Case Study C4-003
Structural Compliance in Patent NMT
Nested Determiner Chains (The "Double-Two" Tangle)

Cédric Stéphany — Technical Translation & AI Alignment Specialist

Case Study Metadata

Dataset ID:	C4-003
Category:	Structural Compliance — Constraint 4
Focus:	Syntactic Linearity / Quantifiers
Model:	Generic NMT
Domain:	Medical Devices / Signal Processing

1 The Context: Recursive Quantifiers

Patent claims often use complex, recursive structures to define subsets of data, such as: *"...a second two of the plurality of amplitude values..."*

This structure contains:

1. **Ordinal:** Second (Sequence)
2. **Cardinal:** Two (Quantity)
3. **Partitive:** Of the plurality (Set membership)

Key Concept

The Linearity Bias:
Generic NMT models process tokens left-to-right. When faced with stacked determiners ([Ordinal] + [Cardinal]), they often fail to construct the correct hierarchical tree in the target language, resulting in a "Word Salad" where adjectives become nouns.

2 The Glitch: "Two Seconds of Values"

In Claim 8, the generic model mistranslated the ordinal "Second" as the time unit "Second," destroying the claim's logic.

2.1 Forensic Evidence (Claim 8)

2.2 Why This Matters

- **Category Error:** The claim defines *Amplitude Values* (Voltage/Pressure), not *Time*. Translating "Second" as a noun (*seconde*) implies a time duration, rendering the claim physically impossible.

Source Phrase (English)	NMT Output (Hallucination)	Golden Rewrite (Correct)
"...a **second two** of the plurality..."	× "...deux **secondes** de la pluralité..." (Meaning: 2 units of time)	"...deux **autres** de la pluralité..." (Meaning: Two others)

Table 1: Syntactic Failure in Nested Quantifiers

- **Reference Failure:** The phrase "a second two" is meant to reference a specific subset of data points distinct from the "first two." The mistranslation breaks this reference chain.
- **Indefiniteness:** A claim that confuses data amplitude with time duration is invalid under 35 U.S.C. § 112 / Art 84 EPC.

3 Alignment Methodology

3.1 Determiner Unpacking Protocol

To fix this, we enforce a **Syntactic Restructuring** rule for stacked quantifiers.

Alignment Methodology

Quantifier Parsing Rule:

1. **Pattern Recognition:** Detect `[Ordinal] + [Cardinal] + [OF]`.
2. **Semantic Tagging:** Tag "Second" as `ORDINAL_ADJ`, not `NOUN_TIME`.
3. **Safety Protocol ("Added Matter"):**
   - While "A second pair" (*Une seconde paire*) is fluent, it introduces the noun "Pair" which is absent in the source. This risks an "Added Matter" objection.
   - **Golden Choice:** We map "A second two" to the functional equivalent: **"Deux autres"** (Two others). This preserves the set logic without injecting new nouns.
4. **Transformation:** `Second Two` → `Deux Autres` (Two others from the set).

4 Key Insights

Key Concept

What This Case Study Demonstrates:

1. **POS Tagging Failure:** The AI misidentified the Part of Speech (POS) for "Second," treating it as a Noun instead of an Adjective.
2. **Structure over Vocab:** The error isn't just about the word "Second"; it's about the model's inability to parse the [Ordinal-Cardinal] hierarchy.
3. **Defensive Translation:** In patent law, a slightly less specific word ("Autres") is safer than an overly specific noun ("Paire") that wasn't in the original text.

Portfolio: Patent Translation AI Alignment Framework
Author: Cédric Stéphany
Specialization: Technical Translation (FR↔EN) — Patents, Telecommunications, Semiconductors
Contact: cedric@tmcwx.com
Last Updated: January 21, 2026