ALPHA

Thought piece

ALPHA LTD
St Andrew's House
St Andrew's Road
Cambridge CB4 1DL
United Kingdom

@thisisalphalive

thisisalpha.com

Speech-to-Speech:
Is the future of translation
moving away from text?

The babelfish looks like it has finally come to life this year, in a number of guises (none of them living creatures, which is both disappointing and somewhat consoling…). Google, Waverly labs and Skype all provide speech-to-speech translation, allowing you to gain almost-instant versions of the speech going on around you in your mother-tongue.

The potential for a world in which we can communicate in real-time, in any language, without the need for human translators or text raises numerous questions about where translation is going. Understanding the future of translation means understanding, first and foremost, how this technology works – and whether it works.

The multilingual earpiece

Waverly labs' wireless translation earpiece – known as the Pilot – has caused a huge stir. An independent firm, Waverley has created a product that promises to let two users converse eye-to-eye, in two different languages. Words from the speaker are translated through the earpiece into your choice of tongue, and read to you through the device.

Pilot In-Ear Language Translator by Waverly Labs

The process uses three major pieces of technology: speech recognition, machine translation, and voice synthesis. The whole is run through its app, which currently requires an internet connection to function. Waverly Labs accept that the translation will not be entirely accurate, but point out that with machine learning, the greater the number of users for each language pair, the greater the accuracy.

The device currently comes equipped with French, Italian, Portuguese, Spanish, and English languages. Further languages will apparently be available in the autumn, but must then be purchased on top of the $249 price of the earbud itself.

The question is whether or not the Pilot will deliver. It's impressive that with a six-man team, the Pilot has been put together – and ahead Google and Skype's offerings. In a Forbes interview back in May, CEO Ochea admitted that the earpiece will not be comparable to Skype speech-to-speech translation tools. "[Pilot] lags much more than Skypes' latest translation engine does...they are using. I'm not saying we built something comparable to that...we're combining technologies that currently do [artificial intelligence], speech translation and wearable technology."[1]

The machine translation used may prove something of an additional issue. Where Google is moving towards neural networks, Waverly labs mention artificial intelligence, but don't assert that their MT engines are run neurally. So it seems likely that what we're seeing is existing machine translation technology, based on a hybrid of statistical and logical translation. The approach will produce the kind of results that will allow basic communication, but is likely to also create some confusion and mistranslation.

[1] https://www.forbes.com/sites/paularmstrongtech/2016/05/17/i-just-spoke-with-pilot-and-i-dont-know-if-its-real/

So perhaps the Pilot will not be the game-changer that has been promised, but it clearly presages technology that will do what it is aiming to. Where it falls down, the Google speech-to-speech app for android seems to succeed, and Skype's on-machine offering is also a realistic option.

The next step is for the bigger players to begin creating wearables, alongside improving both voice recognition and neural translation to the point where the experience becomes workable and accurate for users.

So what does this mean?

Wearable speech-to-speech technology is clearly on the horizon, if not quite here yet. For travellers and those conducting business multinationally, the prospect is an exciting one. And then there are those who have married into families with a different first language, who seem to constitute a large portion of Waverley Lab's backers.

Being able to converse across language barriers will be life-changing for many. But that doesn't mean that it will be industry-changing quite yet. Written content still holds sway over marketing, product information, and the vast majority of online content – even when it comes to video. Speech-to-speech technology will never be instantaneous, and so, when it comes to video, subtitling or voiceover work will still be necessary for a good experience.

The real game-changer is likely to be in AR, at the point when wearing a pair of glasses or contacts lets your eye only see translated text from the first. Though even then, advertising will require a human translator at some point in order to ensure that it actually hooks its audience in.

Translators now, and in the future

As accuracy of machine translation rises, the industry must change. But translators still have – and will continue to have – their place. Neural MT is not a viable alternative when it comes to truly creative translations, and speech-to-speech has a place at a real distance from stylish, accurate, fluent writing of the kind required in most of our interactions with text. But it remains an exciting development, and one that we will be keeping our eyes firmly on.

Found that interesting?

If you're not already signed up to our newsletter, email **marketing@thisisalpha.com** to receive lots more great articles that will help you on your journey to going global.