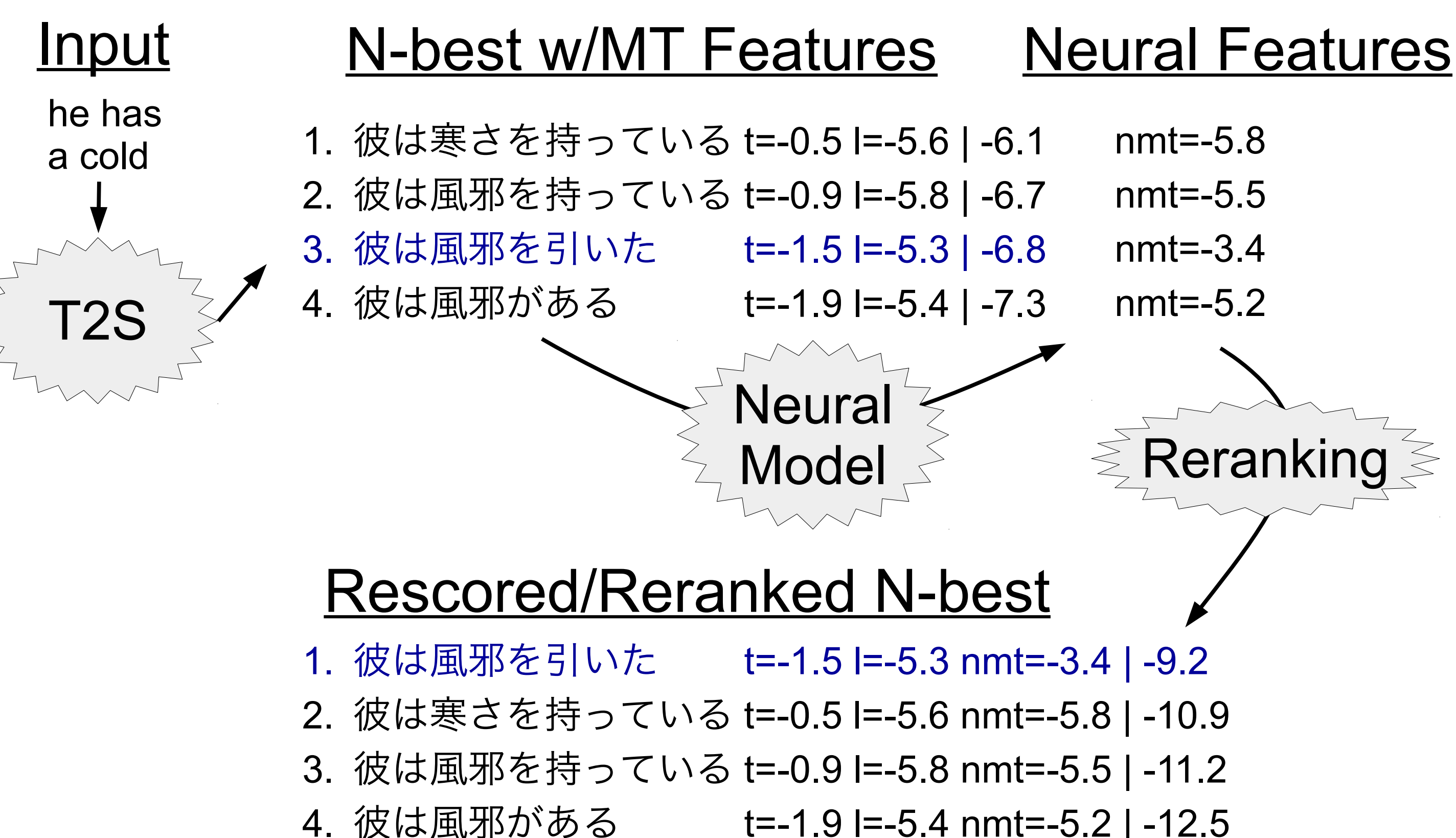


Research Questions

1. Does reranking with neural MT models **improve subjective impressions** of translation results?
2. If so, what are the **qualitative differences** between reranked and non-reranked output?
3. How big of an **n-best list** do we need?

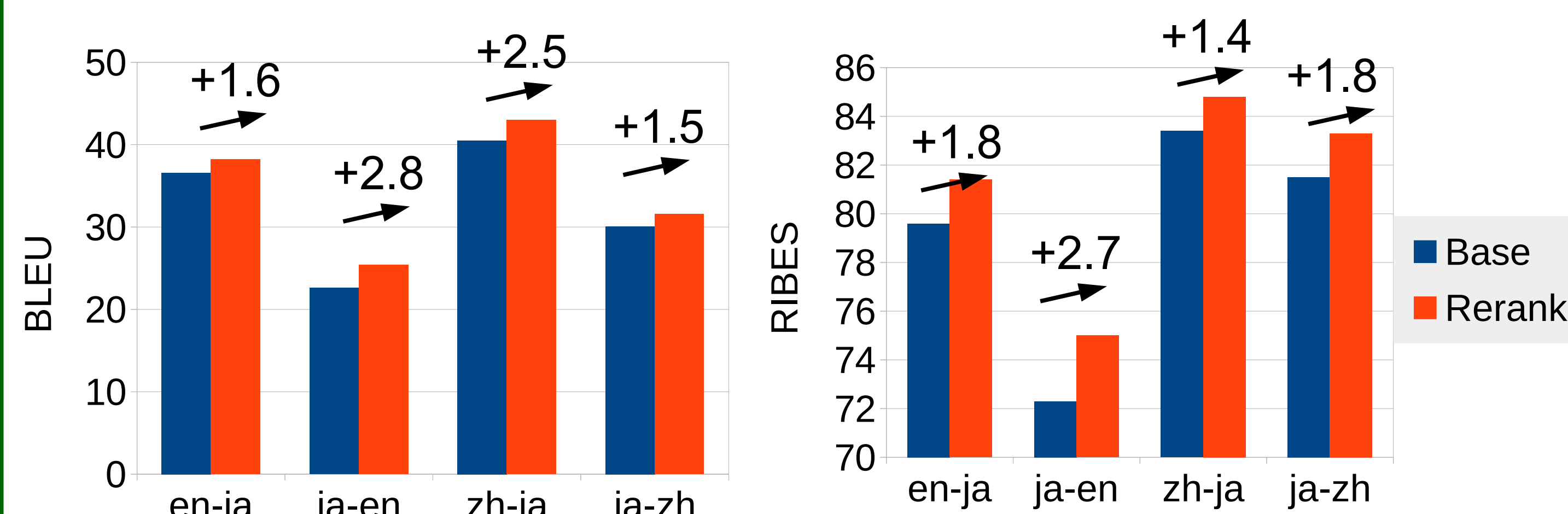
Neural Reranking



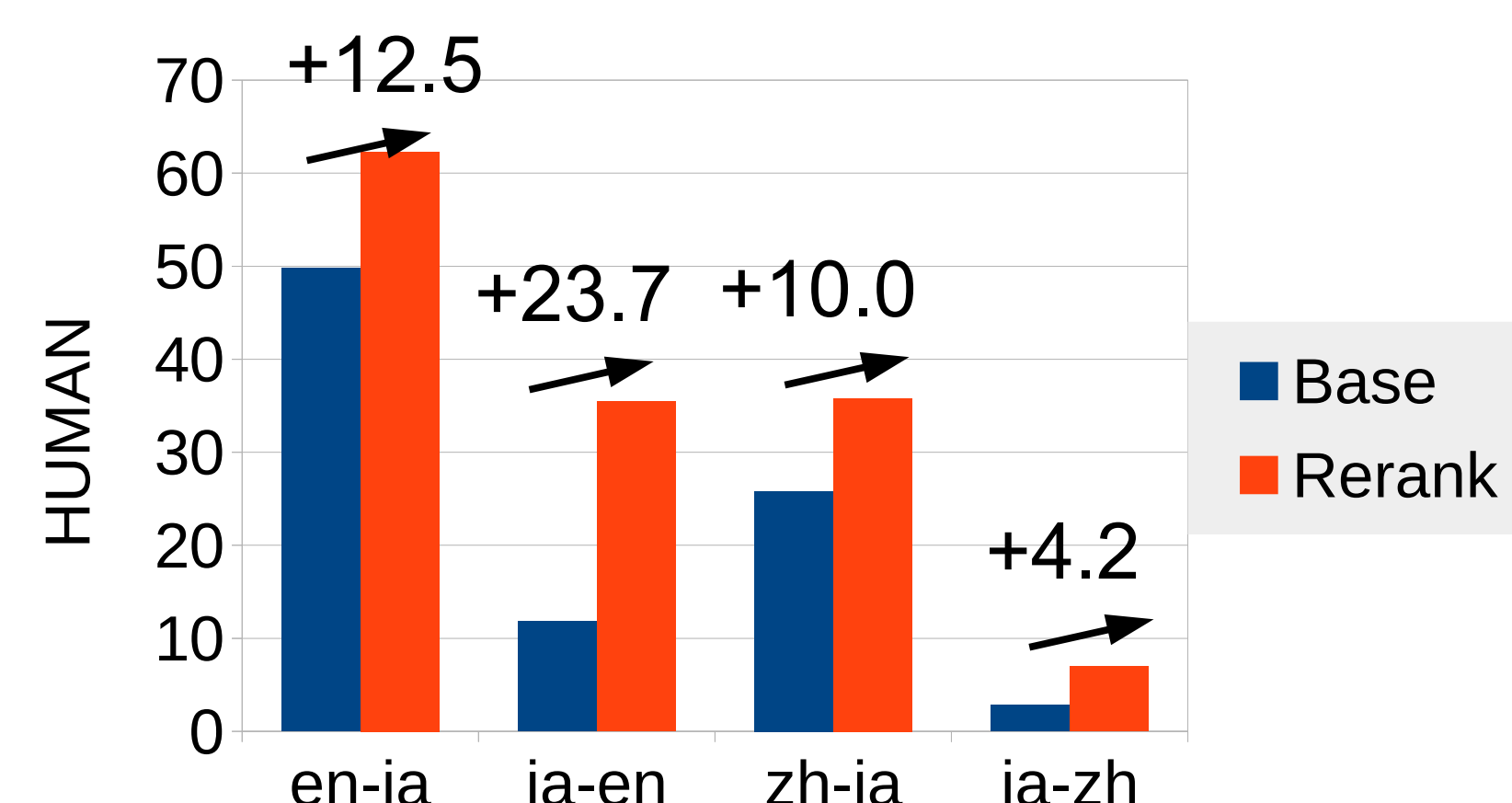
Experimental Setup

- **Data:** ASPEC Scientific Abstracts
 - Japanese ↔ English, Chinese
- **Baseline:** NAIST WAT2014 Tree-to-String System
 - Strong baseline achieving high scores
 - Implemented using Travatar (<http://phontron.com/travatar>)
- **Neural MT Model:** Attentional model
 - Trained ~500k sent., 256 hidden nodes, 2 model ensemble
 - Use words occurring 3+ times (vocab 50,000~80,000)
 - Trained w/ lamtram (<http://github.com/neubig/lamtram>)
- **Automatic Evaluation:** BLEU, RIBES
- **Manual Evaluation:** WAT 2015 HUMAN Score

Quantitative Results

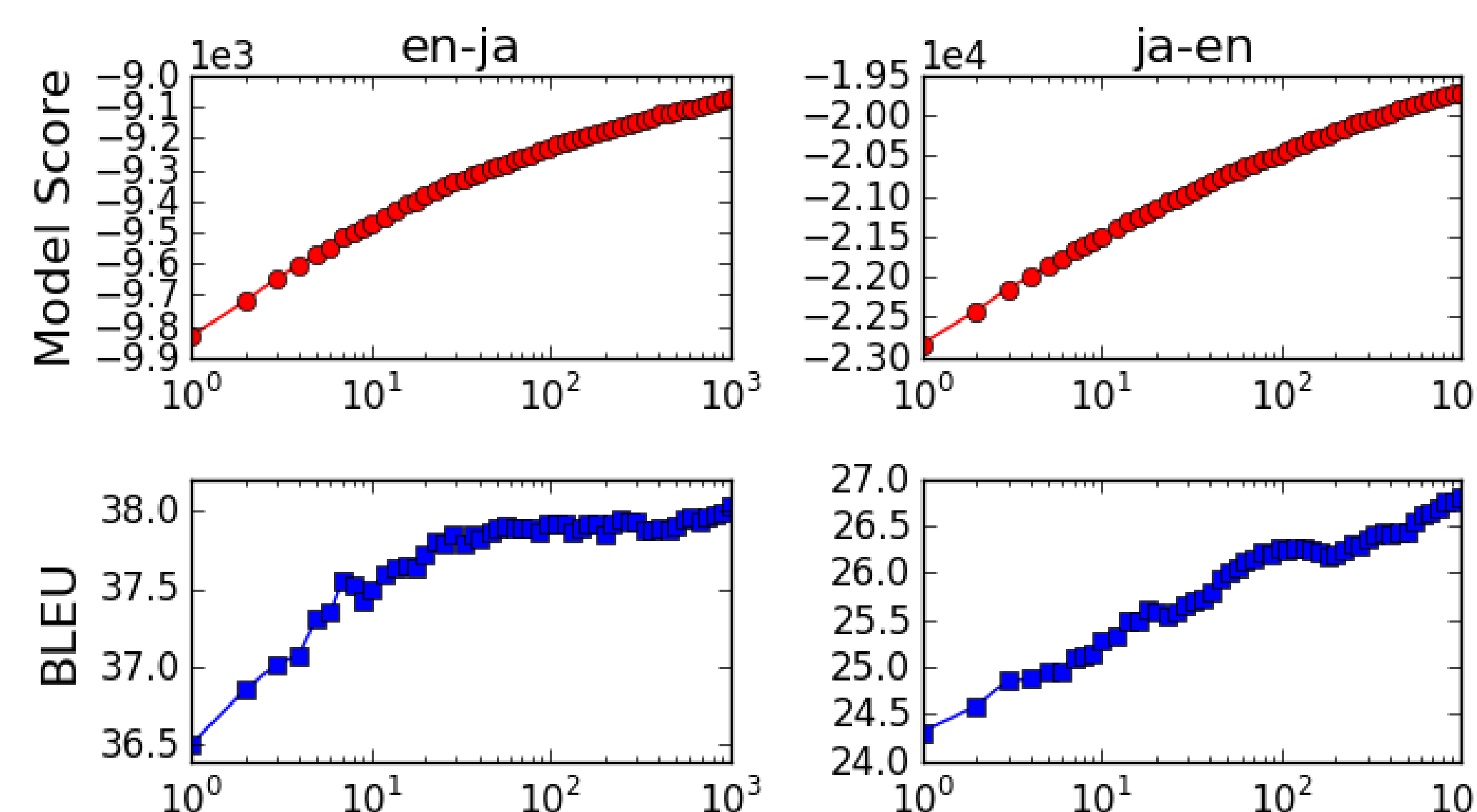


Confirm what we know: Neural reranking helps BLEU.



Show what we didn't know: Also helps manual eval.

Reranking and N-best Size



- All data sets improve **approximately log-linear**
- **Little saturation** even at 100-best

Qualitative Analysis

Type	Improved	Degraded	% Impr.
Reordering	55	9	86%
Deletion	20	10	67%
Insertion	19	2	90%
Substitution	15	11	58%
Conjugation	8	1	89%
Total	117	33	78%

Overall improvements re-confirmed
Particularly reordering, insertion, and conjugation errors

Detailed Analysis

#1 Improvement: Phrasal Reordering (+26, -4)

Source 症例2においては、直腸がんの肝転移に対する化学療法中に、発赤、硬結、皮膚潰瘍を生じた。
Ref In case 2, **reddening, induration, and skin ulcer** appeared during **chemical therapy** for liver metastasis of rectal cancer.
Base In case 2, **occurred** during chemotherapy for liver metastasis of rectal cancer, **flare, induration, skin ulcer**.
Rerank In case 2, **the flare, induration, skin ulcer** was produced during the chemotherapy for hepatic metastasis of rectal cancer.

#2 Improvement: Auxiliary Verb Ins./Del. (+15, -0)

Source これにより得られる支配方程式は壁面乱流のようなせん断乱流にも有用である。
Ref Governing equation derived by this method is useful for turbulent shear flow like turbulent flow near wall.
Base The governing equation **is** obtained by this is also useful for such as wall turbulence shear flow.
Rerank The governing equation obtained by this is also useful for shear flow such as wall turbulence.

#3 Improvement: Coordinate Structures (+13, -2)

Source レーザー加工は高密度光束による局所的な加熱とアブレーションにより行う。
Ref Laser work is done by **local heating and ablation** with high density light flux.
Base The laser processing is carried out by **local heating** by high-density luminous flux and **ablation**.
Rerank The laser processing is carried out by **local heating and ablation** by high-density flux.

#4 Improvement: Verb Agreement (+6, -0)

Source ラングミュア-プロジェクト法や包接化にも触れた。
Ref Langmuir-Blodgett method and inclusion compounds **are** mentioned.
Base Langmuir-Blodgett method and inclusion **is** also discussed.
Rerank Langmuir-Blodgett method and inclusion **are** also mentioned.

What Didn't Work: Terminology (+2, -4)

Source 放射熱を利用する赤外線応用計測が応力解析に役立っている。
Ref Infrared ray applied measurement using **radiant heat** is useful for stress analysis.
Base The infrared application measurement using **radiant heat** is useful in the stress analysis.
Rerank Infrared ray application measurement using **radiation heat** is useful for stress analysis.

Conclusion

- Reranking with neural MT models leads to subjective improvements in MT quality
- Future work includes comparisons with neural language models or neural MT w/o reranking