

Investigating Critical Period Effects in Language Acquisition through Neural Language Models

Ionut Constantinescu, Tiago Pimentel, Ryan Cotterell, Alex Warstadt
(TACL 2025)

(紹介者: フューチャー 森下 睦)

■ バイリンガル懂れますよね

- 子供の頃から2つの言語に触れているとバイリンガルになる（可能性がある）
- 大人になってから言語を学ぼうと思ってもなかなか大変



■ 臨界期を超えると

- 新たな言語獲得が難しくなる
 - 右図: アメリカに移住した年齢が早いほど英語の習熟度が高い
- これまで獲得していた言語を忘れなくなる
(臨界期以前に言語に触れなくなるとその言語を忘れてしまう)
 - 例: 韓国生まれ、8歳まで韓国在住
→ フランスで養子 → 韓国語を忘れてしまう

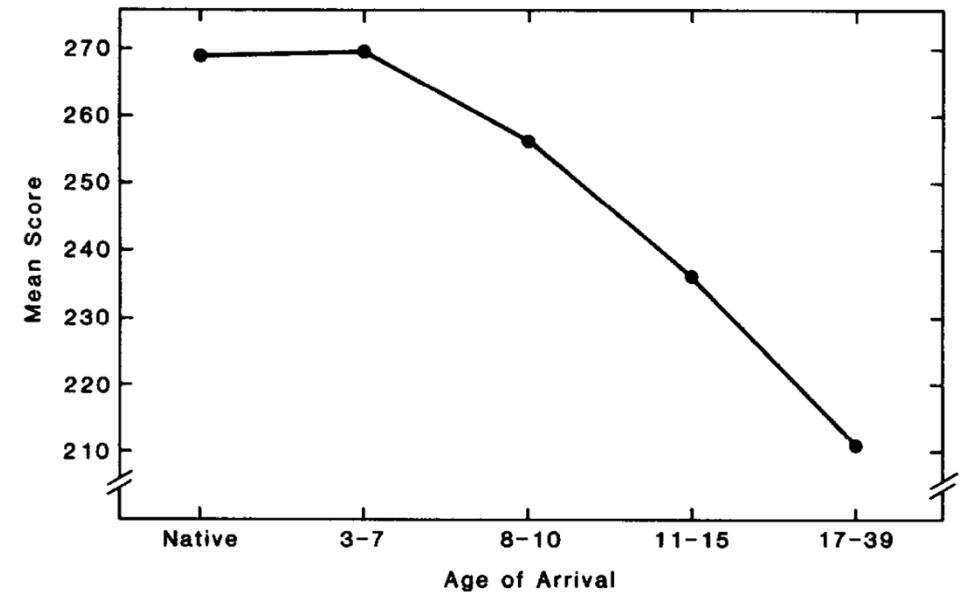


FIG. 1. The relationship between age of arrival in the United States and total score correct on the test of English grammar.

■ なぜ臨界期が存在するのか

- 本当のところは良くわかってない
- 人間の脳の言語中枢における可塑性が成熟に伴って低下することが原因だと考えられている
 - ✓ なぜ可塑性が低下するのか→2つの仮説

■ 生得説

- 生物学的な脳の発達に起因するという考え方
 - 子どもは普遍文法のようなものを備えているが、これが加齢とともに失われる
 - 青年期にかけて使われなくなったニューロン接続が減少
この際に第二言語を容易に獲得するための可塑性が低下 など

■ 経験説

- 可塑性の低下は学習の結果によるものという考え方
 - 学習経験を積みば学習能力が自然に低下するという説

■ 言語学習に関する様々な仮説は検証するのが難しく、議論が進みにくい

- 人間の成長に関してコントロールした実験は行えない

■ 言語モデルはコントロールした実験が容易に行える

- 言語モデルを用いて臨界期に関する仮説にアプローチしたい
- まずは言語モデルと人間の違いを分析
 - 言語モデルに臨界期は存在するのか？
 - 言語モデルに人間の言語学習を模倣させることはできるのか？

- 1 言語モデルにおいて、第二言語学習の臨界期が存在するか
- 2 言語モデルにおいて、第一言語の忘却に関する臨界期が存在するか
- 3 言語モデルの可塑性を低下させると、人間のような臨界期を再現できるか
- 4 言語モデルにおける臨界期の効果は、第一・第二言語の類似性によるか

実験

■ LMに2つの言語を学習させ、その言語能力を測る

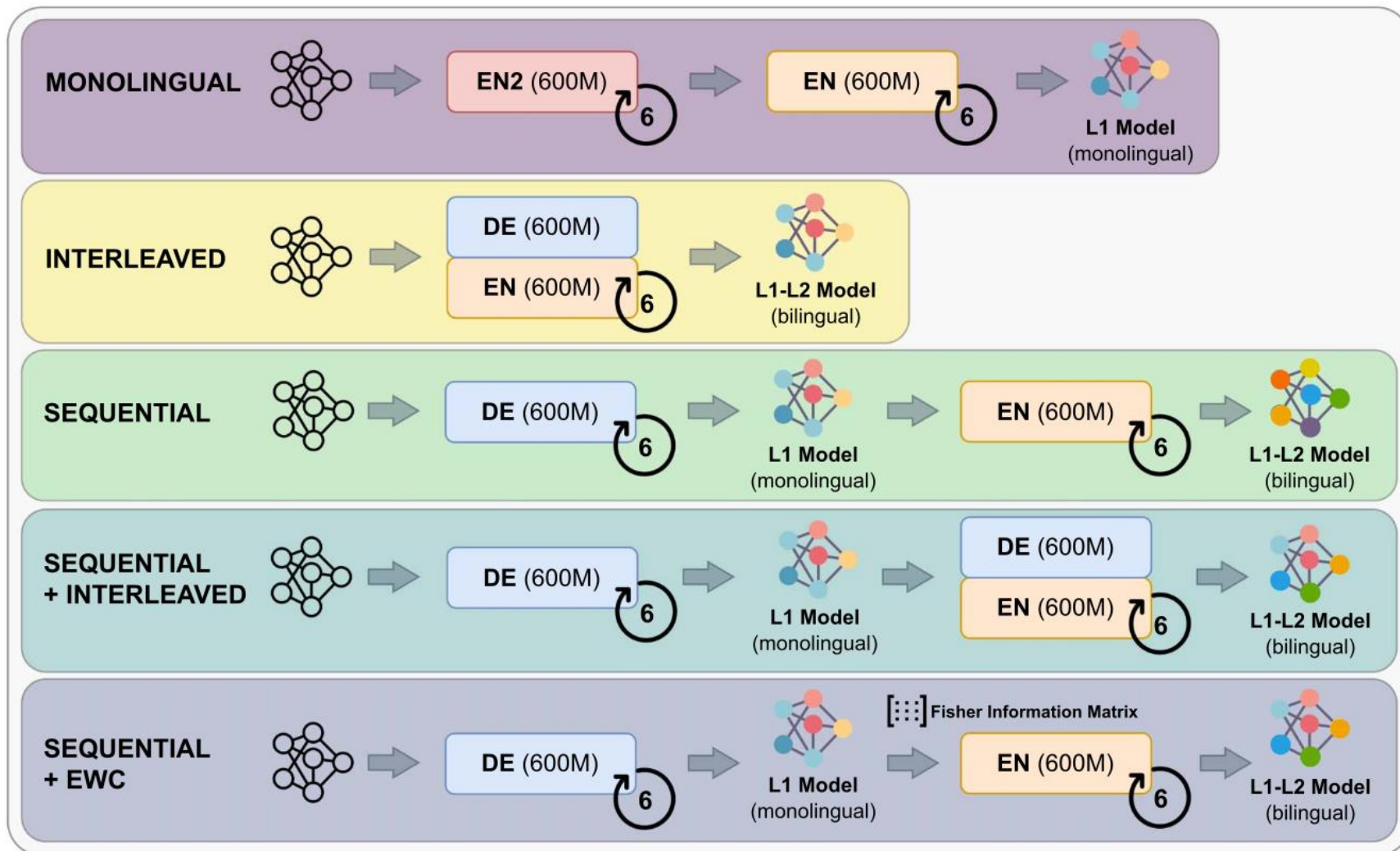
- 第一言語 (L1) → 第二言語 (L2)
- 2つの言語の学習方法は様々 (後述)

■ モデル

- GPT-2 (137M Parameters)
- RoBERTa (125M Parameters)

■ 評価

- 英語能力のみを評価 (評価データが整っているため)
- 評価手法
 - Perplexity
 - BLiMP
 - 文法的に正しい文、正しくない文が与えられ、正しさを判断
 - GLUE
 - 英語の言語理解に関するベンチマーク



→ 英語のみを学習しているネイティブの設定

→ 生まれてすぐに2言語の環境におかれた設定

→ 成長の途中で突然英語圏につれてこれ以降母国語には触れない設定

→ 成長の途中で英語と母国語の両方に触れるようになった設定

→ 成長の途中で突然英語圏につれてこれ以降母国語には触れない設定
モデルは人間的な可塑性の低下を模倣 (L1学習時からモデルパラメータが大きく変わらないよう正則化を追加)

■ 実験1: 英語を第二言語とした学習

- 人間と同様に、言語モデルには第二言語学習における臨界期が存在するのか

■ 実験2: 英語を第一言語とした学習

- 言語モデルは、一度獲得した第一言語を第二言語獲得時にどの程度忘れるのか

■ 実験3: 学習ステップ数を増やす

- 人間と同様に、臨界期を超え十分学んだ第一言語は忘れないのか
- 人間と同様に、十分に学習が進んだLMは新たな言語を獲得しにくいのか

■ 実験4: 様々なL1からの学習

- 第一言語と第二言語の言語的な近さは臨界期効果に影響を及ぼすのか

実験1: 英語を第二言語としての学習

■ 実験目的

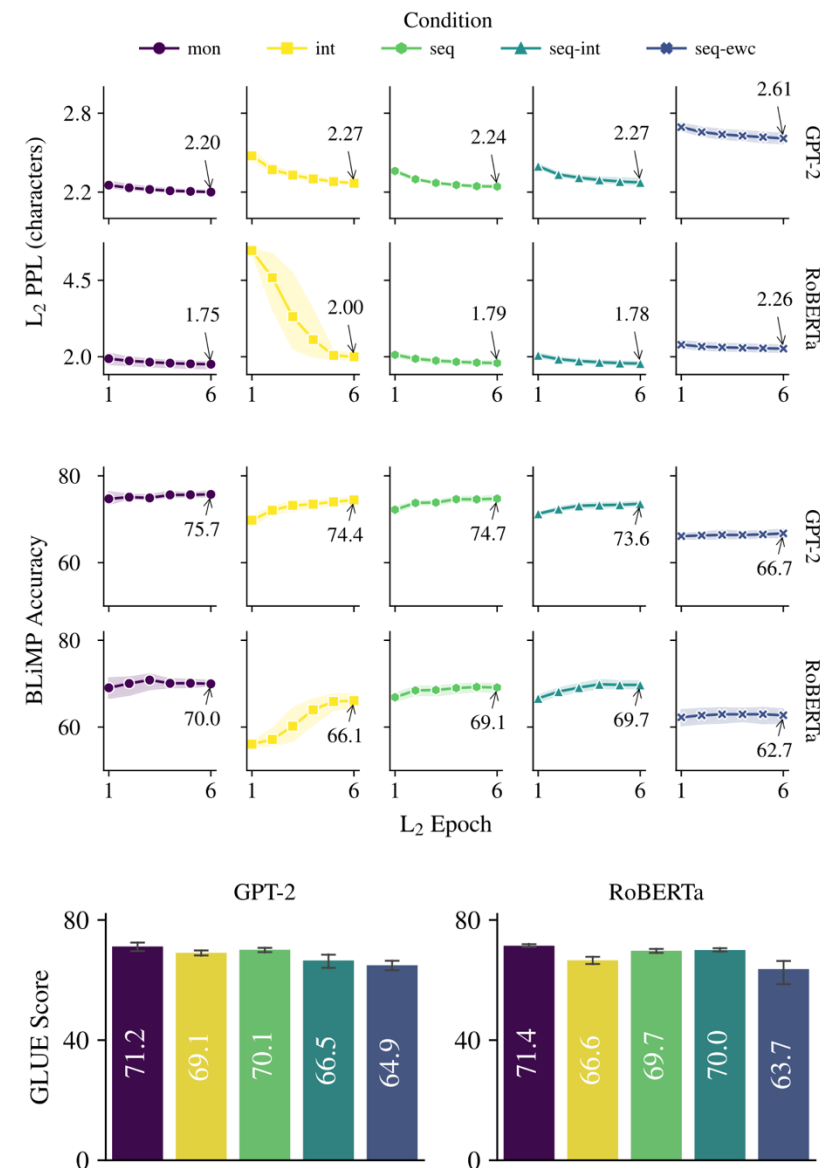
- 人間と同様に、言語モデルには第二言語学習における臨界期が存在するのか

■ 実験設定

- L1: ドイツ語、フィンランド語、L2: 英語

■ 実験結果

- 学習終了時の能力は
Monolingual (英語だけを学ぶ設定) >
Sequential (ドイツ語→英語の順に学ぶ設定) >
Interleaved (ドイツ語、英語を同時に学ぶ設定)
- 人間では Interleaved > Sequential となりLMとは異なる
→ 言語モデルには人間のような臨界期が存在しないことを示唆
- 正則化により可塑性を低下させたモデル (seq-ewc) は
L2の習熟度が大幅低下
→ ある種の人間の模倣と言える？



実験2: 学習順を逆に

FUTURE

■ 実験目的

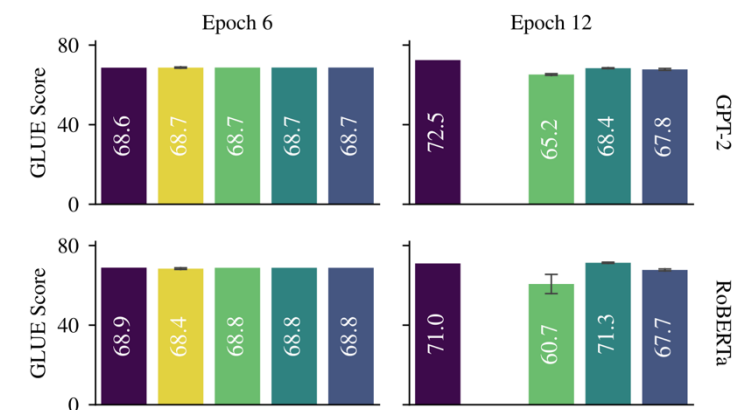
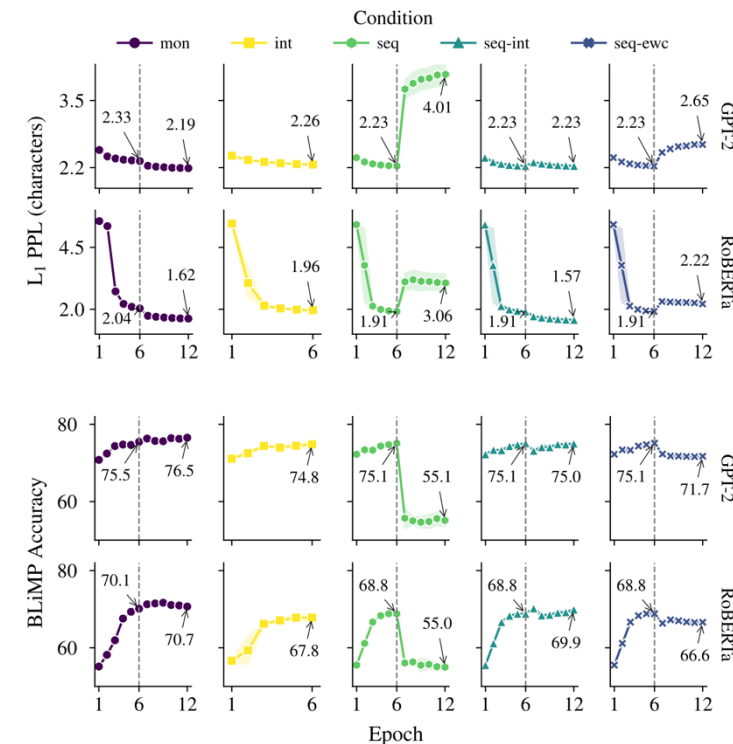
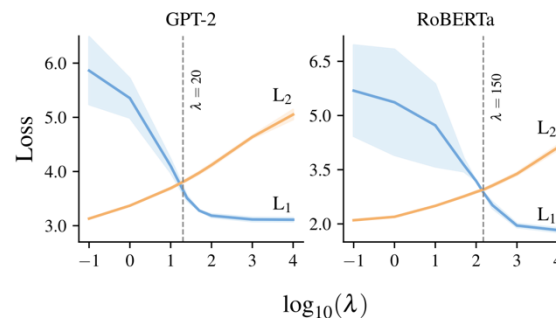
- 言語モデルは、一度獲得した第一言語を第二言語獲得時にどの程度忘れるのか

■ 実験設定

- L1: 英語、L2: ドイツ語、フィンランド語

■ 実験結果

- 英独を順に学習するモデル (seq) はL2学習直後にL1の能力が大幅低下
 - ニューラルネットワークの破滅的忘却
 - L1, L2両方に触れ続けるモデル (int) はL1を忘れない
- 可塑性を低下させたモデル (seq-ewc) はL1を忘れにくい
 - ✓ 一方でL2の学習も行いにくい
 - ✓ 正則化の強さを決めるパラメータによりトレードオフの関係がある ↓



実験3: 学習ステップ数を増やす

■ 実験目的

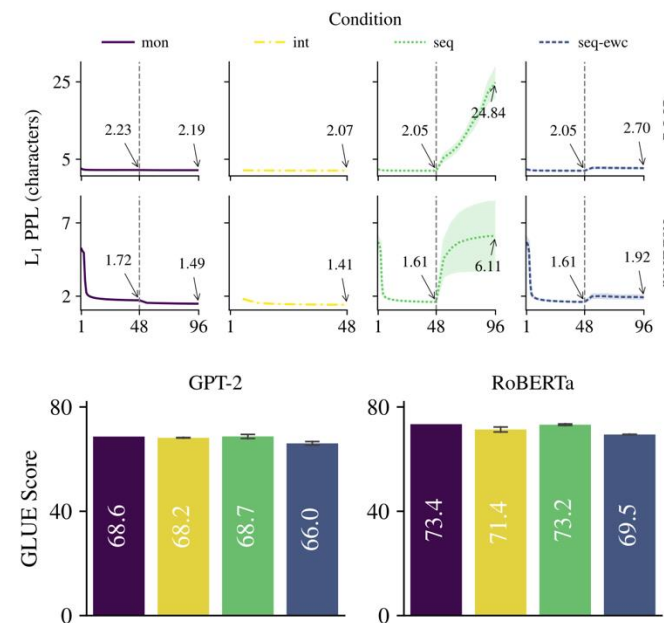
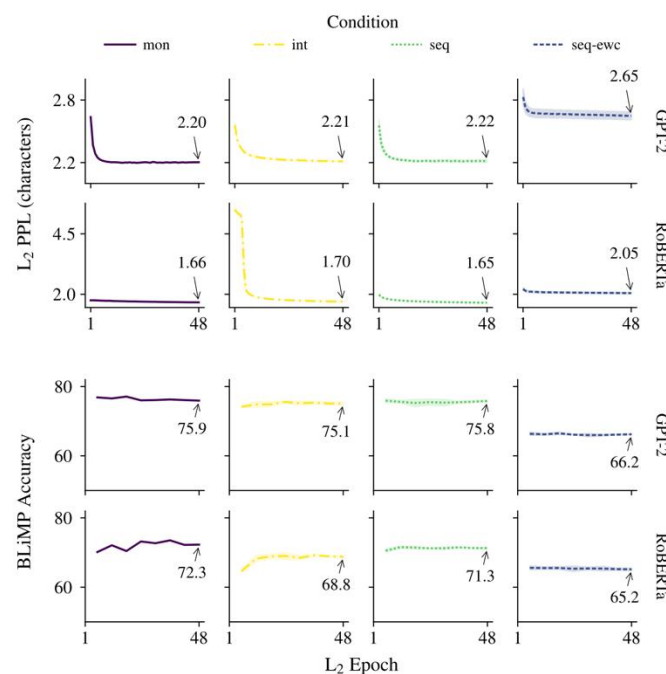
- 人間と同様に、臨界期を超え十分学んだ第一言語は忘れないのか
- 人間と同様に、十分に学習が進んだLMは新たな言語を獲得しにくいのか

■ 実験設定

- L1: ドイツ語、フィンランド語、L2: 英語
- 実験1, 2の8倍エポックを回す

■ 実験結果

- 長期間L1を学習しても、
L2学習直後にL1を忘れてしまう
- 一方でL2の学習は良く進む
- とともに人間の臨界期効果とは
異なる傾向を示した



実験4: 様々なL1からの学習

■ 実験目的

- 第一言語と第二言語の言語的な近さは臨界期効果に影響を及ぼすのか

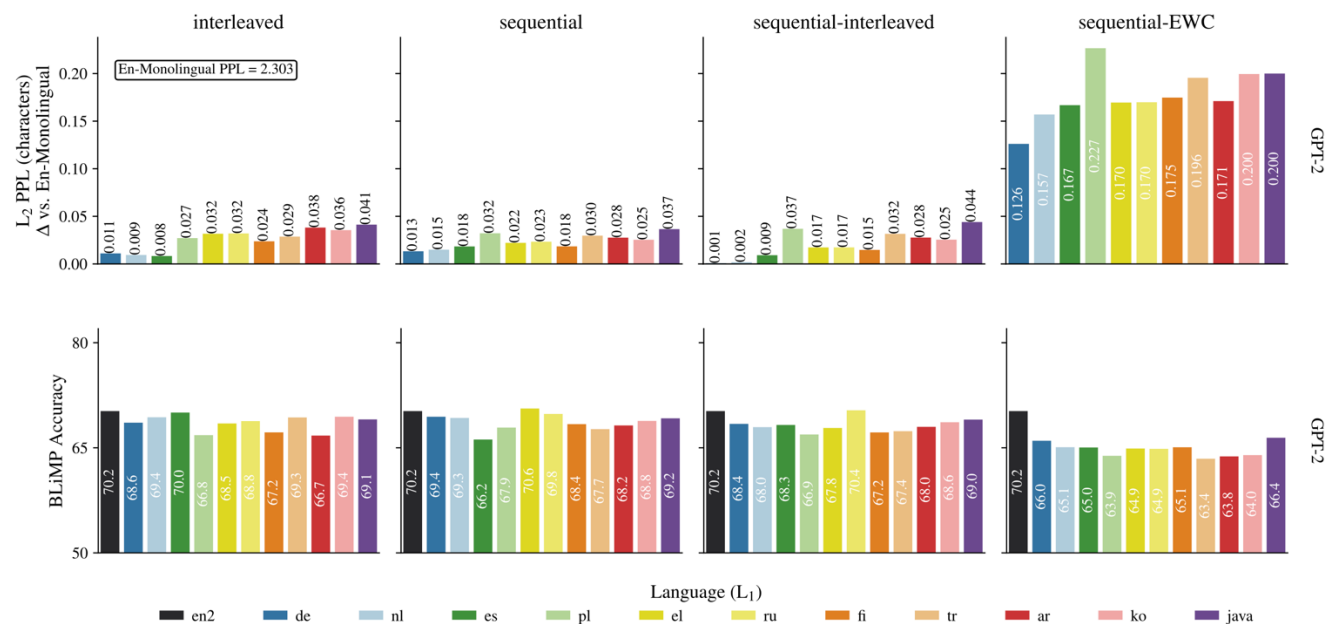
■ 実験設定

- L1
 - 同一言語: 英語2 (en2)
 - ゲルマン語派: ドイツ語 (de), オランダ語 (nl)
 - インド・ヨーロッパ語族、ラテン文字: スペイン語 (es), ポーランド語 (pl)
 - インド・ヨーロッパ語族、非ラテン文字: ギリシャ語 (el), ロシア語 (ru)
 - 非インド・ヨーロッパ語族、ラテン文字: フィンランド語 (fi), トルコ語 (tr)
 - 非インド・ヨーロッパ語族、非ラテン文字: アラビア語 (ar), 韓国語 (ko)
 - プログラミング言語: Java (java)
- L2: 英語



■ 実験結果

- Perplexityについては言語の類似性がL2の獲得に影響を及ぼしている
 - しかし、その差は非常に小さい
- BLiMPの精度については言語の類似性との間に明確な相関はない
 - なんならL1をJavaにするほうが他の自然言語よりも良い・・・
- これらについても人間の直感とは反する結果に



RQ1. 第二言語学習の臨界期が存在するか

- 答え: No.
- L2を英語とした実験において、
L1+英語を同時に学ぶより、**L1→英語と順に学ぶほうが常に結果が良かった**
 - これは**人間と異なる傾向**
- また英語だけを学ぶ設定と、L1→英語と順に学ぶ設定を比較すると、
英語だけを学ぶほうが明らかに結果が良い
 - 人間であれば、生まれた直後から2言語にふれる場合、どちらもネイティブ並みの能力となるため、
これも人間と異なる傾向と言える

RQ2. 第一言語の忘却に関する臨界期が存在するか

- 答え: No.
- LMはどれだけL1を学習しても、他の言語のみが与えられるようになると、
すぐにL1を忘れてしまう
 - ニューラルネットワークの破滅的忘却のため
- L1+L2両方を与える設定にすると、L1も覚えていられるが、
臨界期を超えた人間はL2だけを与えられてもL1を覚えていられる

RQ3. 可塑性を低下させると、人間のような臨界期を再現できるか



- 答え: Yes.
- 正則化により可塑性を意図的に低下させたLMは人間のようにL2の学習が困難になり、L1は忘れにくくなる
 - なお人っぽい動きはしているが、細かいところまで人を模倣できているかは未確認

RQ4. 臨界期の効果は、第一・第二言語の類似性によるか

- 答え: No.
- 一部の評価に対しては（わずかに）影響するが、
全ての評価尺度で同じ傾向を示したわけではない
 - モデルはL1によらず、どんなL2でも学習できる可能性を秘めている

■ 2つの仮説

➤ 生得説

- ✓ 生物学的な脳の発達に起因するという考え方

➤ 経験説

- ✓ 可塑性の低下は学習の結果によるものという考え方

- LMの場合は学習を続けても特に可塑性は低下しなかった
→ 少なくとも経験説に対しては否定的な実験結果となった

- 一方で、著者らもこれが人の臨界期に関する強い証拠になるとは考えていない
 - 論文中に度々 “weak evidence” という表現が出てくる
 - LMの場合はこうでしたくらいの捉え方が正しいと思う

まとめ

■ Research Questions

1

言語モデルにおいて、第二言語学習の臨界期が存在するか
→ No. モデルは学習のどのタイミングにおいても第二言語を習得可能。

2

言語モデルにおいて、第一言語の忘却に関する臨界期が存在するか
→ No. モデルは第二言語のみを見ると、破滅的忘却により第一言語をすぐに忘れる。

3

言語モデルの可塑性を低下させると、人間のような臨界期を再現できるか
→ Yes. 人工的に可塑性を低下させることにより、臨界期のような効果を再現可能。

4

言語モデルにおける臨界期の効果は、第一・第二言語の類似性によるか
→ No. 言語の類似性を与える影響は限定的で、一部の評価指標でのみ相関が見られた。

■ Second Language Acquisition of Neural Language Models

- ACL 2023 Findings、NAIST 大羽さんら
- 様々な言語でL1を学習→L1+L2に対してFine-tuneした際の効果を見る
- 今回の話とかなり近い

■ Developmentally-plausible Working Memory Shapes a Critical Period for Language Acquisition

- ACL 2025 Long、東大 三田さんら
- 人間の言語獲得を模倣する学習法を提案

■ Can LLMs Simulate L2-English Dialogue? An Information-Theoretic Analysis of L1-Dependent Biases

- ACL 2025 Long、メルボルン大 Gaoさんら（共著者にMBZUAI 栗林さん）
- LLMは非母語話者特有の英語使用を模倣するか

■ Is Child-Directed Speech Effective Training Data for Language Models?

- EMNLP 2024 Short, スタンフォード大 Fengさんら
- 子供向けの文章を与えたほうがLLMの学習がより進むか→進まない

- 単なる言語モデルの学習という観点では**そりゃそう的な感想になる部分も多い**
 - L1→L2を順に学習するとL1は破滅的忘却により忘れるとか。そりゃそう。
 - 正則化したら新しい言語を覚えにくくなるよねとか。そりゃそう。

- これだけで人間の臨界期に関する仮説にアプローチできたかという**完璧ではない**
 - **細かいLMの学習設定に依存しない？**と思うことが多い
 - Limitationsにも書かれているがLR Schedulerとかの影響が大きいはず・・・
 - その学習設定は人間を模倣できているんですか？

- 一方で**人間とLMの違いを見ていくのは面白い**
 - 今、LLMは何でもできます的なことを言う人が多い中、個人的には全然そうではないと思っている
 - 人間は（個人的には）LMより良くできているとっていて、**人間を見本に模倣できるところは模倣してほしい**
 - また、現状何が（模倣）できていて、何ができていないのかを明らかにしたい



END