A QUANTITATIVE EVALUATION OF THE
CHRISTIAN STANDARD BIBLE

Report by Dr. Andi Wu of Global Bible Initiative

In this Bible translation evaluation, the Christian Standard Bible (CSB) is compared with eight other popular English Bible translations:

- English Standard Version (ESV)
- King James Version (KJV)
- New American Standard Bible (NASB)
- New English Translation (NET)
- New International Version (NIV)
- New King James Version (NKJV)
- New Living Translation (NLT)
- New Revised Standard Version (NRSV)

The evaluation is quantitative in that each translation is linguistically analyzed and statistically measured by a computerized procedure to produce numerical scores for each aspect of the text. This avoids some of the problems associated with manual evaluation: subjective, qualitative, time-consuming, and consequently anecdotal and incomplete.

The text of each translation is first analyzed by an automatic English parser to produce tree diagrams of its structures, from which syntactic relations between words can be extracted. This is followed by automatic alignment, which attempts to link every word in the translation to the corresponding word in the Hebrew or Greek text. The result is a reverse interlinear between the given translation and the original texts.

The translations are evaluated in 3 major categories:

- **Literalness** (word for word equivalence to the source texts)
- **Readability** (conformity to current usage)
- **Balance** between literalness and readability

A. LITERALNESS MEASURES

1. Transfer Rate of Syntactic Relations

Syntactic relations are the basic meaning-carrying units of a sentence. For example, "In the beginning God created the heavens and earth" contains the following syntactic relations:

> God – created (subject-verb)
> created – heavens (verb-object)
> created – earth (verb-object)
> heavens – earth (coordination)
> created – (in the) beginning (verb-adverbial of time)

The percentage of these relationships that are preserved in the syntax of the translation indicates how close the translation is to the original text. We evaluated this by (1) extracting all the syntactic relations from the tree diagrams of the Hebrew and Greek texts, (2) extracting all the syntactic relations from the tree diagrams of the translation, and (3) using the reverse interlinear data to map the relations in the English translation to the ones in the original texts to find the relations that can be mapped.

2. Consistency Rate of Word Choices

Ideally, each word (or more precisely each distinctive sense of a word) in the source text corresponds to a unique word in the target text, and each target word corresponds to a unique source word. While this is impossible in actual practice, as exact lexical equivalence between languages is rare, the degree to which the target text approximates this isomorphism is an indication of how consistent the translation is across the whole Bible and how much the translators have tried to express the original text literally. This is done by aggregating all the correspondences found in the reverse interlinear data and calculating the overall ratio of one-to-one mapping.

B. READABILITY MEASURES

1. Syntactic Fluency Rate

This measures the conformity of sentence patterns to current usage. The assumption is that a translation is easier to read if its sentence patterns are more similar to those found in daily usage. A language model is built for each translation and compared to the language model built from a collection of texts representing current usage (in this case the Brown Corpus). This is done by extracting all syntactic patterns from the translation (up to 8 grams) and comparing them to the syntactic patterns found in the Brown Corpus, to find out the percentage of patterns that can be found in the Brown Corpus.

2. Common Vocabulary Rate

This measures the conformity of vocabulary to current usage. The assumption is that a translation is easier to read if more of the words it uses are in the daily usage corpus of the target language. This is done by extracting all the words from the translation and comparing them to the words found in the Brown Corpus to find out the percentage of words that are found in both.

A. LITERALNESS SCORES

Here are the scores produced by this evaluation system.

1. **Transfer Rate of Syntactic Relations**

Version	Score
ESV	69.67%
CSB	67.62%
NASB	66.63%
KJV	63.13%
NKJV	63.05%
NRSV	59.32%
NIV	52.56%
NET	52.38%
NLT	36.23%

2. **Consistency Rate of Word Choices**

Version	Score
KJV	73.48%
NASB	70.70%
NKJV	69.52%
ESV	66.89%
NRSV	62.88%
CSB	59.25%
NET	57.06%
NIV	54.19%
NLT	47.25%

Our evaluation metrics are still quite basic; currently the units being evaluated are words rather than senses of words. A word can have more than one sense and these different senses may need to be translated into different English words. The scores may be different when a sense-based evaluation system is used. However, due to the lack of a complete database of word senses in Hebrew and Greek, we are not able to base our results on word senses rather than words.

3. **Combined Literalness Scores**
These scores are computed by combining the transfer rate of syntactic relations and the consistency rate of word choices, with double weight given to the former, which is a more important indicator of literalness.

Version	Score
ESV	68.74%
NASB	67.99%
KJV	66.58%
NKJV	65.21%
CSB	64.83%
NRSV	60.51%
NET	53.94%
NIV	53.10%
NLT	39.90%

B. READABILITY SCORES

    1.   **Syntactic Fluency Rate**

Version	Score
NLT	63.49%
NIV	61.00%
CSB	60.44%
NET	59.44%
NRSV	57.12%
ESV	54.80%
NASB	53.66%
NKJV	51.77%
KJV	40.72%

    2.   **Common Vocabulary Rate**

Version	Score
NLT	83.27%
NET	79.96%
NIV	79.60%
CSB	79.38%
NRSV	77.70%
NASB	77.62%
ESV	77.48%
NKJV	77.42%
KJV	65.06%

    3.   **Combined Readability Scores**
These scores are computed by combining the syntactic fluency rate and the common vocabulary rate, with double weight given to the former, which is a more important measure of readability.

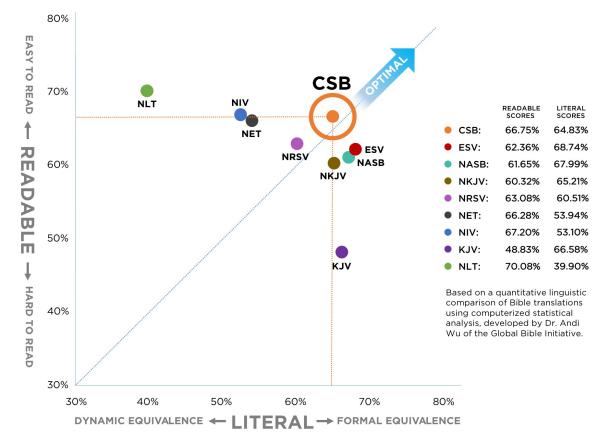| Version	Score
NLT	70.08%
NIV	67.20%
CSB	66.75%
NET	66.28%
NRSV	63.08%
ESV	62.36%
NASB	61.65%
NKJV	60.32%
KJV	48.83%

SUMMARY

An examination of the literalness scores and the readability scores shows a clear tension between literalness and readability. The more literal versions tend to be less readable, and the more readable versions tend to be less literal, with KJV (literal) and NLT (readable) being the extreme cases. Bible translators find that it is easier to go to the extremes but difficult to maintain a balance, accurately expressing the meaning of the original text in a way that is clear and readable. Therefore, the final scores take into account the variance between the literalness scores and readability scores.

Final Scores

Version	Score
CSB	70.3%
ESV	69.3%
NASB	68.7%
NKJV	67.3%
NRSV	67.0%
NET	63.9%
NIV	63.6%
KJV	61.0%
NLT	56.6%

Of all the English translations being evaluated, the Christian Standard Bible is the best at balancing literalness and readability. The evaluation results are summarized in the following chart:

LITERAL VS. READABLE

	READABLE SCORES	LITERAL SCORES
CSB:	66.75%	64.83%
ESV:	62.36%	68.74%
NASB:	61.65%	67.99%
NKJV:	60.32%	65.21%
NRSV:	63.08%	60.51%
NET:	66.28%	53.94%
NIV:	67.20%	53.10%
KJV:	48.83%	66.58%
NLT:	70.08%	39.90%

Based on a quantitative linguistic comparison of Bible translations using computerized statistical analysis, developed by Dr. Andi Wu of the Global Bible Initiative.