

ニューラル機械翻訳のための 日本語の膠着語的性質を考慮したマルチタスク学習

西田悠人 秋葉友良 塚田元
豊橋技術科学大学

{nishida.yuto.ec, akiba.tomoyoshi.tk, tsukada.hajime.hl}@tut.jp

概要

ステミングタスクを補助タスクとしたマルチタスク学習 (MTL) モデルが膠着語に属する言語のニューラル機械翻訳に有効であるという報告があるが、膠着語である日本語の翻訳にも有効であるかは明らかではない。そこで、本研究では日本語-英語および日本語-中国語の双方向翻訳においてステミングを基本としたタスクを補助タスクとした MTL モデルの性能の評価を行った。実験結果より、ステミングタスクを補助タスクとした MTL モデルの有効性は限定的であることと、ステミングタスクと英語文/中国語文をソースとした日本語ステム列生成タスクの両方を補助タスクとした MTL モデルは有効であることを示した。

1 はじめに

膠着語に属する言語は morphologically-rich (形態学的に複雑) であり、それらの言語の翻訳は、特に学習データが少ない場合に、ニューラル機械翻訳 (NMT) モデルにとって依然として困難なタスクであることが知られている [1]。Pan ら [1] はこの問題を解決するために、トルコ語-英語およびウイグル語-中国語の双方向翻訳において、膠着語 (トルコ語、ウイグル語) のステミングタスクを補助タスクとしたマルチタスク学習 (MTL) モデルを提案し、翻訳性能が向上することを示した。しかし、他の言語対 (膠着語-非膠着語の対) において同様の手法が有効であるかは明らかではない。

本研究では、ケーススタディとして、膠着語のなかでも日本語に着目する。日本語-英語および日本語-中国語の双方向翻訳を対象として、ステミングタスクを補助タスクとした MTL モデルの性能を評価することで、Pan ら [1] の手法が日本語の翻訳においても有効であるかを検証する。また、その他のス

テミングを基本としたタスクについても検討し、より有効な補助タスクを探究する。

実験結果より、(1) Pan ら [1] の提案した手法に相当する日本語ステミングタスク $ST_{j \rightarrow j}$ を補助タスクとした MTL モデルの有効性が限定的であることを示し、(2) $ST_{j \rightarrow j}$ と新たに提案した英語文/中国語文をソースとした日本語ステム列生成タスク $ST_{f \rightarrow j}$ の両方を補助タスクとした MTL モデルによって翻訳性能が向上することを示し、(3) 翻訳タスクと補助タスクの混合比を調整することで (2) のモデルに更なる性能改善がみられることを示唆した。

2 関連研究

NMT にマルチタスク学習を導入するための様々な手法が提案されている。翻訳以外の補助タスクを用いて主タスク (翻訳タスク) の性能向上を図る研究として、Luong ら [2] はドイツ語・英語のオートエンコーダと独英翻訳のマルチタスク学習を提案し、Currey ら [3] は英語の句構造解析と英語を原言語とした翻訳 (英露翻訳など) のマルチタスク学習を提案した。あるいは、主タスクと補助タスクを区別せず、複数の翻訳タスクを共同で学習し、その相乗効果によって翻訳性能の向上を図る研究もある。Johnson ら [4] は単一のモデルによる多言語翻訳によって低資源言語の翻訳性能が向上することを報告した。

膠着語を対象とした NMT の研究も多くある。Ataman ら [5] は教師なしの形態学的セグメンテーションモデルによってトルコ語の語彙削減を行うことで、土英翻訳の性能が向上することを報告した。Pan ら [6] は膠着語の語幹と接辞に別の処理を施すサブワード分割手法を提案し、土英翻訳およびウイグル語-中国語翻訳の性能向上を示した。

3 日本語の膠着語的性質を考慮したマルチタスク学習

3.1 マルチタスク学習

タスクごとに固有のタグを各ソース文の先頭に加える方法 [4] により, 単一の seq2seq モデルでマルチタスク学習を行う. これにより, 標準的な NMT アーキテクチャに変更を加えることなく, マルチタスク学習を行うことができる. 本研究では, 主タスク (翻訳タスク) であることを示すタグ<MT>と補助タスクであることを示すタグ<ST>を導入する. 本研究における MTL モデルの概念図を図 1 に示す.

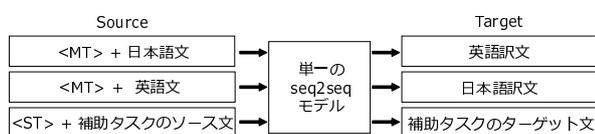


図 1 MTL モデルの概念図

3.2 補助タスク

日本語の膠着語的性質を考慮した補助タスクとして, Pan ら [1] が提案した補助タスクに相当する $ST_{j \rightarrow j}$ と, 新たに提案した $ST_{f \rightarrow j}$ の 2 つを導入する. $ST_{j \rightarrow j}$ は日本語文を入力として, 対応する日本語ステム列を生成するタスクであり, $ST_{f \rightarrow j}$ は英語文/中国語文を入力として, 入力に対する日本語訳文に対応する日本語ステム列を生成するタスクである. なお, 日本語の文節は典型的には 1 つの自立語に 0 個以上の付属語が膠着することで形成されるため, 本研究では日本語文の自立語のみを抽出して原形化することで得られるトークン列を日本語ステム列と定義した.

日本語-英語の双方向翻訳を対象とした際の主タスクである日英翻訳, 英日翻訳と補助タスクである $ST_{j \rightarrow j}$, $ST_{f \rightarrow j}$ の訓練データの例を表 1 に示す.

表 1 タスク毎の訓練データの例 (上: ソース文, 下: ターゲット文)

日英 翻訳	<MT> 実験によってこの結果が検証された。 this result was verified by the experiments .
英日 翻訳	<MT> this result was verified by the experiments . 実験によってこの結果が検証された。
$ST_{j \rightarrow j}$	<ST> 実験によってこの結果が検証された。 実験 この結果 検証 する
$ST_{f \rightarrow j}$	<ST> this result was verified by the experiments . 実験 この結果 検証 する

4 評価実験

4.1 データセット

対訳データとして, KFTT [7] の京都関連文書対訳コーパス全文対 (440,288 文対), IWSLT2017 [8] の英日対訳コーパス全文対 (223,108 文対), Asian Scientific Paper Excerpt Corpus (ASPEC) [9] から ASPEC-JE100k (ASPEC-JE の train1 データの先頭 100,000 文対), ASPEC-JC100k (ASPEC-JC の train データの先頭 100,000 文対) を用いる. 開発データおよびテストデータには, それぞれのコーパスの開発/テストデータを用いた. なお, IWSLT2017 では開発データは dev2010, テストデータは tst2010 を用いた. データの日本語文は MeCab [10] によるトークナイズを行った. MeCab の辞書には IPA 辞書を用いた. 英語文は Moses [11] tokenizer によるトークナイズ, Moses truecaser による truecasing を行った. 中国語文は jieba [12] によるトークナイズを行った.

補助タスクの訓練データは, 対訳データから作成した. $ST_{j \rightarrow j}$ と $ST_{f \rightarrow j}$ の訓練データのターゲット文は対訳データの日本語文に対して自立語抽出と語の原形化を行うことで作成できる. 自立語抽出は, MeCab によって文の各形態素の品詞を解析し, 自立語に相当する品詞に該当する形態素以外を除去することで行った. なお, 自立語に相当する品詞は (1) 名詞-動詞非自立的, 名詞-接尾, 名詞-非自立, 名詞-特殊を除くすべての名詞, (2) 連体詞, (3) 動詞-自立, (4) 形容詞-自立, (5) 副詞, (6) 接続詞, (7) 感動詞, (8) フィラーとした. 品詞の分類は IPA 品詞体系 [13] に基づく. 語の原形化は, 各形態素を MeCab によって得た原形に置換することで行った.

4.2 実験条件

モデルの語彙は, 対訳データの英語文/中国語文と日本語文を結合したうえで作成した. 語彙サイズはソース, ターゲットともに同一とし, KFTT では 100,000, その他では 50,000 とした.

翻訳モデルおよび MTL モデルはオープンソースの NMT ツールキット OpenNMT [14] の Python 実装である OpenNMT-py によって構築した.

モデルの翻訳性能の評価指標には BLEU [15] を用いた. モデルの訓練で得たチェックポイントから翻訳方向ごとに開発データの BLEU が最も高いものを選択し, そのチェックポイントにおけるテストデー

タの BLEU をそのモデルの性能とした。

4.3 ステミングを基本とした補助タスクを用いた実験

本節では、ステミングを基本としたタスクを補助タスクとした MTL モデルの有効性を評価するために、次に示す 4 つのモデルについて訓練および評価を行った。

bi-MT 補助タスクを用いない双方向翻訳モデル。これがベースラインモデルとなる。

bi-MT + ST_{j→j} ST_{j→j} を補助タスクとした MTL モデル。Pan ら [1] の提案したモデルに相当する。

bi-MT + ST_{f→j} 新たに提案した ST_{f→j} を補助タスクとした MTL モデル。

bi-MT + ST_{j→j}&ST_{f→j} ST_{j→j} と ST_{f→j} の両方を補助タスクとした MTL モデル。

MTL モデルの訓練に用いるミニバッチは、各ミニバッチに占める日英(日中)翻訳タスク、英日(中日)翻訳タスク、補助タスクのデータの割合が等しくなるように構成した。すなわち、各ミニバッチの全体(主タスク+補助タスク)は、ミニバッチの主タスクの 1.5 倍となる。また、対訳データとともに用いる補助タスクのデータはその対訳データから作成し、外部データは使用しなかった。

各モデルの各翻訳方向の BLEU (Ja-En/Zh および En/Zh-Ja) とその算術平均 BLEU (Ave.) を表 2 に示す。

表 2 より、bi-MT+ST_{j→j} と bi-MT+ST_{f→j} の平均 BLEU は、IWSLT2017 ではベースライン性能を下回り、その他ではベースライン性能を上回った。また、各翻訳方向でみると、約半数のケースでベースライン性能を下回っていることが確認できる。

対照的に、bi-MT+ST_{j→j}&ST_{f→j} はベースラインに対して一貫して平均 BLEU が向上し、その向上幅は一貫して bi-MT+ST_{j→j} および bi-MT+ST_{f→j} より大きい。また、各翻訳方向でみても、ASPEC-JC100k の Ja-Zh を除いて BLEU が向上したことが分かる。

以上より、日本語-英語および日本語-中国語の双方向翻訳において、Pan ら [1] の提案した補助タスク ST_{j→j} と新たに提案した補助タスク ST_{f→j} は単独では有効性が限定的であることと、両方を組み合わせて用いた MTL モデル bi-MT+ST_{j→j}&ST_{f→j} は有効であることが示された。

4.4 恒等写像タスクとの比較実験

Luong ら [2] は独英翻訳において、ドイツ語文と英語文の恒等写像タスクを補助タスクとした MTL モデルが有効であることを示した。そこで、本研究でも日本語文と英語文/中国語文の恒等写像タスクを補助タスク IM として導入する¹⁾。

Luong ら [2] は翻訳タスクと補助タスクの混合比を 1 : 0.05 としたとき、MTL モデルの翻訳性能が最も高いことを報告している。本研究でも同様の条件で実験を行うために、訓練に用いる各ミニバッチの全体が、ミニバッチの主タスクの 1.05 倍となるようにミニバッチを構成し、MTL モデル bi-MT+IM (1.05x) を訓練した。各訓練データにおけるモデルの BLEU を表 2 に示す。

表 2 より、bi-MT+IM (1.05x) は一貫して平均 BLEU がベースラインを上回っていることが確認できる。このことより、IM を補助タスクに用いることは本研究の実験条件においても有効であることが示された。しかし、bi-MT+ST_{j→j}&ST_{f→j} の平均 BLEU と比較すると、一貫して bi-MT+IM (1.05x) の平均 BLEU のほうが低く、bi-MT+ST_{j→j}&ST_{f→j} のほうがより有効性が高いことが分かる。

4.5 タスク混合比による性能変化

Luong ら [2] は独英翻訳と恒等写像の MTL モデルにおいて、タスク全体に占める恒等写像タスクの割合が高いときにモデルの翻訳性能が低下することを報告している。そのため、前節までのモデルにおいても、タスク混合比を調整することで性能が改善する可能性がある。

そこで、本節では ASPEC-JE100k と ASPEC-JC100k を対象として、bi-MT+ST_{j→j}、bi-MT+ST_{f→j}、bi-MT+ST_{j→j}&ST_{f→j} について、タスク混合比を変えて実験を行い、タスク混合比の調整による性能向上を図る。

4.3 節では、ミニバッチを構成する主タスクと補助タスクの比を 10 : 5 としたが、本節では 10 : 4, 10 : 3, 10 : 2, 10 : 1 について実験を行った。各開発データに対する各翻訳方向の算術平均 BLEU が

1) ただし、Luong ら [2] は恒等写像タスクの訓練データとして外部の単言語資源を用いているのに対して、本研究では外部の単言語資源を用いずに、恒等写像タスクの訓練データを対訳コーパスから作成するという点で異なる。また、本研究とは異なり、Luong ら [2] は MTL モデルとして、複数のエンコーダおよび複数のデコーダから構成されるものを用いている。

表2 ステミングを基本としたタスクまたは恒等写像タスクを補助タスクとして用いた MTL モデルの BLEU

Model	KFTT			IWSLT2017			ASPEC-JE100k			ASPEC-JC100k		
	Ja-En	En-Ja	Ave.	Ja-En	En-Ja	Ave.	Ja-En	En-Ja	Ave.	Ja-Zh	Zh-Ja	Ave.
bi-MT (baseline)	21.33	26.30	23.82	10.48	10.35	10.42	20.54	26.84	23.69	23.17	29.21	26.19
bi-MT+ST _{j→j}	22.11	26.27	24.19	10.43	9.97	10.20	21.52	26.39	23.96	23.27	29.35	26.31
bi-MT+ST _{f→j}	20.71	27.45	24.08	9.64	10.89	10.27	20.66	27.30	23.98	22.76	29.90	26.33
bi-MT+ST _{j→j} &ST _{f→j}	21.87	27.59	24.73	10.67	10.97	10.82	21.19	26.80	24.00	22.92	30.10	26.51
bi-MT+IM (1.05x)	21.50	26.72	24.11	10.18	10.80	10.49	21.16	26.59	23.88	23.28	29.62	26.45

最も高いモデルを 10 : 5 の場合も含めた 5 つのモデルから選択することで、最良のタスク混合比を選択した。実験の結果、ASPEC-JE100k の bi-MT+ST_{j→j} では 10 : 1 が、bi-MT+ST_{f→j} では 10 : 2 が、bi-MT+ST_{j→j}&ST_{f→j} では 10 : 4 が、ASPEC-JC100k の bi-MT+ST_{j→j} では 10 : 4 が、bi-MT+ST_{f→j} では 10 : 4 が、bi-MT+ST_{j→j}&ST_{f→j} では 10 : 2 が最良の混合比であることが分かった。最良の混合比におけるモデルの (テストデータに対する) BLEU を 4.3 節で実験した 10 : 5 の場合のモデルの BLEU とともに表 3 に示す。なお、表中の E/Z は En/Zh の略記であり、 $\Delta Ave.$ は各モデルの平均 BLEU とベースラインモデルの平均 BLEU の差を示す。また、各混合比における BLEU については付録 A.1 に示す。

表3 タスク混合比による MTL モデルの BLEU の変化

Model	Ja-E/Z	E/Z-Ja	Ave.	($\Delta Ave.$)
ASPEC-JE100k				
bi-MT (baseline)	20.54	26.84	23.69	
+ST _{j→j} (1.5x)	21.52	26.39	23.96	(+0.27)
+ST _{f→j} (1.5x)	20.66	27.30	23.98	(+0.29)
+ST _{j→j} &ST _{f→j} (1.5x)	21.19	26.80	24.00	(+0.31)
+ST _{j→j} (1.1x)	21.25	26.49	23.87	(+0.18)
+ST _{f→j} (1.2x)	21.34	26.68	24.01	(+0.32)
+ST _{j→j} &ST _{f→j} (1.4x)	21.54	27.12	24.33	(+0.64)
ASPEC-JC100k				
bi-MT (baseline)	23.17	29.21	26.19	
+ST _{j→j} (1.5x)	23.27	29.35	26.31	(+0.12)
+ST _{f→j} (1.5x)	22.76	29.90	26.33	(+0.14)
+ST _{j→j} &ST _{f→j} (1.5x)	22.92	30.10	26.51	(+0.32)
+ST _{j→j} (1.4x)	23.60	30.48	27.04	(+0.85)
+ST _{f→j} (1.4x)	23.36	31.26	27.31	(+1.12)
+ST _{j→j} &ST _{f→j} (1.2x)	23.63	30.69	27.16	(+0.97)

表 3 より、ASPEC-JE100k の bi-MT+ST_{j→j} を除いて、タスク混合比を調整したモデルの平均 BLEU が 4.3 節のモデルの平均 BLEU を上回っており、タスク混合比を調整することで MTL モデルの性能が向上することが示唆された。また、最良のタスク混合比のモデル同士で比較した場合でも、4.3 節の結果と同様に、bi-MT+ST_{j→j} よりも、bi-MT+ST_{j→j}&ST_{f→j} の平均 BLEU のほうが高いという傾向がみられた。

5 おわりに

本研究では、膠着語のニューラル機械翻訳のケーススタディとして、日本語-英語および日本語-中国語の双方向翻訳を対象とし、(1) Pan ら [1] の提案したモデルが日本語の翻訳においても有効であるかどうかの検証、(2) そのほかのステミングを基本としたタスクの提案、(3) 前項のタスクを補助タスクとした MTL モデルの有効性の評価、(4) MTL モデルのタスク混合比の調整を行った。

実験の結果から、日本語-英語および日本語-中国語の双方向翻訳においては、Pan ら [1] の提案した補助タスクに相当するタスク ST_{j→j} を補助タスクとした MTL モデル bi-MT+ST_{j→j} の有効性は限定的であることが示され、ST_{j→j} と新たに提案した英語文/中国語文をソースとした日本語ステム列生成タスク ST_{f→j} の両方を補助タスクとした MTL モデル bi-MT+ST_{j→j}&ST_{f→j} はベースラインに対する一貫した各翻訳方向の平均 BLEU の改善をもたらすことが示された。また、bi-MT+ST_{j→j}&ST_{f→j} は恒等写像タスクを補助タスクとした MTL モデルより翻訳性能が高いことも実験により確かめた。さらに、双方向翻訳タスクと補助タスクの混合比を調整することによって前述のモデルの性能が更に向上することが示唆された。

今後の課題として、本研究で提案したモデルが他の言語対でも有効であるかの調査や、Pan ら [6] の提案したサブワード分割手法の導入、より翻訳性能を向上させる補助タスクの探究、bi-MT+ST_{j→j}&ST_{f→j} の補助タスクに外部の単言語資源を利用できるように逆翻訳 [16] を用いることが挙げられる。

謝辞

本研究は JSPS 科研費 19K11980 および 18H01062 の助成を受けた。

参考文献

- [1] Yirong Pan, Xiao Li, Yating Yang, and Rui Dong. Multi-task neural model for agglutinative language translation. In **Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop**, pp. 103–110, Online, July 2020. Association for Computational Linguistics.
- [2] Minh-Thang Luong, Quoc V. Le, Ilya Sutskever, Oriol Vinyals, and Lukasz Kaiser. Multi-task sequence to sequence learning. **arXiv preprint arXiv:1511.06114**, 2016.
- [3] Anna Currey and Kenneth Heafield. Incorporating source syntax into transformer-based neural machine translation. In **Proceedings of the Fourth Conference on Machine Translation (Volume 1: Research Papers)**, pp. 24–33, Florence, Italy, August 2019. Association for Computational Linguistics.
- [4] Melvin Johnson, Mike Schuster, Quoc V. Le, Maxim Krikun, Yonghui Wu, Zhifeng Chen, Nikhil Thorat, Fernanda Viégas, Martin Wattenberg, Greg Corrado, Macduff Hughes, and Jeffrey Dean. Google’s multilingual neural machine translation system: Enabling zero-shot translation. **Transactions of the Association for Computational Linguistics**, Vol. 5, pp. 339–351, 2017.
- [5] Duygu Ataman, Matteo Negri, Marco Turchi, and Marcello Federico. Linguistically motivated vocabulary reduction for neural machine translation from turkish to english. **The Prague Bulletin of Mathematical Linguistics**, Vol. 108, pp. 331–342, 2017.
- [6] Yirong Pan, Xiao Li, Yating Yang, and Rui Dong. Morphological word segmentation on agglutinative languages for neural machine translation. **arXiv preprint arXiv:2001.01589**, 2020.
- [7] Graham Neubig. The Kyoto free translation task. <http://www.phontron.com/kfft>, 2011.
- [8] Mauro Cettolo, Marcello Federico, Luisa Bentivogli, Niehues Jan, Stüker Sebastian, Sudoh Katsutho, Yoshino Koichiro, and Federmann Christian. Overview of the IWSLT 2017 evaluation campaign. In **International Workshop on Spoken Language Translation**, pp. 2–14, 2017.
- [9] Toshiaki Nakazawa, Manabu Yaguchi, Kiyotaka Uchimoto, Masao Utiyama, Eiichiro Sumita, Sadao Kurohashi, and Hitoshi Isahara. Aspec: Asian scientific paper excerpt corpus. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, **Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2016)**, pp. 2204–2208, Portorož, Slovenia, may 2016. European Language Resources Association (ELRA).
- [10] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to Japanese morphological analysis. In **Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing**, pp. 230–237, Barcelona, Spain, July 2004. Association for Computational Linguistics.
- [11] Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondřej Bojar, Alexandra Constantin, and Evan Herbst. Moses: Open source toolkit for statistical machine translation. In **Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions**, pp. 177–180, Prague, Czech Republic, June 2007. Association for Computational Linguistics.
- [12] Sun Junyi. jieba. <https://github.com/fxsjy/jieba>. Accessed: 2021-12-25.
- [13] 形態素解析ツールの品詞体系. <https://www.unixuser.org/~euske/doc/postag/#chasen>. Accessed: 2021-12-24.
- [14] Guillaume Klein, Yoon Kim, Yuntian Deng, Jean Senellart, and Alexander Rush. OpenNMT: Open-source toolkit for neural machine translation. In **Proceedings of ACL 2017, System Demonstrations**, pp. 67–72, Vancouver, Canada, July 2017. Association for Computational Linguistics.
- [15] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. BLEU: a method for automatic evaluation of machine translation. In **Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics**, pp. 311–318, Philadelphia, Pennsylvania, USA, July 2002. Association for Computational Linguistics.
- [16] Rico Sennrich, Barry Haddow, and Alexandra Birch. Improving neural machine translation models with monolingual data. In **Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 86–96, Berlin, Germany, August 2016. Association for Computational Linguistics.

A 付録

A.1 各混合比における MTL モデルの翻訳性能

4.5 節で実験した各混合比における MTL モデルの翻訳性能 (BLEU) を表 4 に示す.

Model	ASPEC-JE100k						ASPEC-JC100k					
	Ja-En		En-Ja		Ave.		Ja-Zh		Zh-Ja		Ave.	
	dev	test	dev	test	dev	test	dev	test	dev	test	dev	test
bi-MT+ST _{j→j} (1.5x)	20.63	21.52	26.54	26.39	23.59	23.96	23.23	23.27	30.10	29.35	26.67	26.31
bi-MT+ST _{j→j} (1.4x)	20.81	20.98	26.73	26.24	23.77	23.61	23.19	23.60	30.70	30.48	26.95	27.04
bi-MT+ST _{j→j} (1.3x)	20.78	21.08	27.00	26.47	23.89	23.78	22.92	23.23	30.48	30.57	26.70	26.90
bi-MT+ST _{j→j} (1.2x)	20.87	21.24	26.89	26.73	23.88	23.99	23.14	22.98	30.66	30.28	26.90	26.63
bi-MT+ST _{j→j} (1.1x)	20.85	21.25	27.12	26.49	23.99	23.87	23.16	23.40	30.58	30.63	26.87	27.02
bi-MT+ST _{f→j} (1.5x)	20.88	20.66	27.01	27.30	23.95	23.98	23.26	22.76	30.47	29.90	26.87	26.33
bi-MT+ST _{f→j} (1.4x)	20.93	21.17	27.19	27.21	24.06	24.19	23.58	23.36	31.48	31.26	27.53	27.31
bi-MT+ST _{f→j} (1.3x)	20.99	21.14	27.20	26.93	24.10	24.04	22.79	22.64	31.59	30.85	27.19	26.75
bi-MT+ST _{f→j} (1.2x)	21.12	21.34	27.42	26.68	24.27	24.01	23.38	23.34	31.39	30.99	27.39	27.17
bi-MT+ST _{f→j} (1.1x)	20.70	21.28	27.33	27.21	24.02	24.25	23.06	22.81	30.93	30.87	27.00	26.84
bi-MT+ST _{j→j} &ST _{f→j} (1.5x)	20.96	21.19	27.16	26.80	24.06	24.00	22.96	22.92	30.28	30.10	26.62	26.51
bi-MT+ST _{j→j} &ST _{f→j} (1.4x)	21.02	21.54	27.40	27.12	24.21	24.33	22.97	22.93	31.32	31.13	27.15	27.03
bi-MT+ST _{j→j} &ST _{f→j} (1.3x)	21.06	21.08	27.03	26.90	24.05	23.99	22.89	23.16	31.15	30.65	27.02	26.91
bi-MT+ST _{j→j} &ST _{f→j} (1.2x)	21.04	21.28	26.83	26.29	23.94	23.79	23.31	23.63	31.32	30.69	27.32	27.16
bi-MT+ST _{j→j} &ST _{f→j} (1.1x)	20.77	21.24	27.08	27.03	23.93	24.14	22.96	22.87	31.02	30.57	26.99	26.72