

How good is Skype's instant translation? We put it to the Chinese stress test

Skype recently announced a new translation tool that can interpret live speech in real time, across a number of languages. But the digital calling company's demos and [promo videos](#)—currently only available [as a "preview"](#) on newer versions of Windows—are heavily edited and show people speaking from scripts.

Here at Quartz, where language is an [obsession](#), we decided to give Skype Translator a more realistic stress test.

Watch our video report card above, and read on for more on how we came up with the grades.

The translator preview supports instant audio translation for English, French, German, Italian, Spanish, and Mandarin. We chose to test the combination that is likely to draw the largest number of users, English-Mandarin. Both languages have massive numbers of speakers, but not a lot of overlap; if Skype wants this tool to be useful, it will likely have to perform well bridging the divide.

We started out simple, but moved quickly on to colloquialism, literature, and finally, profanity. Below we analyze each section and give Skype a score based on well it did. Our scoring system aims to measure how much of the original meaning was carried over to the target language: an "A" means the original meaning was translated in full, an "F," not at all.

What Skype is attempting to do is extremely complicated. First, the software has to recognize what we're saying. It has to then figure out what that means, and convey the same meaning as best it can in another language. Finally, it has to speak the resulting translation aloud in a way that a native speaker can understand. Failure at any of these stages makes the translation ineffective.

Stage 1: Baby steps



Chatting over coffee.

We started out with a dead-simple dialogue intended for [beginning learners of English](#). We translated speaker B's lines to Mandarin to see how close Skype could translate back to the original English. To give the translation tool its best chance, we were careful to speak slowly and precisely.

Here's the original:

A: Can I try your coffee?
 B: Sure. Here you go.
 A: Hmm, that's not bad.
 B: There's nothing in it.
 A: What do you mean?
 B: I mean, it's just coffee.
 A: I figured that.
 B: It's not too bitter for you?
 A: It's a little bitter, but it's okay.

And here's what Skype came up with (remember speaker B's lines were originally spoken in Mandarin):

A: Can I try your coffee?
 B: When to you.
 A: That's not bad.

B: There's nothing.

A: What do you mean?

B: I mean, there's only a cup of coffee.

A: I figured that.

B: Don't you think it's too hard?

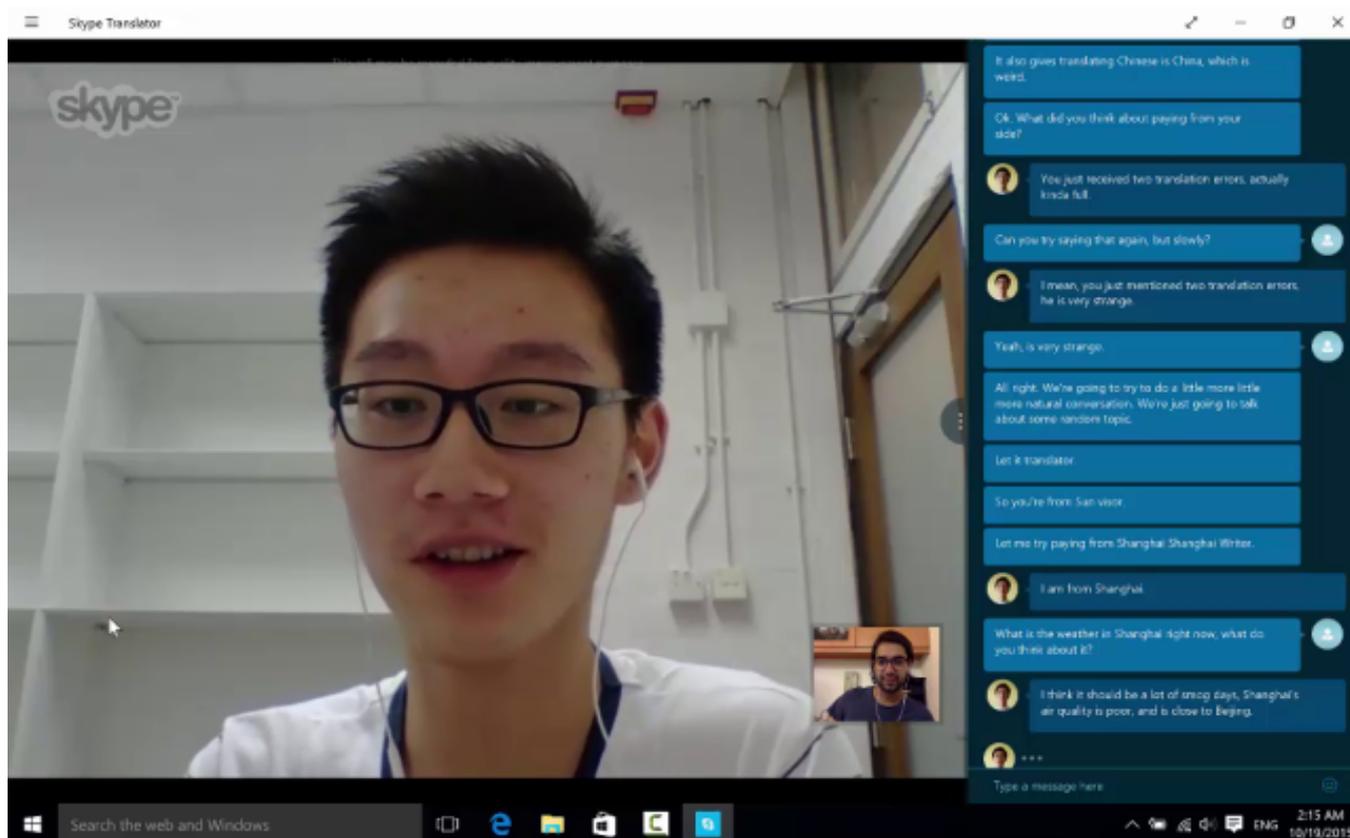
A: It's a little better, but it's ok.

In this basic test, Skype did well in the speech-recognition step for both English and Mandarin. The recording cut out at one point, totally changing the meaning of one Mandarin phrase from "Of course, here you go" to "When to you." It also mis-recognized my English "bitter" as "better," resulting in an incorrect translation. Other than that it transcribed pretty exactly what we said, and did so surprisingly quickly.

The next thing Skype had to do was translate that text. Here it performed pretty well going from English to Mandarin. Almost none of the translations were perfect, but it conveyed the general meaning for the English. But it struggled a lot going from Mandarin to English, despite usually capturing the right words. It translated "bitter" as "hard," and "there's nothing in it" to simply "there's nothing."

Stage 1 score: C+. Our test here shows that with a lot of patience, you could probably have a very basic conversation consisting of simple phrases, especially if the Mandarin speaker were willing to repeat themselves many times or say things in different ways until hitting on something Skype can translate accurately.

Stage 2: Conversation



"So you're from Sun visor."

We advanced from the pre-written dialogue to chatting at a natural pace about whatever occurred to us. Skype continued to do a pretty good job recognizing English and rendering it as Mandarin (even my annoyingly frequent "likes").

It had a hard time with some speech recognition, like "Ping," my colleague's nickname, and would only recognize "Shanghai" when I pronounced it in a way that revealed my nasally Midwestern roots. It did a reasonable job translating my gushing description of Taiwanese scallion pancakes.

Again, not so well on Mandarin to English. Probably the best Mandarin-to-English translation was "Shanghai's air quality is poor." Nearly everything else was incomprehensible. Ping had about 10 seconds of a story about morning runs in smoggy Shanghai translated simply as "According to." When he tried the story again more slowly, the resulting translation made no sense:

I was at the University of memories when I was a Bachelor degree rules running in the morning every morning, and then you cannot wear a mask, when I was about 6:30, you want to go for a run, so a group of people. Above the playground around the round the scene is terrible.

Stage 2 score: D+. Skype did better at this stage than we thought it would. Coming

into it I would have predicted an F. However, it has to get a low score because it was unable to keep up with Mandarin spoken at a normal pace, sometimes not recording the speech at all and other times coming up with incomprehensible translations, so that means one half of the conversation is largely missing.

Stage 3: Academic



Pareto inefficient translation.

We continued anyway, giving Skype some university-level challenges. I read to Ping [a definition of Pareto efficiency](#) that was translated into a mess. Ping read a line from [a novel by Mo Yan](#), a Nobel Prize-winning Chinese writer. It first translated the entire excerpt as simply "This." Then, it came back with something that also made no sense:

It can be said to be the next to Wuhan, such as the eyes and the knife at will go well, Nickels said the master he Xianfeng did such a wonderful. It is a is said to be because the fiscal harm, so prostitute named names.

Ping then said in Mandarin "I don't think it can translate this." It translated that to "My feet are a big fabric."

These are complicated topics that you would need to read to really understand, but the word-for-word translations and mis-translations of some terms made it impossible to

grasp even the general idea.

Stage 3 score: D. We have to give Skype some props for its high-quality English word recognition, though even this did fail on a couple important English words and on many in Mandarin. Skype gets a few points as well for translating the name "Pareto efficiency" correctly. But don't expect it to make English as the scientific *lingua franca* a thing of the past.

Bonus stage: Profanity



Waiting for Skype to translate an f-bomb.

Finally, we switched our profanity filters to "Off" and had a bonus round. It translated a couple English words correctly, like the essential "f—" and "s—." It couldn't translate any of the Mandarin swear words, though, even when Ping used the words for which there were clear English equivalents. It didn't quite get the hidden meaning behind "Netflix and chill."

Bonus stage score: C-. We were impressed that it got essential English cursing, but it couldn't go the other way.

Verdict: Skype Translator is an impressive piece of technology. It is an ambitious attempt to do something very recently thought to exist only in science fiction. Perhaps because of that ambition, though, it largely fails to deliver on its promise of helping humans communicate across languages.

If you have a lot of patience and absolutely no knowledge of the language of your interlocutor, Skype will be a fun way to say hello or deliver a simple message. But for anything else, you're still better off using broken English and body language.

Skype didn't respond when asked to comment on our verdict.

Skype Translator, of course, will only get better, and Microsoft does record snippets of audio from Skype Translator interactions in order to improve translation for the future. We're looking forward to giving it better scores next time.

Zheping Huang contributed reporting and language skills.