Latent Profile Analysis: Understanding Dialect Change and Early Reading Comprehension in African American Children

Rachel Yan
Dept. of Psychology and Dept. of Statistics and Data Science, Smith College, MA, 01060

Introduction: Latent Profile Analysis (LPA) is a person-centered statistical approach that identifies homogenous subgroups within a heterogeneous sample population. It is useful for detecting individual and subgroup patterns of growth or change. Theories and research on children's language development have been dominated by variable-centered analysis methods such as multiple regression and SEM, and few studies have adopted growth mixture techniques such as LPA. This paper provides an overview of LPA and its application to longitudinal data of language and reading development in 200 African American (AA) children in low-income communities (School Readiness Research Consortium, NICHD 2005-12.

BACKGROUND

Data:

- Data came from a curriculum intervention program from The School Readiness Research Consortium.
- The program looked at intervention effects on the academic outcomes of children attending childcare classrooms in low-income areas (N = 200)

Measures:

- During the intervention program, the DELV-ST[1] Score on children's African American English (AAE), and Mainstream American English (MAE) were taken across 4 times (W1-W3-W4-W5).
- The sample size is N=200, except for W4 where n=179.
- Results from previous literatures showed that MAE speakers scored significantly higher on literacy test (Reading Comprehension) than the AAE speakers.
- We investigate this phenomenon by looking at children's status on dialect shift across times.

HYPOTHESIS & MODEL SPECIFICATION

Based on the initial exploration of children's AAE and MAE trends, we suspect that children fall into four distinctive language profiles regarding their dialect shifting status:

- Group 1: High AAE (W1) → High AAE (W5)
- Group 2: High AAE (W1)→ Low AAE (W5) [Dialect Switcher]
- Group 3: Low AAE (W1)→ High AAE (W5)
- Group 4: Low AAE (W1)→ Low AAE (W5)

We used AAE and MAE measures at W1 and W5 to specify a LPA model because there was no missing data. Our Model frame works followed the guidelines stated Berlin et al. (2013).

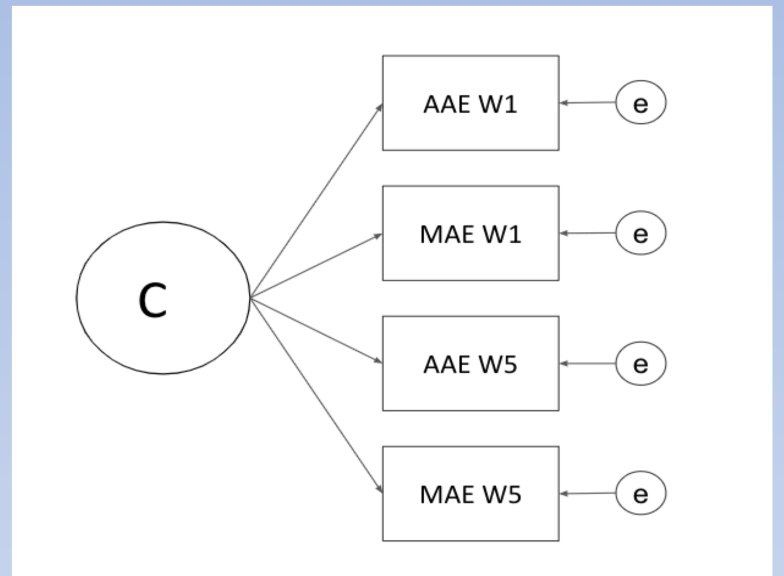

Figure 1. Graphic Representation of LPA Model

MODEL ESTIMATION & SELECTION

- We separately estimated a 2-class, a 3-class, a 4-class and a 5-class model.
- We followed example 7.9-7.10 on Mplus User Guides to conduct our analyses.
- We selected our model based on fit indices in Table. 1 (BIC, ABIC, LMR LRT, Bootstrap LRT and average posterior probabilities) and theoretical background.

Model Fit Indices

LPA Models	BIC (Bayesian Information Criterion)	ABIC (Adjusted BIC)	LMR LRT (Lo-Mendell-Rubin))	Bootstrap LRT
2-class	4007.444	3966.259	208.394 (p<.001: 2-class vs. 1-class)	<.001
3-class	3882.905	3825.879	145.537 (p<.001: 3-class vs. 2-class)	<.001
4-class	3826.191	3753.325	80.179(p=0.0425)	<.001
5-class	3796.011	3707.304	84.721(p=0.0252)	<.001

Sample distribution

Class	1	2	3	4	5
2-class	94(.47)	106(.53)			
3-class	80(.4)	85(.425)	35(.175)		
4-class	31(.155)	56(.28)	60(.3)	53(.265)	
5-class	53(.265)	30(.15)	51(.255)	4(.02)	62(.31)

Table 1. Model Fit Indices and Sample Distribution 2-, 3-, 4-, 5- profile Model.

Mean associated with the 4-profile model

	AAE1	AAE5	MAE1	MAE5
Class 1 (n=31)	9.290323	2.516129	3.451613	11.70968
Class 2 (n=56)	9.285714	9.75	1.767857	3.857143
Class 3(n=60)	10.3	6.616667	2.266667	7.45
Class 4 (n=53)	9.113208	13.49057	1.528302	0.7358491

Average posterior probabilities associated with each profile (profile =4)

Class	1	2	3	4
1	**0.97**	0	0.04	0
2	0	**0.936**	0.028	0.036
3	0.012	0.074	**0.914**	0
4	0	0.04	0	**0.96**

Note: posterior probabilities are the probability that an individual belongs to the assigned profile and to no other profiles.

Table 2. Mean and average posterior probability associated with the 4-profile

MODEL ESTIMATION & SELECTION

We decided to go with a 4-class model because it had:

- lower BIC and ABIC compared to the previous two model
- A significant LMR LRT suggesting that the model indeed explained more variance than a 3-class model.
- An even sample distribution for each profiles.
- A high average posterior probability (>0.90) for each profile

DISCUSSION

- LPA revealed four coherent and distinctive groups of children in the sample based on the trajectory of their dialect change. A one-way ANOVA was also carried out to probe for group differences on children's reading comprehension score. Results showed that children who shifted from high African American English (AAE) at preschool to low AAE and high Mainstream American English (MAE) production on the DELV-ST at first grade had significantly better reading outcomes than children who were high AAE at both preschool and first grade.

- Combined with ANOVAs and regression, LPA adds insight into understanding the reading achievement gap and bi-dialectal development in AA and White children. The current study demonstrates that LPA is an effective technique to determine classes of individuals who share similar developmental trajectories and the extent to which these patterns may relate to other variables of interest.

REFERENCES

Berlin, K. S., Williams, N. A., & Parra, G. R. (2014). An Introduction to Latent Variable Mixture Modeling (Part 1): Overview and Cross-Sectional Latent Class and Latent Profile Analyses. *Journal of Pediatric Psychology*, *39*(2), 174–187. https://doi.org/10.1093/jpepsy/jst084

Ram, N., & Grimm, K. J. (2009). Methods and Measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development*, *33*(6), 565–576. https://doi.org/10.1177/0165025409343765

Asparouhov, T., & Muthen, B. (n.d.). *Using Mplus TECH11 and TECH14 to test the number of latent classes*. 17.