EMNLP 2020

# Learning Music Helps You Read: Using Transfer to Study Linguistic Structure in Language Models

**Isabel Papadimitriou**
Stanford University
isabelvp@stanford.edu

**Dan Jurafsky**
Stanford University
jurafsky@stanford.edu

汇报人：陈妍

学号：51205901044

# Outline

- Introduction
- Motivation
- Contribution
- Method
- Experiment
- Conclusion

# Introduction

This paper proposes **transfer learning** as a method for analyzing the encoding of grammatical structure in neural language models.

将**转移学习**作为一种分析神经语言模型中语法结构编码的方法。

# Motivation

- Understanding **how neural language models learn and represent syntactic structure** is an important analytic question for NLP.

理解神经语言模型如何学习和表示句法结构是NLP的重要分析问题。

- Recent work has directly probed the internal activations of models, or fed them curated inputs that depend on complex syntax, in order to **uncover latent syntactic awareness**.

最近的工作直接探究了模型的内部激活，或向他们提供了依赖于复杂语法的精选输入，以发现潜在的句法意识。

- This paper propose a different approach to measure **the structural awareness of language model**.

而本文提出了一个新的方法，去衡量语言模型学到的结构意识。

# Contribution

Propose a new method - the Test for Inductive Bias via Language Model Transfer (TILT)

TILTs demonstrate the abstract structural notions that LSTMs can learn, rather than probing for the manifestation of a particular known structure, as in most current methods.

TILT证明了LSTM可以学习的抽象结构概念，而不是像大多数当前方法那样探究特定已知结构的表现。

Also contribute to the more general cognitive question of what structural inductive biases a learner needs to be able to easily acquire human language.

我们还为更一般的认知问题做出了贡献，即学习者需要哪种结构归纳偏差才能轻松掌握人类语言。

# Method-TILT

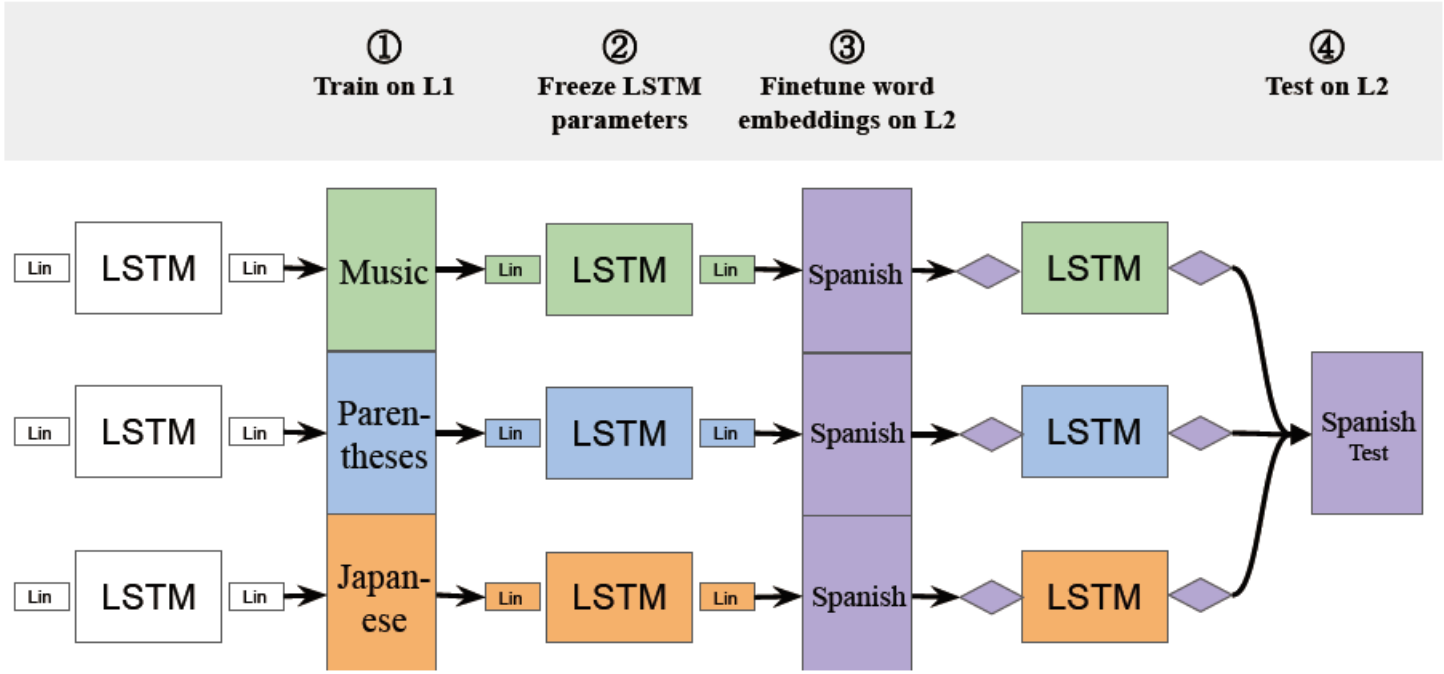



Figure 2: Diagram illustrating our training procedure: $k$ models are trained on $k$ L1 languages, and then their LSTM weights are frozen while their linear layers are finetuned on a common L2 language (in our case, we always use Spanish as the L2). We can then compare their performance on the common L2.

# Experiment

Datasets

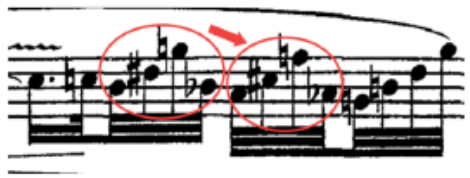
**Random**

**Uniform:** `marroquín jemer pertenecer osasuna formaron citoesqueleto relativismo`

**Zipf:** `en con conocidas y en los victoriano como trabajar ⟨unk⟩ monte * en juegos días en el`

The random corpora are sampled randomly from the Spanish vocabulary. There is no underlying structure of any kind that links words with each other. All words are equally likely to be sampled in the Uniform corpus, while common words are more likely in the Zipfian corpus.

**Music**



The music data is encoded from classical piano performances according to the MAESTRO standard. Music is structured on many levels. The red arrow in the example illustrates how, on a small timescale, each note is linked to its corresponding note when a motif is repeated but modulated down a whole-step.

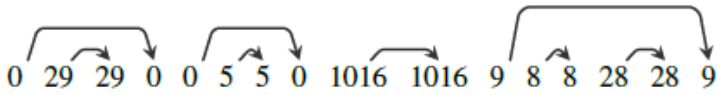
**Code**

```
if (coordFactor == 1.0f)
    return sumExpl
else {
    result = sum * coordFactor
}
```

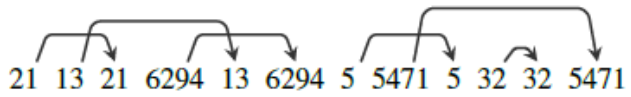
The code corpus is composed of Java code. The above snippet demonstrates some kinds of structure that are present in code: brackets are linked to their pairs, `else` statements are linked to an `if` statement, and coreference of variable names is unambiguous.

**Parentheses**

**Nesting:**




**Flat:**



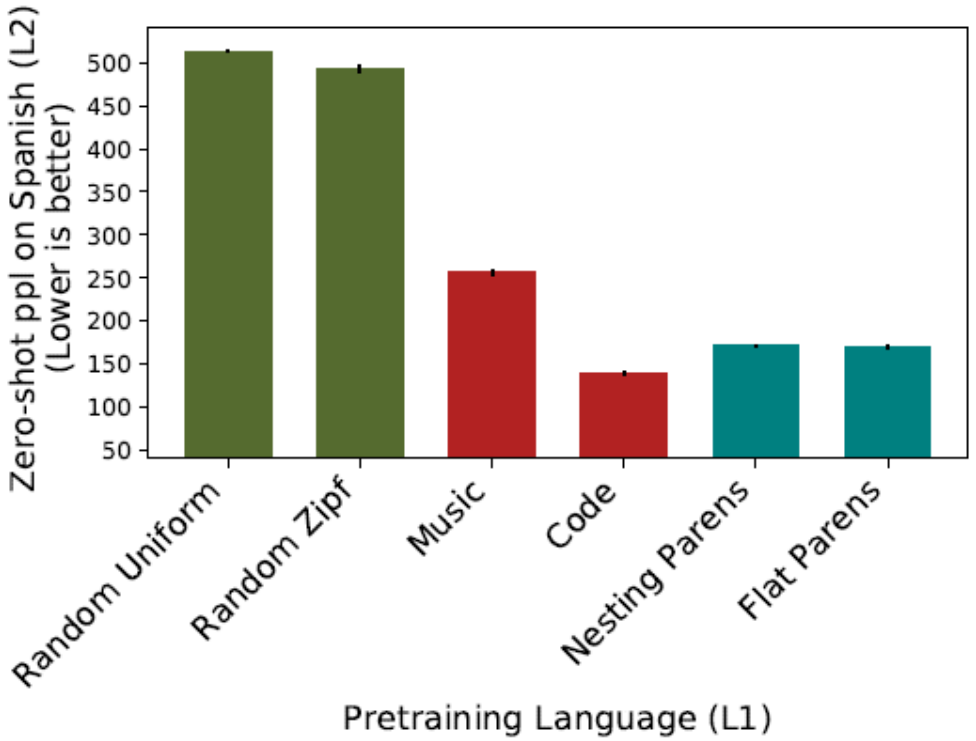

Our artificial corpora consist of pairs of matching integers. In the Nesting Parentheses corpus, integer pairs nest hierarchically and so the arcs do not cross. In the Flat Parentheses corpus, each integer pair is placed independently of all the others, and so the arcs can cross multiple times.

(There is a one-to-one mapping between Spanish words and integers and so these integers are sampled from the same Spanish vocabulary distribution as the Random Zipfian corpus. We visualize these corpora here with integers and the Random corpora with words for simplicity).

Figure 3: Examples illustrating the content of our non-linguistic corpora for Experiments 1-3. All examples are taken from the corpora.

# Experiment

Results



# Experiment

Datasets

| Language | WALS-syntax distance from Spanish (out of a max of 49 features) |
|---|---|
| Spanish (es) | 0 |
| Italian (it) | 0 |
| Portuguese (pt) | 3 |
| English (en) | 4 |
| Romanian (ro) | 5 |
| Russian (ru) | 9 |
| German (de) | 10 |
| Finnish (fi) | 13 |
| Basque (eu) | 15 |
| Korean (ko) | 18 |
| Turkish (tr) | 23 |
| Japanese (ja) | 23 |

Table 1: WALS-syntax distance between Spanish and L1s

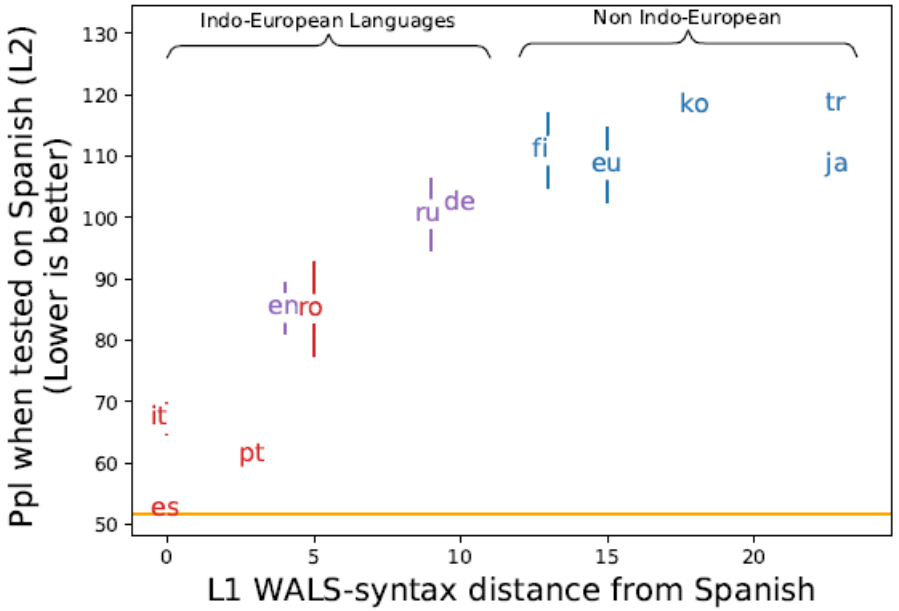
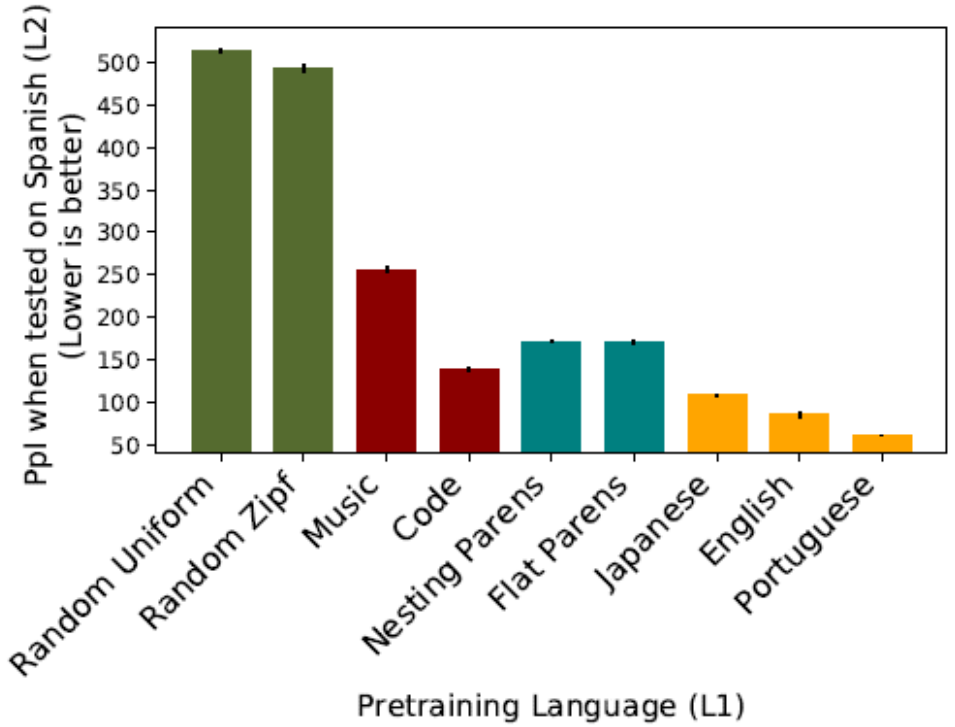



Figure 5: Results of Experiment 4. Transfer is better

# Experiment

Main Results



# Conclusion

- Non-linguistic数据中的潜在结构对于L2的学习有帮助嘛?

有帮助，但没有不同自然语言L1带来的收益大。不过虽然music的帮助是所有实验中的L1中最小的，不过本文对于不同模态语言的潜在结构的迁移的探索是个不错的方向。

- 递归结构对于L2学习影响大嘛?

不大，但是标记之间的配对结构对L2影响比较大。

- 当L1是自然语言时，语言模型可以编码并迁移其中的句法结构嘛?

可以，而且其与L2的句法距离越接近，句法结构的迁移性越好。