Natural Language Understanding Is Needed for Machine Translation

Example English to Spanish translation

Did you see a white cow?

¿Viste una vaca blanca?

Can translation succeed by word or phrase substitution plus some reordering? Problems arise because natural language is ambiguous.

Here is an infamous machine translation from English to Russian and back to English.

The spirit is willing, but the flesh is weak.

The liquor is holding out all right, but the meat has spoiled.
Ambiguities of Natural Language

Lexical Ambiguity

Stay away from the bank.

Structural Ambiguity

He saw that gasoline can explode.
I saw the man on the hill with a telescope.
The chickens are ready to eat.

Semantic Ambiguity

David wants to marry a Norwegian.

Referential Ambiguity

When a bright moon ends a dark day, a brighter one will follow.
She dropped the plate on the table and broke it.

Pragmatic Ambiguity

Don’t you know what day it is?

Multiple Ambiguity

Time flies like an arrow.
Levels of Natural Language

Phonetics: speech sounds, e.g.,
sounds of “k”, “i”, and “t” in “kite”.

Phonology: organization of speech sounds, e.g.,
different “k” sounds in “kite” vs. “coat”,
different “t” and “p” sounds in “top” vs. “pot”.

Morphology: construction of words, e.g.,
use of “-s” to form plurals,
use of “-ed” to form past tense of verbs.

Syntax: combination of words into phrases and sentences, e.g.,
Flying airplanes is dangerous.
Flying airplanes are dangerous.

Prosody: rhythm and intonation of language, e.g., in English, questions usually end with increasing pitch.

Semantics: meaning of language, e.g.,
The pig is in the pen.
The ink is in the pen.
Pragmatics: effect of language on the speaker and listener, e.g.,

   Can you pass the salt?
   Do you know the time?
   I swear to tell the truth ...

World Knowledge: Knowledge of the physical world, social interactions, etc., e.g.,

   The porridge is ready to eat.
   There’s a man in the room with a green hat on.