**Machine Translation Analysis Exercise**

Abuchi Godswill Okeke

School of Computer and Information Sciences, University of the Cumberlands

MSAI532 - M51 Natural Language Processing

Dr. Toni Farley

Feb 21, 2025

For this exercise, I selected the Spanish novel “La Vida Imaginaria” by Mara Torres. The novel invites us to reflect on how our dreams, fantasies, and the stories we tell ourselves can uplift us and blur the lines between real and imagined.

**Selected Sentence and Its English Translation**

The shortest sentence in chapter 1, page 1 is:

**Original Spanish:**

*Nunca lo pienso.*

An official English translation (or a plausible rendition) is:

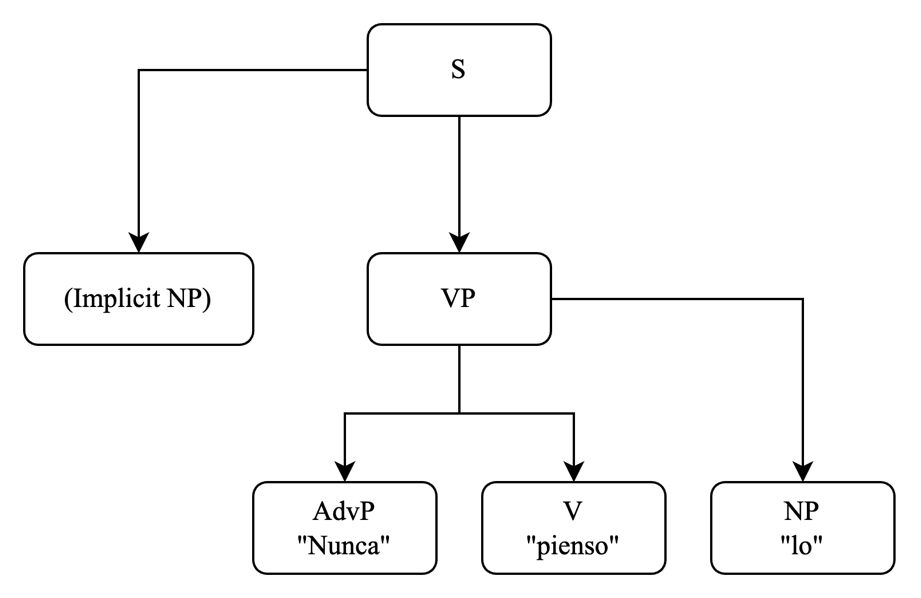
**English:**

*I never think it.*

1. **Parse Trees**

**Figure 1.0**

*Original Spanish Parse Tree*



**Explanation**

**Spanish Sentence – “Nunca lo pienso”**

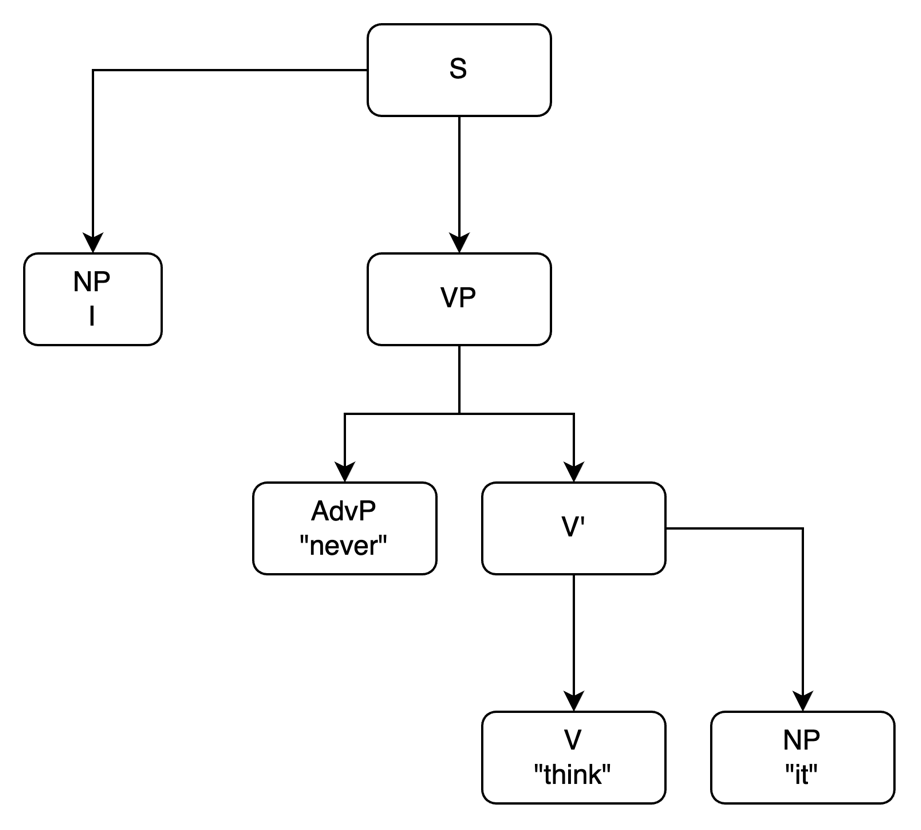
In Spanish, the sentence “Nunca lo pienso” means “I never think it.” Notice that the subject “I” isn’t written because Spanish verbs show who is doing the action. In our parse tree from Figure 1.0, we have:

1. An overall Sentence (S).
2. An implied (Implicit NP) subject (the speaker “I”), even though it isn’t shown.
3. A Verb Phrase (VP) that breaks into three parts:

* An Adverb Phrase (AdvP) with “Nunca” (never),
* The Verb (V) “pienso” (think),
* And a Noun Phrase (NP) for the object “lo” (it).

**Figure 2.0**

*English Parse Tree*



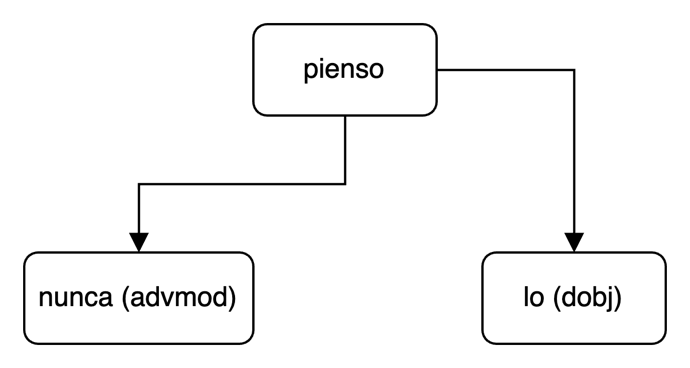
**English Sentence – “I never think it”**

In English, the subject must be explicitly written. The tree in Figure 2.0 shows:

1. A Sentence (S) that splits into a subject (NP: “I”) and a Verb Phrase (VP).
2. Within the VP, there’s an Adverb Phrase (AdvP) with “never,” and an intermediate node that further breaks into:
   * The Verb (V) “think”,
   * And a Noun Phrase (NP) for the object “it.”
3. **Dependency Structures.**

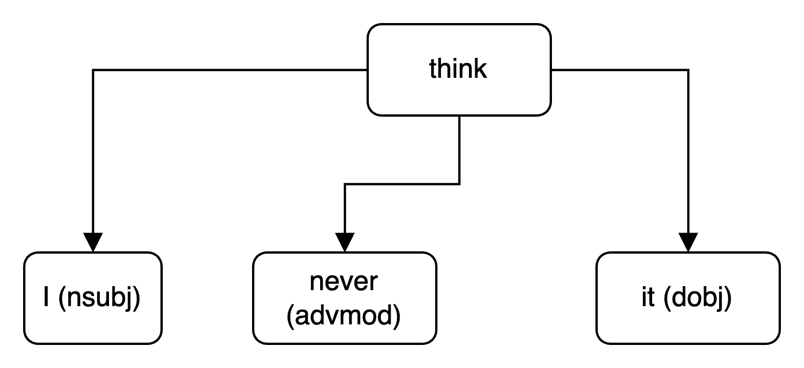
**Figure 3.0**

*Spanish Dependency Structure (original)*



**Figure 4.0**

*English Dependency Structure*



Dependencies in Figure 3.0 and Figure 4.0 show how each word relates to the others.

1. **Spanish Dependency:**  
   The verb “pienso” is the main word (the root). “Nunca” attaches to it as an adverb modifying the action, and “lo” attaches as the object.
2. **English Dependency:**  
   Similarly, in “I never think it,” “think” is the root, “I” is the subject, “never” modifies the verb, and “it” is the object.
3. **Case Structure**

**Figure 5.0**

*Case structure representation of the meaning that the original and translation share.*

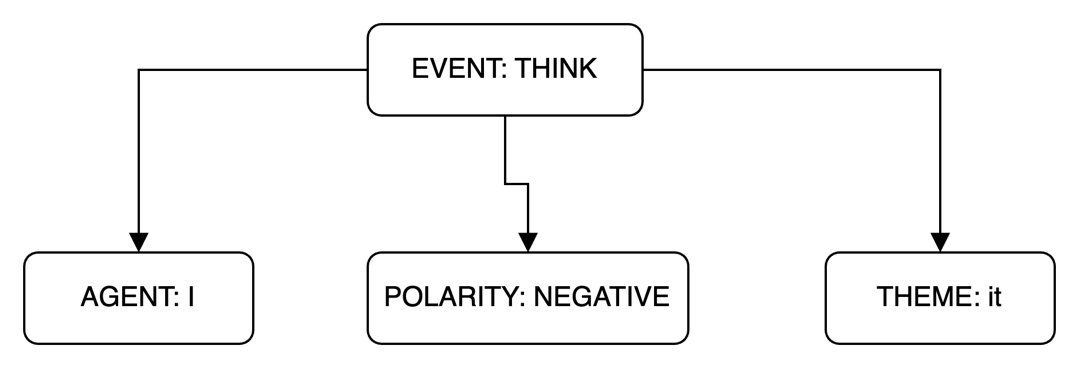


Figure 5.0 abstracts the meaning of the sentence into core roles:

1. Event: The action of thinking.
2. Agent: The person doing the thinking (implicitly “I” in Spanish and explicitly “I” in English).
3. Polarity: The negative aspect (indicated by “never”).
4. Theme: The content of the thought (“it”).

This abstract view is helpful in machine translation because it captures the essence of the sentence independently of language-specific details (Jurafsky & Martin, 2007).

1. **Discussion on Intermediate Representations for Machine Translation**

What this exercise really shows is that even a short sentence like “Nunca lo pienso” hides a surprisingly complex structure beneath its surface. When you compare it to its English counterpart, “I never think it,” you notice that Spanish leaves out the subject because the verb itself implies it, whereas English has to state “I” explicitly. This difference shows why having an intermediate representation is so crucial. It helps capture the underlying meaning (or semantic roles) of the sentence without getting tangled up in the grammatical details specific to each language.

Essentially, both sentences express the same idea: the act of thinking, marked by a negative twist, yet each language arranges that idea in its own unique way. As Jurafsky and Martin (2007) point out, focusing on core elements like the event, agent, and theme can bridge these differences, making it easier for machine translation systems to produce translations that are both faithful to the original meaning and fluent in the target language. In a way, this intermediate layer acts as a common ground where the essence of what’s being said is preserved, regardless of the linguistic packaging, suggesting that more advanced translation models should lean on these deeper representations to capture the true intent behind the words (Jurafsky & Martin, 2007).

**Reference**

Jurafsky, D., & Martin, J. H. (2007). Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition (Draft, October 30, 2007). Pearson.