

Translation Beyond the Sentence

AMTA 2018 • Macduff Hughes • Google

Georgetown - IBM 1954

Мы передаем мысли посредством речи.

We transmit thoughts by means of speech.

Electric Translator Is Latest 'Miracle'

The age of miracles is only just dawning, it appeared today, with the introduction of an electronic "brain" which can translate from one language to another.

The awe-inspiring machine was developed by Georgetown University language experts and the International Business Machine Corp. It was demonstrated yesterday at the I. B. M. offices, 590 Madison Ave., Manhattan.

know a word of Russian herself, then typed out:

"Vyelyichyina ugla opryedelyayetsya otnoshyenyiyem dlyini dugi k radiusu."

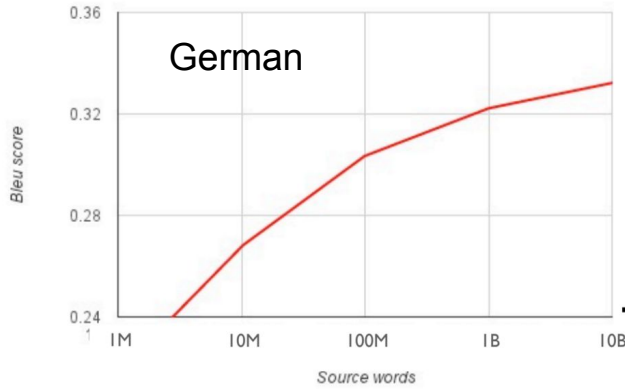
Colored lights blinked rapidly over the "face" of the electronic "brain" and it responded: "Magnitude of angle is determined by the relations of length of arc to radius."

The translator consists of 11 complicated electronic units which look on the outside as

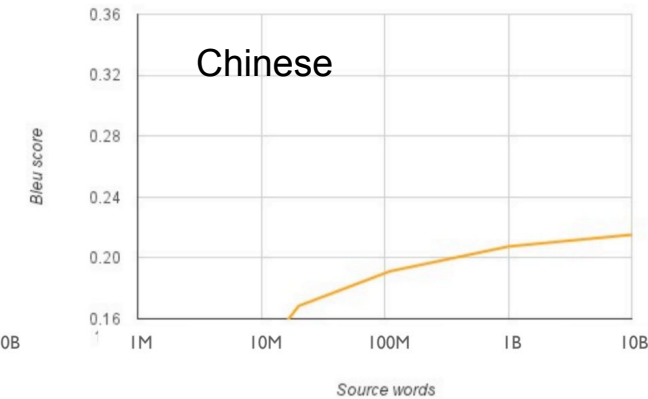
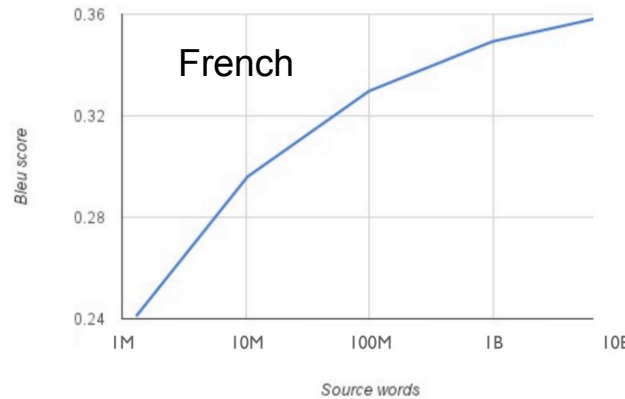
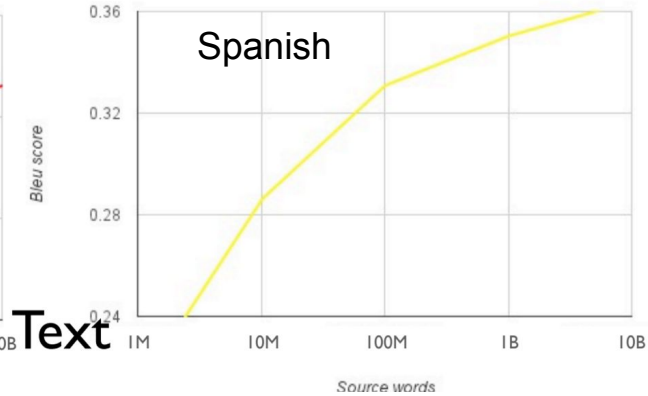
WMT 2017

“The human evaluation (Section 3) involves asking human judges to score sentences output by anonymized systems.”

This task looked incredibly hard



Text



... and some things seemed impossible

- *Word Order*
- *Morphology*
- *Word Sense Disambiguation*
- *etc...*

... and then suddenly solved?

 Research at Google

Home **Publications** People Teams Outreach Blog Work at Google

Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

Venue
CoRR, vol. abs/1609.08144 (2016)

Publication Year
2016

The AI Blog The Official Microsoft Blog Microsoft On the

Microsoft reaches a historic milestone, using AI to match human performance in translating news from Chinese to English

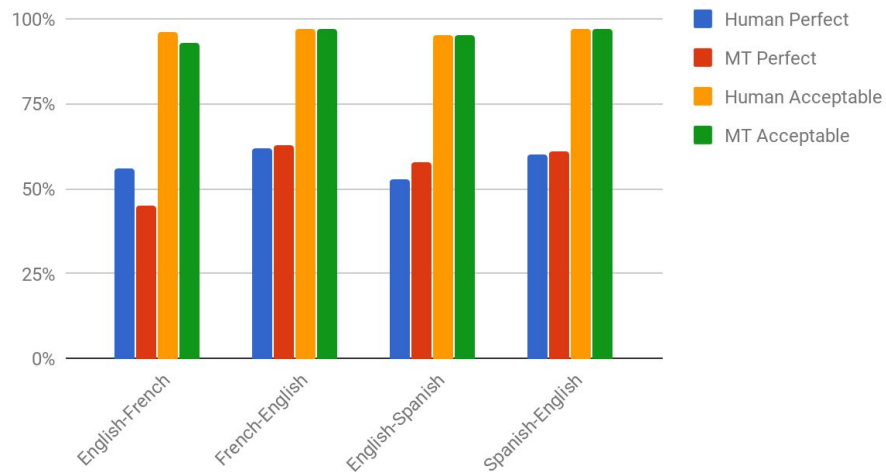
Mar 14, 2018 | [Allison Linn](#)

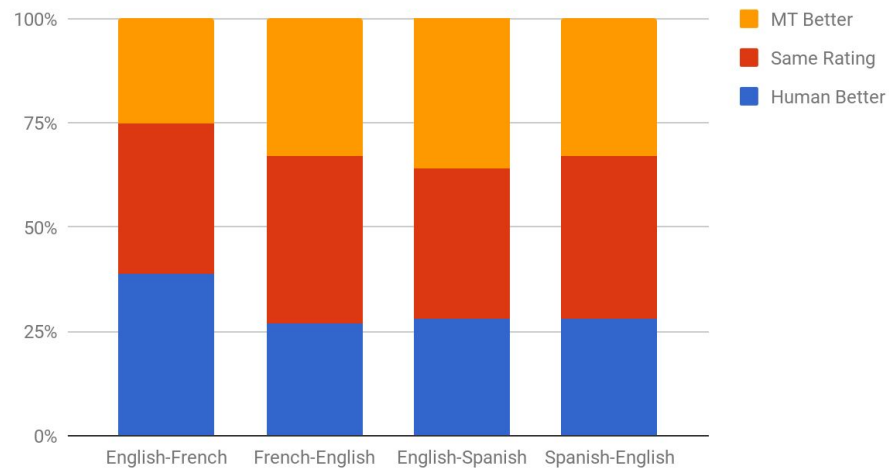


Machine vs. Human: recent internal eval

Single Presentation



Side by Side



Dig deeper?

Source	Human Translation	Machine Translation	Rating	Comment
Mira el avance del primer capítulo de 'La Madame'.	See the progress of the first chapter of 'La Madame.'	Watch the preview of the first chapter of 'La Madame'.	MT Better	HT got wrong sense for "avance".
Este anuncio ha sido marcado para su eliminación.	This announcement was strong for its elimination.	This ad has been marked for deletion.	MT Better	HT translated word for word.
En el pronóstico por regiones, el SMN indicó que la Península de Baja California mantendrá cielo despejado con ambiente caluroso a lo largo del día, y viento del noroeste de 10 a 25 km/h.	In the regional forecast, the SMN reported that the Peninsula of Baja California will have clear skies with hot weather throughout the day, and winds from the Northeast from 10 to 25 km/h.	In the forecast by regions, the SMN indicated that the Baja California Peninsula will maintain a clear sky with a hot environment throughout the day, and a northwest wind of 10 to 25 km / h.	Human Better	MT less fluent: no sense of how a weather forecast should sound.

Who is this human we're comparing to?

“a contractor for a probably large LSP, working at home and at high volumes for less than 10¢/word. They may or may not be tired or distracted today. They may not have time to review some or all of their work for mistakes, and anyway, their incentives are tilted to volume over quality. They know that careful review by clients is unlikely and rejection of their work for mistakes is even less likely.”

Strengths and limitations of the sentence task

Strengths

verifies fundamentals: word choice, syntax.

Limitations

- Humans aren't necessarily great at doing it.
- Humans aren't good at rating it.
- Doesn't account for how translation will be used.
- Doesn't distinguish impact of types of errors.
- Allows as "correct", translations that may be wrong in context.

A milestone, not a destination

Achieving “**human quality**” on this task means it’s time to reflect on where we’re really going.

We have the opportunity, and the obligation, to look at the bigger picture.

Let’s zoom out...

Why do we use machine translation?

To understand

a document, a conversation, a speech...

To express

a thought, an idea, a request...

To learn about a language

Real use cases

- Read a news article
- Read a social media post
- Read the ingredients list on a package in a store
- Listen to a speaker
- Ask someone for help
- Fill out a form
- Make small talk for social or business acquaintance
- Prepare a search index for future use
- Learning: translation to help **practice** for future independent action

Questions to ask

Should **training** for the single sentence be expected to produce good results for this case?

Can **evaluation** of single sentence translations be a good proxy for user satisfaction?

Google observations

About half of Google machine translation volume is for web pages.

In Google Translate web app, close to half of users surveyed say they are trying to learn a language.

Other large uses:

- Communicate with someone
- Work or business
- Travel or holiday

Translation and Context

Types of Context

- Translation and Documents
- Translation and Learning
- Translation and Conversation
- Translation and Society

Translation and Documents

Common issues

- Style and tone
- Word sense disambiguation
“It was a very good play”
- Loss of time information (aspect vs. tense)
Examples follow
- Loss of subject information (pro-drop)
Examples follow

Documents: time information

English Chinese Spanish Detect language ▾

↔ Chinese (Simplified) English French ▾ Translate

He runs to school. ✕
He ran to school.
She is going to school.
She was going to school.
He has finished his work.
He had finished his work.

🔊 🎧 ⌨ ▾ 137/5000

他跑到学校。
他跑到学校。
她正要上学。
她正要上学。
他完成了他的工作。
他完成了他的工作。

☆ 📄 🔊 🔄 ✎

Tā pǎo dào xuéxiào.
Tā pǎo dào xuéxiào.
Tā zhèng yào shàngxué.
Tā zhèng yào shàngxué.
Tā wánchéng le tā de gōngzuò.
Tā wánchéng le tā de gōngzuò.

Documents: time information

English Chinese Spanish Detect language ▾

↔ Chinese (Simplified) English French ▾ [Translate](#)

他跑到学校。 ×

他跑到学校。

她正要上学。

她正要上学。

他完成了他的工作。

他完成了他的工作。 |

🔊 🗣️ 拼 ▾ 47/5000

He ran to school.

He ran to school.

She is going to school.

She is going to school.

He completed his work.

He completed his work.

☆ 📄 🔊 ⏪ ✎

Tā pǎo dào xuéxiào.

Tā pǎo dào xuéxiào.

Tā zhèng yào shàngxué.

Tā zhèng yào shàngxué.

Tā wánchéngle tā de gōngzuò.

Tā wánchéngle tā de gōngzuò.

Documents: pronoun dropping

Llevó una vida poco convencional.

He led an unconventional life.

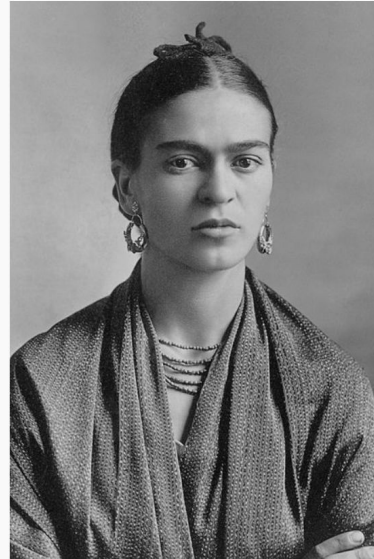
Documents: pronoun dropping

Frida Kahlo

Magdalena Carmen Frida Kahlo Calderón (Coyoacán, 6 de julio de 1907-*ibidem*, 13 de julio de 1954) fue una pintora mexicana.² Su vida estuvo marcada por el infortunio de contraer **poliomielitis** y después por un grave accidente en su juventud que la mantuvo postrada en cama durante largos periodos, llegando a someterse hasta a 32 operaciones quirúrgicas.³ **Llevó una vida poco convencional.**³ Su obra pictórica gira temáticamente en torno a su biografía y a su propio sufrimiento. Fue autora de unas 200 obras, principalmente autorretratos, en los que proyectó sus dificultades por sobrevivir. La obra de Kahlo y la de su marido, el reconocido pintor **Diego Rivera**, se influyeron mutuamente. Ambos compartieron el gusto por el arte popular mexicano de raíces indígenas, inspirando a otros pintores mexicanos del periodo posrevolucionario.

En 1939 expuso sus pinturas en **Francia** gracias a una invitación de **André Breton**, quien intentó convencerla de

Frida Kahlo



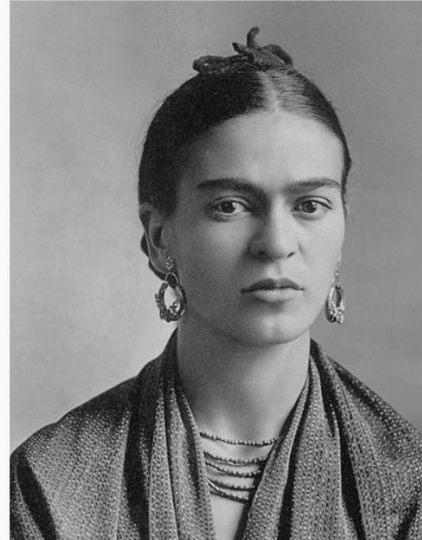
Frida Kahlo en 1932.

Documents: pronoun dropping

Frida Kahlo

Magdalena Carmen Frida Calderon ([Coyoacán](#) , 6 July as as 1907 - *ibidem* , [13 as July as 1954](#)) was [painter](#) [Mexican](#). ² Her life was marked by the misfortune of contracting [poliomyelitis](#) and later by a serious accident in her youth that kept her bedridden for long periods, reaching up to 32 surgical operations. ³ He led an unconventional life. ³ His pictorial work revolves thematically around his biography and his own suffering. She was the author of some 200 works, mainly self-portraits, in which she projected her [difficulties to survive](#). The work of Kahlo and her husband, the renowned painter [Diego Rivera](#) , influenced each other. Both shared a taste for Mexican folk art with indigenous roots, inspiring other Mexican painters of the postrevolutionary period.

Frida Kahlo



3 out of 9 personal pronouns wrong. These are arguably the most severe errors in the translation, but are not counted by sentence metrics.

Translation of Speech and Conversation

Exciting use case for translation

All the problems of documents, and more:

- Anaphora and backward references even more common
- Tangled up with disfluency issues
- References to objects and situation

Speech and Conversation examples

Spanish English Chinese Detect language ▾



English German Spanish ▾

Translate

So I got my degree. Worked for a publisher for a while. Boring but learned a lot!
Did you think about going back to school?
Thought about it, sure. Didn't end up doing it.



171/5000

Also habe ich meinen Abschluss gemacht.
Arbeitete für eine Weile für einen Verlag.
Langweilig, aber viel gelernt!
Hast du darüber nachgedacht, wieder in die Schule zu gehen?
Dachte darüber nach, sicher. Es endete nicht damit.



word choice

third person instead of first; verb lost

Translation and Society

Ethics and social concerns are emerging as major issues for AI and Machine Learning.

Machine Translation is no exception.

What will be (are) the large scale impact of widely available, high quality MT?

Wider impacts of MT

Impact on endangered and minority languages

Communication across cultures

Effect on linguistic change

Representation of minorities and sensitive or historically oppressed groups

What follow are questions, not answers...

Impact on endangered and minority languages

Probably positive

Speakers generally eager for MT products.

Possible risks:

- Could speed the transition of speakers to the more dominant language
- Could MT “flatten” the language - lose nuance, regional variation, etc?

Communication across cultures

We assume more communication is better...

but remember the Douglas Adams quote about the Babel fish.

What are the implications if large groups of people who couldn't understand each other, suddenly can? (Or *think they can.*)

Effect on linguistic change

MT learns from data, which today has typical characteristics...

- Usually from “higher register”, educated writers and speakers
- Limited time window: from dawn of WWW to recent past

and some likely biases

- To modern formal educated speech, away from informal, vernacular, archaic
- Slow to adopt linguistic innovation

Linguistic change example

Spanish gender



Cristina Fernandez de Kirchner ✓

20 de julio de 2017 · 🌐

Uno está recorriendo todo y hay que meter mucha esperanza. Porque la gente está muy preocupada y muy bajoneada. Y la verdad, es que tenemos que darle fuerza y decirle que **todxs unidxs** podemos ponerle un límite a este Gobierno.

Un tema que me preocupa mucho es la caída del consumo de la leche. Porque además del precio, tenemos un problema de seguridad alimentaria. ¿Cómo van a crecer **estxs pibxs** si no les damos leche?

Si el rumbo que lleva el Gobierno trae estos problemas y trae estos dolores, lo que tiene que hacer el Gobierno es cambiar el rumbo. Porque [#AsiNoSePuedeSeguir](#).

81.041 reproducciones

5.046 Me gusta 487 comentarios 2.986 veces compartido

Linguistic change example

Spanish gender

How do you say “everybody” in Spanish?

- **todos:** traditional, masculine+inclusive(?)
- **todos y todas:** more modern, explicitly inclusive
- **tod@s or todxs:** very modern, neutral, but unpronounceable

All are used by some mainstream speakers, all are controversial, all can be found in our training data.

Should MT just let the training data “vote”?

What’s the responsibility of the builder of a system that shows a “right” answer?

Social Concerns in AI/ML

Tech / #NewTech

FEB 26, 2018 @ 12:26 PM

11,137



EDITOR'S PICK

Racist, Sexist AI Could Be A Bigger Problem Than Lost Jobs



Parmy Olson, FORBES STAFF



Inc.

GROWICO
REGISTER NOW

Q SEARCH

NEWSLETTERS

FOLLOW

SUBSCRIBE



INNOVATE

How Do You Design a Robot That Isn't Sexist or Racist? It's Harder Than You Think

A.I. systems are built by human beings with implicit biases. So it's no wonder bots are biased, too. A slew of startups are hoping to do something about that.



By Michelle Cheng *Editorial assistant, Inc.com* [@mbcheng15](#)



Entrepreneur

PROVIDE A BETTER
**CUSTOMER
EXPERIENCE**
REGISTER NOW



TECHNOLOGY / ARTIFICIAL INTELLIGENCE

Why Are Some Bots Racist? Look at the Humans Who Taught Them.

Here are a few things we can do about human biases in machine learning.

Emergent Tech ▶ **Artificial Intelligence**

Fun fact of the day: Voice recognition tech is naturally sexist

Machine learning and AI could make the problem worse

By [Katyanna Quach](#) 14 Mar 2018 at 01:07

64

SHARE ▼

Translation is very much included

FAST COMPANY

Google Translate's Gender Problem (And Bing Translate's, And Systran's...)

Google Translate and other popular translation platforms often provide unintentionally sexist translations where, among other things, *doctors* are men and *teachers* are women. The reason why has to do with a complex mix of algorithms, linguistics, and source materials.

more than gender...

Minority groups

Neutral terms (e.g “homosexual”) translated to offensive slurs



Politically loaded terms

“Illegal alien”, “undocumented immigrant”


Religious terms





One word for “God” vs. several for different religions



...but gender is one of the biggest challenges




English Spanish Chinese Turkish - detected  Spanish English German  [Translate](#)

O bir hemşire.
O bir doktor.

She is a nurse.
He is a doctor. 





   





  29/5000

Turkish English Spanish Detect language   English Turkish Spanish  [Translate](#)

I talked to my doctor.
I talked to my secretary.

Hablé con mi doctor
Hablé con mi secretaria.

    48/5000

Gender: what is the right thing to do?

If ambiguous gender must be resolved for linguistic reasons...

Let the training data vote?

Current behavior, not satisfactory

Adopt gender neutral novelties?

May not meet users needs

Flip a coin?

Much harder than it sounds

Where possible, *change the task!*

- Offer multiple answers with explanation
- Give users more control

Conclusion

The single sentence task is “mostly” solved

Claim

Given enough in-domain data and enough compute power, we can build a system that does the sentence task with similar skill to a typical bilingual human. (No fundamental breakthroughs needed.)

Claim

This has actually been done for some language pairs in some useful domains.

Caveat

Machines will make different types of errors that may be recognizable; hard edge cases that humans can do but computers can't will remain.

It is time to look beyond the single sentence task

The research community needs good contextual **benchmark tasks**. (IWSLT2017 is an example and good start.)

Industry should develop **standard practice** for evaluating MT in context.

Product developers need to consider the **whole workflow** and user journey.

All MT practitioners need to consider **social and ethical implications** of large scale translation.

Thank You!

macduff@google.com