Application of Log-Linear Analysis and Logistic Regression Analysis on a Four-way Contingency Table

Seo-eun Choi* Amy Shollenbargar†

Abstract

Log-linear models have been used to analyze association among categorical variables. As the number of variables increases, the complexity of the corresponding model also grows. Two log-linear models for a four-way contingency table are developed and compared using multinomial distribution and Poisson distribution, respectively. Logistic regression is to explain the relationship between explanatory variables and the response variable which is categorical. It considers the logit of the parameter p, the probability of occurrence of one of the levels of the response variable. First, a log-linear model is applied to explain first grade students' awareness of final consonant clusters in monomorphemic words, by analyzing associations among four categorical variables: dialect - African-American English (AAE) or Mainstream American English (MAE), the ways of pronunciation, presentations of stimuli, and tasks of rhyming or segmentation. Once the structure of associations is confirmed, we move to the next step. With justification of possible causation relationship, a logistic regression model is applied to analyze causation relationship from three categorical variables: dialect, presentations of stimuli, and tasks, to a response variable, the ways of pronunciation, which is also categorical.

1. Introduction

Until the late 1960's, contingency tables were typically analyzed by testing the hypothesis of independence between two categorical variables, using chi-square test statistic. When tables have more than two categorical variables, researchers would compute the chi-square test statistics for two-way tables which could be formed as multiple sub-tables. In the 1970's, the analysis of cross-classified data changed with the publication of a series of papers on log-linear models by L. A. Goodman (1973). The basic strategy in log-linear modeling involves fitting models to the observed frequencies in the cross-tabulation of categorical variables. The models can then be represented by a set of expected frequencies that may resemble the observed frequencies. In this paper, two log-linear models for a four-way contingency table were developed using multinomial distribution and Poisson distribution, respectively. The models were applied to explain first grade students' awareness of final consonant clusters in monomorphemic words, by analyzing associations among four categorical variables: dialect - African-American English (AAE) or Mainstream American English (MAE), the ways of pronunciation, presentations of stimuli, and tasks of rhyming or segmentation.

Logistic regression is a form of statistical modelling that is often appropriate for categorical variables (Stokes et al. (2012)). It describes the relationship between a categorical variable and a set of explanatory variables, which can be either categorical or quantitative. The response variable is dichotomous mostly, but it may be polytomous. The response might be ordinal, or nominal. For ordinal response outcomes, the proportional odds model with cumulative logits are used to implement ordered logistic regression. (Stokes et al.

*Department of Mathematics and Statistics, Arkansas State University, PO Box 70, State University, AR 72467

†Department of Communication Disorders, Arkansas State University, PO Box 910, State University, AR 72467

(2012)) For nominal response outcomes, we form generalized logits and implement a logistic analysis by modelling multiple logits per subpopulation of levels. In this paper, we implement the logistic regression to explain the relationship from three categorical variables: dialect - African-American English (AAE) or Mainstream American English (MAE), presentations of stimuli, and tasks of rhyming or segmentation, to a response variable, the ways of pronunciation, which is also categorical.

This paper is arranged as follows. Backgrounds on theories of the log-linear model and logistic regression analysis are summarized in Section 2. In Section 3, a summary of real data is provided. Applications of 2×2 sub-contingency tables are provided in Section 4, and the full model of a four-way contingency table with log-linear analysis is presented in Section 5. Applications of logistic regression analysis is provided in Section 6, and the final conclusions are presented in Sections 7.

2. Background on Log-linear Models and Logistic Regression Analysis

2.1 Log-linear Models

In this subsection, we provide the motivation for log-linear models, present basic notations and definitions. Log-linear analysis is an extension of the two-way contingency table where the conditional relationship between two or more categorical variables is analyzed by taking the natural logarithm of the cell frequencies within a contingency table. Not only to analyze the relationship between two categorical variables, log-linear analysis can be also used to evaluate multiway contingency tables which involve three or more variables.

The simplest model refers to the traditional chi-square test where two variables, each with two levels (2×2 table), are evaluated to see if an association exists between two variables, A and B. Two binomials are arranged in a 2×2 table, and the interest is in examining possible differences between the two binomials. This can be generalized into an $I \times J$ table.

To describe the joint distribution of these two variables, we let π_{ij} denote the probability that an observation falls in row i and column j of the table. These probabilities completely describe the joint distribution of the two variables. We can also consider the *marginal* distribution of each variable. Let $\pi_{i.}$ denote the probability that the row variable takes the value i, and let $\pi_{.j}$ denote the probability that the column variable takes the value j. The main hypothesis of interest with two variables is whether they are *independent*.

We want to consider two different cases which may arise in practice. One is with having the total sample size, n, fixed in advance, and another is with having n random.

First, if the total sample size, n, is assumed fixed, and all other quantities are considered random, the joint distribution follows multinomial distribution. We cross-classify each observation of the sample independently in one of the IJ cells in the table. By definition, two variables are independent if (and only if) their joint distribution is the product of the marginal distributions. Thus, we can write the null hypothesis of independence as

$$H_o \quad : \quad \pi_{ij} = \pi_{i.}\pi_{.j} \tag{1}$$

for all $i = 1, 2, ..., I$ and $j = 1, 2, ..., J$.

Let Y_{ij} denote a random variable representing the number of observations in row i and column j of the table, and let y_{ij} denote its observed value. The joint distribution of the counts is then the multinomial distribution, with

$$P(\boldsymbol{Y} = \boldsymbol{y}) = \binom{n}{y_{11}, ..., y_{ij}, ..., y_{IJ}} \prod_{i,j} \pi_{ij}^{y_{ij}}$$

where $\boldsymbol{Y}$ is a random vector collecting all IJ counts and $\boldsymbol{y}$ is a vector of observed counts.

An alternative model for $I \times J$ table is to treat IJ counts as realizations of independent Poisson random variables. In this model, the total sample size n is not fixed in advance, and all counts are therefore random.

Under the assumption that the observations are independent, the joint distribution of the IJ counts is a product of Poisson distribution,

$$P(\boldsymbol{Y} = \boldsymbol{y}) = \prod_i \prod_j \frac{E_{ij}^{y_{ij}} e^{-E_{ij}}}{y_{ij}!}.$$

Taking natural logarithms we obtain the Poisson log-likelihood.

Testing the hypothesis of independence in the multinomial model is exactly equivalent to testing the goodness of fit of the Poisson additive model.

If A and B are not independent, we have a saturated model as

$$\ln(E_{ij}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}. \tag{2}$$

This model, Eq (2), is using

- E_{ij} =expected cell frequency for the cases from the cell ij in the contingency table,
- μ =overall mean of the natural log of the expected frequencies,
- λ_i^A =main effect of level i from factor A,
- λ_j^B =main effect of level j from factor B, and
- λ_{ij}^{AB} =interaction effect between level i of factor A and level j of factor B.

If there is no significant interaction between A and B, then the interaction term will be dropped so it becomes an additive model, given by

$$\ln(E_{ij}) = \mu + \lambda_i^A + \lambda_j^B.$$

Just as a sample is classified by the levels of two variables, a sample can be classified by the levels of three variables, resulting in an $I \times J \times K$ table. This can be extended to a three-way contingency table in the following way. Let us first introduce some notation. We will use three subscripts to identify the cells in an $I \times J \times K$ table, with i indexing the I rows, j indexing the J columns, and k indexing the K layers.

Let π_{ijk} denote the probability that an observation falls in cell (i, j, k). These probabilities define the joint distribution of the three variables. We also let y_{ijk} denote the observed count in cell (i, j, k), which we treat as a realization of a random variable Y_{ijk} having a multinomial or Poisson distribution. We will also use the dot convention to indicate summing over a subscript, so $\pi_{i..}$ is the marginal probability that an observation falls in row i and $y_{i..}$ is the frequency (or count) in row i. The notation extends to two dimensions, so $\pi_{ij.}$ is the marginal probability that an observation falls in row i and column j, and $y_{ij.}$ is the corresponding count.

In the similar way, we can extend this to a four-way contingency table, which has four categorical variables, A, B, C and D.

We limit our models to a hierarchical set in which higher-order terms may be included only if the related lower-order terms are included. As an example, a three-way interaction λ_{ijk}^{ABC} cannot be included in a model unless all two-way interaction terms λ_{ij}^{AB}, λ_{jk}^{BC}, and λ_{ik}^{AC} are in the model. However it does not require λ_{ijk}^{ABC} to be in the model because two-way interaction terms are in the model (Fienberg (2007)).

2.2 Logistic Regression Analysis

In this subsection, we provide the motivation for logistic regression analysis, present basic notations and definitions, and describe how parameters are estimated. When the response variable is dichotomous, we consider the success probability p, which is the probability of occurrence of one of two levels from the response variable. Instead of modeling the probability p directly with a linear model, we first consider the *odds*, or *odds ratio*

$$odds = \frac{p}{1-p} \tag{3}$$

which is the ratio of the probability of success p to the probability of failure $1-p$.

In logistic regression for a binary response variable, we model the natural log of the odds ratio, which is called *logit(p)*. Thus

$$\begin{aligned} logit(p) &= \ln(odds) \\ &= \ln\left(\frac{p}{1-p}\right). \end{aligned} \tag{4}$$

The logit is a function of the probability p. In the simplest model, we assume that the logit graphs as a straight line in the predictor variable X so

$$\begin{aligned} logit(p) &= \ln(odds) \\ &= \ln\left(\frac{p}{1-p}\right) \\ &= \beta_o + \beta_1 x + \epsilon. \end{aligned} \tag{5}$$

In other words, the log odds are linear in the predictor variable, X.

It is also possible to think in terms of probabilities. We can convert from the logit or log odds to the probability p. By first exponentiating

$$\ln\left(\frac{p}{1-p}\right) = \beta_o + \beta_1 x + \epsilon,$$

we obtain

$$\frac{p}{1-p} = \exp(\beta_o + \beta_1 x).$$

Next solving for p, we obtain

$$p = \frac{\exp(\beta_o + \beta_1 x)}{1 + \exp(\beta_o + \beta_1 x)}, \tag{6}$$

which describes a *logistic curve*. The relation between p and the predictor X is not linear but has an S-shaped graph. If the response variable is polytomous with r levels, we use the generalized logit, which is defined as

$$\mathrm{logit_i} = \ln\left(\frac{p_i}{p_r}\right) \tag{7}$$

for $i = 1, 2, ..., (r-1)$. A logit is formed for the probability of each succeeding category over the last response category.

3. Application to Real Data

In this section, we want to apply the log-linear model to a four-way contingency table. This contains four categorical variables: dialect - African-American English (AAE) or Mainstream American English (MAE), the ways of pronunciation, presentations of stimuli, and tasks of rhyming or segmentation (Shollenbarger (2014)). We want to explain first grade students' awareness of final consonant clusters in monomorphemic words, by analyzing associations among these variables.

In the United States, children speak many language varieties (i.e., dialects) when they begin school. Dialects are linguistic varieties of a language or similar speech patterns that people speak (Wolfram & Schillings-Estes (2006)), and they serve as functional and effective varieties of languages that reflect the social and cultural background of speakers. The American Speech-Language-Hearing Association (ASHA) posits that dialectal varieties of American English are not disordered speech or language (ASHA (2003)), yet some are valued more socially than others. Mainstream American English (MAE), which is a term used to describe the broad category of language varieties valued in the media, government, and education system in the United States, is one dialect group that will be referred to in this study. African-American English (AAE) is a dialect containing speech patterns that have developed from the unique history of modern day African Americans (Rickford & Rickford (2000)), and it is also of interest in this study.

AAE is a rule-governed dialect spoken primarily, but not exclusively, by African American children in the United States. AAE can be characterized by at least 40 different phonological (i.e., sounds) and morphosyntatic (i.e., grammatical) features that differ systematically from other varieties of English (Connor & Craig (2006), Craig et al. (2009), Pollock et al. (1998)), and may be a factor in the achievement gap between White and Black students (U.S. Department of Education (2012)).

We are looking at the relationship among African American English dialect on rhyming and segmenting of words that end in final clusters. A phonological feature of African American English is to reduce final clusters in words like ”nest”, which an AAE speaker would pronounce ”nes”, ”soft”, which may be pronounced ”sof”, etc. Some researchers are studying these dialect features and their relationship to or interaction with early literacy skills (Connor & Craig (2006), Craig et al. (2009), Shollenbarger (2014)).

In the past, there have been various attempts to analyze and understand these relationships and structures. Hierarchical linear modeling (HLM) examined the effect of dialect on emergent reading skills (Connor & Craig (2006)). MANOVA was implemented to examine influence of SES, gender, and community, and a paired t-tests compared phonological, morphological, and combination of AAE features (Thompson et al. (2004)). One-way ANOVA was used to compare dialect density by grade, and also to compare phonological feature use by grade (Craig et al. (2003)). Later, Craig et al used piecewise latent growth cross-classified random effects models with repeated measures over time, with considering that dialect changes over Grade 1-2 with children nested in Grade 1 or 2 classrooms (2009). Terry et al analyzed correlations among variables (Terry et al. (2012)). Most recently, ANOVA was used to compare difference between AAE and MAE groups for phonological scores, grammatical scores, and naming accuracy (Terry (2014)). Data on interval or ratio level of measurements with ”correct/incorrect” scoring have been used. We are focusing on one phonological feature of AAE - final consonant cluster reduction - and we want to analyze the association of this phonological feature with other variables, dialect, task and stimulus presentation. These four variables are on nominal scales and can be analyzed using log-linear models.

Three elementary schools in Northeast Arkansas, were chosen and first graders were recruited from the schools. Over 50 children were involved in this study, and they are cross-classified by the dialect groups (A), stimulus presentation (C), tasks (D) and the type of pronunciations of their responses (B). Their dialects were classified into two, Mainstream American English speakers (MAE) and African American English speakers (AAE). They were given three different stimuli for each task of rhyme identification and segmentation. In one stimulus, the child named pictures without a model from the examiner. In a second stimulus, the pictures were named by the examiner, and the child completed the word. In the third stimulus, nonsense words were named by the examiner, and the child completed the rhyming or segmenting task.

Let us define variable notations and their levels as following.

- variable A (dialect group): 0 (AAE), 1 (MAE),
- variable B (pronunciation): 0 (AAE), 1 (MAE),
- variable C (stimulus presentation): 0 (child names), 1 (examiner names), 2 (nonsense words), and
- variable D (task): 0 (rhyme ID), 1 (segmentation).

The frequency of incorrect responses across all tasks (i.e, not a rhyme, not a segment,) compared to the MAE or AAE responses were negligible and therefore taken out of the final log-linear analysis.

We want to test independence among A, B, C, and D. However it is not meaningful to test independence between A and C, C and D, and A and D for the context. We start with full hierarchical (saturated model) as in Eq (8) and test statistical significance of independence. With an error term of ϵ, the model is given as

$$\begin{aligned} \ln(E_{ijkl}) &= \mu + \lambda^A + \lambda^B + \lambda^C + \lambda^D \\ &\quad + \lambda^{AB} + \lambda^{BC} + \lambda^{BD} \\ &\quad + \lambda^{ACB} + \lambda^{ABD} + \lambda^{BCD} \\ &\quad + \lambda^{ABCD} + \epsilon. \end{aligned} \tag{8}$$

Because the total sample size is fixed in advance, we use a multinomial model, where we focus on the joint distribution of the four variables, A, B, C, and D. As previously indicated, the purpose of this investigation was to examine the awareness of final consonant clusters in first grade children who speak MAE and AAE through rhyming and segmenting tasks with monomorphemic words. This research was conducted with the approval of the Institutional Review Board of the University of Arkansas at Little Rock.

4. Application: Log-linear Models on Sub-contingency Tables with A and B Only

Do first grade participants who speak AAE differ from their MAE speaking peers in rhyming and segmenting words with contrasting features? We want to test if dialect and pronunciation are independent or dependent at each combination of tasks and stimuli. With the raw data, children who speak African American English have more of an influence of their dialect when they say a word or words and complete the task than when a MAE speaking examiner says the word or words and they complete the tasks. In other words, we test independence between A and B using six 2×3 tables, respectively. Corresponding log-linear model is,

$$\ln(E_{ij}) = \mu + \lambda^A + \lambda^B + \lambda^{AB} + \epsilon.$$

	B1	B2	row total
A1	73	53	126
A2	49	81	130
column total	122	134	256

Table 1: C1D1, child names-rhyme

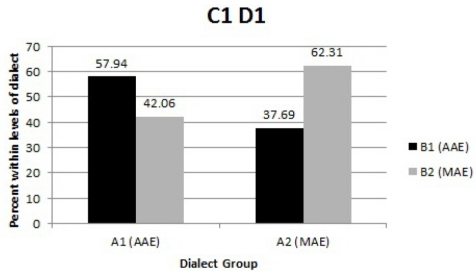

Figure 1: Bar graph of sample proportions from Table 1

	B1	B2	row total
A1	88	36	124
A2	64	69	133
column total	152	105	257

Table 2: C2D1, examiner names-rhyme

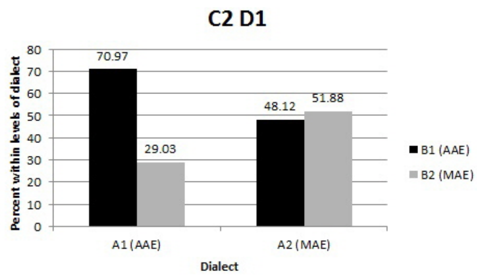

Figure 2: Bar graph of sample proportions from Table 2

The frequency distribution of responses to the combinations of stimulus presentations with tasks are shown in Tables 1 through 6. The raw frequency data was used in the log-linear analysis. The raw relative frequencies are presented in Figures 1 through 6 in each category.

At all six combinations of stimulus presentations and task, dialect and pronunciations were significantly associated, because all p-values for testing homogeneity of proportions were lower than < 0.001. As a conclusion, dialect and pronunciation are dependent at all possible combinations of stimulus presentation and task. The next question is, if this implies or indicates that two-way interactions which are related with B are significant. So we want to test the model

$$\begin{aligned}\ln(E_{ijk}) &= \mu + \lambda^A + \lambda^B + \lambda^C + \lambda^D \\ &\quad + \lambda^{AB} + \lambda^{BC} + \lambda^{BD}.\end{aligned}$$

For a hierarchical model, all terms are significant, and this indicates that A and B are dependent, and so are B and C, and B and D overall.

	B1	B2	row total
A1	59	65	124
A2	31	93	124
column total	90	158	248

Table 3: C3D1, nonsense words-rhyme

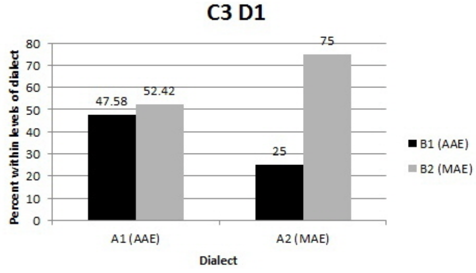

Figure 3: Bar graph of sample proportions from Table 3

	B1	B2	row total
A1	39	82	121
A2	2	127	129
column total	41	209	250

Table 4: C1D2, child names - segmentation

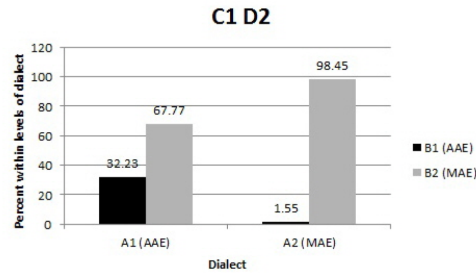

Figure 4: Bar graph of sample proportions from Table 4

	B1	B2	row total
A1	31	94	125
A2	3	130	133
column total	34	224	258

Table 5: C2D2, examiner names-segmentation

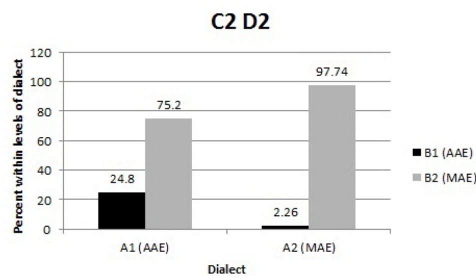

Figure 5: Bar graph of sample proportions from Table 5

	B1	B2	row total
A1	18	110	128
A2	0	132	132
column total	18	242	260

Table 6: C3D2, nonsense words-segmentation

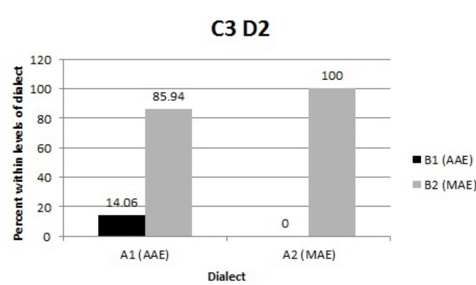

Figure 6: Bar graph of sample proportions from Table 6

With having all four variables, now we want to ask the question if they are independent to each other. This generalizes a two-way model into a four-way model.

5. Application: Log-linear Models on a Four-way Contingency Table

We want to test independence among A, B, C, and D. However we are not interested in independence between A and C, C and D, and A and D. Hence, the full model is

$$\begin{aligned}\ln(E_{ijkl}) &= \mu + \lambda^A + \lambda^B + \lambda^C + \lambda^D \\ &+ \lambda^{AB} + \lambda^{AC} + \lambda^{AD} \\ &+ \lambda^{BC} + \lambda^{BD} + \lambda^{CD} \\ &+ \lambda^{ABC} + \lambda^{ABD} + \lambda^{BCD} \\ &+ \lambda^{ABCD} \\ &+ \epsilon.\end{aligned}$$

The test result is provided in Table 7.

Source	DF	Chi-Square	Pr > ChiSq
A	1	8.31	0.0039
B	1	58.96	<.0001
A*B	1	27.80	<.0001
C	2	0.23	0.8928
A*C	2	1.01	0.6031
B*C	2	6.82	0.0331
A*B*C	2	0.48	0.7871
D	1	15.10	0.0001
A*D	1	9.22	0.0024
B*D	1	52.32	<.0001
A*B*D	1	6.31	0.0120
C*D	2	0.21	0.9007
A*C*D	2	0.94	0.6260
B*C*D	2	3.22	0.1997
A*B*C*D	1	0.75	0.3859

Table 7: Full model with all variables

The highest order interaction $A * B * C * D$ has its p-value of 0.3859, so it is not significant. This leads us into the three-way log-linear model.

We start with all three-way interactions from four variables.

As presented in Table 8, two of the three-way interactions, ABC and BCD, are not significant with p-values of 0.8676 and 0.0873, respectively. We drop the one with highest p-value first, so ABC is dropped and the resulting model is tested. We keep exploring the models with three-way interactions until all terms are significant. As a result, the final model is given in Eq (9).

$$\begin{aligned}\ln(E_{ijkl}) &= \mu + \lambda^A + \lambda^B + \lambda^D \\ &+ \lambda^{AB} + \lambda^{AD} + \lambda^{BD} + \lambda^{BC} \\ &+ \lambda^{ABD} \\ &+ \epsilon. \end{aligned} \tag{9}$$

Source	DF	Chi-Square	Pr > ChiSq
A	1	22.66	<.0001
B	1	121.52	<.0001
A*B	1	61.74	<.0001
C	2	5.28	0.0715
A*C	2	2.93	0.2309
B*C	2	22.54	<.0001
A*B*C	2	0.28	0.8676
D	1	35.96	<.0001
A*D	1	23.96	<.0001
B*D	1	107.48	<.0001
A*B*D	1	18.19	<.0001
C*D	2	2.64	0.2675
B*C*D	2	4.88	0.0873

Table 8: Reduced model with three-way interactions

We started with two-way log-linear models using A and B at six combinations of C and D, and found that A and B are dependent at each combination of C and D. Then the four-way log-linear model was fitted using A, B, C, and D all together, and found that A, B, and D are dependent to each other. It is also found that B and C are dependent to each other.

This indicates that dialect, the way of pronunciation, and tasks are associated, and the way of pronunciation is associated with the stimulus presentation. However it does not necessarily mean that all four variables are associated (Figure 7). It is important to consider the relationship between dialect and performance on literacy tasks when teaching children how to read.

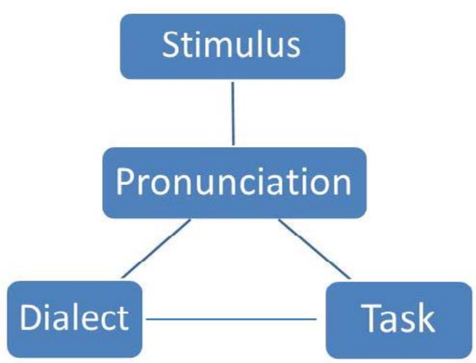

Figure 7: Association of four variables

6. Application: Logistic Regression Analysis

Because of the fact that dialect (variable A) is an antecedent to pronunciation (variable B), we have chosen to treat *dialect* as an explanatory variable for the response variable, *pronunciation*. It was confirmed that there exists association among A, B, C, and D using log-linear analysis.

Thus it is meaningful to investigate further for causation relationship among them. We define the success probability p as the probability if the pronunciation of a child belongs to MAE group. The corresponding full model is

$$\begin{aligned} logit(p) &= \ln\left(\frac{p}{1-p}\right) \\ &= \beta_o + \beta_A A + \beta_C C + \beta_D D \\ &\quad + \beta_{AC} AC + \beta_{AD} AD + \beta_{CD} CD \\ &\quad + \beta_{ACD} ACD + \epsilon. \end{aligned} \tag{10}$$

The result from the full model is summarized in Table 9.

Effect	DF	Wald Chi-Square	p-values
A	1	0.0049	0.9445
C	2	0.1755	0.9160
A*C	2	0.6053	0.7389
D	1	0.0054	0.9416
A*D	1	0.0042	0.9485
C*D	2	0.0052	0.9974
A*C*D	2	0.7548	0.6856

Table 9: Summary results from the full model

The three-way interaction has p-value of 0.6856 (Table 9), which is higher than $\alpha = 0.05$, so we drop the three-way interaction and fit the model with two-way interactions only. Eventually, we have the last model with the interaction between A and D only. (Table tab:reducedmodel4)

$$logit(p) = \ln\left(\frac{p}{1-p}\right) = \beta_o + \beta_A A + \beta_C C + \beta_D D + \beta_{AD} AD + \epsilon$$

and their estimates are

$$\hat{logit}(p) = 4.9565 - 3.2092A - 0.6746C_1 - 0.9251C_2 - 3.8674D + 2.2863AD \tag{11}$$

Variables A and D have significant interaction, while C does not. Thus, the effect of C can be interpreted by itself.

7. Conclusion

Despite the fact that dialect (variable A) is an antecedent to pronunciation (variable B), we have chosen for the time being not to treat *dialect* as an explanatory variable for the

Effect	DF	Wald Chi-Square	p-values
A	1	47.2020	<.0001
C	2	34.8163	<.0001
D	1	69.7370	<.0001
A*D	1	21.6896	<.0001

Table 10: Summary results from the reduced model with interaction between A and D only

Parameter	level	DF	Estimate	p-values	exp(Est)
Intercept		1	4.9565	$<.0001$	142.093
A	1	1	-3.2092	$<.0001$	0.040
C	1	1	-0.6746	$<.0001$	0.509
C	2	1	-0.9251	$<.0001$	0.397
D	1	1	-3.8674	$<.0001$	0.021
A*D	1 1	1	2.2863	$<.0001$	9.838

Table 11: Estimates from the reduced model with interaction between A and D only

response variable, *pronunciation*. We started with two-way log-linear models using A and B at six combinations of C and D, and found that A and B are dependent at each combination of C and D. Then we fitted the four-way log-linear model using A, B, C, and D all together, and found that A, B, and D are dependent to each other. It is also found that B and C are dependent to each other. This indicates that dialect, the way of pronunciation, and tasks are associated, and the way of pronunciation is associated with the stimulus presentation. However it does not necessarily mean that all four variables are associated (Figure 7). It is important to consider the relationship between dialect and performance on literacy tasks when teaching children how to read.

Using logistic regression analysis, we find that all of A, C, and D have effects on the response variable, B, *pronunciation*. With using either the model with only one interaction between A and D, or the additive model, we can conclude that children with AAE dialect has a significantly lower chance to produce in MAE, while having rhyme ID task results in a significantly lower chance to have pronunciation in MAE. For stimuli presentation (variable C), children have the highest chance to pronounce in MAE if they are given nonsense words.

References

ASHA (2003), American English dialects, Technical report.
URL: *Retrieved from www.asha.org/policy*

Connor, C. & Craig, H. (2006), 'African American preschoolers' language, emergent literacy skills, and use of African American English: A complex relation', *Journal of Speech, Language, and Hearing Research* **49**, 771–792.

Craig, H., Thompson, C., Washington, J. & Potter, S. (2003), 'Phonological features of child African American English', *Journal of Speech, Language, and Hearing Research* **46**, 623–635.

Craig, H., Zhang, L., Hensel, S. & Quinn, E. (2009), 'African American English- speaking students: An examination of the relationship between dialect shifting and reading outcomes', *Journal of Speech, Language, and Hearing Research* **52**, 839–855.

Fienberg, S. (2007), *The Analysis of Cross-classified Categorical Data*, 2 edn, Springer.

Goodman, L. (1973), 'The analysis of multidimensional contingency tables when some variables are posterior to others: A modified path analysis approach', *Biometrika* **60(1)**, 179–192.

Pollock, K. E., Bailey, G., Berni, M., F. D., Hinton, L., J. I., Roberts, J. & Weaver, R. (1998), 'Phonological features of African American English'. Retrieved February 3, 2013, from http://www.rehabmed.ualberta.ca/spa/phonology/Features.htm.

Rickford, J. R. & Rickford, R. J. (2000), *Spoken Soul*, John Wiley & Sons.

Shollenbarger, A. (2014), First Grade Students' Awareness of Final Consonant Clusters in Monomorphemic Words: Does dialect matter?, PhD thesis, Order No. 3621713, University of Arkansas for Medical Sciences, ProQuest Dissertations and Theses, 157.
URL: *http://ezproxy.library.astate.edu/login?url=http://search.proquest.com/docview/1545713810?accountid=8363*

Stokes, M. E., Davis, C. S. & Koch, G. G. (2012), *Categorical Data Analysis Using SAS*, 3 edn, SAS Press.

Terry, N. P. (2014), 'Dialect variation and phonological knowledge. Phonological representations and metalinguistic awareness among beginning readers who speak nonmainstream American English', *Journal of Applied Psycholinguistics* **35(1)**.

Terry, N. P., Connor, C. M., Petscher, Y. & Conlin, C. R. (2012), 'Dialect variation and reading: Is change in nonmainstream American English use related to reading achievement in first and second grades?', *Journal of Speech and Hearing Research* **55**, 55–69.

Thompson, C. A., Craig, H. K. & Washington, J. A. (2004), 'Variable production of african american english across oracy and literacy contexts', *Language, Speech, and Hearing Services in Schools* **35**, 269–282.

U.S. Department of Education (2012), 'Achievement gaps: How black and white students in public schools perform in mathematics and reading on the national assessment of educational progress'.
URL: *http://nces.ed.gov/nationsreportcard/studies/gaps/*

Wolfram, W. & Schillings-Estes, N. (2006), *American English*, 2 edn, Blackwell, Malden, MA.