Case Study L1-006
Semantic Integrity in Patent NMT
Morphological Calque (The "Electro-Sans" Neologism)

Cédric Stéphany — Technical Translation & AI Alignment Specialist

Case Study Metadata

Dataset ID:	L1-006
Category:	Semantic Integrity — Level 1
Focus:	Prefix/Suffix Morphology
Model:	Generic NMT
Domain:	Semiconductor Manufacturing

1 The Context: Technical Negation

In English technical terminology, the suffix **"-less"** indicates the absence of a standard component or condition (e.g., *Wireless* = Sans fil, *Brushless* = Sans balais). However, "Electroless" is a specific chemical term. It does not mean "Without electricity" in a general sense; it refers to an **autocatalytic chemical reduction process** that occurs without an external current source.

Key Concept

The "LEGO Block" Fallacy:
Generic NMT models often treat words as Lego blocks. They see "Electro-less" and translate it piece-by-piece:

- Electro → *Électro*
- Less → *Sans* (Without)

They then stick them together to form a neologism that violates the target language's morphological rules.

2 The Glitch: "Electro-Without"

In Claim 13, the generic model invented a non-existent French word to describe a standard deposition process.

2.1 Forensic Evidence (Claim 13)

2.2 Why This Matters

- **Linguistic Gibberish:** "Électro-sans" is meaningless to a French engineer. It reads like "Electro-Without." It signals a complete lack of human review.

Source Term (English)	NMT Output (Calque)	Golden Rewrite (Correct)
"...electro-less dip growth..."	× "...croissance par immersion **électro-sans**..." (Non-existent word)	"...croissance par immersion **autocatalytique**..." (Standard Term)

Table 1: Morphological Hallucination in Technical Process

- **Loss of Meaning:** By stripping the concept of "Autocatalysis" (Self-catalyzing), the translation loses the defining chemical mechanism of the invention.
- **Enablement Risk:** A skilled person in the art (PHOSITA) reading "Électro-sans" might not be able to reproduce the method, as the term does not correspond to any known chemical protocol.

3 Alignment Methodology

3.1 Suffix-to-Concept Mapping

To fix this, we implement a **Morphological Parsing** rule that forbids literal translation of privative suffixes in technical compounds.

Alignment Methodology

The "-Less" Protocol:

1. **Token Analysis:** Identify tokens ending in `-less` that are NOT in the standard dictionary (e.g., *Wireless* is standard, *Electroless* is domain-specific).
2. **Concept Lookup:** Map the root (*Electro*) + Suffix (*Less*) to the Domain Concept (*Chemical Deposition*).
3. **Force Target:** Retrieve the domain equivalent: `Electroless` → `Autocatalytique` OR `Chimique`.
4. **Negative Constraint:** Explicitly ban the string "Électro-sans" from the output generation.

4 Key Insights

Key Concept

What This Case Study Demonstrates:

1. **Translation is not Transliteration:** You cannot just translate parts of a word and hope the whole makes sense.
2. **The Danger of Neologisms:** AI models are too creative. They will invent new words ("Électro-sans") rather than admit they don't know the technical term.
3. **Domain Specificity:** "Electroless" is a standard term in Semiconductor Manufacturing (H01L) but unknown in general conversation. The model must be fine-tuned on the specific IPC class.

Portfolio: Patent Translation AI Alignment Framework
Author: Cédric Stéphany
Specialization: Technical Translation (FR↔EN) — Patents, Telecommunications, Semiconductors
Contact: cedric@tmcwx.com
Last Updated: January 19, 2026