

Linguistics module

Topic 10 Machine Translation

Topic Overview: In this topic students are introduced to Machine Translation (MT). MT or automated translation is a process when a computer software translates text from one language to another.

Key Skills	Communicating Working with others Problem-solving Language awareness Analytical skills	
Expected time	Two hours (not including extra resources)	
Self-assessment (check with students at the end of the topic)	I understand what machine translation is. I understand what a parallel text is I understand some of the reasons machine translation makes mistakes I can analyse the data I was given in the puzzle and draw conclusions. I understand what word sense selection is.	✓ / X

Task 10.1 Description: Students learn about Machine Translation, Parallel texts and why Machine Translation is so difficult.

Learning Intentions	Students will learn what Machine Translation is. Students will understand how the models learn to translate from parallel texts. Students will understand the kinds of errors that MT makes and why
Resources	PPT Machine Translation Slides Worksheet 10.1 Parallel Texts
Task Outline	Teacher shares a PPT with students introducing the topic of Machine Translation. The presentation outlines parallel texts and introduces three exercises. Give students Worksheet 10.1 Parallel Texts to work through. Solutions are in the worksheet and also in the slides.

Task 10.2 Description: The students try a word sense selection MT puzzle on paper

Learning Intentions	Students will understand what ambiguity and word sense selection are and how it is an issue for Neural Machine Translation systems
----------------------------	------------------------------------------------------------------------------------------------------------------------------------

Linguistics module

Resources	AILO Running on MT puzzle and Solution All puzzles mentioned in this module are at: https://ailo.adaptcentre.ie/tymodpuzzles/
Task Outline	One of the most common errors in modern Neural Machine Translation systems is word sense selection: the source language text may contain words which have multiple meanings and the MT system has chosen the wrong one e.g. “angry-legged” instead of “cross-legged”. In this puzzle, students must find the other examples of the wrong meaning being used.

Task 10.3 Description: The students try out a website which investigates gender bias in language

Learning Intentions	The students learn that bias from everyday life is present in our data. This affects the translations they rely on.
Resources	PPT on gender bias in language (in MT slides)
Task Outline	<ol style="list-style-type: none">The students learn that bias from everyday life is present in our data. This affects the translations they rely on. The students should take a look at http://wordbias.umiacs.umd.edu/ and see if there are words that are commonly associated with a particular gender which they hadn't considered before. Example searches: nurse, doctor, pilot, caring, smart, happyUsing the Github link on the slides to real MT data, students should try to translate a few of the test MT sentences using Google Translate into a language with grammatical gender you know (e.g. French? Spanish? German?) Which gender is meant? What issues can you find?

Extra Resource Task 10.4 Description: The students try out an app for translating images

Learning Intentions	The students learn about image and word translation and the issues around it
Resources	PPT on image translation (in MT slides)

Linguistics module

Task Outline

Ask students to think of an object that is shiny/glittery, gold colour, has a loop on one end, reflects the light, and is made of a hard material.

The students write down what they think it is. Play the [YouTube video](#) in the slides and give them the '[Thing translator app](#)'. Let them play with it. Ask if it is working. Some will say it's working fine but someone may say it's really wrong about something and we discuss that word and why it might be wrong.. Discuss correct words, versus wrong words, why it might be. Machine Learning and Machine Translation need good data, and lots of data to learn. You can return to the item you described, what did they write down? Then show it's a christmas decoration not a necklace / earring.. The image app takes in description data but does not always get it right. Other issues are that the image 'guess' is wrong, your translation into another language will be wrong.

Extra Resources

We encourage every school and student to take part in the 2022/23 [All Ireland Linguistics Olympiad \(AILO\)](#). The preliminary round is held in your own school at the end of January with the National Final in Dublin City University in March 2023.

Register at <https://ailo.adaptcentre.ie/enter/> and try other puzzles at <https://ailo.adaptcentre.ie/puzzles/>.