Translation Patterns

OshonSoft Project
www.oshonsoft.com

Four models and their results

Four fresh 50,613-parameter MLPs learned from 20,000 examples each. A and AX generate translated Subject, Predicate and Object lines before assembling the final Translation. B and BX give only the final translation. A/B share fixed-pattern input sentences, with one adjective before each noun. AX/BX share a second corpus in which either adjective can be absent.

Model	Selected epoch	Whole reply + EOT	Translation	Correct EOT
A	10	900/900	900/900	900/900
AX	7	888/900	893/900	900/900
B	10	900/900	900/900	900/900
BX	10	888/900	888/900	900/900

The four experiments have been repeated with the current Lab training and generation code. The download includes their initial and selected models and complete saved replies. These runs reproduced the reported learning curves and final scores. The final questions had already been evaluated in the original study, so this repetition checks reproduction of that result.

Understand the word-order task

The fixed dictionary has 160 nouns, 60 adjectives and 30 verbs. English articles disappear. The target uses Serbian base forms and Subject-Object-Verb order. Each noun precedes its adjective. This controlled exercise does not teach natural Serbian agreement, inflection or sentence style; random combinations may have unusual meanings.

In A/AX, intermediate Subject and Object translations retain the original English phrase order. The final line reverses the words of each translated phrase and appends the translated Predicate unchanged. The model learns every intermediate token; a wrong intermediate line can affect later generation. The reference below illustrates the two answer formats.

```
USER Translate: the quiet dog sees the sweet deer.

A reference:
Subject: tih pas
Predicate: videti
Object: sladak jelen
Translation: pas tih jelen sladak videti.

B reference:
pas tih jelen sladak videti.
```

AX/BX train on four patterns, 5,000 examples each: no adjectives, object adjective only, subject adjective only, and both adjectives. A/B use 20,000 examples of the last pattern. A missing adjective has no placeholder. Evaluation contains new sentences within trained patterns, not an unseen grammatical structure.

Learning progression

Each entry counts exact complete replies, including EOT, out of 450 Training check questions.

Epoch	A	AX	B	BX
1	441	65	387	36
2	450	299	447	199
3	450	403	450	326
4	450	418	450	381
5	450	436	450	411
6	450	444	450	427
7	450	450	450	433
8	450	449	450	438
9	450	448	450	442
10	450	449	450	443

Project	Targets per epoch	Ten-epoch updates
A	460000	4600000
AX	420000	4200000
B	140000	1400000
BX	120000	1200000

One epoch visits each selected next-token target once. A has 23 targets per answer including labels, line breaks and EOT. AX has 19, 21 or 23; B has seven; BX has five, six or seven. Compare both epoch and update axes. The website includes all per-line scores and 176 saved reply views. B first reached perfect checks at epoch 3; AX reached perfect whole replies at epoch 7 and then scored 449, 448 and 449. BX reached 443/450 at epoch 10.

Selection, final errors and interpretation

The rule was fixed before training: complete ten epochs, choose the most exact whole Training check replies, and choose the later checkpoint on a tie. Each choice was locked before evaluating its 900 final questions. A/B share final input sentences. AX/BX share another set with 225 questions per pattern.

Pattern	AX whole reply	BX whole reply
Neither adjective	224/225	225/225
Object adjective only	220/225	225/225
Subject adjective only	223/225	222/225
Both adjectives	221/225	216/225

AX's final line scores were Subject 900/900, Predicate 897/900, Object 895/900 and Translation 893/900. Its complete reply was correct in 888 cases. Both AX and BX stopped correctly in all 900 cases. For one saved AX error, `Translate: the shark needs the brave moon.`, the generated Object was `hrabar seme` instead of `hrabar mesec`, and the incorrect noun carried into the final translation. The download preserves all twelve AX and all twelve BX errors in the complete Final evaluation answers.

A/B and AX/BX are paired comparisons of answer format with matching input sentences, dimensions, initialization and rate. Intermediate answers change the vocabulary and total update budget. B versus BX also changes the corpus, pattern exposure, question set and epoch target count. There is one seed per configuration. These results show learning in the measured recipes, not an isolated causal estimate of word position or guaranteed repeatability.

Files and settings

Each project has one `experiment.json`, a `training/` folder and a `questions/` folder. Final questions are under `questions/final/`. The released snapshots belong in `models/initial.osai` and `models/selected.osai`. A new execution creates its own `runs/run-000001/`, with settings, frozen inputs, checkpoints and complete Prompt/Gold/Raw answers. The download contains those inputs and both model snapshots. Supplied replies are grouped in `answers/update-000000000/`, followed by one folder for each measured update; `answers/final/` contains all final replies. Each file keeps the original question filename and the complete Prompt/Gold/Raw records.

Use a writable folder such as Documents/OshonSoft AI Language Model Lab. The application stays in Program Files. Training results are written to the project folder and do not require administrator rights. Larger research projects are separate downloads; they are not part of the ordinary-example installer.

Use `translation_patterns/A`, `AX`, `B` or `BX`. Preserve `sentences_01.txt` through `sentences_10.txt`, `dev_random.txt`, `train_probe.txt`, `questions/final/test_random.txt` and `lexicon.json`. The known-training probe contains 90 examples; it is not an evaluation of all 20,000 training replies.

Setting	Value
Architecture	MLP: embeddings, concatenation, tanh, softmax
V / C / D / H	520 / 38 / 16 / 37
Parameters	50613
Active tokens	A/AX 516; B/BX 511

Setting	Value
Tokenizer	Words and punctuation; Preserve line breaks
Initialization	Wider embeddings +/-1.0; seed 12345
SGD / objective	0.01; MODEL body and EOT
Scheduler	Global epochs
Epoch target / checks	10 / every epoch
Generation	Greedy; reply limit 40
Data seeds	A/B 4309021; AX/BX 4309031
Longest prediction prefix	A/AX 36; B/BX 20 tokens

Repeat the training and try a model

On an AMD Ryzen 7 7735HS laptop, the repeated runs took about 3 h 48 min (A), 3 h 25 min (AX), 1 h 4 min (B) and 1 h 4 min (BX). These times include training and all configured checks while several projects shared the CPU. Time spent paused is excluded. Other hardware and workloads will give different timings.

1. Extract the chosen project into the writable application folder. Open the installed Lab and select the **Project** page.

2. Click **Choose project...** and select the folder containing `experiment.json`. The settings fill the controls as a draft. Review them before preparing the model.

3. Click **Prepare and initialize model**. Read the actual training-target count. The prepared recipes use globally shuffled epochs and MODEL body plus EOT targets.

4. Review **Total target**, its **Epochs/Updates** unit, **Learning rate**, **Reply limit** and **Check every N epochs / update list**. Start a fresh run. Update zero and the endpoint are included in the saved checks.

5. Follow the saved checkpoints and answers with **Open run folder** and **Read replies...**. Use **Pause + save** before closing. To recover, choose the same `run.json`, click **Open / recover run**, then **Resume**. Resume keeps the original target.

6. Apply the study's checkpoint-selection rule. Load that checkpoint with **Use checkpoint...**, then choose **Final evaluation**. Keep final questions outside the ordinary checkpoint checks. Record every wrong reply as well as the correct ones.

Use ten Epochs and checks every epoch. Retain the original selection rule, including later-on-tie. The selected checkpoints in the completed study were A 4,600,000 updates, AX 2,940,000, B 1,400,000 and BX 1,200,000. Load the supplied selected.osai through Persistence, set Greedy and reply limit 40, and clear the conversation before each sentence. A/AX produce four lines; B/BX produce one.

Other target languages

German, French, Spanish and Italian input editions contain four projects each, with 20,000 training examples, 450 check questions, 900 separate final questions and 90 known probes per project. They share the English prompts and semantic split identities with the Serbian editions. They are prepared materials, with no trained language models or accuracy results.

The folders use `de_`, `fr_`, `es_` and `it_` before A, AX, B or BX. V520, D16 and H37 remain fixed; each project's saved context and target counts account for multiword dictionary entries. Load its own experiment.json rather than copying the Serbian settings.

The mechanical ordering rule also applies inside multiword translations. For example, French `en bois pomme de terre` becomes `terre de pomme bois en`. The task uses dictionary forms and does not teach natural French grammar. Shared spellings and multiword equivalents mean the language editions are not equally difficult by construction.