

学習者コーパス研究における マルチレベル順序ロジットモデルの活用

小林雄一郎
日本大学

Multilevel Ordinal Logistic Regression Analysis for Learner Corpus Research

KOBAYASHI, Yuichiro
Nihon University

Abstract

言語習得研究では、学習者の習熟度を客観的に予測する指標が大きな関心を集めている。そして、そのような発達指標の特定には、学習者コーパスの活用が有効である。コーパスを用いた発達指標の研究では、慣習的に仮説検定が用られることが多いが、近年の言語研究では、回帰分析に基づく言語習得のモデル化が提唱されている。そのような流れの中で、本稿では、(a) 言語習得における母語やタスクの影響を統制できる、(b) 習熟度のような順序尺度のデータを統計的に適切な形で取り扱える、(c) 集団レベルだけではなく個人レベルの分析も行うことができる、などの利点を持つマルチレベル順序ロジットモデルを、実際のデータに基づく活用事例とともに紹介する。また、Rでマルチレベル順序ロジットモデルを実行するためのコードも示す。

Keywords: 学習者コーパス, 言語習得, マルチレベル分析, 順序ロジットモデル, 母語の影響

1. はじめに

定量的な言語習得研究において、学習者の発達を客観的に予測するための指標は、1970年代から現在にいたるまで大きな関心を集め続けている (e.g., Kobayashi et al., 2022; Verspoor et al., 2021)。そして、そのような発達指標の特定にあたっては、学習者が産出したライティングやスピーキングを電子的に集積した学習者コーパスの活用が極めて有効である。具体的には、異なる習熟度の学習者の書き言葉や話し言葉を定量的に比較し、様々な言語項目（語彙、文法、統語、談話など）の頻度データを分析することで、習熟度が上がるにつれて変化する頻度パターンをモデル化し、

それぞれの習熟度の学習者に特徴的な言語項目を特定することができる。

学習者コーパス研究において、異なる習熟度の学習者を比較する場合、カイ2乗検定、分散分析、Kruskal–Wallis 検定などの仮説検定を用いることが伝統的に多い (e.g., Ionin & Díez-Bedmar, 2021; Takahashi, 2018; Thewissen, 2013)。しかしながら、近年の言語研究では、単純な仮説検定による比較ではなく、回帰分析に基づく言語習得のモデル化が提唱されている (e.g., Gries, 2015a; Gries & Deshors, 2021)。そのような流れの中で、本稿では、(a) 言語習得における母語やタスクの影響を統制できる、(b) 習熟度のような順序尺度のデータを統計的に適切な形で取り扱える、(c) 集団レベルだけではなく個人レベルの分析も行うことができる、などの利点を持つマルチレベル順序ロジットモデルを、実際のデータに基づく活用事例とともに紹介する。

2. マルチレベル順序ロジットモデル

2.1 回帰分析

最近の学習者コーパス研究では、回帰分析を用いた統計モデリングの注目を集めている (e.g., Gries & Deshors, 2014; Gries & Wulff, 2013; Wulff & Gries, 2015, 2019, 2021)。回帰分析を用いることで、言語習得に影響を与える（と予想される）複数の変数を効率的に分析することが可能になり、個々の変数が言語習得に与える影響の度合いを明らかにすることもできる。また、回帰分析にも様々なモデルがあり、説明変数（言語的特徴など）と目的変数（習熟度など）の間に線形の関係性を仮定しない非線形モデルが存在する。そして、言語習得の過程は非線形となり得るため、一般化（混合）加法モデルのような非線形回帰分析を用いる研究もある (e.g., Murakami, 2016; Verspoor et al., 2021)。

2.2 マルチレベル分析

言うまでもなく、学習者のライティングやスピーキングに見られる言語的特徴は、学習者の母語や年齢、タスクやトピックなどの要因による影響を受けることがある。そのような場合は、マルチレベル分析を用いて、それらの要因を統制した上で、言語習得のモデル化を行う。たとえば、分析対象とする学習者データが「母語–学習者」のような階層構造を持っていると仮定することで、母語の影響を統制する形で言語習得の過程を分析することができるようになる。また、「学習者–測定時期」のような階層構造を仮定することで、言語習得の経年的分析を行うことが可能になる。

多くの学習者コーパスは、母語、産出モード、タスク、トピックなどの階層構造を持っている。このような階層的データに対して通常の回帰分析を行うと、(a) 標準誤差を不当に小さく見積もってしまうことで第1種の過誤を犯す可能性がある、(b) 回帰係数の中に集団の性質と個人の性質が混在することで分析結果の解釈が困

難になる, などの問題が生じる (e.g., Heck & Thomas, 2020; Snijders & Bosker, 2012)。従って, 階層構造を持つデータを適切に分析するには, 通常の実帰分析ではなく, マルチレベル分析を用いなければならない。

以上のような問題意識に基づき, 近年の言語研究では, 階層的データに対するマルチレベル分析の使用が推奨されており (e.g., Gries, 2015b, 2021; Schäfer, 2020), 活用事例も徐々に増えている (e.g., Gries & Deshors, 2015; Kyle et al., 2021; Murakami, 2016; Paquot et al., 2021)。

2.3 順序ロジットモデル

実験データを用いる言語習得研究では, 分析対象とする要因が反応時間などの量的変数に与える影響の度合いを調査することが多い。その一方, コーパスを用いる言語習得研究では, 言語テストなどで判定された習熟度のような質的変数に与える影響の度合いを分析することが多い。しかしながら, 学習者コーパス研究において, 順序尺度 (3 値以上) のデータを扱うマルチレベル分析は, Kobayashi (2021) 以外に見つけられない。

そこで本稿では, Kobayashi (2021) における分析の一部を例に, マルチレベル順序ロジットモデルを用いた学習者コーパス研究の方法を説明する。この手法は, 順序尺度の変数 (3 つ以上) を分析するための順序ロジットモデルをマルチレベルに拡張したものである (Fullerton & Xu, 2016)。また, 順序ロジットモデルは, 類似手法である順序プロビットモデルと比べて, 外れ値を含むデータに対して頑健であるとされる (Liao, 1994)。なお, 以下で紹介する分析事例の結果に言及する場合は, 本稿ではなく, Kobayashi (2021) を引用されたい。

そして, 分析結果に加えて, R (R Core Team, 2021) を用いたマルチレベル順序ロジットモデルの実行方法を簡潔に示す。¹ ただ, R の基本的な使い方まではカバーできないため, 必要に応じて, R の入門書 (e.g., 小林 他, 2020) を合わせて参照されたい。

3. 活用事例

3.1 分析目的

本研究の目的は, 母語の影響を統計的に統制した上で, 英語学習者の習熟度が上がるにつれてメタ談話標識の使用頻度がどのように変化するかを明らかにすることである。具体的には, (1) どのようなメタ談話標識が学習者の習熟度を予測する発達指標となり得るのか, (2) 学習者の母語は言語習得にどの程度の影響を及ぼすのか, という2つのRQに答える。なお, メタ談話標識とは, 読み手に文章の命題内容や展開, 書き手の立場を伝える機能を担う表現である (Hyland, 2005)。

3.2 分析データ

本研究の分析データは、International Corpus Network of Asian Learners of English (ICNALE; Ishikawa, 2011) における日本人 (JPN), タイ人 (THA), 台湾人 (TWN) の英語学習者が書いたライティングである。個々の学習者のデータには、A2, B1.1, B1.2, B2 という4段階の Common European Framework of Reference (CEFR) に基づく習熟度情報が付与されている。分析にあたっては、これらのレベルを1~4の値に変換した (A2 が1, B2 が4)。表1は、分析データの概要をまとめたものである。

表1

分析データの概要

CEFR level	Numbers	JPN	THA	TWN	Total
1	learners	154	119	29	302
	(words)	(34,959)	(26,866)	(6,491)	(68,316)
2	learners	179	179	87	445
	(words)	(40,463)	(40,899)	(20,658)	(102,020)
3	learners	49	100	61	210
	(words)	(11,315)	(23,267)	(14,636)	(49,218)
4	learners	18	2	23	43
	(words)	(4,281)	(514)	(5,856)	(10,651)
Total	learners	400	400	200	1,000
	(words)	(91,018)	(91,546)	(47,641)	(230,205)

そして、表1の1,000本のライティングから、メタ談話標識の10種類のカテゴリごとの相対頻度 (100語あたり) を集計した。頻度集計にあたっては、自作のPerlプログラム (Kobayashi, 2020) を用いた。

4. Rによる分析と結果

まず、R 4.2.0を用いて、1,000本のライティングからメタ談話標識の相対頻度を集計した頻度表 (ここでは、タブ区切りのテキストファイル) を読み込み、分析に使用する。そして、本研究の分析に用いる3つのメタ談話標識のカテゴリ (TRA,

SEM, BOO), 学習者の母語 (L1) と習熟度 (CEFR) の列を指定する。また, 後段で (マルチレベル) 順序ロジットモデルの目的変数とする習熟度を順序尺度のデータに変換する。

```
> # パッケージの読み込み (必要に応じて, 事前にインストール)
> library(dplyr)
> library(ordinal)
> # 頻度表 (タブ区切りのテキストファイル) の読み込み
> dat <- read.delim(file.choose(), header = TRUE)
> # 分析に用いる情報を抜き出し (分析に合わせて, 適宜変更)
> dat.2 <- select(dat, c(TRA, SEM, BOO, L1, CEFR))
> # データの冒頭を確認
> head(dat.2)
      TRA      SEM      BOO  L1  CEFR
1 2.427184 12.135922 3.398058 JPN    2
2 4.891304  1.630435  1.630435 JPN    3
3 5.213270  2.843602  1.895735 JPN    1
4 4.326923 11.057692  1.923077 JPN    4
5 6.542056  0.467290  0.467290 JPN    4
6 3.982301  0.884956  2.212389 JPN    2
> # 目的変数とする列を順序尺度のデータに変換
> dat.2$CEFR <- ordered(dat.2$CEFR)
> class(dat.2$CEFR)
[1] "ordered" "factor"
> # データの概要を確認
> summary(dat.2)
      TRA      SEM      BOO
Min.   : 0.4566   Min.   : 0.000   Min.   :0.000
1st Qu.: 3.9024   1st Qu.: 1.342   1st Qu.:1.038
Median : 4.8309   Median : 3.119   Median :1.909
Mean   : 4.9131   Mean   : 4.534   Mean    :1.998
3rd Qu.: 5.9059   3rd Qu.: 7.288   3rd Qu.:2.807
Max.   :10.1852   Max.   :16.372   Max.    :9.223

  L1      CEFR
JPN: 400    1:302
THA: 400    2:445
```

TWN: 200 3:210
4: 43

次に、ICC パッケージの ICCest 関数を用いて、メタ談話標識のカテゴリごとの級内相関係数 (Finch et al., 2019) を計算し、それぞれのカテゴリの使用傾向に母語の影響が見られるかどうかを調べる。級内相関係数の値が大きいほど、観測値の独立性が低く、マルチレベル分析を行う必要性が高い。一般的に、級内相関係数が 0.1 以上 (あるいは 0.05 以上) の場合、データに階層性があるとみなされる (Hox et al., 2018)。

```
> # パッケージの読み込み (必要に応じて、事前にインストール)
> library(ICC)
> # 級内相関係数の計算 (計算結果の一部のみを表示)
> ICCest(as.factor(L1), TRA, data = dat.2, alpha = 0.05,
CI.type = "Smith") [1 : 3]
$ICC
[1] -0.001535097

$LowerCI
[1] -0.004757873

$UpperCI
[1] 0.00168768

> ICCest(as.factor(L1), SEM, data = dat.2, alpha = 0.05,
CI.type = "Smith") [1 : 3]
$ICC
[1] 0.2517476

$LowerCI
[1] -0.1325053

$UpperCI
[1] 0.6360005

> ICCest(as.factor(L1), BOO, data = dat.2, alpha = 0.05,
```

```
CI.type = "Smith")[1 : 3]
```

```
$ICC
```

```
[1] 0.1084263
```

```
$LowerCI
```

```
[1] -0.09179082
```

```
$UpperCI
```

```
[1] 0.3086435
```

各カテゴリーの級内相関係数 (\$ICC) を確認すると、SEM の値 (0.2517476) と BOO の値 (0.1084263) が 0.1 以上であり、分析データに階層性 (集団的類似性) があるとみなされる。これは、2 つの言語項目の使用傾向に母語の影響が見られることを意味している。

そこで本研究では、階層性を持つ分析データを適切に取り扱うために、マルチレベル順序ロジットモデルを用いた分析を行う。その際、3 つのカテゴリーの相対頻度 (TRA, SEM, BOO) を説明変数、学習者の習熟度 (CEFR) を目的変数とし、マルチレベル分析における集団レベルの変数を母語 (L1) とする。また、マルチレベル分析の有効性を確認するために、通常のシングルレベルの順序ロジットモデルも行い、2 つのモデルを比較する。R で (マルチレベル) 順序ロジットモデルを実行する方法は複数存在するが、ここでは、ordinal パッケージの `clm` 関数と `clmm` 関数を使用する。

```
> # シングルレベルの順序ロジットモデル (母語の情報を使用しない)
```

```
> model.1 <- clm(CEFR ~ TRA + SEM + BOO, data = dat.2)
```

```
> summary(model.1)
```

```
formula: CEFR ~ TRA + SEM + BOO
```

```
data: dat.2
```

```
link threshold nobs logLik AIC niter max.grad
```

```
logit flexible 1000 -1172.02 2356.04 5(0) 4.87e-07
```

```
cond.H
```

```
2.7e+03
```

```
Coefficients:
```

```
Estimate Std. Error z value Pr(>|z|)
```

```

TRA -0.08146  0.03846  -2.118  0.034174 *
SEM -0.05194  0.01565  -3.320  0.000901 ***
BOO -0.11488  0.05357  -2.144  0.032005 *
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
      Estimate Std. Error z value
1|2  -1.7216     0.2351  -7.323
2|3   0.2379     0.2280   1.044
3|4   2.2763     0.2641   8.618
>
> # マルチレベル順序ロジットモデル (母語の情報を使用する)
> model.2 <- cglmm(CEFR ~ TRA + SEM + BOO + (1 | L1), data =
dat.2)
> summary(model.2)
Cumulative Link Mixed Model fitted with the Laplace
approximation

formula: CEFR ~ TRA + SEM + BOO + (1 | L1)
data:    dat.2

link threshold nobs logLik AIC niter max.grad
logit flexible 1000 -1152.72 2319.45 424(863) 2.91e-04
cond.H
1.6e+03

Random effects:
Groups Name Variance Std.Dev.
L1 (Intercept) 0.2369 0.4868
Number of groups: L1 3

Coefficients:
      Estimate Std. Error z value Pr(>|z|)
TRA -0.07840  0.03874  -2.024  0.0430 *

```

```

SEM -0.04044    0.01699   -2.380    0.0173  *
BOO -0.09567    0.05466   -1.750    0.0801  .
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
      Estimate Std. Error z value
1|2  -1.7850    0.3705  -4.818
2|3   0.2469    0.3660   0.674
3|4   2.3368    0.3891   6.005

```

上記の分析結果には多くの情報が含まれているが、ここでは、textreg パッケージを用いて、特に重要な情報を1つの表に整理する。

```

> # パッケージの読み込み (必要に応じて、事前にインストール)
> library(textreg)
> # モデルの比較 (Model 1 がシングルレベル, Model 2 がマルチレベル)
> texreg::screenreg(list(model.1, model.2))

```

```

=====
              Model 1          Model 2
-----
TRA              -0.08 *          -0.08 *
                (0.04)           (0.04)
SEM              -0.05 ***         -0.04 *
                (0.02)           (0.02)
BOO              -0.11 *           -0.10
                (0.05)           (0.05)
1|2              -1.72 ***         -1.78 ***
                (0.24)           (0.37)
2|3               0.24             0.25
                (0.23)           (0.37)
3|4               2.28 ***          2.34 ***
                (0.26)           (0.39)
-----

```

AIC	2356.04	2319.45
BIC	2385.49	2353.80
Log Likelihood	-1172.02	-1152.72
Num. obs.	1000	1000
Groups (L1)		3
Variance: L1: (Intercept)		0.24
=====		
*** p < 0.001; ** p < 0.01; * p < 0.05		

まず、3つの言語項目の係数 (TRA, SEM, BOO) を確認すると、シングルレベルのモデル (Model 1) では3つの言語項目全てが有意であるのに対して、マルチレベルのモデル (Model 2) では TRA と SEM だけが有意である。そして、シングルレベルのモデルでは SEM のアステリスクが3つであるのに対して ($p < 0.001$), マルチレベルのモデルでは1つに減っている ($p < 0.05$)。この結果は、母語の影響を考慮しないシングルレベルのモデルでは、習熟度の予測における言語項目の寄与度が (不当に) 高く見積もられていることを示している。ちなみに、線形モデルと違って、順序ロジットモデルの係数は、目的変数に対する説明変数に与える影響の度合いとして解釈することができないため、限界効果 (Agresti, 2018) などを別途計算する必要がある (順序ロジットモデルの係数は、モデルによって推定される潜在変数に対する影響の度合いを表している)。また、線形モデルの決定係数に相当する指標が必要な場合は、Nagelkerke の擬似決定係数 (Nagelkerke, 1991) などを計算する。

続いて、シングルレベルのモデルとマルチレベルのモデルの赤池情報量基準 (AIC) を確認すると、それぞれ 2356.04 と 2319.45 である。情報量基準を参照する場合は、値が低いほど、そのモデルの予測力が高いことを示している。特に、情報量基準の値に 10 以上の差がある場合、値の大きい方のモデル (ここでは、シングルレベルのモデル) を選択するべきではない (Burnham & Anderson, 2004)。

ちなみに、anova 関数を用いてモデルを比較することも可能である。その場合にも、シングルレベルのモデル (model.1) と比べて、マルチレベルのモデル (model.2) が有意に優れていることが示される ($p < 0.001$)。

```
> # モデルの検定 (model.1 がシングルレベル, model.2 がマルチレベル)
```

```
> anova(model.1, model.2)
```

```
Likelihood ratio tests of cumulative link models:
```

```
formula:
```

```
link: threshold:
```

```

model.1 CEFR ~ TRA + SEM + BOO                logit flexible
model.2 CEFR ~ TRA + SEM + BOO + (1 | L1)      logit flexible

      no.par      AIC    logLik  LR.stat  df  Pr(>Chisq)
model.1      6  2356.0  -1172.0
model.2      7  2319.4  -1152.7   38.593   1  5.22e-10 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

最後に、`effects` パッケージの `Effect` 関数を用いて、マルチレベルのモデルで係数が有意であった説明変数に関して、学習者の習熟度に与える影響を可視化する。図 1 と図 2 では、ある要因（ここでは、TRA と SEM の相対頻度）が変化した場合の目的変数（ここでは、習熟度）の確率に与える影響が可視化されている。なお、変化させる要因以外は、平均値に固定している。

```

> # パッケージの読み込み（必要に応じて、事前にインストール）
> library(MASS)
> library(effects)
> # モデルの可視化（マルチレベルのモデルで有意な説明変数のみ）
> eff.1 <- Effect(focal.predictors = "TRA", model.2)
> plot(eff.1, rug = FALSE)
> eff.2 <- Effect(focal.predictors = "SEM", model.2)
> plot(eff.2, rug = FALSE)

```

図 1

モデルの可視化 (TRA)

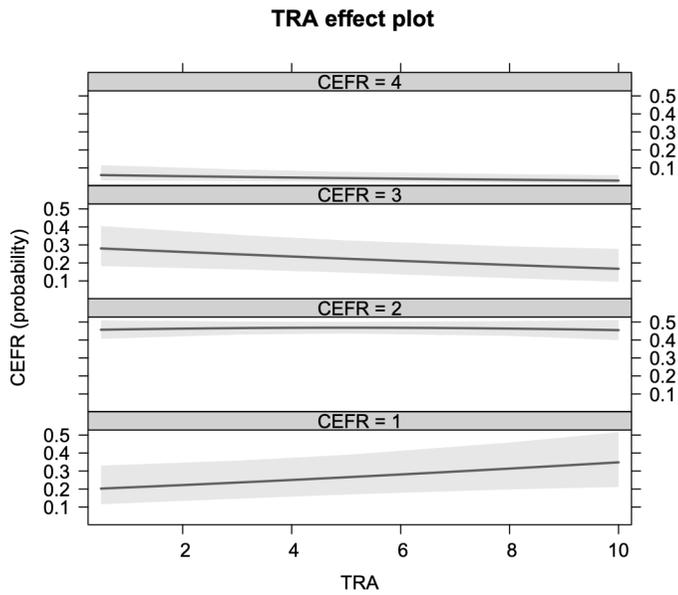


図 2

モデルの可視化 (TRA)

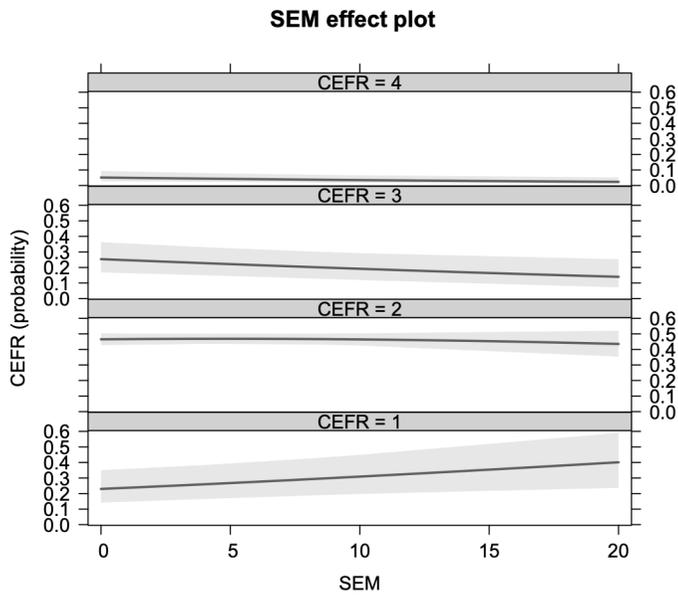


図1と図2を見ると、TRAとSEMは、習熟度が低い学習者(CEFR = 1~2)に顕著な言語項目である。習熟度が上がるにつれて、平均文長が発達し、構文が多様化していくために、これら2つの言語項目の頻度が相対的に減少していくと思われる。この点を検証するために、今後の質的な分析が求められる。

前述のように、本研究のRQは、(1)どのようなメタ談話標識が学習者の習熟度を予測する発達指標となり得るのか、(2)学習者の母語は言語習得にどの程度の影響を及ぼすのか、という2つであった。そして、(1)TRAとSEMが学習者の習熟度を予測する発達指標となり得る、(2)母語の影響を統制せずにメタ談話標識の習得を適切にモデル化することはできない、という結論を得た。

5. おわりに

本稿では、学習者コーパスとマルチレベル順序ロジットモデルを用いて、母語の影響を統制しつつ、習熟度のような順序尺度のデータに対する言語的特徴の影響をモデル化した。ただし、マルチレベル分析には様々な手法があり、本稿で用いたものは、シンプルな仮定に基づくモデル(ランダム切片モデル)である。今後、中心化などの前処理、より複雑な仮定(ランダム傾きモデル、ランダム切片・ランダム傾きモデル)、ベイズ推定(e.g., Bürkner & Vuorre, 2019; Liddell & Kruschke, 2018)などを用いることで、言語習得に関して、より深い知見を得られる可能性がある。また、RQの設定によっては、母語やタスクなどの名義尺度を予測するマルチレベル多項ロジットモデルが有効となる場合もある。

謝辞

本研究は、JSPS 科学研究費補助金(科研費)21K00660の助成を受けたものである。また、本稿の一部は、外国語教育メディア学会関西支部メソドロジー研究部会2021年度第1回研究会(2021年7月10日)で発表したものである。発表時に有益な情報をくださった方々に、心より感謝を申し上げる。

注

1. 本稿のコードに何らかの誤りが含まれている可能性、将来的に正しく動作しなくなる可能性がある。従って、コードを実行する場合は、自己責任で利用されたい。

参考文献

- Agresti, A. (2018). *An introduction to categorical data analysis* (3rd ed.). Wiley.
- Bürkner, P., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial.

- Advances in Methods and Practices in Psychological Science*, 2(1), 77–101.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33, 261–304.
- Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel modeling using R* (2nd ed.). CRC Press.
- Fullerton, A. S., & Xu, J. (2016). *Ordered regression models: Parallel, partial, and non-parallel alternatives*. CRC Press.
- Gries, S. Th. (2015a). Statistics for learner corpus research. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 159–182). Cambridge University Press.
- Gries, S. Th. (2015b). The most under-used statistical method in corpus linguistics: Multi-level (and mixed-effects) models. *Corpora*, 10(1), 95–125.
- Gries, S. Th. (2021). (Generalized linear) mixed-effects modeling: A learner corpus example. *Language Learning*, 71(3), 757–798.
- Gries, S. Th., & Deshors, S. C. (2014). Using regressions to explore deviations between corpus data and a standard/target: Two suggestions. *Corpora*, 9(1), 109–136.
- Gries, S. Th., & Deshors, S. C. (2015). EFL and/vs. ESL? A multi-level regression modeling perspective on bridging the paradigm gap. *International Journal of Learner Corpus Research*, 1(1), 130–159.
- Gries, S. Th., & Deshors, S. C. (2021). Statistical analyses of learner corpus data. In N. Tracy-Ventura & M. Paquot (Eds.), *The Routledge handbook of second language acquisition and corpora* (pp. 119–132). Routledge.
- Gries, S. Th., & Wulff, S. (2013). The genitive alternation in Chinese and German ESL learners: Towards a multifactorial notion of context in learner corpus research. *International Journal of Corpus Linguistics*, 18(3), 327–356.
- Heck, R. H., & Thomas, S. L. (2020). *An introduction to multilevel modeling techniques: MLM and SEM approaches* (4th ed.). Routledge.
- Hox, J. J., Moerbeek, M., & van de Schoot, R. (2018). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge.
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.
- Ionin, T., & Díez-Bedmar, M. D. (2021). Article use in Russian and Spanish learner writing at CEFR B1 and B2 levels: Effects of proficiency, native language, and specificity. In B. L. Bruyn & M. Paquot (Eds.), *Learner corpus research meets second language acquisition* (pp. 10–38). Cambridge University Press.
- Ishikawa, S. (2011). A new horizon in learner corpus studies: The aim of the ICNALE project. In G. Weir, S. Ishikawa, & K. Poonpon (Eds.), *Corpora and language*

- technologies in teaching, learning and research* (pp. 3–11). University of Strathclyde Press.
- Kobayashi, Y. (2020). Rhetorical preferences in L2 writings: A contrastive analysis of metadiscourse markers. *Journal of Modern Languages*, 30(2), 1–24.
- Kobayashi, Y. (2021). Identifying L2 developmental indices while controlling for L1 effects: A multilevel ordinal logistic regression analysis. *Journal of Pan-Pacific Association of Applied Linguistics*, 25(2), 87–104.
- Kobayashi, Y., Abe, M., & Kondo, Y. (2022). Exploring L2 spoken developmental measures: Which linguistic features can predict the number of words? *English Corpus Studies*, 29, 1–18.
- 小林雄一郎・濱田彰・水本篤 (2020). 『Rによる教育データ分析入門』 オーム社
- Kyle, K., Crossley, S., & Verspoor, M. (2021). Measuring longitudinal writing development using indices of syntactic complexity and sophistication. *Studies in Second Language Acquisition*, 43(4), 781–812.
- Liao, T. F. (1994). *Interpreting probability models: Logit, probit, and other generalized linear models*. Sage.
- Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Psychology*, 79, 328–348.
- Murakami, A. (2016). Modeling systematicity and individuality in nonlinear second language development: The case of English grammatical morphemes. *Language Learning*, 66(4), 834–871.
- Nagelkerke, N. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78, 691–692.
- Paquot, M., Naets, H., Gries, S. Th. (2021). Using syntactic co-occurrences to trace phraseological complexity development in learner writing: Verb + object structures in LONGDALE. In B. L. Bruyn & M. Paquot (Eds.), *Learner corpus research meets second language acquisition* (pp. 122–147). Cambridge University Press.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Online. <http://www.r-project.org/>
- Schäfer, R. (2020). Mixed-effects regression modeling. In M. Paquot & S. Th. Gries (Eds.), *A practical handbook of corpus linguistics* (pp. 535–561). Springer.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Sage.

- Takahashi, Y. (2018). A corpus-based study on Japanese EFL learners' use of relative clause constructions: CEFR criterial feature and error analysis. *English Corpus Studies*, 25, 57–78.
- Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: Insights from an error-tagged EFL learner corpus. *The Modern Language Journal*, 97(S), 77–101.
- Verspoor, M., Lowie, W., & Wieling, M. (2021). L2 developmental measures from a dynamic perspective. In B. L. Bruyn & M. Paquot (Eds.), *Learner corpus research meets second language acquisition* (pp. 172–190). Cambridge University Press.
- Wulff, S., & Gries, S. Th. (2015). Prenominal adjective order preferences in Chinese and German L2 English: A multifactorial corpus study. *Linguistic Approaches to Bilingualism*, 5(1), 122–150.
- Wulff, S., & Gries, S. Th. (2019). Particle placement in learner language. *Language Learning*, 69(4), 873–910.
- Wulff, S., & Gries, S. Th. (2021). Exploring individual variation in learner corpus research: Methodological suggestions. In B. L. Bruyn & M. Paquot (Eds.), *Learner corpus research meets second language acquisition* (pp. 191–213). Cambridge University Press.