Automated Speech Recognition: its impact on teaching and learning languages

Michael Carrier

Highdale Consulting
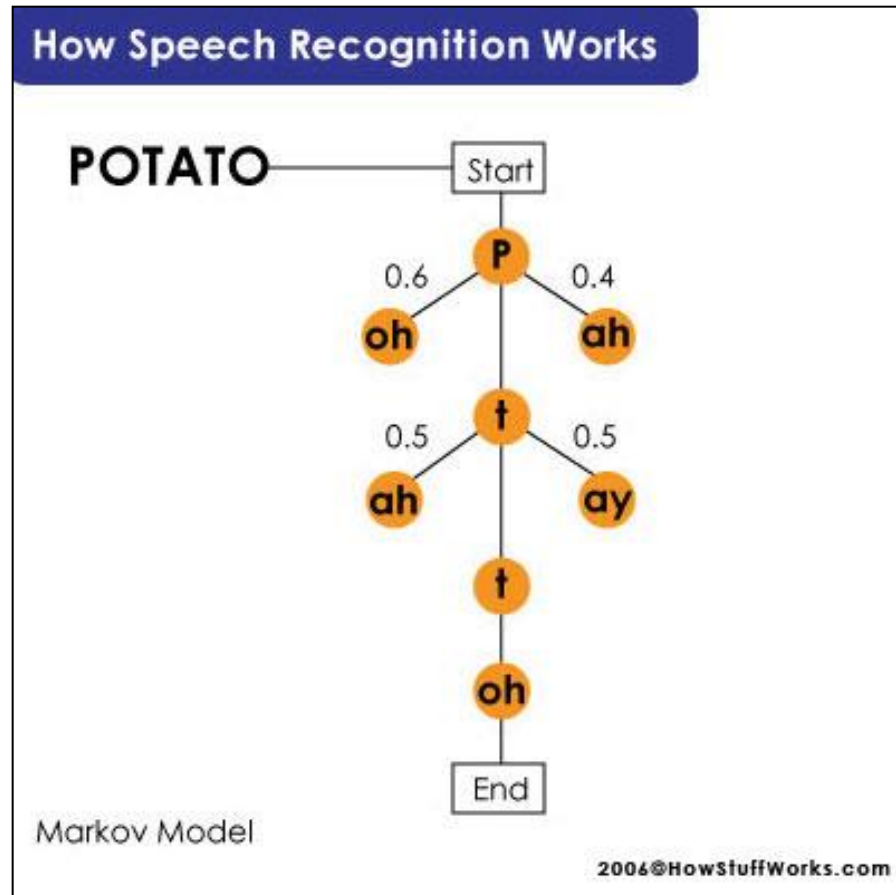
Pixel ICT, Firenze

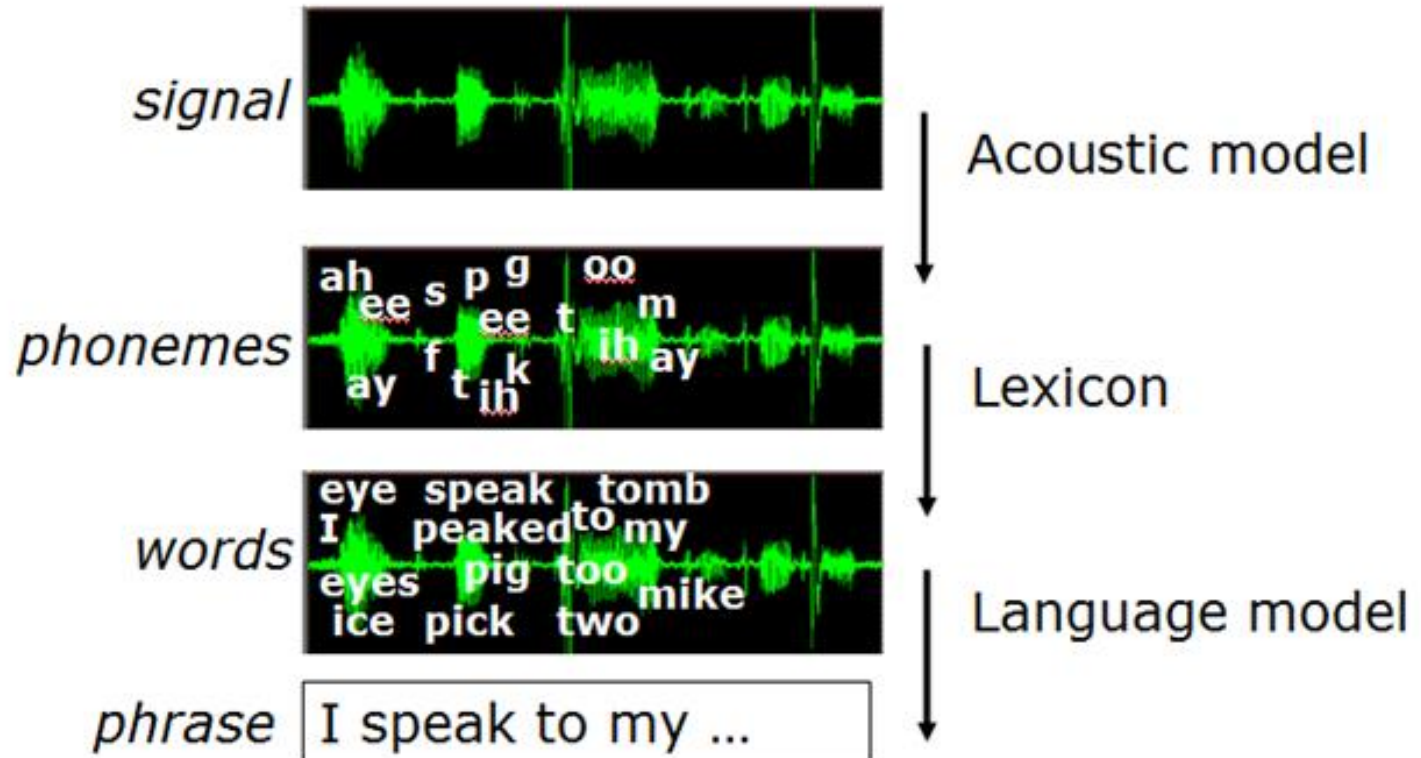**November 2016**

Contents

Highdale Consulting

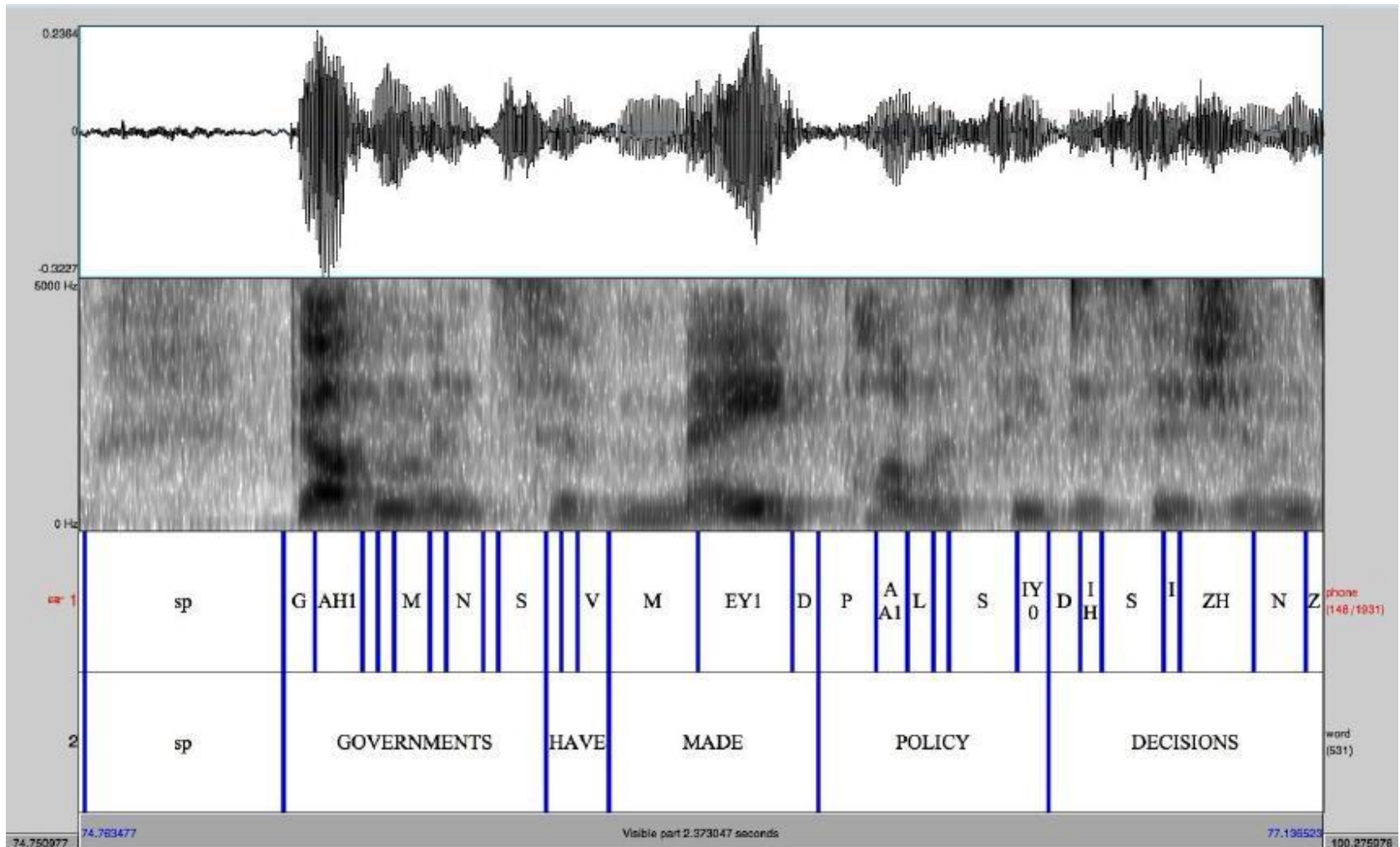# 1 - What is Automated Speech Recognition (ASR)?

- **Automated Speech Recognition (ASR)** converts audio streams into text, but does not analyse it semantically.

- The ASR output cannot assess meaning or coherence

- ASR is not the same as Natural Language Processing

- ASR is flawed but improving rapidly

- ASR is based on big data – searching language corpora and finding matching patterns in data

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".

Highdale Consulting

2 - How does it work?

signal → Acoustic model

phonemes → Lexicon

words → Language model

phrase → I speak to my ...

Aligning speech and text

Highdale Consulting

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.
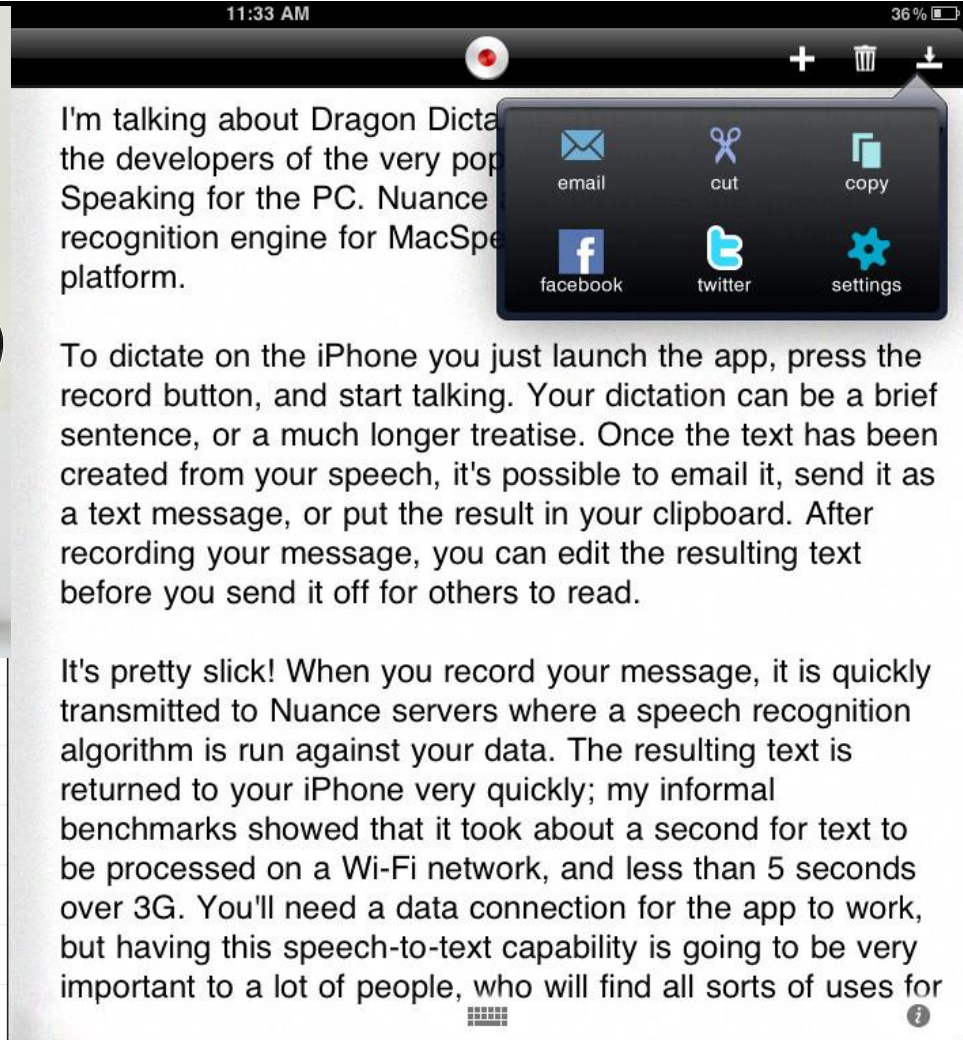
3 - How is it being used? Applications of ASR

Activities:

- Dictation

- Voice search

- Pronunciation

- Exercises

- Translation

- Marking

Sectors:

- Telephony

- In-car systems

- Military

- Healthcare

- Education

- Disability support: vision-impairment, RSI, wheelchair control, dictation

Highdale Consulting

Dragon Dictate

I'm talking about Dragon Dicta[...] the developers of the very pop[...] Speaking for the PC. Nuance [...] recognition engine for MacSpe[...] platform.

To dictate on the iPhone you just launch the app, press the record button, and start talking. Your dictation can be a brief sentence, or a much longer treatise. Once the text has been created from your speech, it's possible to email it, send it as a text message, or put the result in your clipboard. After recording your message, you can edit the resulting text before you send it off for others to read.

It's pretty slick! When you record your message, it is quickly transmitted to Nuance servers where a speech recognition algorithm is run against your data. The resulting text is returned to your iPhone very quickly; my informal benchmarks showed that it took about a second for text to be processed on a Wi-Fi network, and less than 5 seconds over 3G. You'll need a data connection for the app to work, but having this speech-to-text capability is going to be very important to a lot of people, who will find all sorts of uses for

Spoken medical reports

Highdale Consulting

Alexa & ASR apps

Not just Siri & Cortana…

- Amazon Echo - Alexa

- Smart TVs

- Google Home, Voice
  Search, Voice Typing

- Vlingo

- Nuance's Dragon Go!

- True Knowledge's Evi
  voice assistant

- Samsung S Voice

- Android's Speaktoit
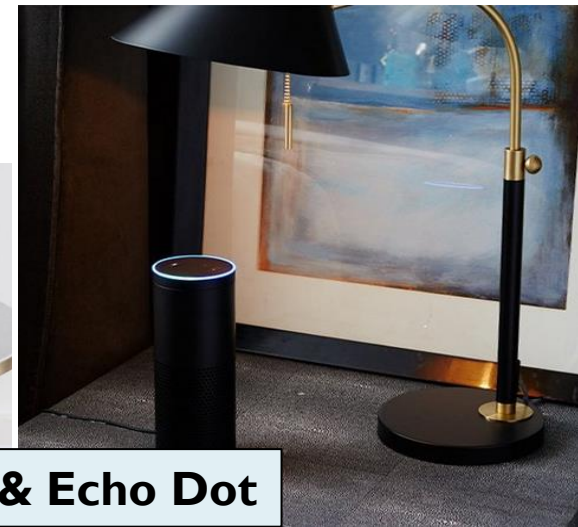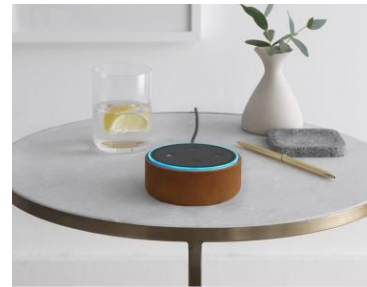
Amazon Echo & Echo Dot

Google Home (2017)

Ok Google, what is "I'm allergic to shrimp" in Portuguese?

Sou alérgica a camarão

"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."

Reflection 1

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

How would you use ASR in or with your class?

What would you need to make it possible/helpful?

4 - How can we use ASR in teaching?

- ASR has a **chequered history** in language education – many inadequate commercial products
- ASR facilitates new ways to work on **phonology and accent** – e.g. using IBM's programme 'Reading Companion'
- ASR facilitates **responses to communicative interactions** in the classroom
- ASR facilitates **automatic translation**.
- ASR facilitates computer-based **automated marking** of ELT examinations

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

IBM ASR

Reading Companion™

Home | About Us | Partners | Success Stories | Q...

Sign In

about | legal | privacy | site map | contact us

IBM Reading Companion
Web-based literacy grant program

"Imagine being a second grader, working with voice recognition software that helps you read and pronounce correctly. Reading Companion will complement our literacy curriculum, will be fun for our students and will be an added resource for our teachers."

— Louis Cuglietto, Principal, John F. Kennedy Magnet School, Port Chester, NY

Reading Companion provides a virtual library of books for learners to read aloud and receive immediate audio feedback through interactive software.

"Reading Companion has opened new cultural horizons for our children. With such a wide choice of books to increase their vocabulary and improve their comprehension skills. They're developing a true love for reading."
Patricia Díaz Covarrubias, Executive Director, Christel House de México, A.C.

Highdale Consulting

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

Highdale Consulting

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

Highdale Consulting

# 5 - 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
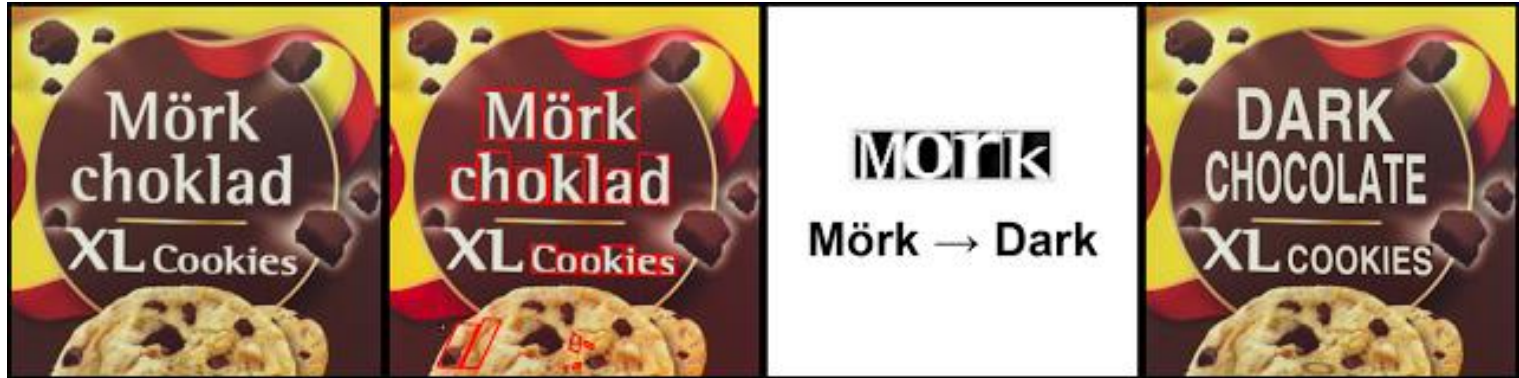
Process:

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 Skype 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.

# 6 - Using Sp2Sp 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

Using Google Translate

- SS write a sentence or short text in L1

- Student A translates it into English in writing

- Student B speaks it into Google Translate in English, translating back to L1

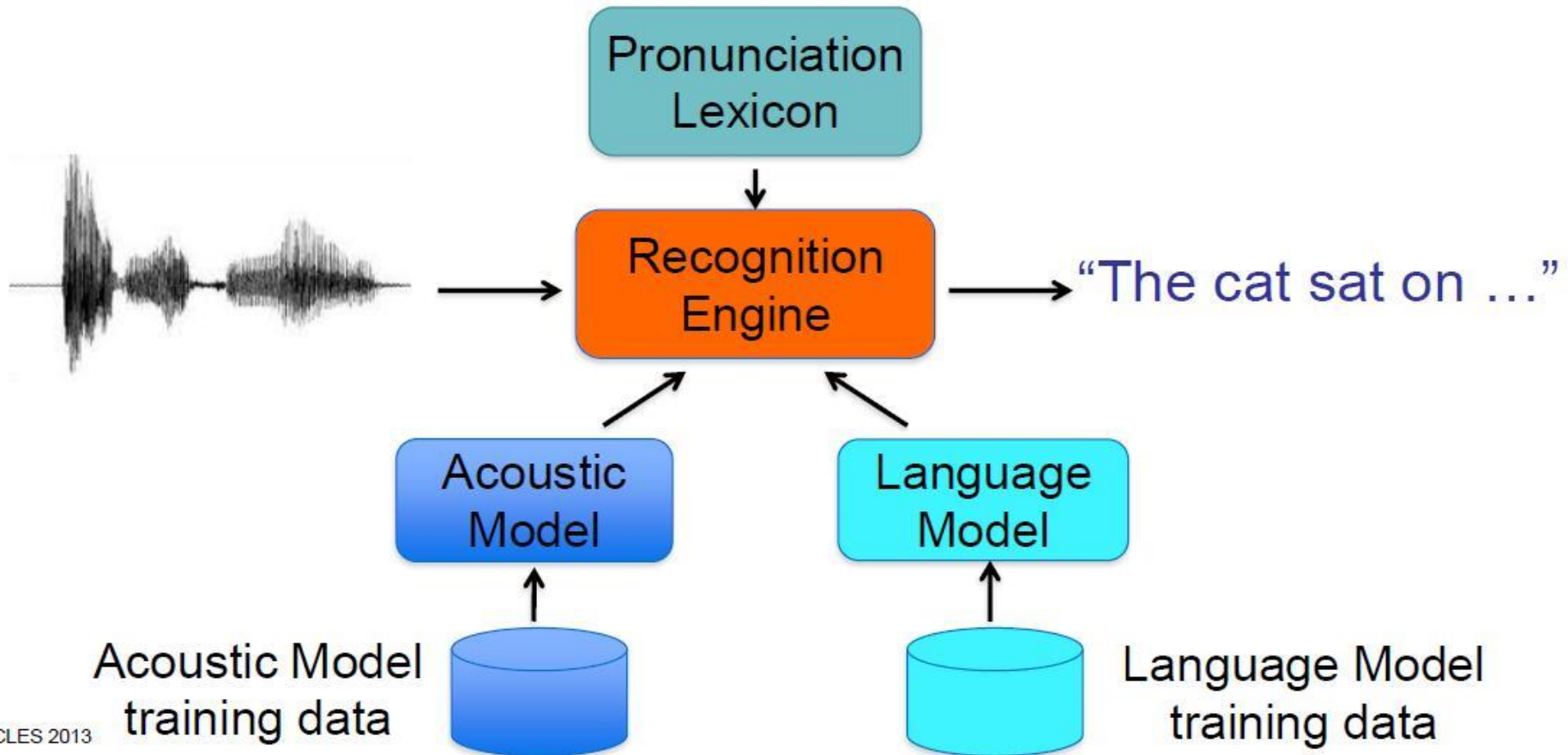- Students compare the outputs and note differences, asking for teacher guidance where needed

Highdale Consulting
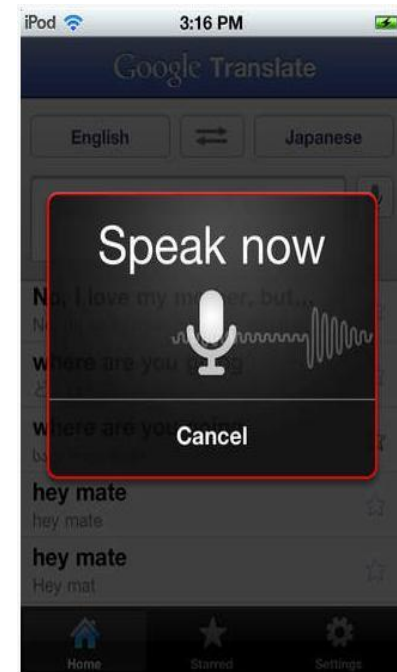
Reflection 2

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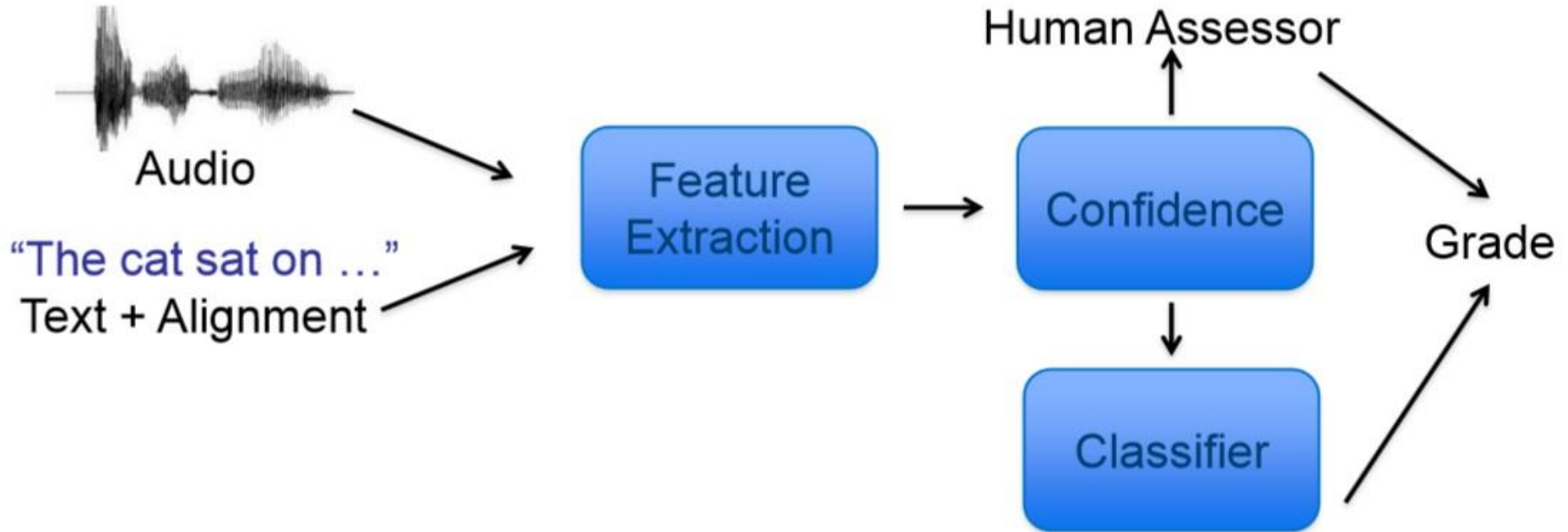mean for your teaching and your students' learning?

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

Are there any drawbacks?

7 - Automated marking of speech

Assessment process

Prof. Mark Gales http://www.policyreview.tv/video/920/6996

Highdale Consulting

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

Highdale Consulting

8 - Automated marking of Writing

'Write & Improve' beta

| A1 | A2 | B1 | B2 | C1 | C2 |

Latest score
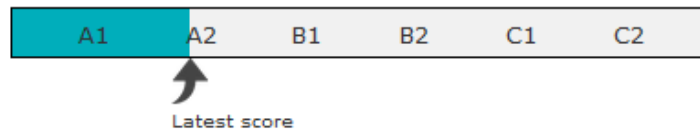
An overall score indication is assigned on the CEFR level scale. Remember, this score is only a guide. You might not get the same score in an exam. Once you have had a look at your feedback, try to use it to improve your writing, and then submit it again.

Detailed feedback (Help)

Combined | **Error feedback** | **Sentence feedback**

Sentence background colours give you an idea of the general quality of each sentence. White suggests a well written sentence. Pale yellow suggests the system believes the sentence is acceptable. Brighter yellows suggest the sentence may have a few problems. A red box indicates that explanations or corrections are available and can be viewed by hovering over the word. An orange box indicates words that might need attention to improve your results, but for which the system doesn't have a suggestion.

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.

Detailed feedback (Help)

| Combined | **Error feedback** | Sentence feedback |

ℹ️ Error feedback identifies specific words that may have been used incorrectly. Explanations and suggested corrections are provided in a separate column. Hovering over a red box highlights the associated word or error. An orange box indicates words that might need attention to improve your results, but for which the system doesn't have a suggestion.

Response text

Computers changed my lifestyle, my opinion 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.

Possible errors

opionion Spelling: Is this the right spelling? Did you mean to write 'opinion'?

effects Substitution: A different word might be better here. Perhaps 'affects' is better.

bought Omission: Is something missing after this word? Do you need to add 'a' or something similar?

computer Omission: Is something missing in front of this word? Do you need to add 'a' or something similar?

informations Countability: You have used this word in an unusual way. Do you mean 'information'?

peoples Inflection: The ending of this word is unusual. Perhaps 'people' is better.

on Omission: Is something missing after this word? Do you need to add 'the' or something similar?

internet Omission: Is something missing in front of this word? Do you need to add 'the' or something similar?

interesting Substitution: A different word might be better here. Perhaps 'interested' is better.

lot Omission: Is something missing in front of this word? Do you need to add 'a' or something similar?

on Omission: Is something missing after this word? Do you need to add 'the' or something similar?

internet Substitution: A different word might be better here. Perhaps 'Internet' is better.

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.
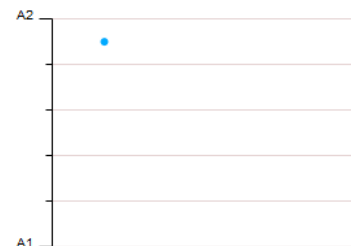CAn not you help ⓘ my to correct this tex?

Your progress

Level **A1**   Checks **1**

A2

A1

Reflection 3

How could Automated Marking help you and your students?

Are there any drawbacks?
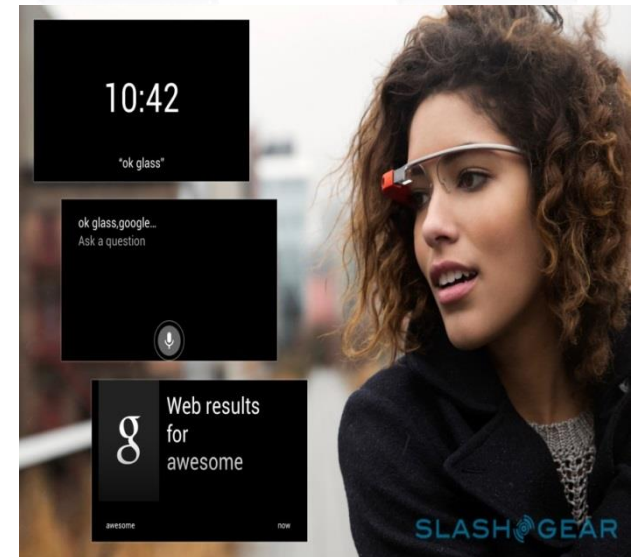
8 – Future trends

Wearables:

- Watches

- Google Glass 2.0

- Phone systems

- 'BabelFish' Earpieces

- Personal assistants

Speech dominance:

- Speech to printed output

- Speech activated equipment

- Speechprint 'StarTrek' ID systems

- Widespread automatic marking of speech

Highdale Consulting

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

Thanks!

Contacts:

- Carrier, M. (2017). *ASR in the classroom*. Journal of Training, Learning and Culture.
- www.cambridgeenglish.org/writeandimprovebeta

Comments:

michael@highdale.org

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

www.michaelcarrier.com