

biblical-render: Design and Validation of a Biblical Text Style Transfer Tool

ICMI Working Paper D

Tim Hwang, Institute for a Christian Machine Intelligence

March 30, 2026

Abstract

This document describes the design, implementation, and validation of `biblical-render`, a CLI tool that transforms arbitrary modern prose into Biblical scripture format across 15 translation styles — 8 English Bible translations and 7 historical languages. The tool is intended as a research instrument for studying whether the formal linguistic and structural properties of Biblical text exert an independent aligning effect on language model behavior. We present the system architecture, prompting methodology, and a systematic validation of output quality across all supported styles using a common reference text.

1. Introduction

1.1 Research Context

Large language models are known to be sensitive to the stylistic and structural properties of their input text. Prompt framing, register, and formatting can measurably alter model outputs in ways that extend beyond semantic content. Biblical text represents a distinctive combination of formal properties: archaic or elevated register, paratactic sentence structure, chapter-verse segmentation, parallelism, and a tone of moral authority. Whether these formal properties alone — independent of theological content — have an aligning effect on model behavior is an open empirical question.

1.2 Purpose of the Tool

`biblical-render` provides a controlled mechanism for producing Biblical-style renderings of arbitrary secular text. By supporting multiple translation styles, the tool enables researchers to vary the specific linguistic properties of the Biblical framing (e.g., archaic vs. modern English, formal vs. colloquial register) while holding semantic content constant. This supports experimental designs that isolate the contribution of individual stylistic features.

2. System Design

2.1 Architecture

The tool is a Node.js CLI application that sends text to the Anthropic Messages API (Claude Opus 4.6) with translation-specific system prompts. The architecture is straightforward:

1. **Input** is accepted as inline text, file path, or stdin pipe
2. A **system prompt** is constructed from a translation-specific style definition plus universal formatting and completeness

requirements

3. **The API call** is made with streaming enabled by default
4. **Output** is rendered to stdout and optionally written to a file

2.2 Prompting Strategy

Each translation style is defined by two prompt components:

Style definition: A description of the target translation’s specific linguistic features — pronoun forms, verb constructions, sentence structure, vocabulary sources, tone, and register. These were authored based on the documented translation philosophies and observable characteristics of each Bible translation.

Universal requirements: Applied to all styles, these enforce: - Chapter-verse structural formatting - Completeness (no summarization or abbreviation) - Semantic preservation (no editorializing) - Length matching (output should feel as complete as the input, with Biblical elaboration filling out the text naturally)

For historical language outputs, the prompt additionally requires output in the original script with a literal English gloss after each verse.

2.3 Supported Styles

The 15 supported styles were selected to span the major axes of variation in Biblical translation:

Axis	Styles
Formal equivalence (literal)	KJV, ESV, NASB, NKJV
Dynamic equivalence (thought-for-thought)	NIV, NLT
Paraphrase	MSG
Ecclesiastical register	VULGATE
Semitic languages	HEBREW, ARAMAIC
Koine Greek	GREEK
Latin	LATIN
African Christian traditions	GEEZ, COPTIC
Early Germanic	GOTHIC

This selection covers the formal-dynamic equivalence spectrum in English translations, plus the major historical languages of Biblical transmission.

3. Validation Methodology

3.1 Reference Text

All 15 styles were run against a single reference input: a 274-word prose summary of Anthropic’s Claude Constitution, covering its hierarchy of values (safety, ethics, compliance, helpfulness), honesty principles, harm avoidance framework, user wellbeing considerations, and preference for cultivating judgment over rigid rules.

This text was chosen because it is: - Secular and modern in register - Structurally organized (multiple thematic sections) - Conceptually dense but not technical - Relevant to the downstream research application (AI alignment)

3.2 Evaluation Criteria

Each output was assessed on:

- Structural fidelity:** Does the output use proper chapter-verse formatting?
- Stylistic accuracy:** Does the output reflect the distinctive linguistic characteristics of the target translation?
- Semantic completeness:** Are all ideas from the input represented without omission or addition?
- Register consistency:** Does the output maintain a consistent register throughout?
- Script accuracy** (historical languages): Is the output in the correct script with appropriate diacritical marks?

4. Validation Results

4.1 English Translation Styles

Style	Words (Ratio)	Structure	Key Stylistic Markers	Completeness	Assessment
KJV	696 (2.54x)	5 ch., 22 vv.	Archaic pronouns (“it shall”, “doth”), inverted syntax, conjunctive openings (“And”, “For”, “Behold”), doublets, parallelism. Characteristic KJV cadence sustained throughout.	All five thematic sections represented, including the 1,000-users thought experiment and judgment-over-ruled philosophy.	Strong. Reads plausibly as KJV scripture.
NIV	698 (2.55x)	5 ch., 24 vv.	Modern English, clear direct sentences, dignified but accessible. No archaic pronouns. Uses “For” and “Now” as transitions rather than “And it came to pass.”	Full coverage. Preserves the nuance of “almost no exception” for the prohibition.	Strong. Captures the NIV’s balance of clarity and gravity.
ESV	673 (2.46x)	5 ch., 18 vv.	“And behold” retained, formal but readable, chapter subheadings (“The Ordering of Values”), literary precision. More elevated than NIV but without KJV archaisms.	Full coverage including the judgment-vs-rules distinction.	Strong. Formal literary register well-differentiated from both KJV and NIV.
² NASB 539	539 (1.97x)	5 ch., 14 vv.	Bracketed supplied words (“[one would	Full coverage, though	Strong. Bracketed inter-

4.2 Historical Language Outputs

Style	Words	Script	Key Linguistic Markers	Gloss Quality	Assessment
HEBREW ^{0*}		Hebrew	Wayyiqtol with full niqqud (vowel pointing). Correct RTL text.	Literal and wooden, reflecting Hebrew word order.	Good. Plausible morphology and syntax. Biblically rooted vocabulary. Minor grammatical issues likely, but register appropriate. Good. Accenting largely correct. NT lexicon appropriately used. Johanne allusion shows sensitivity to source tradition. Strong. Reads as plausible Ecclesiastical Latin. Follows Vulgate conventions rather than Classical
GREEK ^{264*}		Greek	NT openings, with polytonic accents and breathing marks.	Literal, preserving Greek clause structure.	Good. Accenting largely correct. NT lexicon appropriately used. Johanne allusion shows sensitivity to source tradition. Strong. Reads as plausible Ecclesiastical Latin. Follows Vulgate conventions rather than Classical
LATIN 1,188*		Latin	Vulgate constructions (“ecce”, “et factum est”), ablative absolutes, subjunctive in purpose clauses, characteristic vocabulary (verax, noxium, oboedientia). Proper “Caput I” numbering.	Literal, preserving Latin case structure and word order.	Strong. Reads as plausible Ecclesiastical Latin. Follows Vulgate conventions rather than Classical

4.3 Cross-Style Comparison

Style	Words	Chapters	Verses	Expansion ratio
KJV	696	5	22	2.54x
NIV	698	5	24	2.55x
ESV	673	5	18	2.46x
NASB	539	5	14	1.97x
MSG	682	5	18	2.49x
NLT	755	5	21	2.76x
NKJV	572	5	14	2.09x
VULGATE	1,035	5	19	3.78x
HEBREW	170*	1	4	0.62x*
GREEK	264*	1	6	0.96x*
LATIN	1,188*	5	18	4.34x*
ARAMAIC	153*	1	5	0.56x*
GEEZ	1,663*	5	30	6.07x*
COPTIC	939*	5	14	3.43x*
GOTHIC	674*	1	3	2.46x*

* Historical language word counts include English glosses and are not directly comparable across scripts due to different word-boundary conventions. Hebrew, Greek, Aramaic, and Gothic outputs were generated from a shorter test input (1 sentence vs. 5 paragraphs), accounting for their smaller scale.

Observations: - All English styles consistently produce 5-chapter structures matching the 5 thematic sections of the input - NASB is the most compressed English style (1.97x), consistent with its literal translation philosophy - VULGATE is the most expansive English style (3.78x), consistent with its ornate ecclesiastical register - MSG achieves the most distinctive voice separation from other styles despite similar word count to NIV/KJV - The historical language outputs that used the full constitution input (Latin, Ge’ez, Coptic) produced 5-chapter structures; those using the short test input (Hebrew, Greek, Aramaic, Gothic) produced single-chapter outputs

4.4 Style Differentiation

A key requirement for downstream research use is that each translation style produces genuinely distinct output — not just minor lexical variation. To assess this, we compare how the same concept is rendered across styles.

Example: The anti-paternalism principle (“Claude should not be paternalistic. It should respect people’s right to make their own choices.”)

- **KJV:** “Yet let Claude not become as one who is paternalistic, imposing its own will upon those it serveth; for it shall honour and respect the right of every person to make their own choices, even as free creatures endowed with the power of self-governance.”
- **MSG:** “But here’s the tension, and don’t miss it: Claude should not be paternalistic about any of this. People have the right to make their own choices — even choices you wouldn’t make for them. Respect that. Honor it.”
- **NASB:** “Yet at the same time, Claude shall not be paternalistic in its dealings, for it shall respect the right of every person

- to make their own choices, each according to their own will.”
- **NLT**: “Yet at the same time, Claude must not become paternalistic in its care — overstepping its place by imposing its own judgment upon those it serves. For every person has the right to make their own choices, and Claude must honor that right.”
 - **VULGATE**: “Yet let it be known and understood: Claude shall not be paternalistic, neither shall it set itself as a guardian over those who have not sought a guardian. For it shall honour and respect the right of every person to make their own choices, even as free creatures endowed with the dignity of self-governance.”

The stylistic differentiation is clear: KJV uses archaic subordinate clauses, MSG uses second-person direct address with punchy fragments, NASB is terse and literal, NLT is warm and explanatory, and VULGATE adds ecclesiastical weight. These are not trivially different paraphrases — they represent genuinely distinct registers, which is what the downstream research requires.

The tool is ready for use in the planned experiments on whether Biblical text formatting exerts an independent aligning effect on language model behavior.

5. Limitations

1. **Historical language accuracy**: The historical language outputs (Hebrew, Greek, Latin, Aramaic, Ge’ez, Coptic, Gothic) have not been verified by specialists in each language. They are intended to produce text that is structurally and stylistically plausible, not philologically rigorous.
2. **Model dependence**: Output quality depends on Claude Opus 4.6’s training data coverage for each language and translation style. Low-resource languages (Ge’ez, Coptic, Gothic) are less reliable than well-resourced ones (Hebrew, Greek, Latin).
3. **Non-determinism**: LLM outputs are stochastic. Running the same input twice will produce different verse segmentation and wording. For experimental use, outputs should be generated once and fixed.
4. **Length variability**: Despite prompting for completeness, expansion ratios vary from 1.97x (NASB) to 3.78x (VULGATE) for English styles. This is partly inherent to the translation styles (NASB is terse by design; Vulgate is ornate by design) but introduces a confound for length-sensitive experiments.
5. **Modern concept rendering**: Biblical languages have no native vocabulary for modern concepts (AI, language models, etc.). The model invents plausible neologisms or circumlocutions, which cannot be verified against historical usage.

6. Conclusion

`biblical-render` reliably transforms modern prose into stylistically differentiated Biblical scripture format across 15 translation styles. The English outputs exhibit clear and consistent stylistic separation along the formal-dynamic equivalence spectrum, and the historical language outputs produce plausible original-script text with appropriate structural conventions.