Automated Speech Recognition: its impact on teaching and learning languages

Michael Carrier
Highdale Consulting

ICC Conference
ECML Graz, April 2017

Highdale Consulting

Contents

1-What is ASR?

2-How does it work?

3-How is it being used?

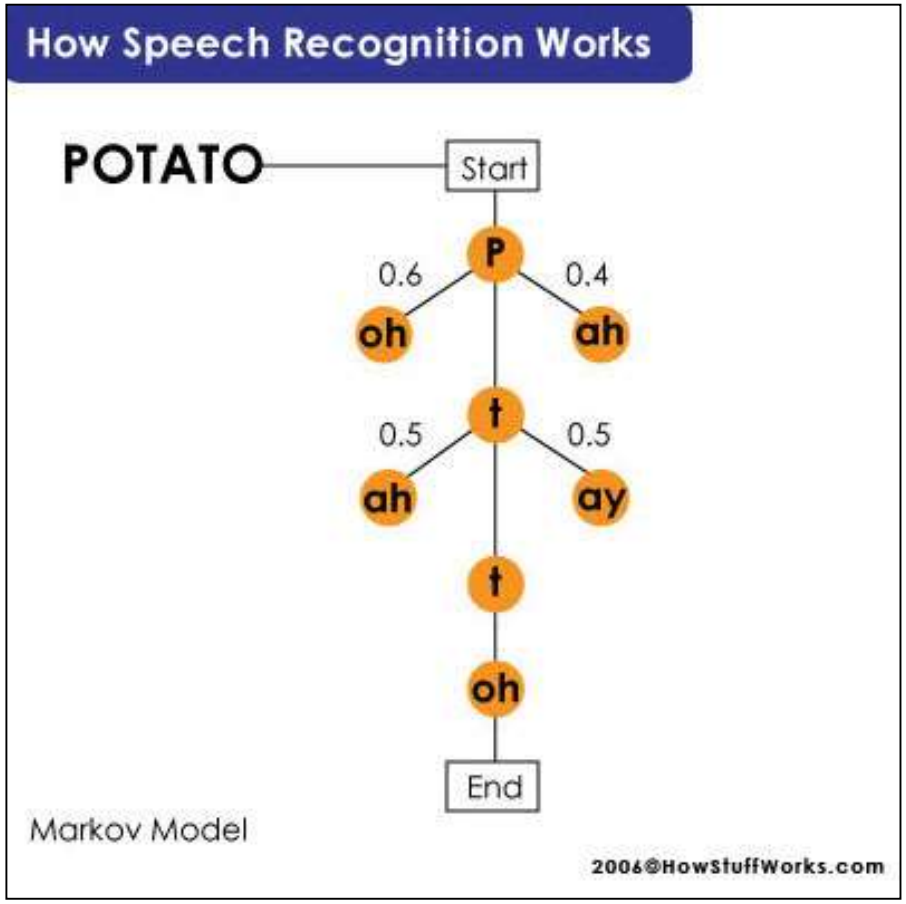
4-How can we use it in class?

5-ASR and Speech-to-Speech (Sp2Sp) translation

6-Using Sp2Sp in class

7-Automated marking of speech & writing

8-Future trends



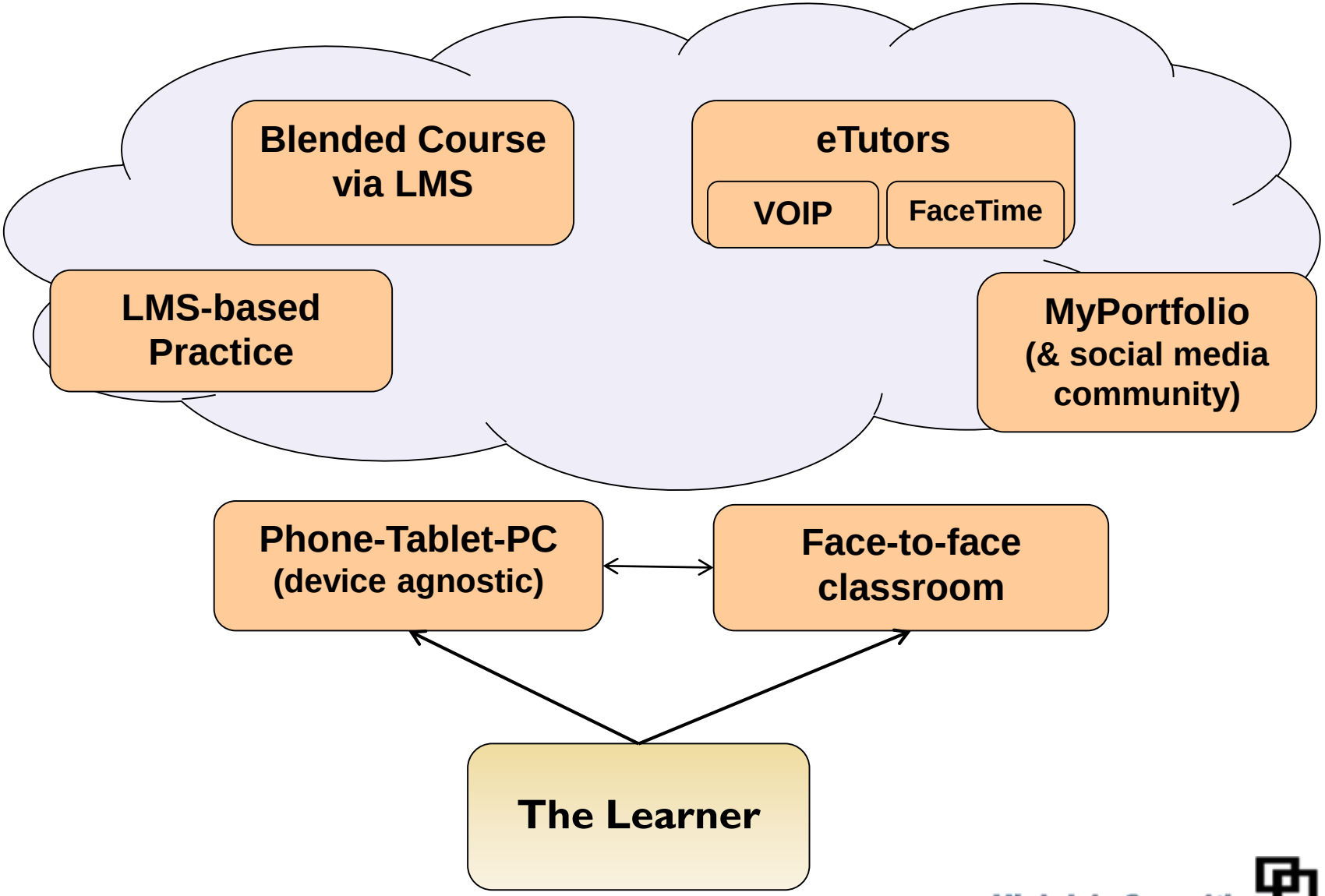
Highdale Consulting

….to MOOCS: Digital learning is here to stay

Advantages:

- Access
- Equity
- Ubiquity
- Autonomy
- Reach
- Sustainability

Digital Learning ecosystem



Highdale Consulting

Do you speak digital?

- BYOD
- VLE
- Flipped
- ASR
- IWB
- VR
- AR
- SMS
- Airplay
- Bluetooth

Reflection

Score yourself out of 10 for concept recognition

Digital speech

The rise of speech-enabled tech

Meet Pilot

The world's first smart earpiece which translates between users speaking different languages

Highdale Consulting

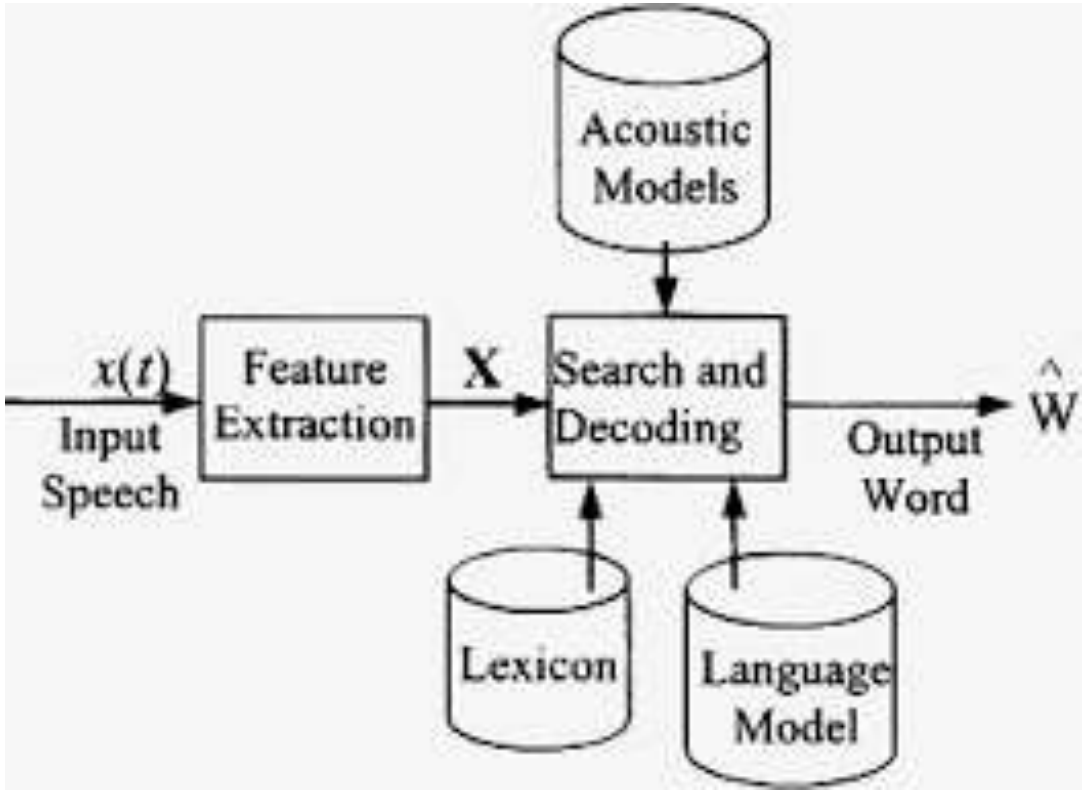
1 - What is Automated Speech Recognition (ASR)?

- **Automated Speech Recognition (ASR)** converts audio streams into text, but does not analyse it semantically
- The ASR output cannot assess meaning or coherence (yet)
- ASR is not the same as parsing – it doesn't apply grammatical rules
- ASR is based on big data – searching language corpora and finding matching patterns in data
- ASR is flawed but improving rapidly

Speech recognition, also referred to as **speech-to-text** or **voice recognition**, is a technology that recognizes speech, allowing voice to serve as the "main interface between the human and the computer".

2 - How does it work?

ASR process



Speech recognition engines need:

an acoustic model

which is created by taking audio recordings of speech and their transcriptions (taken from a speech corpus), and 'compiling' them into a statistical representations of the sounds that make up each word (through a process called 'training').

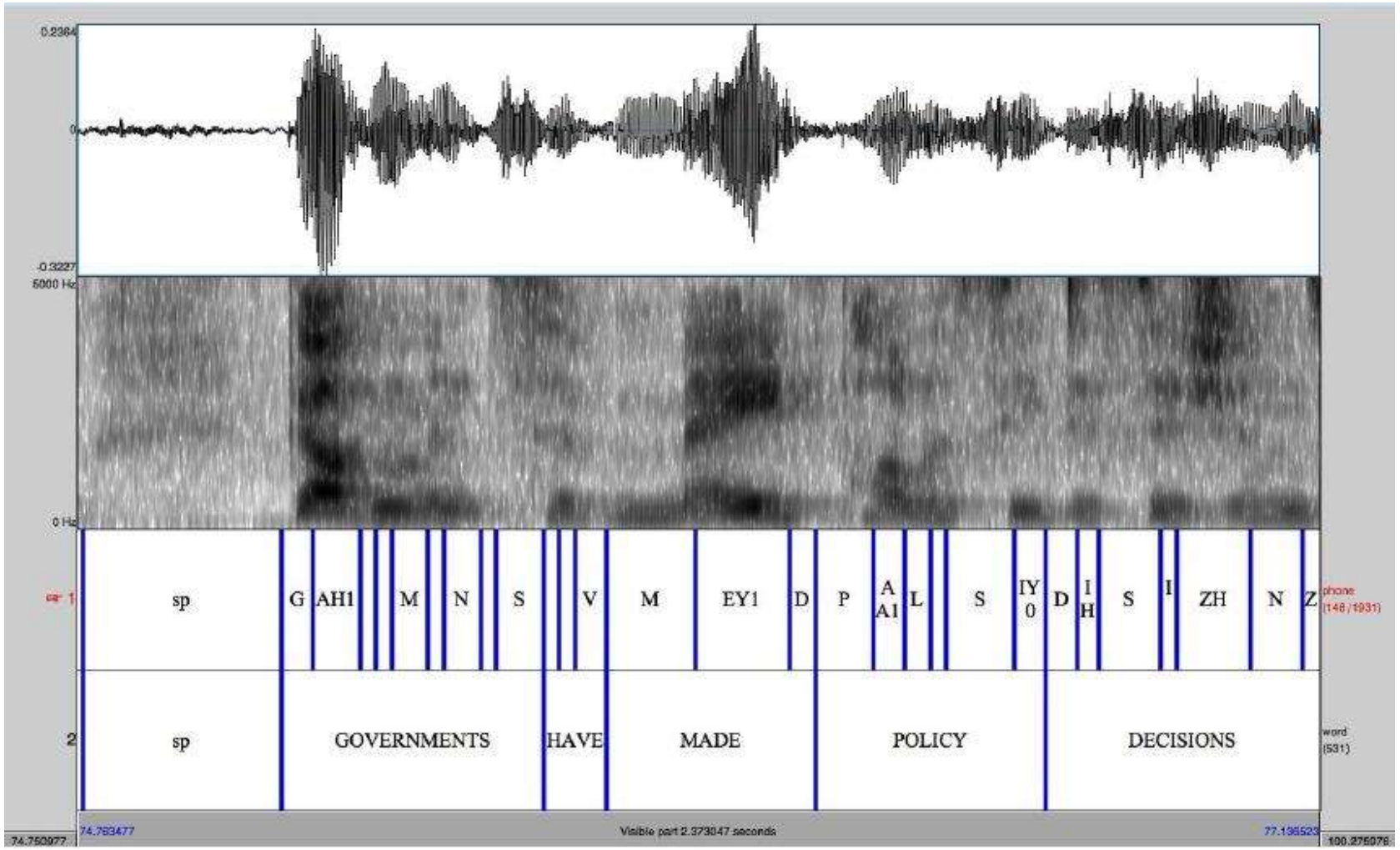
a language model

A language model is a file containing the probabilities of sequences of words.

a grammar

is a much smaller file containing sets of predefined combinations of words.

Aligning speech and text



Highdale Consulting

Key concepts….

- Hidden Markov models
- Neural networks
- Fourier transforms
- Vocabulary base
- Corpora
- Language modelling
- Context dependency

Accuracy criteria:

- Vocabulary size and confusability
- Speaker dependence vs. independence
- Isolated, discontinuous, or continuous speech
- Task and language constraints
- Read vs. spontaneous speech
- Adverse conditions

Part-of-speech tags used:

MD	modal auxiliary (can, should, will)
NC	cited word (hyphenated after regular tag)
NN	singular or mass noun
NN$	possessive singular noun
NNS	plural noun
NNS$	possessive plural noun
NP	proper noun or part of name phrase
NP$	possessive proper noun
NPS	plural proper noun
NPS$	possessive plural proper noun
NR	adverbial noun (home, today, west)
OD	ordinal numeral (first, 2nd)
PN	nominal pronoun (everybody, nothing)
PN$	possessive nominal pronoun
PP$	possessive personal pronoun (my, our)
PP$$	second (nominal) possessive pronoun (mine, ours)
PPL	singular reflexive/intensive personal pronoun (myself)
PPLS	plural reflexive/intensive personal pronoun (ourselves)
PPO	objective personal pronoun (me, him, it, them)

How Siri works…

1 - The sounds of your speech are encoded into a compact digital form

2 - The signal from your phone is relayed back to a server in the cloud.

3 - Simultaneously, your speech is evaluated locally, on your device. Siri decides whether it can handle the request locally -- eg you asked it to play a song -- or if it must connect to the network.

4 - The server compares your speech against a statistical model to estimate the **phonemes** spoken. The highest-probability estimates get the go-ahead.

5 - Your speech - now understood as a series of vowels and consonants - is run through a language model, which estimates the **words** in your utterance. The computer then creates a list of possible meanings for the sequence of words in your speech.

6 – As a result, the computer determines that your intention is clear – you want to send an SMS to Erica , her phone number should be pulled from your phone's contact list, and the rest of your speech is your message to her – and this text message appears on screen.

3 - How is it being used?

Applications of ASR

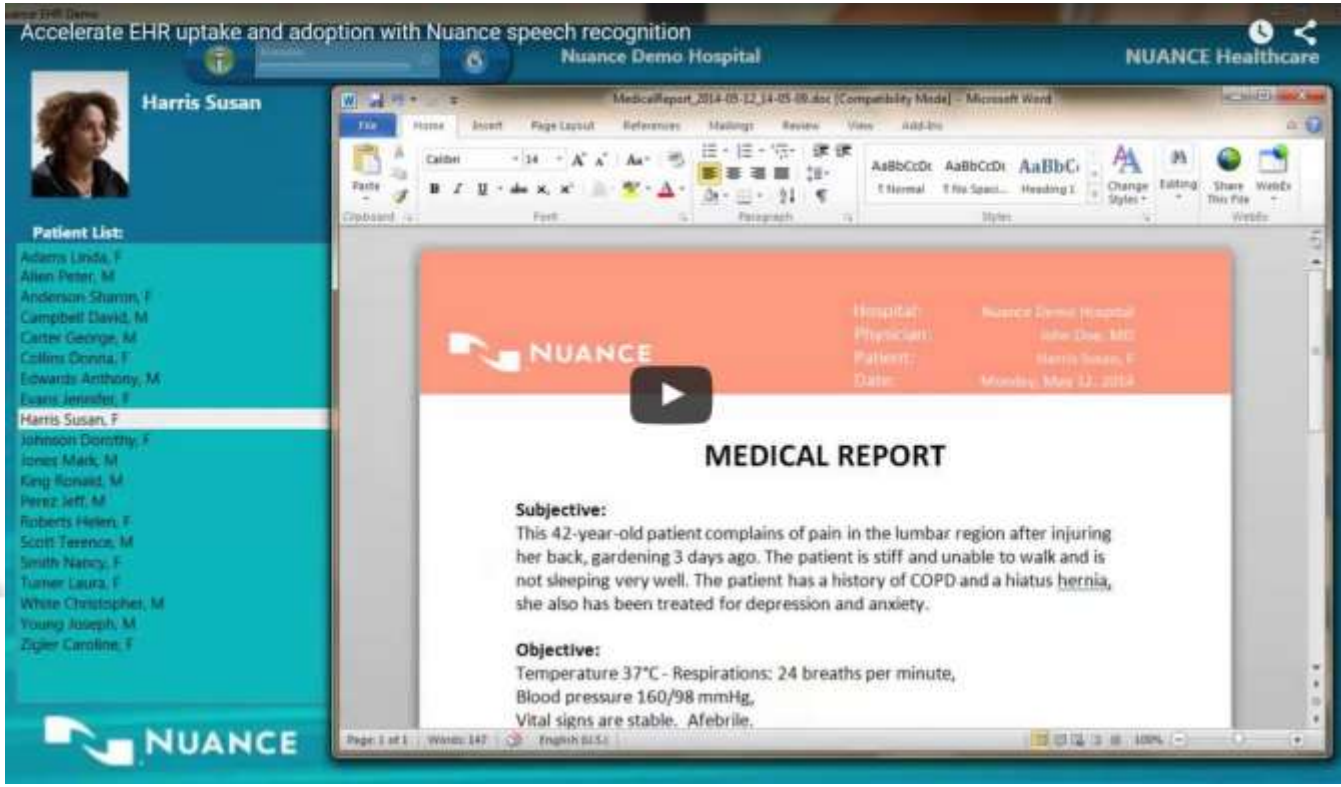
Activities:

- Dictation
- Voice search
- Pronunciation
- Exercises
- Translation
- Marking

Sectors:

- Telephony
- In-car systems
- Military
- Healthcare
- Education
- Disability support: vision-impairment, RSI, wheelchair control, dictation

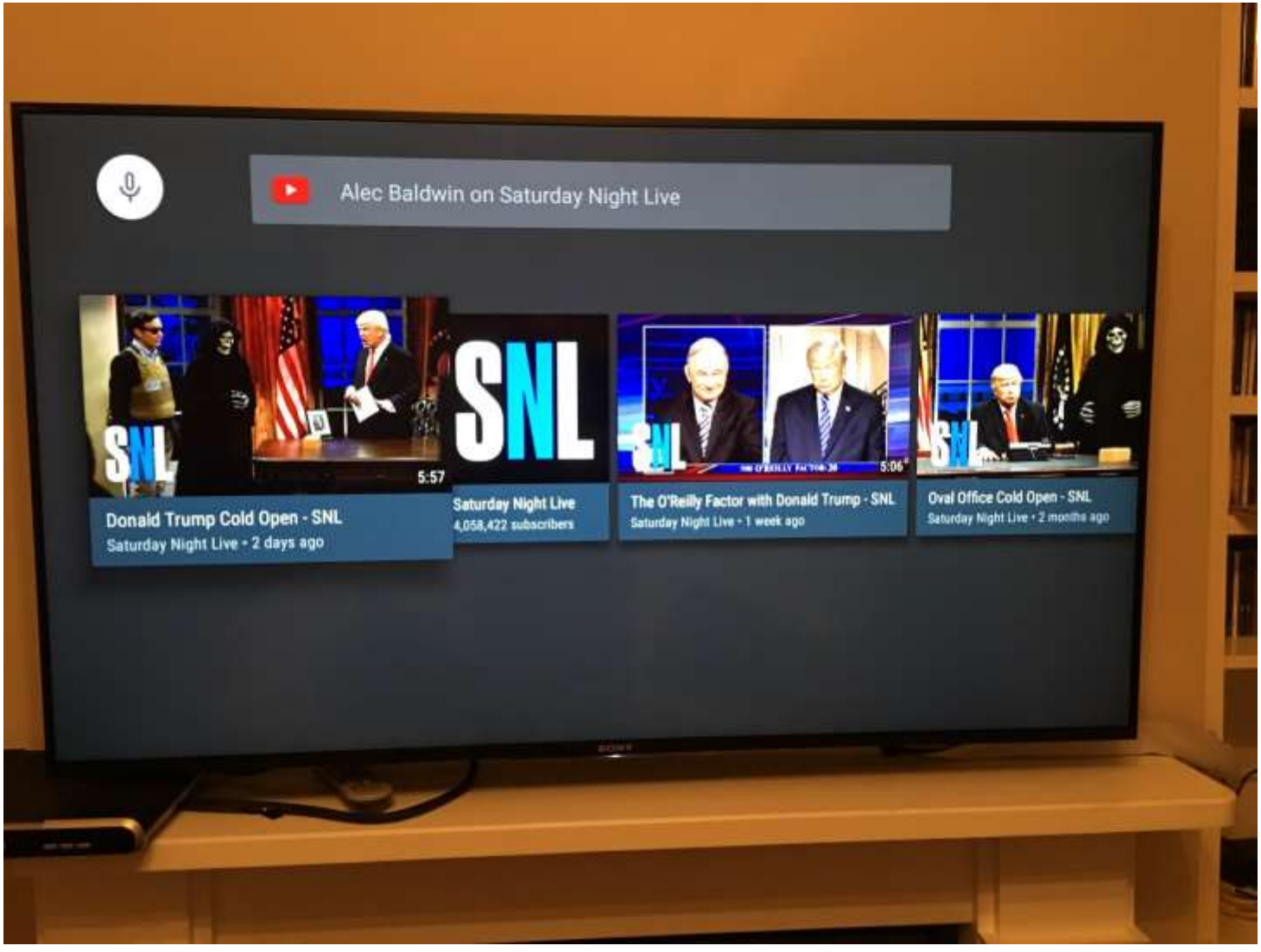
Speech to Text - Dragon Dictate



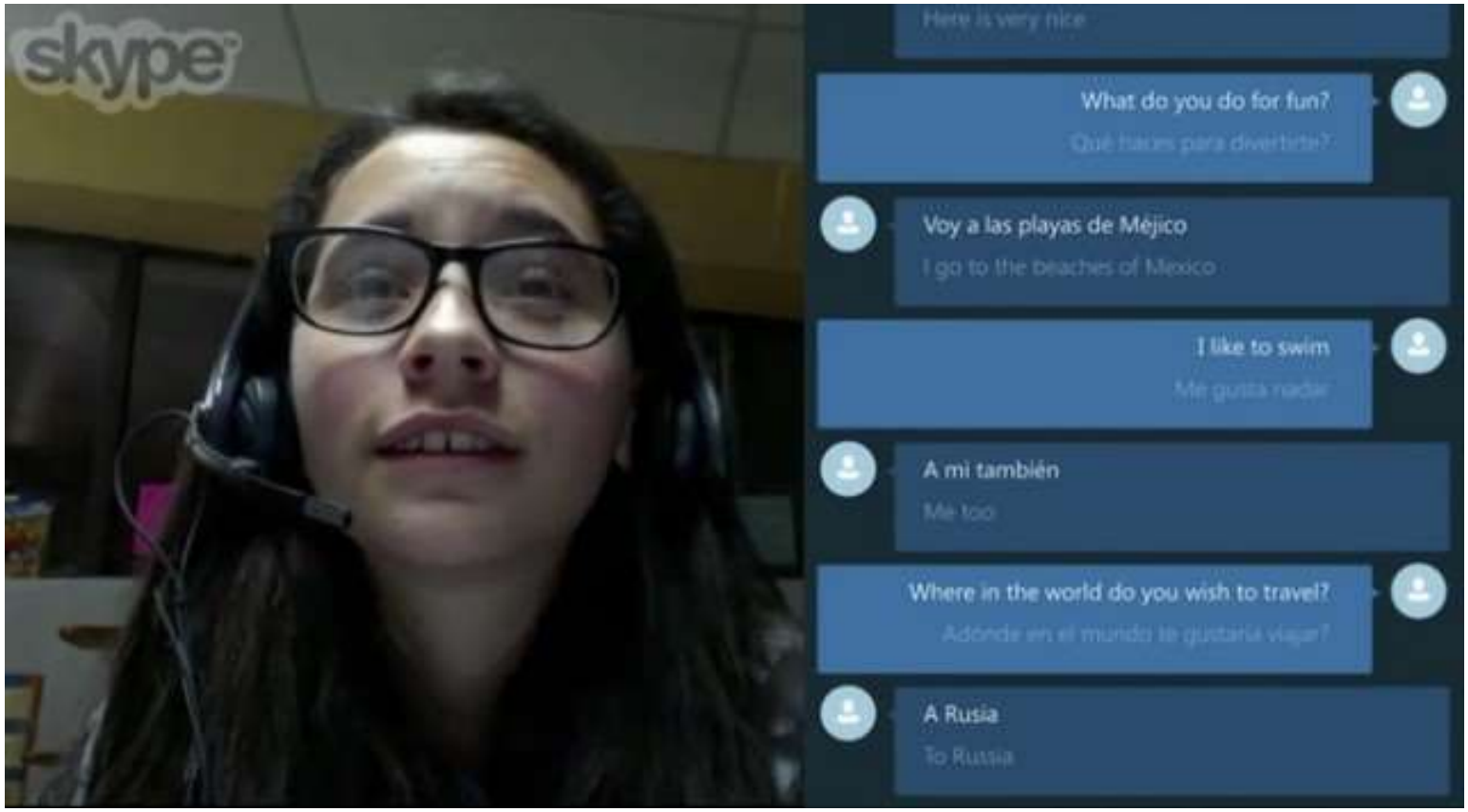
Calisto

Highdale Consulting

Speech control: Internet of Things



Speech-enabled translation (SET)



Highdale Consulting

Cortana

The personal assistant

Alexa, Siri & ASR personal assistants

Google Home

"Alexa, open Just Eat and ask for my last order."

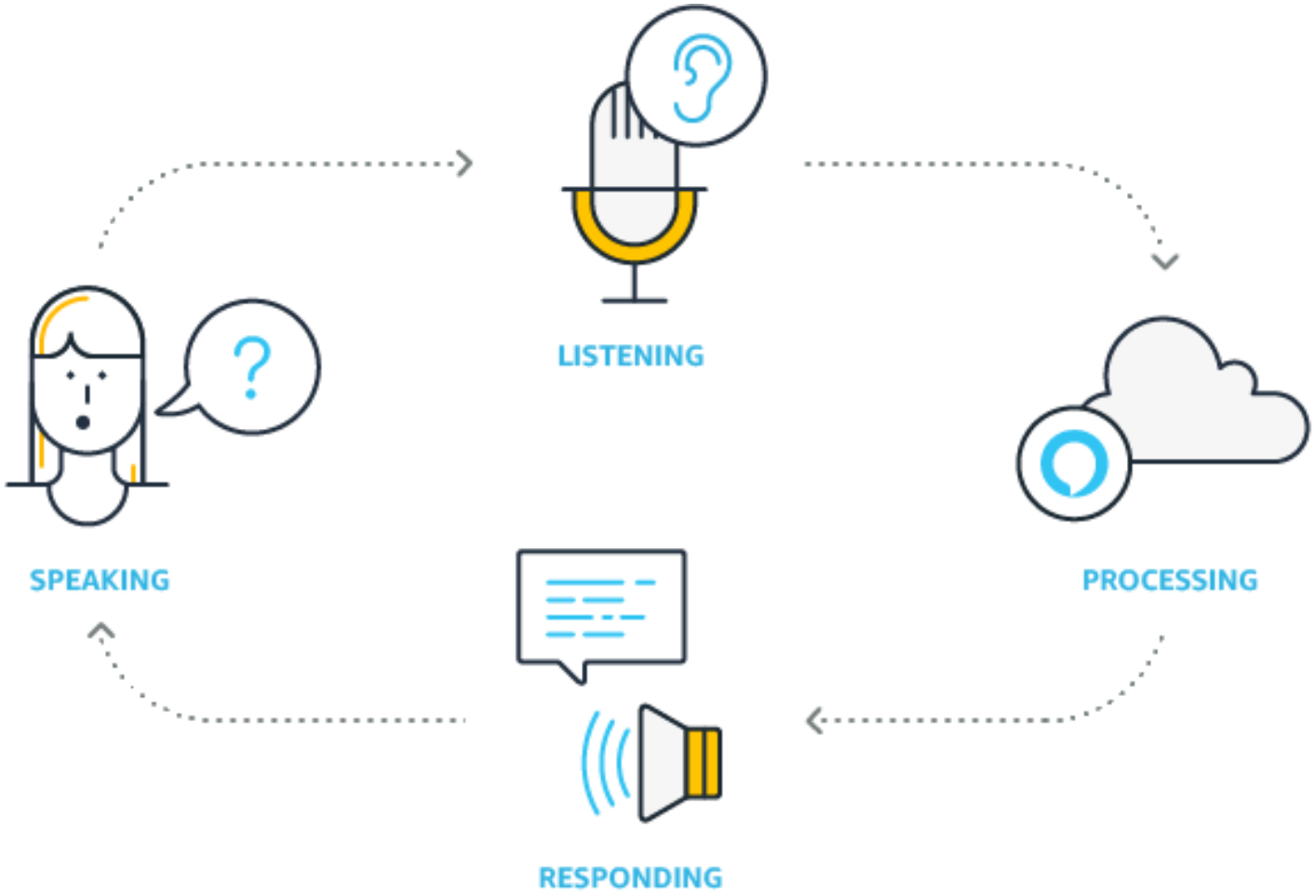
"Alexa, ask The Guardian for headlines."

"Alexa, ask Jamie Oliver for a recipe."

"Alexa, ask National Rail to check my commute."

Highdale Consulting

AVS – Amazon Voice Services



“Alexa….”

- What’s the news?
- Open the Telegraph
- Launch National Rail
- Play radio 4
- What’s the weather in X?
- Set a timer for 15 minutes
- What's on my calendar for tomorrow?
- Add 'go to the grocery store' to my to-do list
- What’s on my to-do list?
- How many degrees Fahrenheit in 22 degrees Celsius?

Amazon Echo & Echo Dot

Alexa will play music, provide information, deliver news and sports scores, tell you the weather, control your smarthome and even allow Prime members to order products they've ordered before.

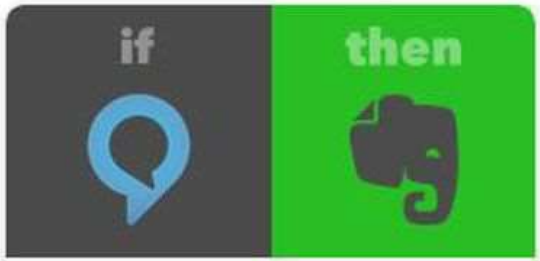
She updates through the cloud automatically and learns all the time.The more you use Echo, the more Alexa adapts to your speech patterns, vocabulary and personal preferences.

Add your new Alexa To Dos to Todoist.

by amazon_alexa 100 10

Email yourself your Shopping List when you ask Alexa's what's on your Shopping List.

by amazon_alexa 655 12



Add your Alexa To Dos to Evernote.

by amazon_alexa 314 6

Receive a notification when an item on your To Do List is edited.

by amazon_alexa 37 1

Receive an iOS notification when an item is added to your Alexa To Do list.

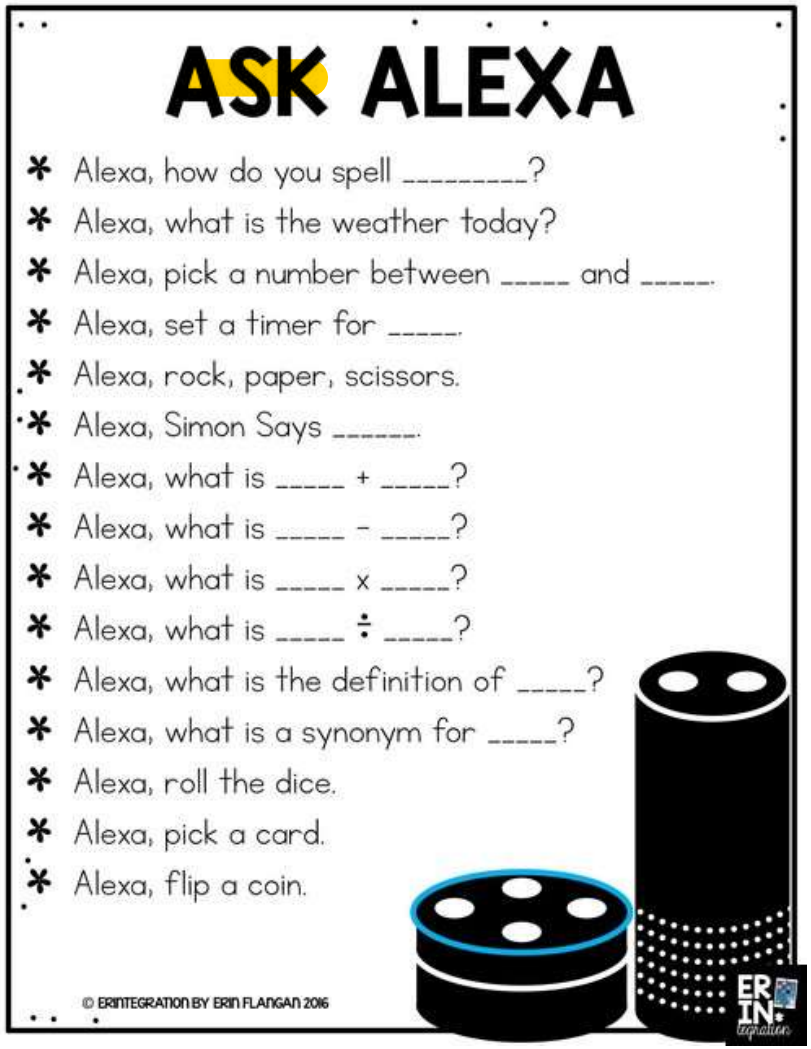
by amazon_alexa 59 1

Add task to Toodledo when item added to Alexa To Do List.

by amazon_alexa 19 0

Alexa in the classroom

Imagine a classroom where the teacher can work closely with a small group of students while others in the room are able to get their questions answered without interrupting the group. "Alexa does 7+7=14?" "Alexa does the word tomato have an e at the end?" "Alexa when was Abraham Lincoln born?" The Echo can be used for timers throughout the room. I can have the Echo read books for students as a modern-day listening station. Does a student with reading difficulties need to do research (making traditional Internet or book research difficult?) *https://www.donorschoose.org/project/extend-our-learning-with-an-amazon-echo/1591237/*



Students use devices in classrooms to access the internet to answer fairly simple questions. What if you could ask a device that sits in the classroom that question? What year was the Battle of Hastings? How long do you bake a Mushroom for? For example, asking 'How to cook a Mushroom?' doesn't get an answer. But ask 'How long should I cook a Mushroom for?' does get the desired answer.

I like the idea that facts are just a question away, and could see it being useful in lessons where skills are more important to learn than just straight facts.

Reflection

How would you use Alexa in the classroom?

How would you use Speech-enabled translation with your class?

Other ASR apps

Not just Siri & Cortana…

- Google Voice Search
- Google Voice Typing
- Vlingo
- Nuance's Dragon Go!
- True Knowledge's Evi voice assistant
- Samsung S Voice
- Android's Speaktoit

4 – ASR in ELT

Classroom Applications

"The learner interacts with English Tutor in short, real-life dialogs where the user controls the conversation flow…. English Tutor is able to provide **instant feedback** on the student's speaking performance"

Highdale Consulting

How can we use ASR in teaching?

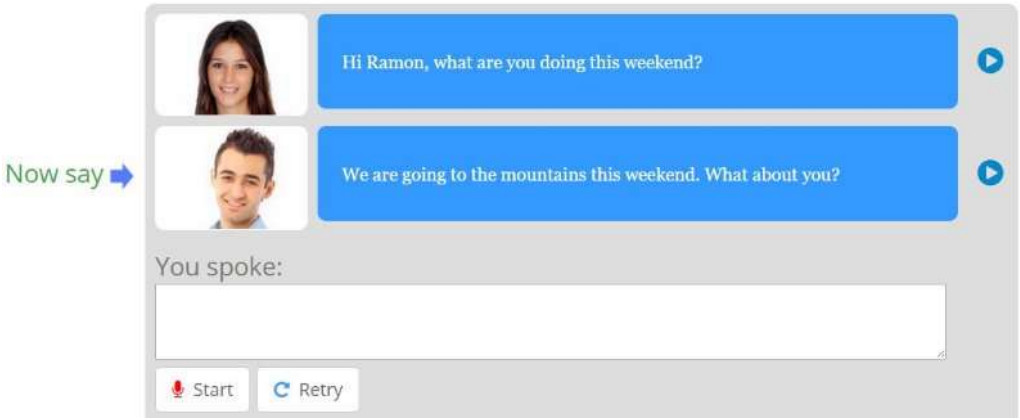
ASR has a **chequered history** in language education – many inadequate commercial products. But with higher quality we can revisit it for 4 uses:

1. ASR facilitates new ways to work on **phonology and accent**
2. ASR facilitates **responses to communicative interactions** in the classroom
3. ASR facilitates **automatic speech-to-speech translation**
4. ASR facilitates computer-based **automated marking** of ELT examinations

ASR in ELT

Computer-assisted pronunciation (CAPT)

Interactive practice



Speech-to-speech translation (SET)

CAPT – pronunciation

Carnegie Speech

- ‘’By comparing student voice recordings against a balanced and extensive range of English speakers, Carnegie Speech Assessment uses patented pinpointing technology to identify exactly what was correct and incorrect in the student's speech.’’
- ‘’ NativeAccent delivers pronunciation training, operational and financial returns by personalizing the English pronunciation training experience for each user. With pronunciation tutoring technologies pinpointing the exact error in English pronunciation, and demonstrating via text, audio and graphics exactly what must be done to correctly pronounce English words and phrases, training operations are streamlined while English pronunciation skills rapidly improve.’’

Duolingo

- This new feature checks how close your pronunciation is to a native speaker and then provides some basic feedback.

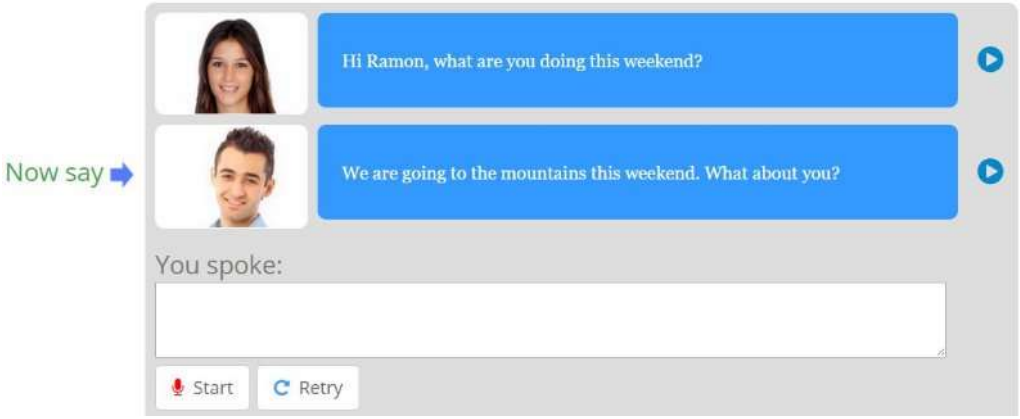
EduSpeak

- Recognizes adult and child voices
- Works for native and non-native speakers
- Human-calibrated pronunciation scoring

Interactive practice

Speaking Pal

- short interactions in which the learner is required to speak with a virtual character. ASR can highlight words which are spoken incorrectly



Duolingo

- Less sensitive ASR
- cf. Reviews on ELTJAM

Other providers:

- Sanako, SpeechAce,

  Dexway, Rosetta Stone etc

Results:
Works best if structured
Can be over-accepting
Caveat Emptor!

But…. better than no speaking practice at all, for some autonomous students

ASR in the classroom

Story tasks:

If students have the ASR app, they tell a story by dictating to device:

- One student has dictating role for user accuracy
- Group edits the resulting transcribed text and checks accuracy / appropriateness, correcting where necessary

Conversation tasks:

SS write a dialogue

Perform it as dictation

Correct written output

SS Initiate free conversation

Take in turns to dictate response to previous student

Check accuracy via converted text

ASR self-study

Solo speaking:

-Teacher gives text or dialogue to practice outside class

-Student practises dictating it – checking output matches the teacher model (listening to comparative audio if available)

Phonology:

Practise speaking and gaining feedback at pronunciation, stress, word level – for example:

- Reading Companion
- Carnegie Speech
- SpeakingPal
- EduSpeak

Writing:

-Use dictation app to give descriptions or tell stories orally

-Email results to teacher / peers

Carnegie Speech:

Phonology diagnostics - students practise at home, where they speak into the microphone and get feedback on pronunciation, stress & intonation performance

Automated Speech Recognition

Try it out in your classroom

Idea 1

Turn pronunciation work into a team game. Learners download and open Google Docs, then change the input mode by clicking the microphone icon on the on-screen keyboard. Teams then race to dictate ten sentences accurately into the document. The fastest team is the winner!

Idea 2

Speed up essay writing – or report writing. If working on a computer, learners open a new Google Doc and turn on 'Voice typing' (from the Tools menu). They'll need to clean up their writing later, but they'll get their ideas down more quickly.

Idea 3

Add a twist to webquests and other project work: learners can only search with their voices. They'll need to ask a digital assistant like Siri (Apple), Cortana (Microsoft) or Google Assistant … in English, of course.

Idea 4

Need to focus on a specific pronunciation area like uncle/ankle, think/sink, fit/feet? Get learners to practise dictating minimal pairs to their smartphones.

Idea 5

Learners will love interactive games where they control characters with only their voices. Search your app store for 'voice controlled game'.

CAMBRIDGE ENGLISH
Language Assessment
Part of the University of Cambridge

thedigitalteacher.com

Highdale Consulting

Reflection

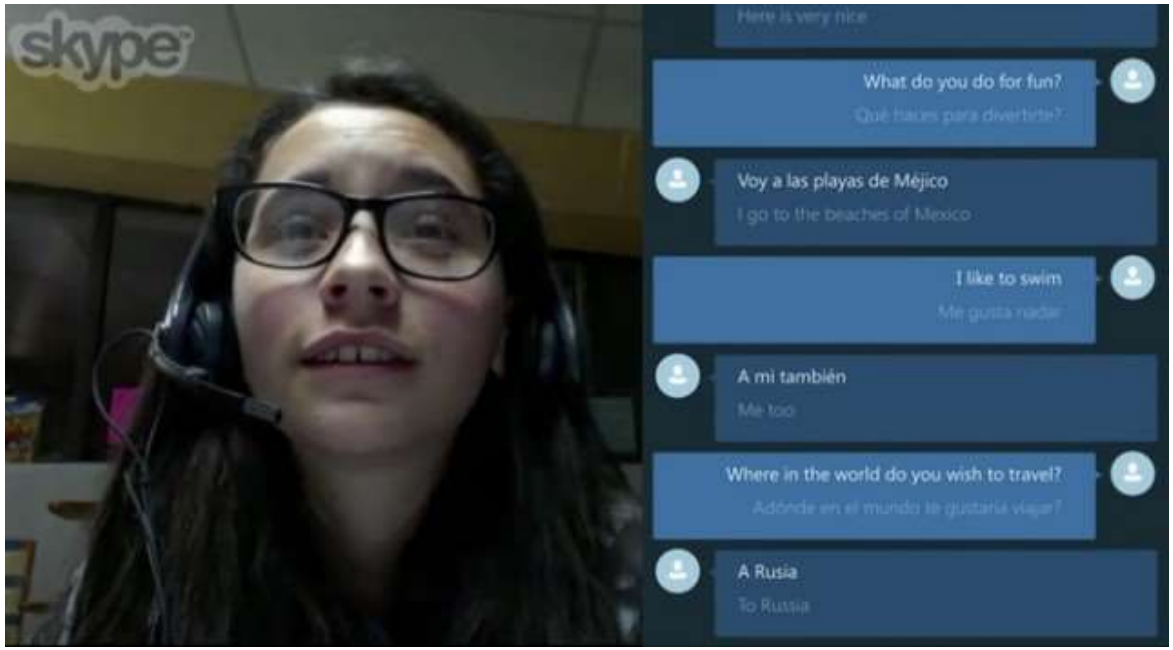
What might be the impact of this for teachers & learners in the classroom?

How would you use ASR with your class?

What would you need to make it possible/helpful?

5 - Speech-to-speech translation

Speech-enabled translation (SET)



Highdale Consulting

The Babel Fish

Meet Pilot

The world's first smart earpiece which translates between users speaking different languages

The Way It Works

PILOT EARPIECE

Using designed noise-canceling microphones, the Pilot earpiece filters out ambient noise from the speech of someone talking

PILOT APP

The translation process occurs using speech recognition, machine translation and machine learning as well as speech synthesis technologies

PILOT EARPIECE

The second earpiece returns the translation. This all happens simultaneously without interruption, as each person speaks to one another

Highdale Consulting

ASR and Sp2Sp translation

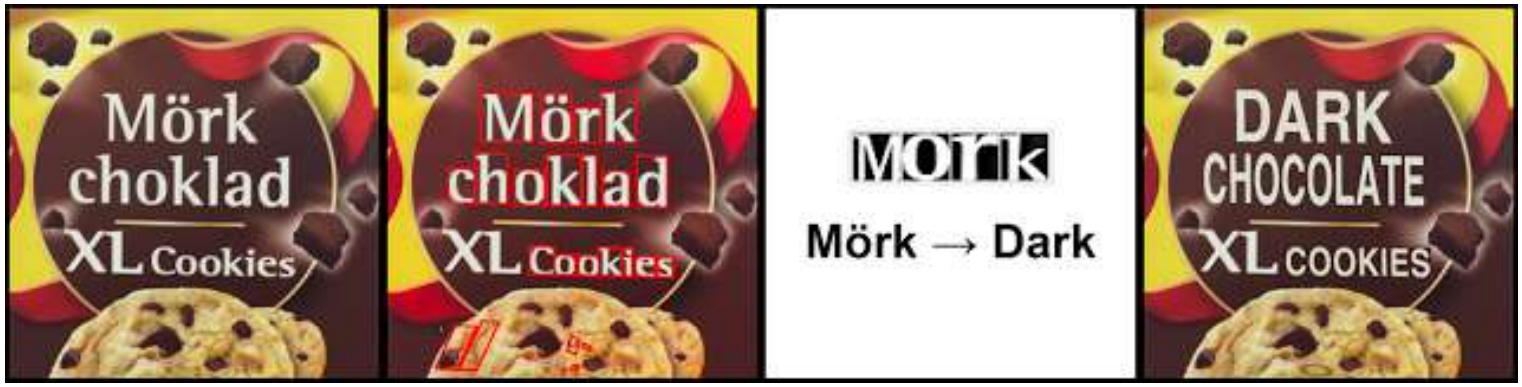
How Google Translate works

“When Google Translate generates a translation, it looks for patterns in hundreds of millions of documents to help decide on the best translation for you.

By detecting patterns in documents that have already been translated by human translators, Google Translate can make intelligent guesses as to what an appropriate translation should be.

This process of seeking patterns in large amounts of text is called **"statistical machine translation**". Since the translations are generated by machines, not all translations will be perfect.. “

WordLens - > Google



Process:

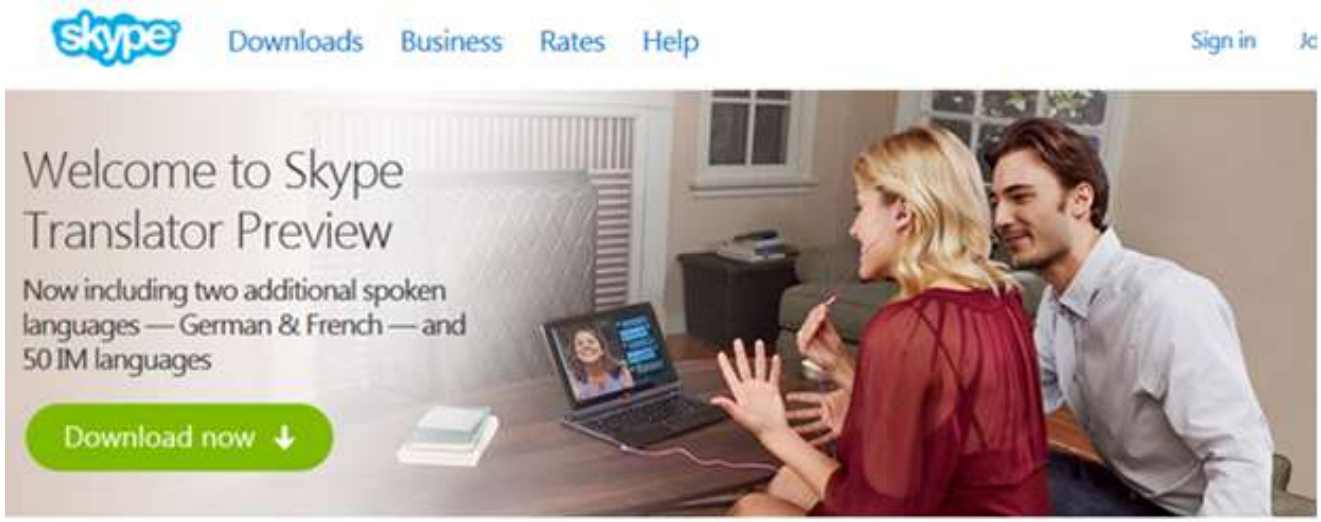
1 – find the letters in the image

2 – recognise what each letter actually is

3 - put the letters together and look up in a dictionary for a translation

4 – replace the L1 letters in the image with the L2 letters

Highdale Consulting



Skype Translator helps you communicate across language barriers, bringing people closer together. Simply use almost any Skyp enabled device to make a free Skype Translator video or voice call with someone who speaks another language — and start talk Translator is currently available in English, Spanish, French, German, Italian and Mandarin. More languages are coming soon.

Highdale Consulting

6 - Using translation in class

Pros & Cons?

- It is happening – people are using it already, so should we make space for it in our pedagogical approach?

Process?

- Learn
- Speak/record in pairwork
- Check meaning via Sp2Sp translation
- Discuss differences in group/with teacher

Highdale Consulting

SET lesson ideas

- SS write a dialogue in L1; then each translate their side into English and write it down, creating an English dialogue
- Student A speaks her side in English and writes down spoken translation into L1
- Student B speaks her side in English and writes down spoken translation into L1
- Students compare the outputs and note differences from their original L1 dialogue

- SS write a dialogue in L1
- Student A speaks her side in L1 and writes down spoken translation into English
- Student B speaks her side in L1 and writes down spoken translation into English
- Students compare the outputs and note differences, asking for teacher guidance where needed

Reflection

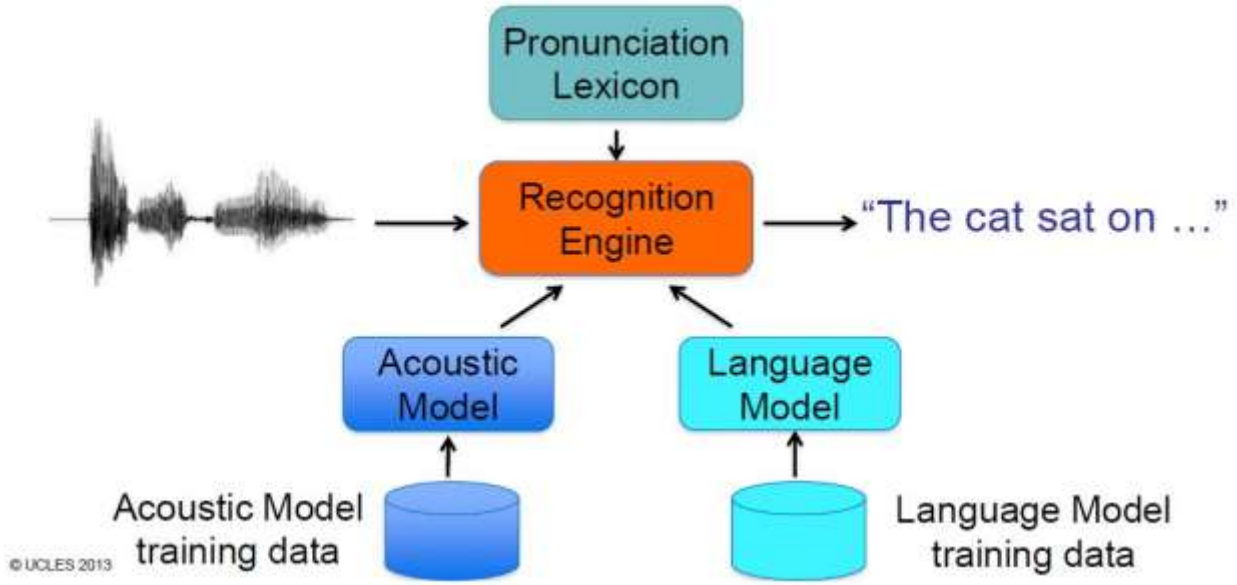
What does the instant availability of on-demand speech-to-speech translation
mean for your teaching and your students' learning?

How could 'Speech-2-Speech' auto-translation tools help you and your students, in and out of class?

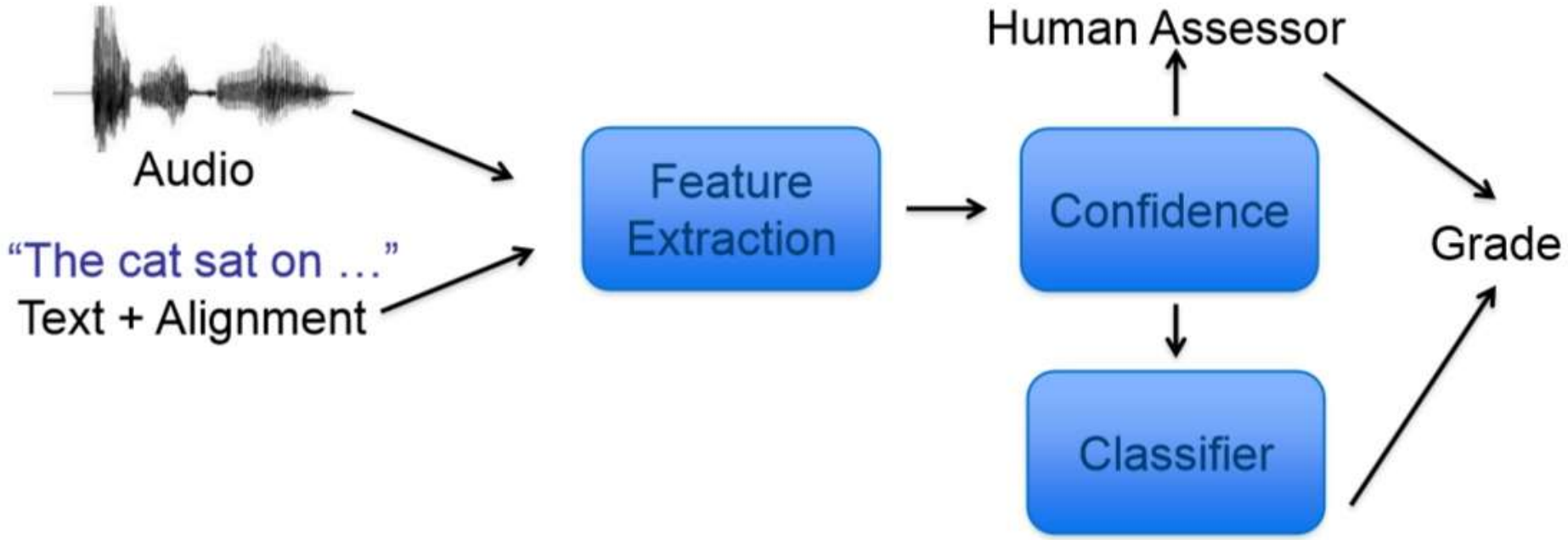
Are there any drawbacks?

7 - Speech-enabled marking

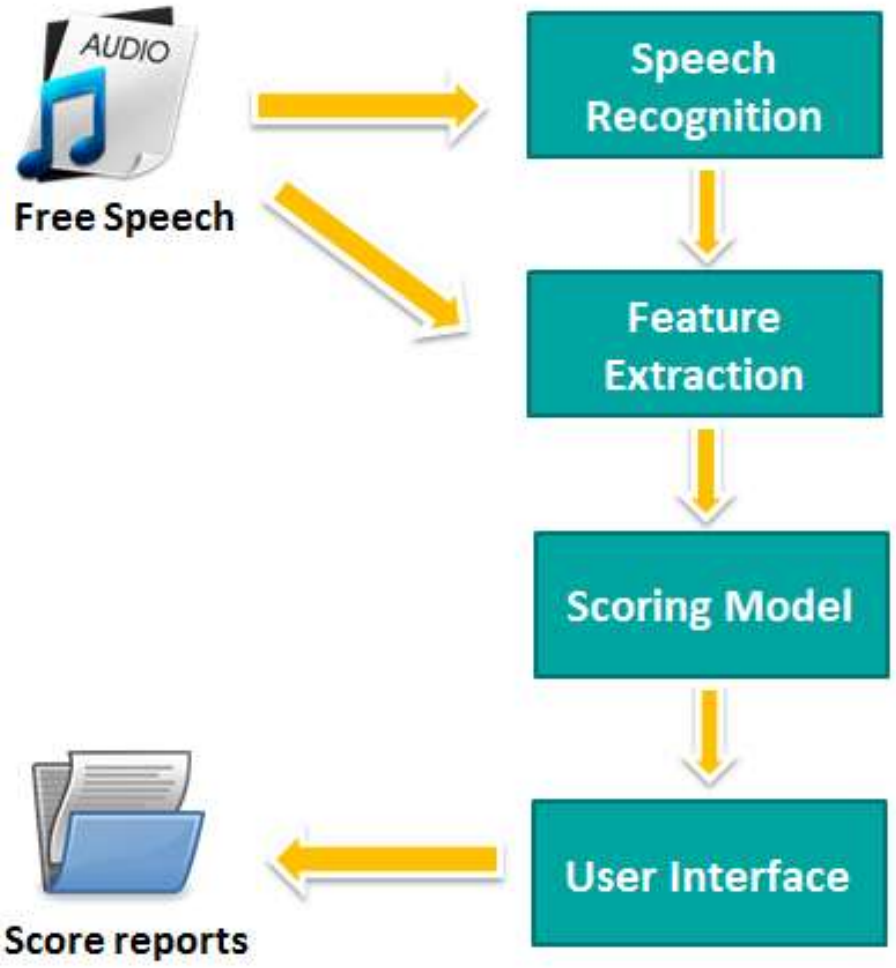
Automated marking of speech



Automated marking of speech



Linguaskill - Speaking Auto-marker



Highdale Consulting

ALTA


Cambridge ALTA
Home
Who we are
Projects
Opportunities
Useful links
Contact
Research in computer systems and platforms, corpus linguistics, computational linguistics, speech processing and machine learning
Learn more
The Cambridge University Institute for Automated Language Teaching and Assessment (ALTA) is a virtual institute which brings together teams from computing, engineering, linguistics and language assessment to investigate new ways of using technology to enhance language learning and to develop cutting-edge approaches to assessment which will benefit learners and teachers worldwide. It forms part of Cambridge University's interdisciplinary Language Sciences initiative.
Sponsorship from Cambridge English Language Assessment will provide the funding for a team of PhD students and post-doctoral researchers and will focus on a range of topics, including:
• text and speech processing
• machine learning
• corpus development and analysis
• security, platforms and deployment
The initial period of funding is for 5 years beginning in October 2013. However, the ALTA institute is intended to be a rolling research programme with funding renewed for a further 5 years towards the end of the current period.

Writeandimprove.com

Your English teacher told you about Write and Improve and said you should use it to practise and improve your English writing. Your teacher has asked you to write **a report** about your experience with Write and Improve, describing what you liked or didn't like about it and why. Finish by saying whether you would recommend your classmates to try it.

Write your **report**.

Start again

Computers changed my lifestyle, my opionion about the world.
I think it effects life more than you thinking.
I first bought computer as I was in university.
Internet was something amazing for me.
I use Google and Facebook.
I like I can have every informations I need whenever I want.
I like I can contact with anyone who is from any country when I want
I meet a lot of peoples on internet and I am interesting in them.
I learn lot about their culture by speaking them on internet.
I wasn't imagined this before.
CAn not you help my to correct this tex?

104 words entered. For this task you should enter between 140 and 190 words. Try to write more.

Saved

Check again →

Back

This is a good start! Now improve your writing. Read the feedback. Make changes and click Check again!

Computers changed my lifestyle, my ! opionion about the world.
I think it ! effects life more than you thinking.
I first → bought ▲ computer as I was ☆ in university.
Internet was something amazing for me.
I use Google and Facebook.
I like I can have every ! informations I need whenever I want.
I like I can contact with anyone who is from any country when I want
I meet a lot of ! peoples → on ▲ internet and I am ! interesting in them.
I learn ▲ lot about their culture by speaking them → on ! internet .
I wasn't imagined this before.
CAn not you help ! my to correct this tex?

Your progress

Level A1

Checks 1

Reflection

How could
Automated Marking
help you and your students?

Are there any drawbacks?

Impact of ASR on language teachers

Impacts?

- Changing role of teachers?
- Changing perception of status of teachers?

Teacher Development Needs?

- Digital literacy development for teachers
- Digital pedagogy workshops for teachers
- ASR-related lesson plans & resources

Pros and cons

ASR-based systems:

- can evaluate pronunciation and compare to L1 speaker models
- can evaluate fluency (hesitations, pauses, speed, partial words)
- cannot assess meaning, coherence of topic discussed

BUT:

- constantly improving quality and correlation to human assessors

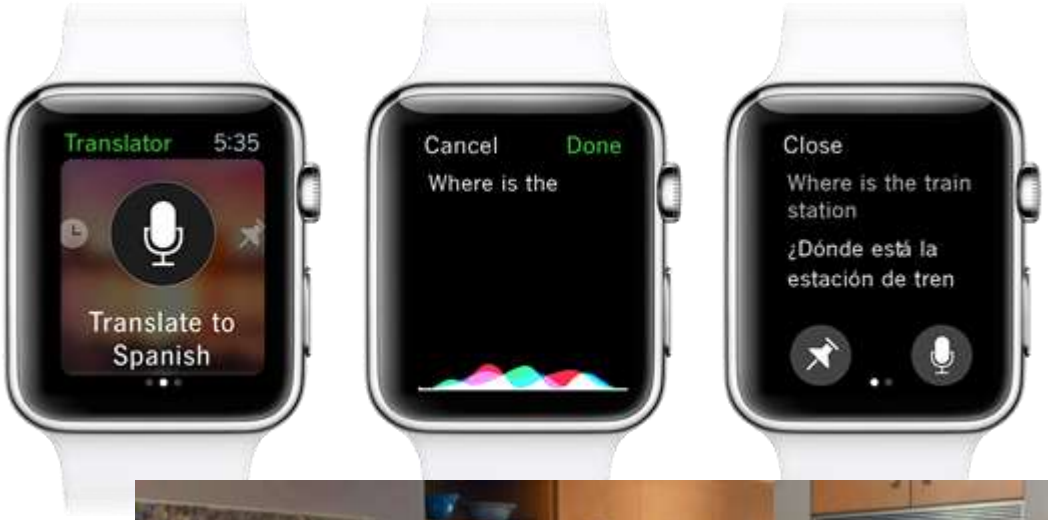
Useful for:

- Detection of mispronunciation
- diagnostic evaluation
- feedback loops for learners
- low stakes practice assessments
- training & evaluating human assessors

8 – Future trends

Wearables:

- Watches
- Hololens/AR
- Google Glass 2.0
- ‘BabelFish’ earpieces
- Personal assistants



Speech dominance:

- Speech-activated equipment
- Speech to printed output
- Speechprint ‘StarTrek’ ID systems
- Widespread automatic marking of speech

Highdale Consulting

ASR articles

Automated Speech Recognition in language learning: potential models, benefits and impact

Michael Carrier

The paper considers automated speech recognition in language learning arguing that speech recognition has reached a level of accuracy where it is powering automatic translation and testing. The author considers the impact of ASR technology on language teaching, describes the process of the development of appropriate pedagogical models, and explains how to prepare teachers for their application. This paper will give a critical analysis of the pedagogical uses and dangers of ASR technology and address how ASR can be used to automate language assessment.

KEYWORDS: *Automated Speech Recognition, ASR, ELT, speech-to-speech translation, translation software.*

INTRODUCTION

The technology of Automated Speech Recognition (ASR) is rapidly becoming more sophisticated and is becoming part of everyday life.

The aim of this paper is
impact on English lang
the use of automated s
(ASR) technology.

I will discuss the nature
works and how it can b
I will also touch on the
speech to speech transl
ASR and speech synthe
how this may also impa
motivation and teacher
Finally, I will address ho
to automate language

1. INTRODUCTION - W

Automated Speech Rec
converts audio streams
written text. ASR is still
improving rapidly in ter
in recognising spoken
transcribing it into writ

ASR is based on big dat
language corpora and t

References

Auralog. Talk To Me. Software programme. Montigny-le-Bretonneux, France: Auralog.

Adams, D. (1979). *Hitchhiker's Guide to the Galaxy*. London, Pan Books.

Bouillon, P., Cervini, C., Mandich, A., Rayner, E., & Tsourakis, N. (2011). *Speech recognition for online language learning: Connecting CALL-SLT and DALIA*.

Briscoe, E. (2016). Retrieved from http://www.cam.ac.uk/research/features/computer-tutor#sthash.NmmbdTSh.dpuf

ALTA Institute. (2016). Retrieved from http://alta.cambridgeenglish.org/

Carrier, M. (2016). *Speech recognition in language teaching: its impact on teachers and students*. In: Conference Proceedings. ICT for Language Learning 2016. Florence: LibreriaUniversitaria Edizioni.

ELTjam. (2016). Retrieved from http://eltjam.com/speech-recognition-apps-for-elt-speakingpal-english-tutor/

Escudero-Mancebo, D., Camara-Arenas, E., Tejedor-Garcia, C., Gonzalez-Ferreras, C., and Cardenoso-Payo, V. (2015). *Implementation and test of a serious game based on minimal pairs for pronunciation training*. In: Proceedings of SLaTE 2015. Leipzig: ISCA Special Interest Group SLaTE

Gales, M. (2013). *Automatic assessment of spoken English*. Retrieved from http://events.cambridgeenglish.org/cambridge-english-centenary-conference/index.html

Gales, M., Litman, D. et al. (2016). *Towards Using Conversations with Spoken Dialogue Systems in the Automated Assessment of Non-Native Speakers of English*. In: Proceedings 17th Annual SIGdial Meeting on Discourse and Dialogue, pp. 270-275, Los Angeles, CA, September. Retrieved from: https://www.aclweb.org/anthology/W/W16/W16-3635.pdf

IBM. (2012). Reading Companion. Retrieved from http://www.ibm.com/ibm/responsibility/2012/communities/education-communities.html

Saville, N. (2016). Retrieved from http://www.cam.ac.uk/research/features/computer-tutor#sthash.NmmbdTSh.dpuf

Thornbury, S. (2005.) *Awareness, appropriation and autonomy*. English Teaching Professional. Issue 40, Sept. 2005.

Ye, H., & Young, S. J. (2005, September). *Improving the speech recognition performance of beginners in spoken conversational interaction for language learning*. In INTERSPEECH (pp. 289-292, 2005).

Zinovjeva, N. (2005). *Use of speech technology in learning to speak a foreign language*. Retrieved from http://www.speech.kth.se/~rolf/NGSLT/gslt_papers_2005/Natalia2005.pdf

Thanks!

Contacts:

- Carrier, M. (2017). *ASR in the classroom*. Journal of Training, Learning and Culture **www.icc-languages.eu/tlcjournal**

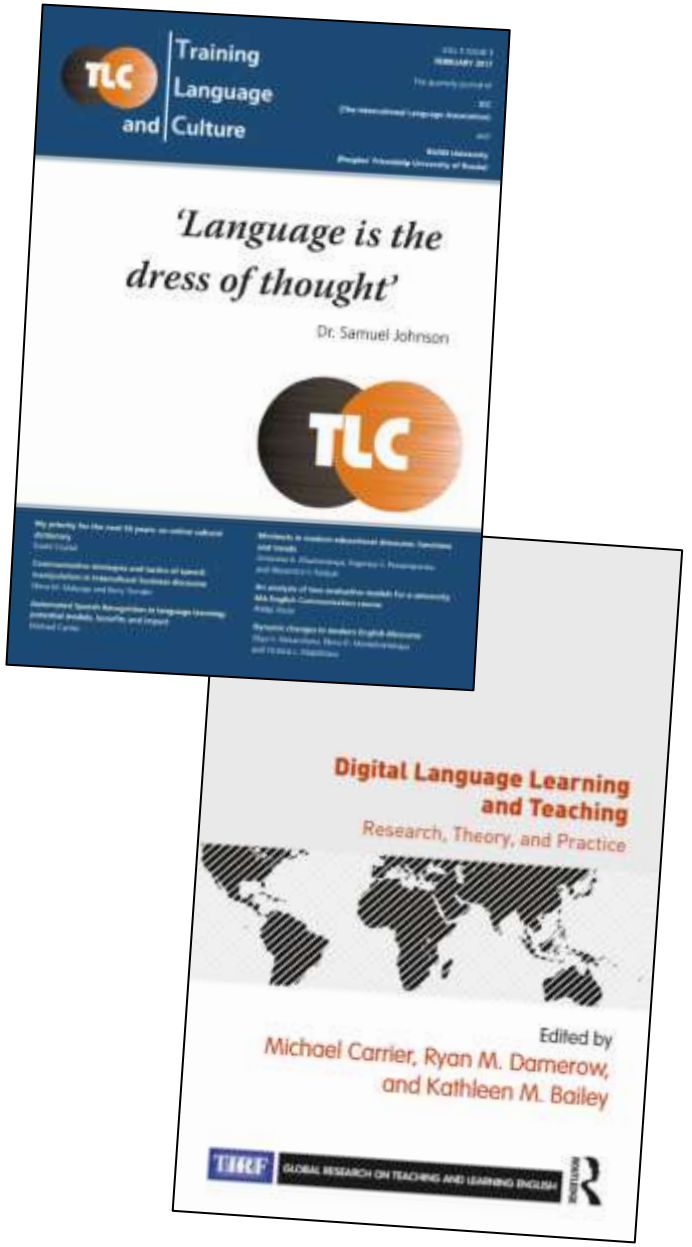
- **alta.cambridgeenglish.org**
- **www.writeandimprove.com**

Comments:

michael@highdale.org

If you would like a copy of the presentation & references:

www.michaelcarrier.com



Highdale Consulting