How Multiple Regression Models Can Be Written to Better Reflect the Complexity of First and Second Language Acquisition Research: An Attempt to Limit Type VI Error

Kyle Perkins **Isadore Newman**
Florida International University

The purpose of this paper is to show how multiple regression models can be written to investigate and to better reflect the complexity of four major research themes in first and second language acquisition research in order to limit Type VI error. The four major themes were selected because of their importance to their respective fields, their multidimensional nature, and the interdisciplinary research influence on them. Each theoretical perspective examined in this paper brings with it its own research methodology, and; thus, the need for a paper that demonstrates how multiple regression models can be written to investigate four of the major foci of first and second language acquisition research to better reflect the complexity of language acquisition research.

Second language acquisition is an interdisciplinary, multidimensional field whose principal foci include the study of how second language learners create a new language with scant exposure to the target language and the description of what is learned of the second/target language and what is not learned (Gass, 2013). Researchers from a broad spectrum of disciplines (e.g., education, linguistics, psychology, sociology, psycholinguistics, sociolinguistics, discourse analysis, neurolinguistics, and conversational analysis) conduct second language acquisition research. The first language acquisition topics shown below are also multidimensional and have been investigated with interdisciplinary approaches.

First and second language acquisition data were chosen as examples for which multiple regression models could be constructed and shown in this paper because first and second language acquisition involve non-linear, multidimensional, and multiplicative phenomena. The following examples are offered to illustrate this point. First and second language differences can predict the amount of avoidance that learners show in using particular linguistic structures. First and second language similarities can result in differential learning rates, different learning paths, an overproduction of first language-influenced forms, and predictability/selectivity. Semantic and conceptual facts about the first language can affect a learner's choice of forms in the second language. Some individuals are more successful in learning a second language than others.

Some factors that may be responsible for these differences include age (i.e., the Critical Period Hypothesis), aptitude, motivation (which may change over time), attitude, affect, anxiety, social distance (perceived or real), learning strategies, extroversion and introversion. English Language Learners develop hypotheses (consciously or unconsciously) regarding the rules of the second language and create rules that may vary according to the context of use. Linguistic principles reflect the infinite possibilities of human language creation and the limits of human language variation. English Language Learners create their own language systems (i.e., an Interlanguage), which is composed of elements from the learner's first language, the second/target language, and elements that do not originate from either the target language or the native language, and a learner's Interlanguage is in constant flux. There is individual variation in learner data and individual data can be obscured with/within group data. The paths of acquisition are not always identical for learners from different language groups/language families. The second language research topics listed above are discussed in Gass (2013) where more information on each of these topics can be found.

Language Acquisition Themes

Major themes in first and second language acquisition research include:

1. A general theory of second language acquisition should account for language acquisition by learners with a variety of characteristics in a variety of social and instructional environments.
2. There are aspects of language acquisition that are common to all second language learners and contexts.

3. There is interest in predicting differences in the success of second language acquisition using predictor variables such as the second language learners' personalities, their general and specific intellectual abilities, their motivation, or their age.
4. Second language learners pass through sequences of development, some of which are predictable while others are not.
5. Second language learning in an instructional environment is different from second language learning in a non-instructional setting.
6. There is interest in identifying specific variables that may affect learning similarly in different environments and measuring those effects.
7. Second language acquisition/learning does not exhibit monotonic, linear development because new language forms and patterns must be integrated into an existing system, called the learner's interlanguage, and those forms and patterns must be restructured until all of the pieces fit.
8. Interlanguage is composed of diverse elements from the second language learner's native language and target language and from other sources.
9. Some second language acquisition researchers use multiple studies to address a common question.
10. There is interest in the extent to which both the first language and the second language are activated when second language learners use their second language.
11. Both functional and typological approaches are used to investigate second language acquisition.
12. Second language acquisition researchers are interested in the extent to which the variation in a second language learner's output is linguistically based or socio-linguistically determined.
13. Second language acquisition researchers investigate and describe the nature and function of input, interaction, and output.
14. There are numerous factors that may be responsible for some second language learners being more successful than others, and some of those factors include age; aptitude; attitude; anxiety; motivation, both instrumental and integrative; socio-psychological influences; intelligences; and learning strategies (Gass, 2013).
15. Form-focused instruction focuses the learner's attention to the forms and structures of the target language within the context of communicative interaction; the instructor may give metalinguistic information, highlight the form in question, or provide corrective feedback. Some studies have indicated that form-focused instruction and corrective feedback are more effective in promoting second language learning than programs that are limited in their scope to emphases on accuracy, comprehension, or fluency *per se*.
16. Constrained skills theory (Paris, 2005) is a reconfiguration of the reading developmental process with a continuum of skills ranging from highly constrained to unconstrained skills.
17. Chenoweth and Hayes' (2001) model of written language production consists of three levels and the model can be used to differentiate between novice writers and expert writers.
18. Research has shown that there are reader variables and text variables that affect the nature of second language reading comprehension (Alderson, 2000).
19. Response to Intervention (RTI) is a means of providing early, systematic assistance to first and second language children who are experiencing academic difficulty. The hallmarks of RTI are early intervention, frequent progress measurement, and research-based instructional intervention for those children who continue to have difficulty (National Association of State Directors of Special Education, 2005).

According to Fuchs and Vaughn (2012), RTI was codified into federal law through the reauthorization of the Individuals with Disabilities Improvement Act of 2004 as a method for learning disability identification and, as a result, has become a major presence of education reform. Presently, 47 states are delivering some form of RTI (Berkeley, Bender, Peaster, & Saunders, 2009). Twenty nine states use the essential components to meet the special needs of students with disabilities, and 14 states use essential RTI components to meet the needs of English Language Learners. The 2002 reauthorization of the Elementary and Secondary School Education Act expanded accountability for improving schools and

the attained proficiency of all students having difficulty mastering mathematics and reading (McInerney & Elledge, 2013; National Center on Response to Intervention, 2010).

Purpose

While it is the case that such a variety of expertise from a plethora of disciplines enhances the fields of first and second language acquisition, the downside is that each of the different theoretical perspectives naturally brings with it its own research methodology. In addition, many studies consider only one or two variables and then generalize to the whole. Thus, there is the need for conducting research based on first and second language acquisition theories and the multiple variables associated with them. This paper is a theoretical, methodological piece whose purpose is to demonstrate how multiple regression models can be written to investigate some of the major foci of first and second language research, and to better reflect the complexity of some of the major findings in order to limit Type VI error. Type VI error has been noted in detail in the literature by Newman, Dietchman, Burkholder, Sanders, and Ervin (1976), and defined in research situations by Newman, Fraas, Newman, and Brown (2002), where "A Type VI error occurs when an inconsistency exists between the study's research question and the analytic technique and/or research design used in the study" (p. 138).

Methods

Four major themes were selected from the sampler of research topics. The selected topics are amenable for the development of multiple linear regression models and they illustrate that learning is neither additive nor linear. Research methodology must reflect these realities, which multiple linear regression models can capture. The models presented in this study are not a component of a statistical package. Although the models presented here are related to first and second language examples, the applicability of these models is multidisciplinary (i.e., they could be utilized in psychology and other behavioral sciences, medicine, public health, social work, and other social sciences).

In the presentation of various multiple regression models in the body of the paper, the authors will refer to testing the Full Model versus the Restricted Model. A Full Model allows a treatment group to have its own achievement mean, which is accomplished through the use of dichotomous predictor variables and least squares weighting coefficients calculated to minimize the sum of the squared values in the error vector (McNeil, Newman, & Fraas, 2012). A Full Model is necessary in order to ascertain whether the data collected by the researcher support a research hypothesis. A research hypothesis always has a corresponding statistical hypothesis, the null hypothesis that states the condition contained in the research hypothesis is not true. In this study, the authors refer to a Restricted Model which, when compared to the Full Model, reflects the condition being tested.

The formula used to calculate the F-statistic for testing the difference between the R^2 values of the Full and Restricted Model was:

$$F(df_n, df_d) = \frac{(R_F^2 - R_R^2)/df_n}{(1 - R_F^2)/df_d} \quad (1)$$

The R^2_F is the proportion of observed criterion variance accounted for by the Full Model. The R^2_R is the proportion of observed criterion variance that the Restricted Model explains. The difference between these two R^2 values ($R^2_F - R^2_R$), which is located in the numerator of the F test, is the proportion of unique variance in the criterion variable that the deleted predictor variable(s) explain. This value ($R^2_F - R^2_R$) is divided by the degrees of freedom of the numerator (df_n), which is the number of linearly independent vectors used to account for the proportion of variance difference between R^2_F and R^2_R. That is, df_n is equal to the difference between the number of linearly independent vectors in the Full Model and the number of linearly independent vectors in the Restricted Model. The denominator of the F test contains the value equal to 1 minus R^2_F (i.e., $1 - R^2_F$), which is the proportion of variance unexplained by the Full Model (error variance). This value ($1 - R^2_F$) is divided by the degrees of freedom of the denominator (df_d). The df_d value is equal to the number of observations (N) minus the number of linearly independent vectors in the Full Model. In essence, df_d is the number of observations that are free

to vary after coefficients for each of the linearly independent vectors in the Full Model have been calculated (McNeil et al., 2012, pp. 68-69).

For further information on multiple regression models, the authors recommend Bottenberg and Ward (1963), Byrne (1974), Cohen, Cohen, West, and Aiken (2002), Hair, Black, Babin, Anderson, and Tatham (2006), Jennings (1967), Kerlinger and Pedhazur (1973), McNeil et al. (2012), McNeil and Newman (1996), McNeil, Newman, and Kelly (1996), and Ward and Jennings (1973).

Multiple Regression Models

The authors will now demonstrate how multiple regression models can be written to better reflect the complexity of first and second language acquisition research topics; the general topics include form-focused instruction, restructuring in second language learning, the interaction of test-takers and tasks in second language reading comprehension, and RTI.

Form-Focused Instruction

Research on native French-speaking students enrolled in English as a Second Language classes in Canada has shown that although many learners developed good listening comprehension, communicative confidence, and fluency in English, they continued to experience problems with linguistic accuracy and complexity (Day & Shapson, 1991; Ellis, 2001; Lightbown & Spada, 1990, 1994; Lyster, 1994; White, 1991; Williams, 1999). Pedagogical theory suggests that learners will benefit from form-focused instruction that is defined as the instructor's focusing the learners' attention on the forms and structures of the language within the context of communicative interaction by providing metalinguistic information, highlighting the form in question, or providing corrective feedback.

Research has suggested that form-focused instruction and corrective feedback within communicative and content-based second and foreign language classes can assist learners in augmenting their knowledge and use of particular grammatical features (Lyster, 1994). Form-focused instruction may be useful in classroom situations where learners share the same first language, and transfer from that same first language may result in errors that are unlikely to lead to global errors that completely impede communication (Ellis, 2001). In situations such as these, learners are ill equipped to discover/notice errors on their own. Focused instruction, therefore, will help learners to notice targeted features in subsequent input and interaction (not in a statistical sense), although the effects of instruction may not be long-lasting and learners at the lower end of the proficiency continuum may experience difficulty attending to language form (Lightbown & Spada, 2006).

If researchers want to estimate how much the treatment (i.e., form-focused instruction) accounts for variance in the dependent score (a test of linguistic and/or communicative competence), independent of person differences (person vectors), they could use the following model to control for individual differences. Testing Model 1 (the Full Model) versus Model 2 (the Restricted Model) estimates how much the treatment accounts for a significant amount of variance in the dependent score, independent of person differences (person vectors).

Full Model 1: $$Y = b_0U + b_1P_1 + b_2P_2 + b_3P_3 + ... + b_N - b_{N-1} + b_N\,(\text{Treatment}) + E \quad (2)$$

Restricted Model 2: $$Y = b_0U + b_1P_1 + b_2P_2 + b_3P_3 + ... + b_N - b_{N-1} + E \quad (3)$$

where

Y = the stacked dependent variable: Subject 1 at Time 1, Subject 1 at Time 2, etc.

$b_0 - b_{n-1}$ = partial regression weights

U = a unit vector and has a 1 for each subject in the sample

P_1 = 1 if the score came from Person 1; otherwise, 0

P_2 = 1 if the score came from Person 2; otherwise, 0

P_n = 1 if the score came from Person $_n$; otherwise, 0

Treatment = 1 if the person was in Treatment; otherwise, 0

E = Error Vector (residuals)

McNeil et al. (1996) and McNeil et al. (2012) explain how regression models can be written to reflect dependent *t*-tests and repeated measures when the Y variable is stacked: Subject 1 at Time 1, Subject 1 at

Time 2…Subject 2 at Time 1, Subject 2 at Time 2, etc. In this case, N is the number of replicates and not the number of subjects, *ergo*, there is not a problem with inverting the matrix and a full rank solution. This is how the General Linear Model could be used to test repeated measures questions.

Restructuring in Second Language Learning

Second language competence is best viewed as a multi-variable, rather than a unitary variable construct, and second language competence does not develop in a linear pattern in all of its component parts (Perkins, Brutten, & Gass, 1996). Learners reorganize, reconfigure, and combine their existing knowledge structures in new ways, and, as a result, ability levels change differentially. Learners may reconfigure their previous knowledge by "chunking" information to reduce memory load and by developing strategies and models that help them to discern when and how facts and skills are relevant. Restructuring can also refer to new learning as well, where

> Changes that reflect restructuring are discontinuous or qualitatively different from a previous stage. Learning means the inclusion of additional information that must be organized and structured. Integrating new information into one's developing L2 [second language] system necessitates changes to parts of the existing system, thereby restructuring, or reorganizing, the current system and creating a (slightly) new L2 system. (Gass, 2013, p. 256)

Furthermore, according to Lightbown (1985),

> [restructuring] occurs because language is a complex hierarchical system whose components interact in non-linear ways. An increase in one area may reflect an increase in complexity or accuracy in another, followed by overgeneralization of a newly acquired structure, or simply by a sort of overload of complexity which forces a restructuring or at least a simplification in another part of the system. (p. 177)

The result of restructuring is often reflected in what is known as U-shaped behavior (Kellerman, 1985; Lightbown, 1983). The U is a graphic representation of the accuracy of learner forms. In an earlier state, at Time 1, a learner might produce, for example, the correct form of an irregular strong verb (i.e., Sarah ate an apple). At a later stage, at Time 2, the learner appears to have digressed and to have lost what he or she knew at Time 1, producing a form like "Sarah eated an apple." At an even later stage, Time 3, the learner's linguistic behavior resembles the Time 1 behavior (i.e., Sarah ate an apple).

If researchers wanted to investigate the stages of learning during restructuring, they could use the following multiple regression model. Testing Model 3 versus Model 4 examines if a second degree relationship (i.e., an inverted U shaped relationship) accounts for a significant amount of unique variance over and above what can be accounted for by a linear relationship and individual differences.

Full Model 3: $$Y = b_0U + b_1Stage_1 + b_2Stage_1^2 + b_3P_1 + \ldots + b_nP_n + E \quad (4)$$

Restricted Model 4: $$Y = b_0U + b_{n+1}(Stage_1)^1 + b_{n+2}(P_1) + \ldots + b_{n+n}(P_n) + E \quad (5)$$

where

$b_o - b_{n-1}$ = partial regression weights
U = a unit vector and has a 1 for each subject in the sample
Stage = 1 if the score was from a person in Stage 1
Stage = 2 if the score was from a person in Stage 2
Stage = 3 if the score was from a person in Stage 3
$Stage^2$ = Stage*Stage
P_1 = If the score came from Person 1, regardless of the stage s/he was in
P_2 = If the score came from Person 2, regardless of the stage s/he was in
P_n = If the score came from $Person_n$, regardless of the stage s/he was in
E = Error Vectors (residuals)

What has been described in the previous paragraph can be operationalized as relatively abrupt changes in the relative difficulty of test items for a person who moves from one stage of second language development to the next stage. Multiple regression models can indicate who and what changed during that developmental process.

Testing Model 5 against Model 6 permits the researcher to ascertain if there is a significant difference in the test scores between the three stages for second language development students. This model also

allows for repeated measures in which more than one test score on the dependent variable can come from the same subject in different stages.

Full Model 5: $Y = b_0U + b_1Stage_1 + b_2Stage_2 + b_3Stage_3 + b_4(P_1) + \ldots + (b_nP_{n)} + E$ (6)

Restricted Model 6: $Y = b_0U + b_{n+1}(P_{n+1}) + b_4(P_1) + \ldots + (b_nP_n) + E$ (7)

where

U = a unit vector and has a 1 for each subject in the sample
Stage 1 = 1 if the dependent score came from a person in Stage 1; otherwise, 0
Stage 2 = 1 if the dependent score came from a person in Stage 2; otherwise, 0
Stage 3 = 1 if the dependent score came from a person in Stage 3; otherwise, 0
$b_0 - b_{n-1}$ = partial regression weights
P_1 = 1 if the dependent score came from Person 1; otherwise, 0
P_2 = 1 if the dependent score came from Person 2; otherwise, 0
P_3 = 1 if the dependent score came from Person 3; otherwise, 0
P_n = 1 if the dependent score came from $Person_n$; otherwise, 0
E = Error Vector (residuals)

If Full Model 5 proves to be statistically significant, then it must be followed with a test of multiple comparisons.

Interaction of Test-Takers and Tasks in Reading Comprehension

Research in second language reading comprehension research has shown that factors within the reader and factors associated with the text being read affect the nature of second language reading (Alderson, 2000). A few example factors are listed below:

Reader variables:

- Content background knowledge
- Knowledge of the language
- Knowledge of the genre/text type
- Metalinguistic knowledge and metacognition
- Knowledge of the world
- Cultural knowledge
- Reader skills and abilities
- Reader purpose
- Reader motivation/interest
- Reader affect

Text variables:

- Text topic and content
- Text type and genre
- Text organization
- Text readability

Both first and second language reading comprehension researchers can expect changes in person ability and in item difficulty as readers advance across the attained reading comprehension proficiency continuum for the following reasons. Reading tasks vary in their difficulty according to the text, the test-tasks, and the reader as well as the interaction of text, test-tasks, and the reader. Texts vary by content (e.g., level of abstraction, information/propositional density, theme, text form or type, contextualization, and cultural conventions) and by the writer's style (e.g., use of vocabulary and structures, cohesion and coherence, and use of redundancy) (Alderson, 2000).

The multiple components involved in reading comprehension interact in different ways according to the proficiency of the reader and the characteristics of the text. Interaction can also involve the integration of meaning across words, sentences, and passages in addition to the reader and text variables presented above. There are demands of working memory at the word level (i.e., recall and retention of semantic meaning), at the sentence-level (i.e., merging of the syntactic and semantic cues to create a proposition),

and at the text level (i.e., synthesizing propositions into a coherent idea) (Paris & Hamilton, 2009). Over time, there may be changes in the reader's cognitive load processing ability, and all of these factors lead to the need to identify who changed and what changed as readers' progress across the attained reading comprehension proficiency continuum.

The following regression models can help researchers to identify who changed and what changed as readers become more proficient. Testing Model 7 versus Model 8 allows one to determine if the text accounts for a significant amount of unique variance in predicting the dependent variable when controlling for task and individual differences (person vectors), which allows and controls for the effects of each person having more than one score on the dependent variable. Testing Model 7 versus Model 9 allows one to determine if the task accounts for a significant amount of unique variance in predicting the dependent variable when controlling for text and individual differences (person vectors). Testing Model 10 versus Model 7 provides a test for interaction to determine if there is an interaction between task and text when the model controls for individual differences.

$$\text{Model 7: } Y = b_0U + b_1\text{Text}_1 + b_2\text{Text}_2 + b_3\text{Task}_1 + b_4\text{Task}_2 + b_5P_1 + b_6P_2 + \ldots + b_nP_n + E \quad (8)$$

$$\text{Model 8: } Y = b_0U + b_{n+1}\text{Task}_1 + b_{n+2}\text{Task}_2 + b_{n+n}P_1 + b_{n+n+1}P_2 + \ldots + b_{n+n+n+2}P_n + E \quad (9)$$

$$\text{Model 9: } Y = b_0U + b_{n+n+1}\text{Text}_1 + b_{n+n+2} + \text{Text}_2 + b_{n+n+n+1}P_1 + b_{n+n+n+2}P_2 + \ldots + b_{n+n+n+3}P_n + E \quad (10)$$

$$\text{Model 10: } Y = b_0U + b_1(\text{Task}_1*\text{Text}_1) + b_2(\text{Task}_1*\text{Text}_2) + b_3(\text{Task}_2*\text{Text}_1) + b_5(P_1 + \ldots + b_{N+1}(P_n) + E \quad (11)$$

where

U = a unit vector and has a 1 for each subject in the sample

$b_0 - b_{n-1}$ = partial regression weights

Text = 1 if narrative; 0, otherwise

Text = 1 if expository; 0, otherwise

Task = 1 if literal comprehension; 0, otherwise

Task = 1 if inferential comprehension, 0, otherwise

E = Error Vector (Residuals)

RTI

RTI is "a major theoretical and practical shift in [US] federal policy and law that affects both regular and special education classrooms" (Reutzel & Cooter, 2011, p. 44). The No Child Left Behind Act, the Individuals with Disabilities Act, and schools attempting to meet the criterion/threshold of Annual Yearly Progress use elements of RTI.

The RTI tiered model is being used for assessment, progress monitoring, and screening across a broad spectrum of content subject areas, including English language arts, and it is designed to provide early intervention for students who are struggling in the general education curriculum. RTI entails the provision of scientifically-based intervention directly geared to individual student needs, and; thus, requires assessing learning rates longitudinally and changes in level of performance in order to make data-based decisions.

A description of an RTI multi-tiered approach follows, "Tier I consists of the research-based core classroom instruction intended for all students. Tier II consists of supplemental instruction and interventions that are provided in addition to and in alignment with effective core instruction" (Bureau of Exceptional Education and Student Services, 2008). Decisions as to who is in need of Tier II interventions are determined through ongoing classroom assessments. Typically, Tier II interventions take place in small group situations where students receive a minimum of 30 minutes per day three times a week of "evidence-based instruction designed to supplement the core literacy program" (Reutzel & Cooter, 2011, p. 47 as cited in Govoni, 2011). The following is a three-tiered RTI model for English Language Learners:

Tier I: All students receive high quality and appropriate instruction, behavioral support as well as regular progress monitoring (universal screening). All students are provided additional appropriate instruction or support in the general education setting.

Tier II: Students who do not make progress in Tier I are provided more intensive support through intensive interventions (double dose) still as part of general education.

Tier III: Students who need intensive individualized interventions either in small groups or one-on-one. (Govoni, 2011, p. 235)

The authors propose that the regression-discontinuity design (RD) can be used to analyze RTI assessment data, "The traditional approach in RD is to test for slope differences between the control and treatment groups. However, an alternative approach in RD is presented that tests for slope differences, intercept differences, and permits examination of change in individuals" (Newman & Schumacker, 2012, p. 16).

RD would be used where the eligibility for a program is determined by a person's scoring above or below a certain point on the eligibility criterion. Tier I represents general education where every student, regardless of ability, receives high quality instruction. Students in Tier I who are not keeping up may need extra support (i.e., they would be eligible for Tier II). This would be a subset of students (approximately 20% to 30%) who have received effective instruction but may require more intensive intervention to meet their learning needs. Based on progress monitoring, researchers may find that some students (approximately 5% to 8%) who have had systematic, effective intervention yet do not respond and are, therefore, eligible for Tier III, which may include special education services.

The RD design could compare outcomes for persons in Tier II or III (i.e., those eligible for the program or the treatment group) with those persons who were not eligible (i.e., those in Tier I who do not need special treatment or an intervention). In this context, "discontinuity" refers to a difference in the slope of the regression line and the intercepts of the regression line for the two treatments, and a difference in the regression line for the two groups suggests a program (i.e., intervention) effect. There are tests for slope differences between the control (Tier I) and the treatment groups (Tiers II and III). Tests for intercept differences would indicate whether the before treatment group intercept is different from the after treatment group intercept. One could then compare the mean gain from the first of many repeated measures to the last observation in the after treatment (intervention) regression line controlling for individual differences.

Newman and Schumacker (2012) provided the following generic RD model in Figure 1. The basic RD equation is:

$$Y_{Post} = b_0+b_1Z+b_2X_{Slope}+E \quad (12)$$

where

Y_{Post} = post measures

Z_1 = before treatment (1 = if score comes before treatment; 0, otherwise)

Z_2 = after treatment (1= if score comes after treatment; 0, otherwise)

X_{Slope} = Common Slope

E = Error Vector (residuals)

b's = estimated partial regression weights

Specifically, X_{Slope} is a common slope, and it is used to test to ascertain if the slope in Line 1 is different from the slope in Line 2. To accomplish this, one must set X_{Slope1} to equal X_{Slope2} to determine if a significant difference exists.

Testing Model 11 (a Full Model) against Model 12 (a Restricted Model) is a test of slope differences, independent of intercepts.

Full Model 11: $Y\ (post) = b_0U + b_1(Z_1) + b_2(X_{Slope1}) + b_3(Z_2) + b_4(X_{Slope2}) + E$ (13)

where

U = a unit vector and has a 1 for each subject in the sample

$b_0 – b_{n-1}$ = partial regression weights

Z_1 = 1 if a subject score occurred before intervention; 0, otherwise

Z_2 = 1 if a subject score occurred after intervention; 0, otherwise

X_{SlopeC} = common slope ($X_{Slope1} = X_{Slope2)}$

X_{Slope1} = slope for subjects before intervention

X_{Slope2} = slope for subjects after intervention

E = Error Vector (residuals)

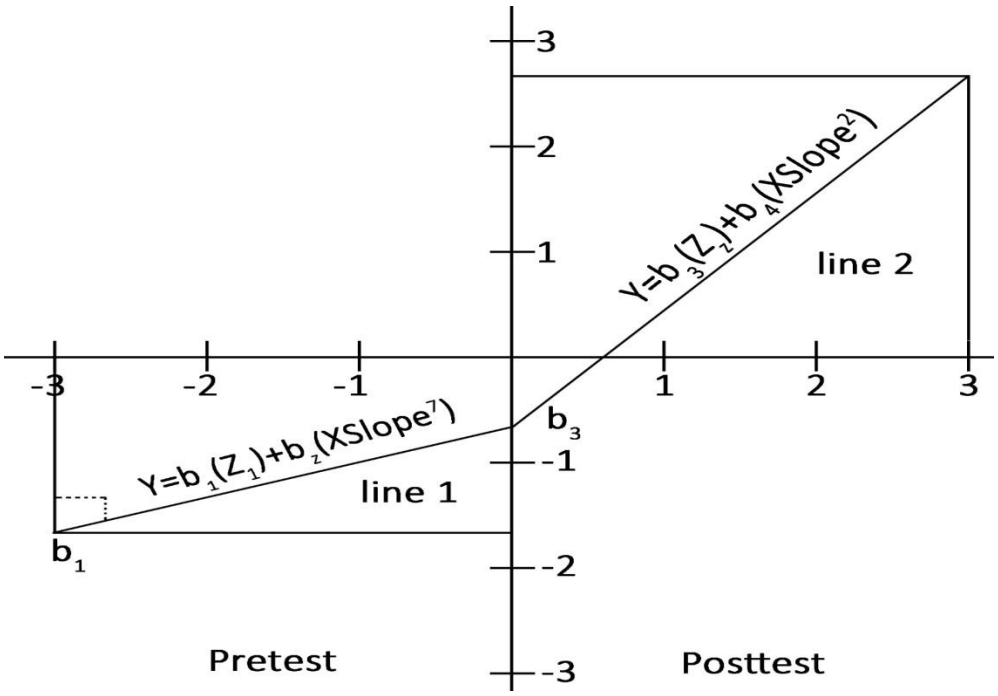

Figure 1. Regression discontinuity model.

In Full Model 11, b_1 is the intercept for line 1; b_2 is the slope for line 1; b_3 is the intercept for line 2; and b_4 is the slope for line 2. In $Y = a_0U + b_2(X_{Slope1}) + b_4(X_{Slope2})$, if there is no significant difference between lines 1 and 2, the R^2 is not statistically significantly different. We can also test to ascertain if there is a difference in intercepts that are independent of the slope differences. If there is no statistically significant difference between the intercepts, the intercepts are not statistically significantly different above and beyond what can be accounted for by slope differences.

Restricted Model 12: $Y = b_0U + b_5(Z_{1b}) + b_6(Z_{1B}) + b_7 + (X_{SlopeC}) + E$ (14)

Testing Model 11 against Model 13 allows one to test for intercept differences, independent of slope differences.

Model 13: $Y = b_0U + b_8(X_{Slope1}) + b_9(X_{Slope2}) + E$ (15)

One can use a regression equation to calculate the predicted score in a test of gain score differences where t_1 is the last time point for before treatment (Z_1) and t_2 is the last time point for after treatment (Z_2). See Newman and Schumacker (2012) for more details on how to do this.

Conclusions

The purpose of this study was to show how multiple regression models can be written to investigate and to better reflect the complexity of four major research themes in first and second language acquisition research in order to reduce Type VI error. The models presented in this paper are examples of what have to be developed in first and second language acquisition research to decrease the likelihood of making a Type VI error. These models were developed to preclude "practices that fail to distinguish between statistical analysis and research design issues, and analyze a research question that involves practical significance with an analytical technique that fails to do so" (Newman et al., 2002, p. 138). Although the models presented in this study are related to first and second language examples, the applicability of these models is multidisciplinary (i.e., they could be utilized in psychology and other behavioral sciences, medicine, public health, social work, and other social sciences).

The four major themes were selected because of their importance to their respective fields and because each of them is multidimensional and has been influenced by interdisciplinary research, thus, making them prime candidates for a study of this nature. Each theoretical perspective examined in this

study brings with it its own research methodology. Thus, the need for a demonstration related to how multiple regression models can be written to investigate four of the major foci of first and second language acquisition research, and to demonstrate how multiple regression models can be written to better reflect the complexity of first and second language acquisition research.

References

Alderson, J. C. (2000). *Assessing reading.* New York: Cambridge University Press.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of Response to Intervention: A snapshot of progress. *Journal of Learning Disabilities, 42,* 85-95.

Bottenberg, R. A., & Ward, J. H. (1963). *Applied multiple linear regression.* Lackland Air Force Base, TX: Aerospace Medical Division, AD 413128.

Bryne, J. (1974). The use of regression equations to demonstrate causality. *Multiple Linear Regression Viewpoints, 5,* 11-22.

Bureau of Exceptional Education and Student Services. (2008). *Florida Department of Education statewide response to instruction/intervention (RtI) implementation plan*. Tallahassee, FL: Clearinghouse Information Center.

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing. Generating text in L1 and L2. *Written Communication, 18,* 80-98.

Cohen, J., Cohen, C., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). New York: Routledge.

Day, E., & Shapson, S. (1991). Integrating formal and functional approaches to language teaching in French immersion: An experimental approach. *Language Learning, 41,* 25-58.

Ellis, N. C. (2001). Memory for language. In P. Robinson (ed.), *Cognition and second language instruction* (pp. 33-68). Cambridge: Cambridge University Press.

Fuchs, L., & Vaughn, S. (2012). Responsiveness-to-Intervention: A decade later. *Journal of Learning Disabilities, 45,* 195-203.

Gass, S. M. (2013). *Second language acquisition: An introductory course* (4th ed.). New York: Routledge.

Govoni, J. M. (Ed.). (2011). *Preparing the way. Teaching ELLs in the K-12 classroom.* Dubuque, IA: Kendall Hunt Publishing Company.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate statistics* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Jennings, E. E. (1967). Fixed effects analysis of variance by regression analysis. *Multivariate Behavioral Research, 2,* 95-108.

Kellerman, E. (1985). If at first you succeed… In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 345-371). Rowley, MA: Newbury House Publishers.

Kerlinger, F. N., & Pedhazur, E. J. (1973). *Multiple regression in behavioral research.* New York: Holt, Rinehart & Winston.

Lightbown, P. M. (1983). Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. Seliger & M. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 217-243). Rowley, MA: Newbury House Publishers.

Lightbown, P. M. (1985). Great expectations: Second language acquisition and classroom teaching. *Applied Linguistics, 6,* 173-89.

Lightbown, P. M., & Spada, N. (1990). Focus on form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition, 12,* 429-48.

Lightbown, P. M., & Spada, N. (1994). An innovative program for primary ESL in Quebec. *TESOL Quarterly, 28,* 563-79.

Lightbown, P. M., & Spada, N. (2006). *How languages are learned* (3rd ed.). Oxford: Oxford University Press.

Lyster, R. (1994). The effect of functional-analytic teaching on aspects of French immersion students' sociolinguistic competence. *Applied Linguistics, 15,* 263-87.

McInerney, M., & Elledge, A. (2013, May). *Implementing ESEA flexibility plans. Using a Response to Intervention framework to improve student learning. A pocket guide for state and district leaders.* Washington, DC: American Institutes for Research.

McNeil, K., & Newman, I. (1996). Categorical or continuous interaction? *Mid-Western Educational Researcher, 6,* 32-34.

McNeil, K., Newman, I., & Fraas, J. W. (2012). *Designing linear models to test research hypotheses.* Lanham, MD: University Press of America, Inc.

McNeil, K., Newman, I., & Kelly, F. J. (1996). *Testing research hypotheses with the general linear model.* Carbondale, IL: Southern Illinois University Press.

National Association of State Directors of Special Education. (2005). *Response to Intervention: Policy considerations and implementation.* Alexandria, VA: Author.

National Center on Response to Intervention. (2010, March). *Essential components of RTI-A closer look at Response to Intervention.* Washington, DC: Author.

Newman, I., Deitchman, R., Burkholder, J., Sanders, R., & Ervin, L. (1976). Type VI error: Inconsistency between the statistical procedure and the research question. *Multiple Linear Regression Viewpoints, 6,* 1-19.

Newman, I., Fraas, J. W., Newman, C., & Brown, R. (2002). Research practices that produce Type VI errors. *Journal of Research in Education, 12,* 138-144.

Newman, I., & Schumacker, R. (2012). Regression-discontinuity designs in medical settings: An alternative to quasi-experimental designs in testing treatment effectiveness. *Multiple Linear Regression Viewpoints, 38,* 16-25.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40,* 184-202.

Paris, S. G., & Hamilton, E. E. (2009). The development of children's reading comprehension. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 32-53). New York: Routledge.

Perkins, K., Brutten, S. R., & Gass, S. (1996). An investigation of patterns of discontinuous learning: Implications for ESL measurement. *Language Testing, 13,* 63-82.

Reutzel, D. R., & Cooter, Jr., R. B. (2011). *Strategies for reading assessment and instruction: Helping every child succeed.* (4th ed.). Boston: Pearson Education, Inc.

Ward, J., & Jennings, E. (1973). *Introduction to linear models.* Englewood Cliffs, NJ: Prentice Hall.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research, 7,* 33-61.

Williams, J. N. (1999). Learner-generated attention to form. *Language Learning, 49,* 583-625.

Send correspondence to:
Kyle Perkins
Florida International University
Email: perkinsk@fiu.edu