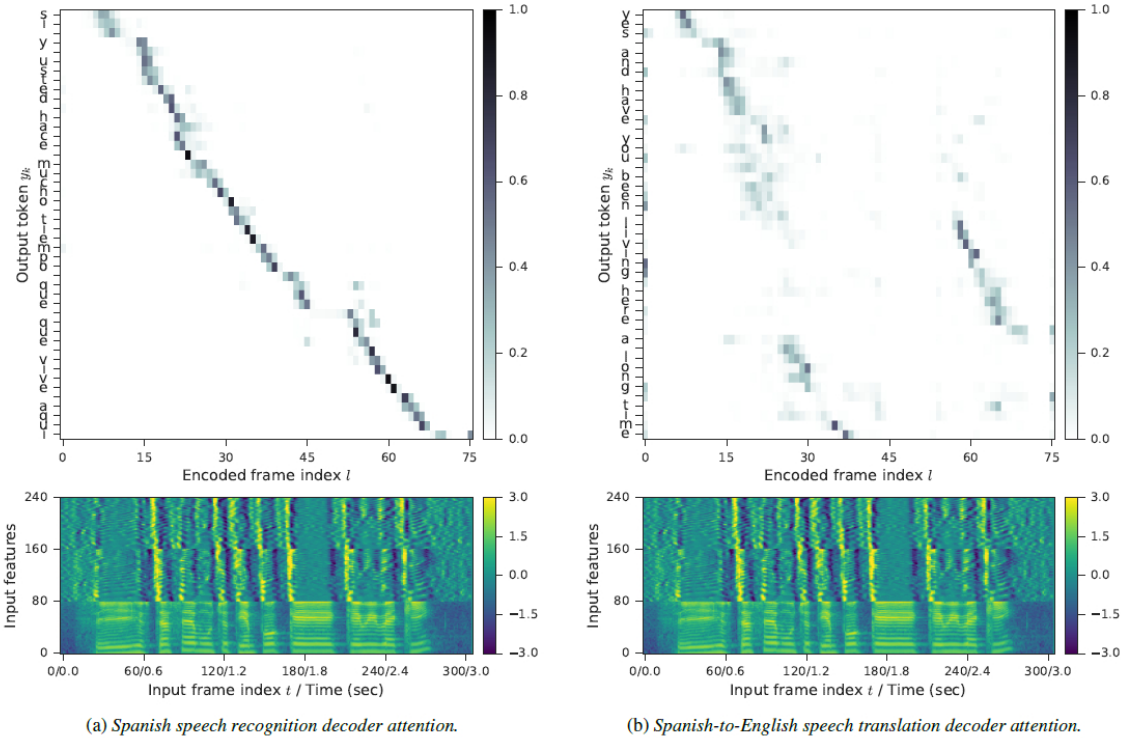


# Google Brain posse takes neural network approach to translation

April 8 2017, by Nancy Owano



Credit: arXiv:1703.08581

(Tech Xplore)—The closer we can get a machine translation to be on par with expert human translation, the happier lots of people struggling with translations will be.

Work done at Google Brain is drawing interest among those watching

for signs of progress in [machine translation](#).

*New Scientist* said, "Google's latest take on machine translation could make it easier for people to communicate with those speaking a different language, by translating [speech](#) directly into text in a language they understand."

Machine translation of speech normally works by converting it to text, then translating that into text in another language. Any error in [speech recognition](#) will lead to an error in transcription and a mistake in the translation, the report added.

The Google Brain team cut out the middle step. "By skipping transcription, the approach could [potentially](#) allow for more accurate and quicker translations."

The researchers have authored a paper, "Sequence-to-Sequence Models Can Directly Transcribe Foreign Speech" and it is on *arXiv*. Authors are Ron Weiss, Jan Chorowski, Navdeep Jaitly, Yonghui Wu and Zhifeng Chen.

The authors in their paper described their approach involving an encoder-decoder [deep neural network](#) architecture that directly translates speech in one language into text in another.

"We present a model that directly translates speech into text in a different language. One of its striking characteristics is that its architecture is essentially the same as that of an attention-based ASR neural system." ASR stands for [automatic speech recognition](#).

What did the authors do for testing? Matt Reynolds in *New Scientist*: "The team trained its system on hundreds of hours of Spanish audio with corresponding English text. In each case, it used several layers of neural

networks – computer systems loosely modelled on the human brain – to match sections of the spoken Spanish with the written translation."

Reynolds said they analyzed the waveform of the Spanish audio to learn which parts seemed to correspond with which chunks of written English.

"When it was then asked to translate, each neural layer used this knowledge to manipulate the audio waveform until it was turned into the corresponding section of written English."

Results? The team reported 'state-of-the-art performance' on the conversational Spanish to [English](#) speech translation tasks, said *The Stack*.

The model could outperform cascades of speech recognition and machine translation technologies.

The team used the BLEU score, which judges machine translations based on how close they are to that by a professional human. BLEU stands for bilingual evaluation understudy. According to *Slator*, "BLEU has become the de facto standard in evaluating machine translation [output](#)."

Using BLEU, the proposed system recorded 1.8 points over other translation models, said *The Stack*.

"It learns to find patterns of correspondence between the waveforms in the source language and the written text," said Dzmitry Bahdanau at the University of Montreal in Canada (who wasn't involved with the work), quoted in *New Scientist*.

Moving forward, the authors in their paper wrote that "An interesting extension would be to construct a multilingual speech [translation](#) system

following in which a single decoder is shared across multiple languages, passing a discrete input token into the network to select the desired output [language](#)."

In other words, as Reynolds said in *New Scientist*, "The Google Brain researchers suggest the new speech-to-text approach may also be able to produce a system that can translate multiple languages."

**More information:** Sequence-to-Sequence Models Can Directly Transcribe Foreign Speech, arXiv:1703.08581 [cs.CL]  
[arxiv.org/abs/1703.08581](https://arxiv.org/abs/1703.08581)

### **Abstract**

We present a recurrent encoder-decoder deep neural network architecture that directly translates speech in one language into text in another. The model does not explicitly transcribe the speech into text in the source language, nor does it require supervision from the ground truth source language transcription during training. We apply a slightly modified sequence-to-sequence with attention architecture that has previously been used for speech recognition and show that it can be repurposed for this more complex task, illustrating the power of attention-based models. A single model trained end-to-end obtains state-of-the-art performance on the Fisher Callhome Spanish-English speech translation task, outperforming a cascade of independently trained sequence-to-sequence speech recognition and machine translation models by 1.8 BLEU points on the Fisher test set. In addition, we find that making use of the training data in both languages by multi-task training sequence-to-sequence speech translation and recognition models with a shared encoder network can improve performance by a further 1.4 BLEU points.

Citation: Google Brain posse takes neural network approach to translation (2017, April 8)  
retrieved 13 August 2024 from <https://techxplore.com/news/2017-04-google-brain-posse-neural-network.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.