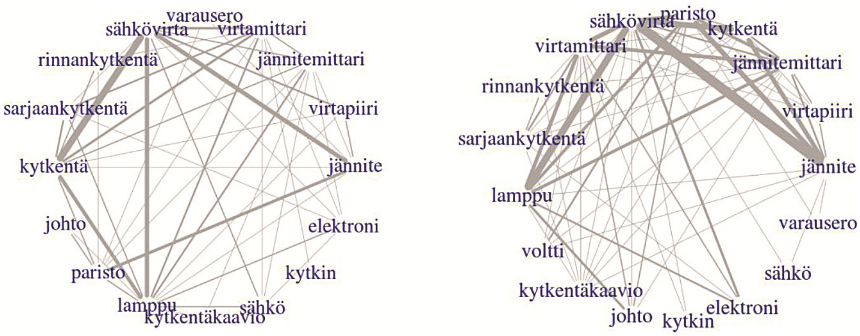


**Fig. 2.** Conceptual network from Automatic (left) and Manual (right) Transcription for Finnish class.

## 5 Conclusions and Future Work

The analysis and results shown in this paper could help teachers to have a better understanding of his/her class without much effort which, without the technology could be extremely high time consuming and expensive. We do not expect replacing other ways of feedback and training, but we think this information can enhance teaching analysis. We neither want to have a tool that evaluates the quality of teaching of teachers.

Regarding the future work, we still need to improve the automatic analysis and collect feedback from teachers to get new visualizations like names of the students for instance or appearance of content in each section of the class (start, middle and end). Finally, we have to propose metrics to compare the different networks as for instance in [1].

**Acknowledgements.** Funding from PIA-CONICYT Basal Funds for Centers of Excellence Project FB0003 is gratefully acknowledged and to the AKA-EDU-01 grant from CONICYT. The authors are also thankful for the funding provided to project No. 294218 by the Academy of Finland

## References

1. Helaakoski, J., Viiri, J.: Content and content structure of physics lessons and students' learning gains. In: Fischer, H.E., Labudde, P., Neumann, K., Viiri, J. (eds.) Quality of Instruction in Physics. Comparing Finland, Germany and Switzerland, pp. 93–110. Waxmann, Münster (2014)
2. Klieme, E., Pauli, C., Reusser, K.: The pythagoras study. Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In: Janík, T., Seidel, T. (eds.) The Power of Video Studies in Investigating Teaching and Learning in the Classroom, pp 137–160. Waxmann, Münster (2009)
3. Donnelly, P., Blanchard, N., Olney, A., Kelly, S., Nystrand, M., D'Mello, S.: Words matter: automatic detection of teacher questions in live classroom discourse using linguistics, acoustics, and context. In: Proceedings of the Seventh International Learning Analytics and Knowledge Conference, pp. 218–227. ACM (2017)
4. Ikeda, N., Takeuchi, Y., Matsumoto, T., Kudo, H., Ohnishi, N.: Support system for lecture captioning using keyword detection by automatic speech recognition. In: Miesenberger, K., Bühler, C., Penaz, P. (eds.) ICCHP 2016. LNCS, vol. 9759, pp. 377–383. Springer, Cham (2016). doi:10.1007/978-3-319-41267-2_53
5. Mirzaei, M.S., Meshgi, K., Kawahara, T.: Automatic speech recognition errors as a predictor of L2 listening difficulties. In: CL4LC 2016, pp. 92–201 (2016)
6. Shadiev, R., Hwang, W.-Y., Chen, N.-S., Huang, Y.-M.: Review of speech-to-text recognition technology for enhancing learning. J. Educ. Technol. Soc. **17**(4), 65–84 (2014)
7. Google: Google Cloud Speech API. https://cloud.google.com/speech/