

## メタ分析の手続きについて — 大切なことは全て効果量が教えてくれるか? —

亘理 陽一  
静岡大学

---

### 概要

メタ分析とは、ある分野・領域の複数の研究結果を統計的手法によって統合する、系統的レビューの一種である。本稿は、このメタ分析について、先行する代表的なメタ分析研究においても十分明らかにされていない分析手続き上の諸問題を検討することを目的とする。第二言語習得・外国語教育研究におけるメタ分析の妥当性や再現性にかかわって、Norris and Ortega (2000)を主たる具体例として、1) 測定法や尺度の異なるものを一緒にしていいのか、2) 比較の仕方が異なるものを一緒にしていいのか、3) 誰もがメタ分析を実際に行うことができるのかについて論じる。

**Keywords:** メタ分析, 効果量, 測定法, 調整変数, Norris and Ortega (2000)

---

### 1. はじめに

メタ分析には、一次研究をふるいにかけて、悪質な研究を切り捨てていくようなイメージがある（実際、一次研究に対する波及効果もメリットとして挙げられている。印南, 2012; Norris & Ortega, 2006; Oswald & Plonsky, 2010などを参照）。実際には、「過去に行われた複数の独立した研究結果を統合するための（統合できるか否かの検討も含めた）統計解析」と説明される（丹後, 2002, p. 1; 浦野・亘理, 2013）。例えば Mizumoto, Urano, and Maeda (2014)は、全国英語教育学会の紀要 *ARELE* に掲載された450論文の分析を通じて、量的手法を用いた仮説生成型の実証研究が多く、統制群と処置群の比較を行っている28論文63事例の効果量の平均が中程度 (Hedge's  $g = 0.76$ , 95% CI [0.59, 0.93]) である一方で、十分な検定力を持つものは27事例に留まること（十分な検定力で同じ効果量を得るためには、サンプルサイズは最低で29人以上必要であること）などを明らかにしている。

第二言語習得・外国語教育研究におけるこのようなメタ分析研究は、過去十数年の間に急激な増加を見せている（Oswald & Plonsky, 2010, p. 87）。Plonsky (n.d.)には2014年9月時点で187の研究が挙げられており、ここ数年に至っては「ブーム」と言っても過言ではない様相を呈している。この状況に、言わばメタ・メタ分析とでも呼ぶべき、メタ分析についての分析を行う研究も見られる。例えば Plonsky and Brown (2014)は、corrective

feedback (CF)に関する研究を含む (18 のメタ分析研究の内 CF の効果のみを算出することが可能な) メタ分析研究が 13 もあるにもかかわらず, その効果量の平均が Cohen's  $d$  で 0.16 から 1.16 に渡ることを指摘し, CF のタイプやその操作的定義の違いから原因を分析している。あるいは Plonsky (2012)は, 導入・文献レビュー6項目, 研究手法 10項目, 結果と考察 6項目 (各項目 0~2点, 44点満点) からなる, 第二言語習得・外国語教育研究におけるメタ分析を評価する枠組みを提示している。これらは, メタ分析研究の粗製濫造を防ぎ, 次の段階へと進める上で重要な作業だと言える。

しかしそれ以前に, そもそもメタ分析の手続きには, 先行するメタ分析研究においても十分明瞭になっているとは言い難い点がある。以下, 第二言語習得・外国語教育研究におけるメタ分析研究の嚆矢となった Norris and Ortega (2000)を中心に据えて, 具体的に検討してみたい。

## 2. 測定法や尺度の異なるものを一緒にしていいのか

第二言語の指導の効果を分析した Norris and Ortega (2000)によれば, 1980年から1998年までに出版された 250 を超える論文の内, 基準を満たす 40 研究の明示的指導 ( $k = 71$ ) の効果量の平均 ( $d = 1.13, 95\% \text{ CI } [0.93, 1.33]$ ) は, 19 研究の暗示的指導 ( $k = 29$ ) の効果量の平均 ( $d = 0.54, 95\% \text{ CI } [0.26, 0.82]$ ) を上回る。明示的指導の方が暗示的指導より有効だという結論がここから導き出されている。このメタ分析において Norris and Ortega (2000)は, 演繹的にメタ言語的な説明が与えられるか, 特定の言語形式に注意を向けたり規則に到達するよう指導されたりする場合を「明示的」(explicit)と定義し, 明示的指導では, Focus on Forms (FonFs)よりも Focus on Form (FonF) のほうが僅かに効果量大きいことも示している (Explicit FonFs が  $d = 1.08, 95\% \text{ CI } [0.80, 1.36]$ で, Explicit FonF が  $d = 1.22, 95\% \text{ CI } [0.91, 1.53]$ )。

しかし, このメタ分析に含められた一次研究を見てみると, 実験のデザインや測定法の異なる研究が多数混在していることが分かる。例えば表 1 に示す 2 つの論文は, 対象言語と明示的文法指導の下位分類において異なる, 類似の文法概念の指導を意図した研究とみなせる。Leow (1998)は, 文法形式への接触量の違い (一度か, 3 週間後に同じ手続きをもう一度か) と指導法の違い (教師中心か学習者中心か) の組み合わせによる 4 グループ 88 名に対して, 4 択の文完成テストと与えられた動詞の語形を変化させて空欄を埋めるテストを行い, それぞれについて正答数を得点として平均と標準偏差を報告している。それに対して DeKeyser (1997)は, 指導順序・項目ごとの練習量が異なる 3 グループ 61 名に対して写真・イラストと文を対応させるテストを行い, 反応時間 (ミリ秒) と誤答数について平均と標準偏差を報告している。メタ分析においてはその効果量, すなわち標準化した平均差を求めているとはいえ, Shin (2010)でも指摘されている通り, このような研究の結果を一緒にするべきではないだろう。

表 1

Norris and Ortega (2000)の一次研究の例

論文	Dekeyser (1997)	Leow (1998)
指導法の分類	Explicit FonFs	Explicit FonF
比較群	Explicit FonFs	Explicit FonFs
対象言語	人工言語	スペイン語
対象項目 <sup>a</sup>	人称・性・数	人称・性・数

Note. <sup>a</sup> 対象項目の分類は Watari (2014)に基づく。

このことは、メタ分析における“apples and oranges problem”（多様な研究がごちゃ混ぜにされていること）、あるいは“garbage in, garbage out”（ゴミのような研究からはゴミのような結果しか出てこないこと）として古くから指摘されている問題が関係している（山田・井上, 2012, pp. 29-30）。ゴミのような研究でなかったとして、第二言語習得・外国語教育研究の現状においては、異なる関心・環境・対象・条件等からなる複数の独立した一次研究をメタ分析によって集約する場合、その結果をどの程度一般化し得るかという外的妥当性のほうが重視されていると言える。しかし仮に、メタ分析に医学・薬学研究における特定の治療法や薬剤の効果を明らかにするような水準を望むのであれば、研究の内的妥当性、つまり独立変数と従属変数との間の因果関係の厳密さを重視すべきである。少なくとも、そもそも生データ同士で比較できるのは、「結果が意味のある尺度に基づいて報告されており、分析に含めた研究が全て同じ尺度を用いている場合」であることは念頭に置いておきたい（Borenstein, Hedges, Higgins, & Rothstein, 2009, Chapter 4, Section 2, para. 1）。

例えば Norris and Ortega (2000)に含められた一次研究の中で、文法性判断テスト (Grammaticality Judgment Tests, GJT)を結果の測定に用いている研究のみで、ランダム効果モデルによるメタ分析を実施すると図 1 のような結果が得られる ( $k = 11, g = 1.07, 95\% \text{ CI } [0.60, 1.53]$ )。Norris and Ortega (2000)によれば、全体の効果量の平均は  $d = 0.96, 95\% \text{ CI } [0.78, 1.14]$ であるから、文法性判断テストは効果をやや強めに出していることになる。

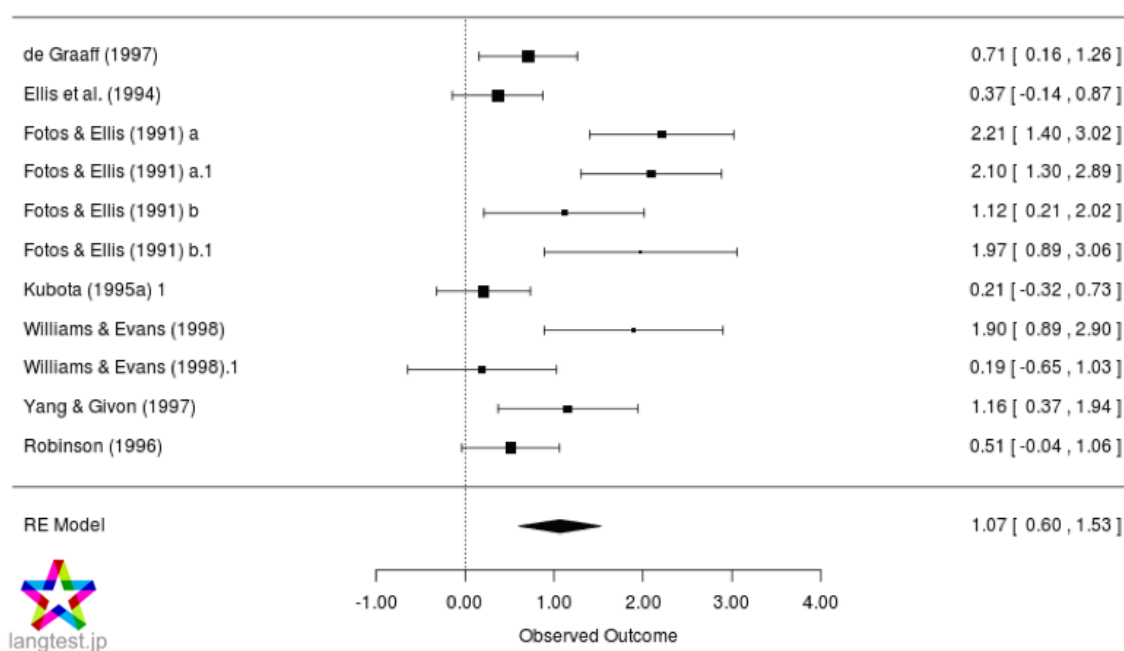


図 1. GJT に限定した Norris and Ortega (2000)の再分析のフォレストプロット

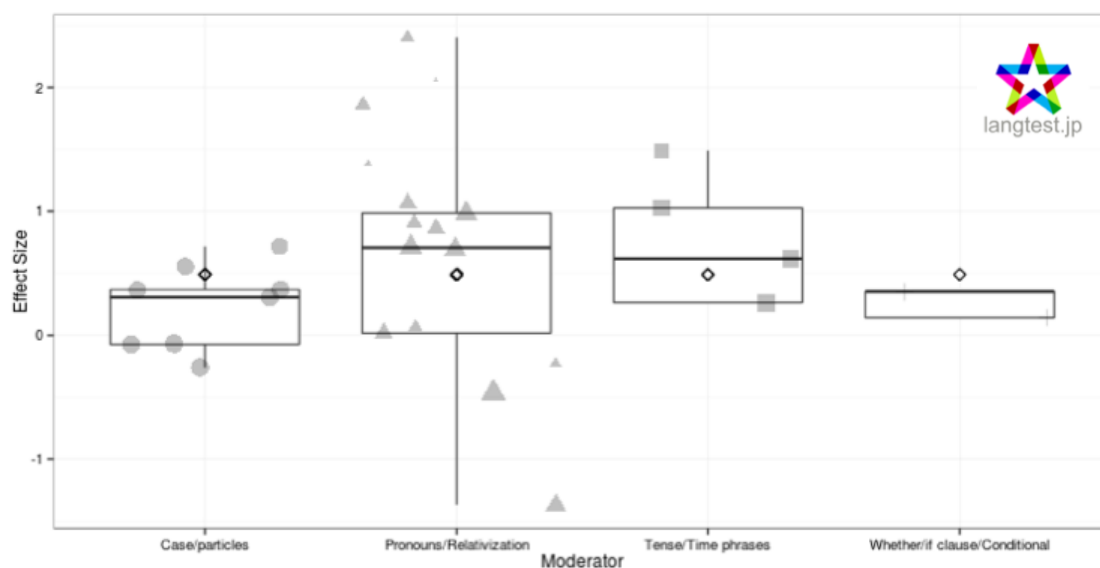


図 2. VanPatten and Cadierno (1993)およびその追試のメタ分析 (対象項目別)

内的妥当性を高めるために、当該分野の追試研究(replication)を蓄積し、限られた範囲でメタ分析を行うという選択もある (Lee, Jang, & Plonsky, 2014; Shintani, in press)。Norris and Ortega (2000)にも、Tomasello and Herron (1989)の追試 (とみなせる) 研究 (Ellis, Rosszell, & Takashima, 1994; Fotos & Ellis, 1991; Kubota, 1995a) や、VanPatten and Cadierno (1993)の追試 (とみなせる) 研究 (Cadierno, 1995; DeKeyser & Sokalski, 1996; Kubota, 1996;

Salaberry, 1997; VanPatten & Oikkenon, 1996; VanPatten & Sanz, 1995) が含まれている。後者 (テスト結果全てを分析に含めると  $k = 29$ ) について同様にメタ分析を実施すると,  $g = 0.55$ , 95% CI [0.27, 0.83]という平均効果量が得られ, さらに Watari (2014)に基づいて対象項目を調整変数とする分析を行うと, 図 2 のような結果が示された。つまり Norris and Ortega (2000)に含められた一次研究に関する限り, 明示的指導全体の効果量との比較ではいわゆる *processing instruction* の効果は弱く, さらに効果が見られるのは代名詞・関係詞節化 ( $g = 0.68$ , 95% CI [0.30, 1.07]) と時制・時間表現 ( $g = 0.84$ , 95% CI [0.13, 1.56]) の指導であって, 格・分詞 ( $g = 0.24$ , 95% CI [-0.26, 0.74]) および *Whether/if* 節・条件節 ( $g = 0.25$ , 95% CI [-0.80, 1.30]) の指導に効果があると言うことはできないことが分かる。

他方, 外的妥当性の点から見て重要なことは, そのメタ分析研究の括りに対して当該分野の研究者から納得を得られるかどうか, 意味のある調整変数(*moderator variables*)を設定して一次研究のみからでは明らかにならない知見をもたらす得るかどうかであると考ええる。例えば In'nami and Koizumi (2009)は, テストのフォーマットの影響について「幹の等価性」(冒頭の, 何を問うているかを説明する部分が同じまたは類似であるかどうか)という調整変数を導入することで, 10,000 以上の論文の内では基準を満たす 37 の研究に対するメタ分析から, 幹が等価であれば多肢選択式問題のほうがオープンエンドの問題よりも易しいが, 等価でない場合はオープンエンドの問題のほうが易しくなる (多肢選択から見ると, 幹が等価の場合  $g = 0.49$ , 95% CI [0.10, 0.88]で, 等価でない場合  $g = -1.02$ , 95% CI [-1.46, -0.58]) という結論を導き出している (In'nami & Koizumi, 2009, p. 234; 印南, 2012)。

### 3. 比較の仕方が異なるものを一緒にしていいのか

メタ分析で実際に用いられることの多い効果量には, (a) 平均 (連続変数) に基づくもの (Cohen's  $d$ , Hedge's  $g$ ), (b) 二値変数に基づくもの (オッズ比, リスク比), (c) 相関に基づくもの (Fisher's  $z$ ) がある (Borenstein et al., 2009; 山田・井上, 2012)。これらは相互に変換することが可能だが, 重要なことは, 少なくとも a)と b)は常に 2 つの群の比較によって算出されるものだけということである。第二言語習得・外国語教育研究のメタ分析で用いられることの多い(a)は, 平たく言えば「両群は, 標準偏差あたりで見てどのくらい離れているのか」あるいは「両群の分布はどの程度重なっているのか, いないのか」という指標である (浦野, 2013)。

Norris and Ortega (2000)は事後(immediate post)テストの結果を基本として, 各処置群に対する Cohen's  $d$  を算出しているが, 2 つの群のパターンは以下の 5 つに及ぶ (Norris & Ortega, 2000, pp. 446-447)。

1. 統制群と 1 つ以上の処置群とを事後テストの結果で比べてを算出 ( $n = 10$ )
2. 比較群と処置群とを事後テストの結果で比べて算出 ( $n = 20$ )
3. (人工言語などの) 前提知識なしとして統制群を設けていないものについては, で

きるだけ偶発的な条件をベースラインの比較群とし、処置群と事後テストの結果で比べて算出 ( $n = 4$ )

4. 統制群も比較群もないが、事前・事後の結果を報告しているものについては、各処置群について事前・事後で算出 ( $n = 5$ )
5. 統制群も比較群もなく事前についての報告もないものについては、できるだけ偶発的な条件をベースラインの比較群とし、処置群と事後テストの結果で比べて算出 ( $n = 10$ )

多くの研究の知見をメタ分析に含めるためと断って上記のパターンを明示しているとは言え、ここまで異なるデザインの研究の結果を一緒にすることは果たして妥当と言えるだろうか。実際 Watari (2014)は、明示的指導を含む研究の全てのテスト結果に基づく Norris and Ortega (2000)の再分析において、いくつかの比較のタイプ (Explicit FonFs vs. Implicit FonF, Explicit FonFs vs. Implicit FonFs, Explicit FonF vs. Explicit FonF) は信頼区間に 0 を含んでおり効果が認められないこと、統制群との比較に限定した場合、(重み付き)平均効果量は Norris and Ortega (2000)が各指導法について報告しているよりもいくぶん小さいこと (Explicit FonFs が  $g = 0.95$ , 95% CI [0.65, 1.25]で、Explicit FonF が  $g = 1.06$ , 95% CI [0.76, 1.37]) を明らかにしている。逆に言えば、少なくとも Norris and Ortega (2000)の時点では、第二言語習得・外国語教育研究の当該領域においてメタ分析の基準を満たす、統制群との比較による研究は 10 しかなかったということを意味しており、少なくとも今後は、そのような状況でメタ分析を行うこと、あるいは当該分野について何らかの一般化を行なうことが適当なのか考える余地があるように思われる。

さらなる問題は効果量の計算方法である。(a)は両群のサンプルサイズと平均・標準偏差から計算するのが一般的であるが、Norris and Ortega (2000)の一次研究においてそれが可能だった (つまり両群のサンプルサイズと平均・標準偏差、ないしは元データがきちんと報告されていた) のは 58%に過ぎず、残りは両群のサンプルサイズと  $t$  値・自由度、もしくは両群のサンプルサイズと  $F$  値・誤差項自由度から効果量を計算している (Norris & Ortega, 2000, pp. 443-444)。しかし Norris and Ortega (2000)の場合、一次研究の多くは指導法の異なる複数の処置群を持つものとして分析されており、 $F$  値で効果量を計算した場合、その違いは見えなくなってしまう。例えば Cadierno (1995)は、processing instruction (Explicit FonF)と traditional instruction (Explicit FonFs)および統制群に対する事前テストと 3 回の事後テストの結果を繰り返しのある分散分析を用いて分析しているが、各群の標準偏差が報告されていない。この研究において processing instruction 群は解釈タスクにおいて事前テストからのひときわ高い伸びを見せているが、 $F$  値で計算された効果量では traditional instruction 群との区別ができなくなってしまう。

この問題は、「効果量はどのレベルで一つにまとめるのが適切か」ということにも関連している。In'nami and Koizumi (2009)は、各研究の効果量間の非独立性によるバイアス

(同一の実験協力者が受けたテストが複数あって、それがそのままメタ分析に含められると、その研究の結果が全体の平均効果量に強い影響を持つことになるということ)を減らすため、研究ごとに1つの効果量を算出するのが適切であると述べる一方で、重要かもしれない変数を混ぜることになり、調整変数に関する情報が失われるかもしれないと指摘している (In'nami & Koizumi, 2009, p. 229-230)。

Lipsey and Wilson (2001)も同様に、分析の単位は効果量ではなく個別の研究であり、効果量はひとつにまとめるか、代表の(ないしはランダムに)効果量を選択すべきであるとしている。Norris and Ortega (2006)はその議論を紹介しつつ、(a) 研究ごとに一つの効果量のみ含める、(b) 各研究の全ての効果量を含めるが、慎重な記述に努め、それ以上の(因果関係の)推測的な分析は避ける、(c) 観測値の独立性からの逸脱を処理する統計手法を用いる、という選択肢があり、ひとつにまとめて情報が失われることや、元データを集めることの困難さを考慮すると、言語学習・教育研究のメタ分析においてさしあたり明らかにしたいことから見て(b)が最善ではないかと述べている(Norris & Ortega, 2006, p. 30)。実際 Norris and Ortega (2000)は、研究ごとだけでなく、処置のタイプ(FonF/FonFs, Explicit/Implicit)や従属変数のタイプごと、処置の長さごと、処置のタイプおよび群ごとの事前・事後の変化、遅延事後の平均効果量も算出・報告している(前述の通り、比較のパターンと計算方法の問題は残ったままであるものの)。この問題に対する回答は「メタ分析を通じて知りたいことによる」というものであるのかもしれない。

#### 4. 誰もがメタ分析を実際に行なうことができるのか

メタ分析は、ナラティブ・レビューと比べてのメリットが頻繁に説かれる (Gass, Behney, & Plonsky, 2013; 印南, 2012; Norris & Ortega, 2006; Oswald & Plonsky, 2010)。例えば Jeon and Yamashita (2014, p. 176)に端的に示されているように、どういう範囲で検索し、どういう基準で当該の研究をメタ分析に採用したかが原理的かつ追跡可能だという点で、当該領域の先行研究の網羅性と絞り込みにおいては間違いなくナラティブ・レビューより客観性・再現性が高いと言える(検索のソースとなるデータベースの問題については In'nami & Koizumi, 2010 を参照)。

しかし、必要な論文、あるいはデータに実際にアクセスできるかどうかと言えば、必ずしもそうとは言えない現状がある。これはいわゆる“file drawer problem”(公刊される研究は有意なものに偏り、公刊されない研究は注目されないこと)にも関連しているが、それ以前の現実的な問題として、電子ジャーナルや書籍へのアクセス環境が十分に整っていない学生・院生・研究者がメタ分析を行うのは難しい場合が少なくない。「ブーム」の様相を呈しているだけに、メタ分析の実行可能性(feasibility)という問題にももう少し注意が向けられても良いと思われる。

理論的により深刻な問題として、検索とデータの入手に関して「著者に連絡を取る」

ことが認められているということがある。例えば印南 (2012) は、メタ分析の手続きとして「該当分野の研究者に連絡をとり、収集に漏れがないかを確認」することを推奨し、「著者に連絡をとり必要な情報を送ってもらうことで、研究者間のネットワークが広がることをメタ分析の意義に挙げているが、客観性・再現性の点で問題があるとは言えないだろうか (印南, 2012, pp. 228, 248)。客観性・再現性がメタ分析の優れた点であるならばなおさら、論文・データ・計算過程へのアクセスは透明になっていなければならないはずである。

例えば Norris and Ortega (2000) の一次研究にも、記載されている情報だけでは一部あるいは全ての効果量が算出できず、著者から直接データを得たことが推察される論文が含まれているが、そのことに関する記述が無い限り、追試や再分析は不完全なものになってしまう。したがって、メタ分析研究は少なくとも (a) 著者やその他の個人・団体にコンタクトを取ることによって初めて得られたデータがある場合、その事実を論文中に明記し、(b) どうすれば同じデータが得られるのかどうかを可能な限り示すべきである。あるいは、(c) 必要な情報が記載されていない論文は一切メタ分析に含めないという選択もある。一次研究に対する波及効果を考えればむしろ、こちらのほうが第二言語習得・外国語教育研究に必要な選択なのではないだろうか。

## 謝辞

本稿は、メソドロジー研究部会 2014 年度第 2 回研究会 (2014 年 9 月 15 日、信州大学) における口頭発表の内容にもとづきます。発表に対してご意見・ご質問をお寄せくださったみなさん、分析手法について詳しい解説をしてくださった芝浦工業大学の印南洋先生、資料を提供してくださった The University of Auckland の新谷奈津子先生、Northern Arizona University の Luke Plonsky 先生、草稿に対してコメントをくださった北海学園大学の浦野研先生に感謝いたします。特に関西大学の水本篤先生は、langtest.jp や分析手法の解説のみならず、草稿の詳細なチェックにも多大なご協力をいただきました。この場を借りて深く感謝の意を表します。

## 参考文献

- \*は、Norris & Ortega (2000) でメタ分析に含められた研究の内、本論のメタ分析に含めた研究を示す。
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. West Sussex, U.K.: John Wiley & Sons.
- \*Cadierno, T. (1995). Formal instruction from a processing perspective: An investigation into the Spanish past tense. *The Modern Language Journal*, 79, 179-193.



- \*De Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19, 249-276.
- DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19, 195-221.
- \*DeKeyser, R. M., & Sokalski, K. J. (1996). The differential role of comprehension and production practice. *Language Learning*, 46, 613-642.
- \*Ellis, R., Rosszell, H., & Takashima, H. (1994). Down the garden path: Another look at negative feedback. *JALT Journal*, 16, 9-24.
- \*Fotos, S., & Ellis, R. (1991). Communicating about grammar: A task-based approach. *TESOL Quarterly*, 25, 605-628.
- Gass, S. M., Behney, J., & Plonsky, L. (2013). *Second language acquisition: An introductory course* (4th ed). New York: Routledge.
- 印南洋 (2012). 「メタ分析」平井明代(編)『教育・心理系研究のためのデータ分析入門』(pp. 224-251) 東京図書.
- In'nami, Y., & Koizumi, R. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing*, 26, 219-244. doi: 10.1177/0265532208101006
- In'nami, Y., & Koizumi, R. (2010). Database selection guidelines for meta-analysis in applied linguistics. *TESOL Quarterly*, 44, 169-184. doi: 10.5054/tq.2010.215253
- Jeon, E. H., & Yamashita, J. (2014). L2 reading Comprehension and its correlates: A meta-analysis. *Language Learning*, 64, 160-212. doi: 10.1111/lang.12034
- \*Kubota, M. (1995). The Garden Path technique: Is it really effective? *Working Papers of Chofu Gakuen Women's Junior College*, 27, 21-48.
- \*Kubota, M. (1996). The effects of instruction plus feedback on Japanese university students of EFL: A pilot study. *Bulletin of Chofu Gakuen Women's Junior College*, 18, 59-95.
- Lee, J., Jang, J., & Plonsky, L. (2014). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36, 1-23. doi: 10.1093/applin/amu040
- Leow, R. P. (1998). The effects of amount and type of exposure on adult learners' L2 development in SLA. *The Modern Language Journal*, 82, 49-68.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Mizumoto, A. (n.d.). *Langtest.jp Web App*. Retrieved from [http://langtest.jp:3838/meta/citation\("metafor"\); citation\("meta"\)](http://langtest.jp:3838/meta/citation()
- Mizumoto, A., Urano, K., & Maeda, H. (2014). A systematic review of published articles in ARELE 1-24: Focusing on their themes, methods, and outcomes. *ARELE*, 25, 33-48.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and

- quantitative meta-analysis. *Language Learning*, 50, 417–528. doi: 10.1111/0023-8333.00136
- Norris, J. M., & Ortega, L. (2006). The value and practice of research synthesis for language learning and teaching. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 3–50). Philadelphia: John Benjamins.
- Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics*, 30, 85–110. doi: <http://dx.doi.org/10.1017/S0267190510000115>
- Plonsky, L. (2012). Replication, meta-analysis, and generalizability. In G. Porte (Ed.), *A guide to replication in applied linguistics* (pp. 116–132). New York: Cambridge University Press.
- Plonsky, L. (n.d.). *Meta-analysis in applied linguistics*. Retrieved from [http://oak.ucc.nau.edu/ldp3/bib\\_metaanalysis.html](http://oak.ucc.nau.edu/ldp3/bib_metaanalysis.html)
- Plonsky, L., & Brown, D. (2014). Domain definition and search techniques in meta-analyses of L2 research (Or why 18 meta-analyses of feedback have different results). *Second Language Research*, 1–12. doi: 10.1177/0267658314536436
- \*Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search and instructed conditions. *Studies in Second Language Acquisition*, 18, 27–68. doi: <http://dx.doi.org/10.1017/S0272263100014674>
- \*Salaberry, M. R. (1997). The Role of input and output practice in second language acquisition. *The Canadian Modern Language Review*, 53, 422–51.
- Shin, H. W. (2010). Another look at Norris and Ortega (2000). *Working Papers in TESOL and Applied Linguistics*, 10, 15–38.
- Shintani, N. (in press). The effectiveness of processing instruction on L2 grammar acquisition: A meta-analysis. *Applied Linguistics*, 36.
- 丹後俊郎 (2002). 『メタ・アナリシス入門: エビデンスの統合をめざす統計手法』朝倉書店.
- Tomasello, M., & Herron, C. (1989). Feedback for language transfer errors. *Studies in Second Language Acquisition*, 11, 385–395.
- 浦野研 (2013, August). 「有意性と効果量についてしっかり考えてみよう」外国語教育メディア学会 (LET) 第 53 回全国研究大会ワークショップ, 東京 (文京学院大学)  
Retrieved from <http://www.urano-ken.com/blog/2013/08/05/let2013-workshop/>
- 浦野研・亘理陽一 (2013, October). 「英語教育研究における追試 (replication) の必要性」外国語教育メディア学会 (LET) 関西支部メソドロジー研究部会 2013 年度第 2 回研究会, 秋田 (大学コンソーシアムあきた) . Retrieved from <http://www.urano-ken.com/blog/2013/10/25/methoken-in-akita/>

- \*VanPatten, B., & Cadierno, T. (1993). Input processing and second language acquisition: A role for instruction. *Modern Language Journal*, 77, 45-57.
- \*VanPatten, B., & Oikkenon, S. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition*, 18, 495-510.
- \*VanPatten, B., & Sanz, C. (1995). From input to output: Processing instruction and communicative tasks. In F. Eckman, D. Highland, P. Lee., J. Mileham, & R. Weber (Eds.), *SLA theory and pedagogy* (pp. 169-185). Hillsdale, NJ: Lawrence Erlbaum.
- Watari, Y. (2014, August). What does "explicit" mean? A methodological consideration in explicit grammar teaching research. Poster presented at the International Association for Applied Linguistics (AILA), Brisbane, Australia.
- \*Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139-155). Cambridge University Press.
- 山田剛史・井上俊哉 (2012). 『メタ分析入門: 心理・教育研究の系統的レビューのために』東京大学出版会.
- \*Yang, L., & Givón, T. (1997). Benefits and drawbacks of controlled laboratory studies of second language acquisition. *Studies in Second Language Acquisition*, 19, 173-194.