

Case Study C2-003

Structural Compliance in Patent NMT

Morpho-Syntactic Alignment of French Method Claims

Cédric Stéphany — Technical Translation & AI Alignment Specialist

Case Study Metadata

Dataset ID: C2-003

Category: Structural Compliance — Constraint 1

Focus: Verb Nominalization

Model: Generic NMT

Domain: Photonics / Optics

1 The Context: Parallelism is Non-Negotiable

In patent claims, particularly method claims involving multiple steps, **grammatical parallelism** is a strict requirement. If the first step of a method is nominalized (converted to a noun), all subsequent steps linked by "and" or semicolons must follow the exact same grammatical pattern.

Key Concept

The "Sibling Rule":

A method claim cannot switch voices mid-sentence.

- **Correct:** ...comprising: **applying** (Noun)...; and **applying** (Noun)...
- **Incorrect:** ...comprising: **applying** (Noun)...; and **to apply** (Infinitive)...

This mixed structure is not only stylistically poor but can trigger objections for lack of clarity (EPC Art. 84), as it obscures the hierarchical relationship between the steps.

2 The Glitch: "Infinitive Drift" (Context Loss)

Generic NMT models often exhibit a "short attention span." In this case study, the model correctly identifies the nominalization requirement for the *first* step immediately following the trigger word "comprising."

However, as the sentence lengthens, the model loses the structural context. By the time it reaches the second step (separated by text and a semicolon), it reverts to its default training behavior: translating the English gerund ("applying") as a French infinitive ("appliquer").

Critical Issue

The Error: The model generates a "Frankenstein" claim:

1. Step 1: *l'application de...* (Correct Noun)
2. Step 2: *...et appliquer de la chaleur...* (Incorrect Infinitive)

This inconsistency signals to the patent examiner that the document was machine-translated and lacks professional review.

3 The Alignment Challenge

3.1 The Translation Failure

Source (English)	AI Hallucination (Failure)	Golden Rewrite (Correct)
"...comprises: applying a kerf pattern...; and applying heat..."	<p>× Structural Drift:</p> <p>"...comprend : l'application d'un motif...; et appliquer de la chaleur..."</p> <p>(Mixed Forms: Noun + Infinitive)</p>	<p>Parallel Compliance:</p> <p>"...comprend : l'application d'un motif...; et l'application de la chaleur..."</p> <p>(Consistent Nominalization)</p>

Table 1: Parallelism Failure in Multi-Step Method Claims

4 Alignment Methodology

To fix this, we trained the model to enforce **Long-Range Dependency** checks using Label Studio relations.

Alignment Methodology

Annotation Process:

1. **Sibling Linking:** Annotators do not just link the verb to the main trigger ("comprising"). They explicitly link the **second verb** to the **first verb**.
2. **Relation Type:** Step_1 (Noun) → ENFORCES_PARALLELISM → Step_2 (Candidate).
3. **Constraint Logic:** "If Sibling A is a Noun, Sibling B must be a Noun." This explicitly penalizes the model for switching grammatical forms after a coordinator ("and").

This teaches the model that "and" acts as a **grammatical mirror**, reflecting the structure of the preceding element rather than starting a new sentence.

5 Results & Impact

5.1 Quality Assurance Metrics

- **Consistency Score:** 100% parallelism achieved in test sets containing lists of 3+ method steps.
- **Readability:** Improved flow and professional tone that matches human-drafted patent applications.
- **Risk Reduction:** Elimination of "mixed-mode" syntax that frequently triggers clerical objections during formalities examination.

Portfolio: Patent Translation AI Alignment Framework

Author: Cédric Stéphany

Specialization: Technical Translation (FR↔EN) — Patents, Telecommunications, Semiconductors

Last Updated: January 5, 2026