

과목명: AD 빅데이터기초분석과 정보활용

담당교수: 박현숙

과제명: 기말평가

e-mail: hungry324@naver.com

학과: 데이터과학융합스쿨

전공: 디지털금융정보전공

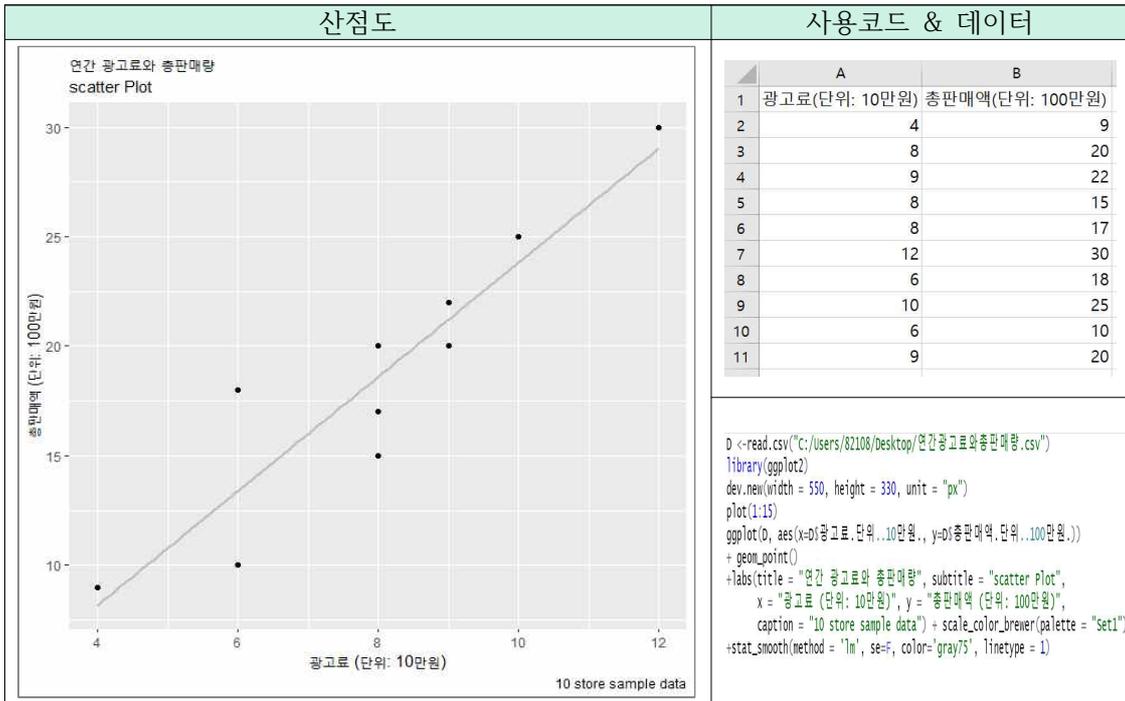
학번: 20183254

이름: 최윤수

I. 어떤 특수한 종류의 상품을 팔고 있는 상점을 중심으로 광고가 판매량에 미치는 관계를 알아보기 위하여 비슷한 여건하에 있는 많은 상점 중에서 10개의 상점을 표본으로 추출하여, 이 상점들의 연간 광고료와 총판매량을 알아보니 그 자료가 다음과 같다.

광고료(단위: 10만원)	4	8	9	8	8	12	6	10	6	9
총판매액(단위: 100만원)	9	20	22	15	17	30	18	25	10	20

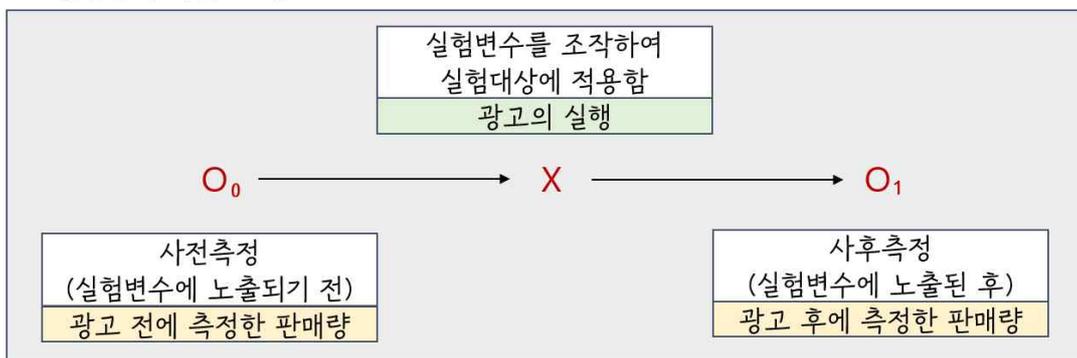
1) 산점도를 그려라.



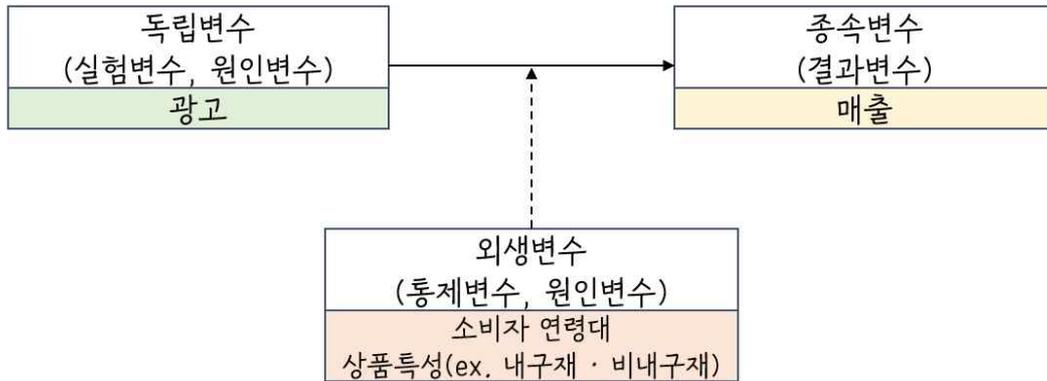
2) 광고가 판매량에 미치는 관계를 알아보기 위해 타당한 모형을 설정하라.

□ 광고가 판매량에 미치는 관계 (=영향)

- 실험설계 기본 모형



- 실험설계 기본 모형 (2)

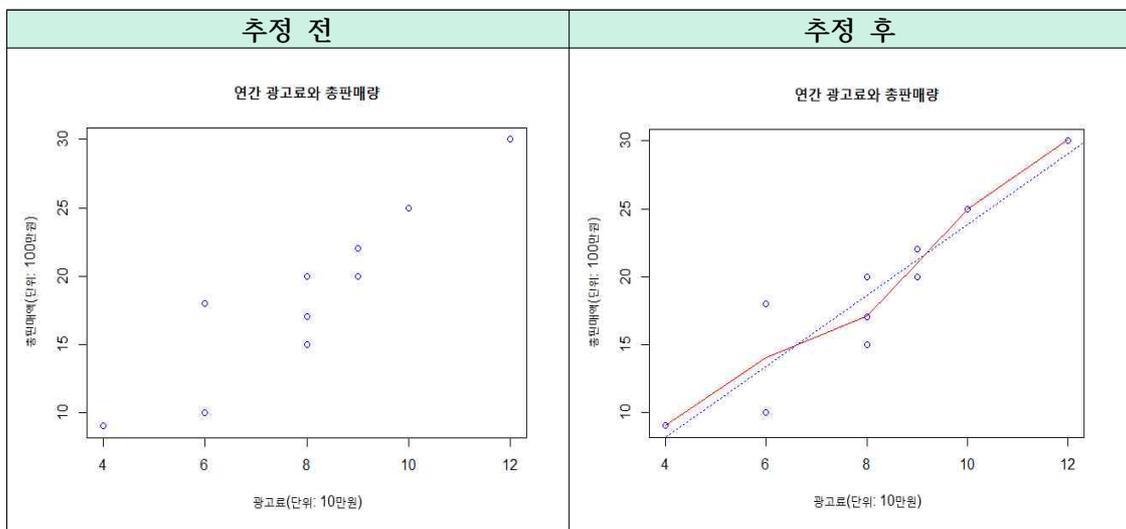


이미지 제작: 20183254 최윤수

3) 위 자료에 적합을 위해 모형이 갖는 가정을 전제하여야.

→ 가설: 광고는 제품 판매량 증대에 긍정적인 영향을 미친다.

4) 최소제곱법에 의한 대표할 LOWESS를 추정하고, 1)번 그림에 그려 넣어라.



파란 점선은 기존의 linear regression라인이고 빨간선인 lowess에 의한 회귀곡선이다

```
D <-read.csv("C:/Users/82108/Desktop/연간광고료와총판매량.csv")
plot(D)
lines(lowess(D))
library(MASS)
out=lm(D$총판매액.단위..100만원.~D$광고료.단위..10만원.,data=D)
s1=lowess(D$총판매액.단위..100만원.~D$광고료.단위..10만원.)
plot(D$총판매액.단위..100만원.~D$광고료.단위..10만원.,data=D,col="blue",
      xlab="광고료(단위: 10만원)",ylab="총판매액(단위: 100만원)",
      main="연간 광고료와 총판매량")
lines(s1,col="red")
abline(out,lty="dotted",col="blue")
```

5) 상관계수는? 광고와 총판매량은 선형적 관계가 있다고 말할 수 있는가?

```
> summary(lm(D$총판매액.단위..100만원.~D$광고료.단위..10만원., D))

Call:
lm(formula = D$총판매액.단위..100만원. ~ D$광고료.단위..10만원.,
    data = D)

Residuals:
    Min       1Q   Median       3Q      Max
-3.600 -1.502  0.813  1.128  4.617

Coefficients:
            Estimate Std. Error t value
(Intercept)   -2.2696     3.2123  -0.707
D$광고료.단위..10만원.  2.6087     0.3878   6.726

Pr(>|t|)
(Intercept)    0.499926
D$광고료.단위..10만원. 0.000149 ***

---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.631 on 8 degrees of freedom
Multiple R-squared: 0.8497, Adjusted R-squared: 0.831
F-statistic: 45.24 on 1 and 8 DF, p-value: 0.0001487

> cor.test(D$총판매액.단위..100만원., D$광고료.단위..10만원., method="pearson")

Pearson's product-moment correlation

data: D$총판매액.단위..100만원. and D$광고료.단위..10만원.
t = 6.7261, df = 8, p-value = 0.0001487
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6963398 0.9816763
sample estimates:
      cor
0.9218123
```

위에 나온 0.9218123 상관계수 값을 통해서 광고와 판매량은 양적 선형관계를 이루고 있다고 볼 수 있다

II. 다음의 비유창성 데이터를 분석하시오.

◎ 목적 : 유창성장애라는 것은 말의 전진적 진행이 운동신경적인 잘못으로 인해 말소리, 음절, 또는 낱말의 산출이 방해 받는 것이라 정의한다. 일반적으로 말의 흐름, 즉 산출이 부드럽지 않고 순조롭지 않은 현상을 가지고 있는 장애를 유창성장애라고 하는데, 이런 유창성장애를 가진 사람들은 발화시 정상인과는 다른 비유창성 유형 및 빈도를 보이게 된다. 소리와 음절의 산출과 이해에 관한 연구결과에 따르면 화자가 산출하는 비유창성 유형과 비유창성의 빈도가 그 사람이 말더듬인지 아닌지를 판단할 수 있는 지표가 된다고 한다(Zebrowski, 1994). Zebrowski(1994)는 화자의 비유창성 빈도가 높을수록 청자는 화자의 비유창성 정도가 심각하다고 판단하게 된다고 한다. 또한 유창성장애를 가지고 있는 사람 정상인의 말속도 측정은 많은 논란이 있어왔으나 현재 임상에서는 진단시 말속도의 측정을 포함하도록 하고있다(Guitar, 1998; Zebrowski, 1994). 이는 말속도가 말더듬의 정도와 관계가 있으며 의사소통에 영향을 준다고 보고있기 때문이다(신문자, 2000).

◎ 연구문제 :

가. 유창성장애 아동과 정상아동에게서 나타나는 비유창성의 유형 및 빈도를 각각 분석하고 아래의 조건에서 비교하시오.

- (1) 놀이 상황에서, 유창성장애 아동과 정상아동의 비유창성 유형 및 빈도는 유의한 차이를 보이는가?
- (2) 과제제시 상황에서, 장애 아동과 정상아동의 비유창성 유형 및 빈도는 유의한 차이를 보이는가?
- (3) 유창성장애 아동은 놀이상황과 과제제시 상황에서 나타나는 비유창성 유형 및 빈도에 유의한 차이를 보이는가?

나. 유창성장애 아동과 정상아동의 전체말속도(overall speech rate)를 측정하고 비교하시오.

(1)-1. 놀이 상황에서, 유창성장애 아동과 정상아동의 전체말속도(overall speech rate)는 유의한 차이를 보이는가?

(1)-2. 과제제시 상황에서, 유창성장애 아동과 정상아동의 전체말속도(overall speech rate)는 유의한 차이를 보이는가?

(1)-3 유창성장애 아동은 놀이상황과 과제제시 상황에서 나타나는 전체말속도(overall speech rate)에 유의한 차이를 보이는가?

◎ 비교집단에 따른 연구방법 : 서울 및 경기지역에 거주하는 만 4세, 5세 아동을 대상으로 실시하였다. 연구에 참여한 아동은 총 20명으로 유창성장애 아동 10명, 유창성장애 아동과 생활연령을 일치시킨 정상아동 10명이다. 유창성장애 아동의 평균연령은 58.7개월(연령범위 48개월-70개월)이었으며, 정상아동의 평균연령은 58.9개월(연령범위 50개월-70개월)이었다.

비유창성 데이터를 활용하여, 데이터 요약(탐색), 데이터분석(평균값차이, 분산분석 등)을 하여, 서론(문제제기), 본론(제기된 문제와 데이터 설명, 그리고 데이터 탐색에 따른 2차 문제제기, 분석방법 제시, 분석결과표), 결론(정상아동, 장애아동의 비교분석 결과를 결론)에 따라 분석보고서를 작성하시오.

유창성장애 아동 · 정상 아동의 비유창성 말속도에 따른 비교 분석보고서

목 차

· 서론	-----	1
· 본론	-----	2
- 데이터설명		
- 분석결과		
· 결론	-----	6

2020.06.24.

작성자: 데이터과학융합스쿨 20183254 최윤수

I 서론

유창성장애는 말의 전진적 진행이 운동신경적인 잘못으로 인해 말소리, 음절 또는 낱말의 산출이 방해를 받는 것이라 정의한다. 즉, 일반적으로는 말더듬이라 불리는 말장애의 한 종류라 볼 수 있다. 말의 흐름에 산출과정이 부드럽지 않고 순조롭지 않은 특징을 가지고 있다. 이런 유창성장애를 가진 사람들은 발화시 정상인과는 다른 비유창성 유형 및 빈도를 보이게 된다. 소리와 음절의 산출과 이해에 관한 연구결과에 따르면 화자가 산출하는 비유창성 유형과 비유창성의 빈도가 그 사람이 말더듬인지 아닌지를 판단할 수 있는 지표가 된다고 한다. (Zebrowski, 1994). Zebrowski(1994)는 화자의 비유창성 빈도가 높을수록 청자는 화자의 비유창성 정도가 심각하다고 판단하게 된다고 한다.

여기서 비유창성 유형과 유창성장애의 대표적인 특징으로 보고 있는 말속도에 관한 연구가 활발하게 진행되고 있다. 그 이유는 유창성장애가 2세-5세 사이에 가장 많이 발생하게 되는데, 이 시기가 사람의 언어발달이 이루어지는 시기이기 때문이다. 따라서, 아동에게서 발생하는 비유창성이 발달적인 과정에서의 비유창성인지 문제가 되는 말더듬이 현상인 것인지 구별할 수 있는 기준이 필요하다.

유창성장애를 특징짓는 비유창성 유형은 단어를 기준으로 작은 단위의 요소들로 비유창성을 판단한다. Zebrowski(1994)는 정상인에게도 보이는 전형적 비유창성을 단어간 비유창성으로, 유창성장애인들이 주로 보이는 비전형 비유창성을 단어내 비유창성으로 분류하였다.

유창성장애를 가지고 있는 사람 정상인의 말속도 측정은 많은 논란이 있어왔으나 현재 임상에서는 진단시 말속도의 측정을 포함하도록 하고 있다. (Guitar, 1998; Zebrowski, 1994). 이는 말속도가 말더듬의 정도와 관계가 있으며 의사소통에 영향을 준다고 보고있기 때문이다.

정상 아동과 유창성장애 아동의 조음과 언어, 말속도, 비유창성 특징간의 연구 관계와 더불어 상황에 따른 말속도와 유창성 측정과의 관계 또한 살펴보아야 한다.

선행 연구 결과에 따르면 말속도와 비유창성의 발생을 단순히 발화시간의 문제로 해석할 수 없다는 결론을 내릴 수 있기 때문이다.

이에 본 연구에서는 유창성장애 아동과 정상 아동에게 나타나는 비유창성 유형을 전형적 비유창성과 비전형적 비유창성으로 분류하여 놀이상황과 과제상황에서 데이터를 비교하여 상황에 따라 유의한 차이를 보일 수 있는 지를 데이터를 통해 알아보려고 한다.

II 본론 - 데이터설명

실험과정에서 피험자가 서로 다른 유형의 대화 상황에 처해졌을 때, 이러한 대화 상황이 외생 변수로서 작용될 수 있다고 문제를 제기한 뒤, “유창성장애 아동과 정상 아동의 전체말속도와 조음속도를 측정하고 제시상황에 따라 유의한 차이를 보인다”라는 가설에서 출발하여 본 연구를 진행한다.

1. 연구대상

서울 및 경기지역에 거주하는 만 4세, 5세 아동을 대상으로 실시하였다. 연구에 참여한 아동은 총 20명으로 유창성장애 아동 10명, 유창성장애 아동과 생활연령을 일치시킨 정상아동 10명이다. 유창성장애 아동의 평균연령은 58.7개월(연령범위 48개월-70개월)이었으며, 정상아동의 평균연령은 58.9개월(연령범위 50개월-70개월)이었다.

2. 데이터 설명

정상아동 그룹과 장애아동그룹을 비유창성 분류유형 방식으로 분석하여 비유창성의 유형 및 빈도를 계산하였다.

<표1 비유창성 분류유형>

유형		설명
전형적	주저	1초 또는 그 이상의 침묵.
	삽입	의미전달과 관계없는 중모음, 낱말, 구를 발화에 포함하는 것.
	수정	발화 시 전달내용, 문법형태나 낱말의 발음 등을 바꿔 말하는 것.
	미완성	낱말이나 발화를 완성하지 않고 끝내는 것. 보통 수정이 뒤따름.
	구반복	두 개 이상의 완성된 낱말에서의 반복
	단어반복	낱말 전체의 반복. 일음절 단어도 포함.
비전형적	음절반복	소리반복과 낱말반복의 중간으로 음절만 반복.
	소리반복	한 음소나 이중모음의 일부만을 반복.
	연장	음운이나 모음의 한 요소가 부적절하게 길게 지속되는 것.
	막힘	음운을 시작하거나 파열음 발화 시 생기는 부적절한 시간차로 인해 소리가 막히는 것.

말속도 분석 데이터의 경우, 발화 단위로 나눈 후 놀이상황과 과제상황에서 아동별로 각각 발화 10개씩을 선택하고, 전체말속도와 조음속도를 문장을 끝낸 시간(분) 단위로 계산하였다.

데이터 탐색과정에서 긍정이나 부정을 나타내는 1어절 발화(ex. 네, 싫어요, 아니요 등)은 아동이 자주 사용함으로써 유창하게 발화하여 유창성을 인위적으로 확대 해석시킬 수 있는 문제가 발생하므로 해당 데이터는 분석에서 제외한다.

- 2차 문제제기

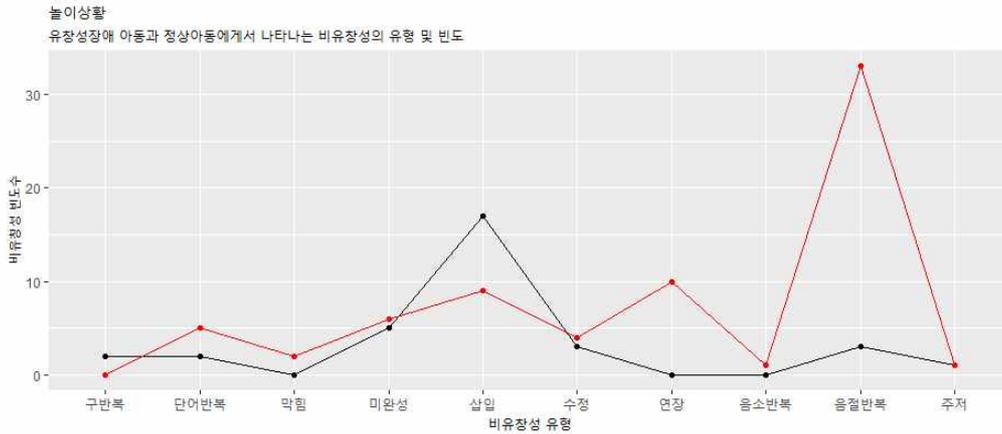
단, 정상 아동과 유창성장애 아동의 말속도는 개인차로 인한 다양성이 존재한다고 볼 수 있어 말속도가 유창성장애와 직접적인 관련이 있다는 주장은 뒷받침될 수 없다.

II 본론 - 분석결과

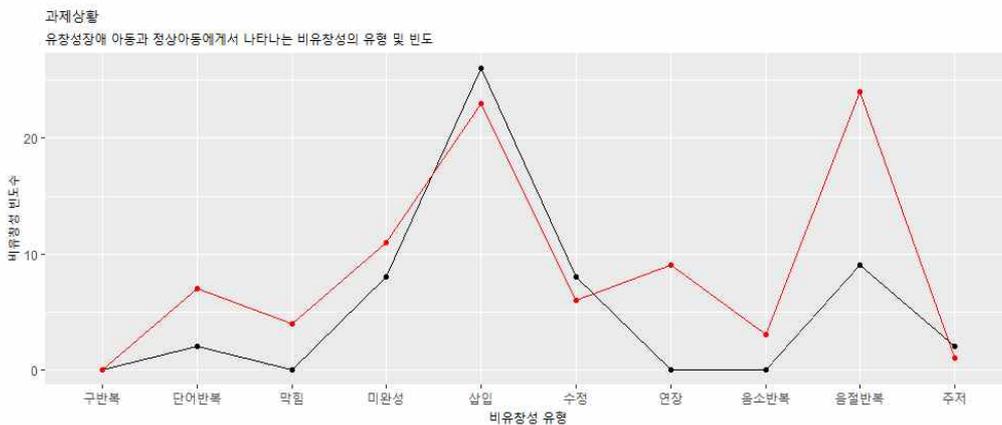
1. 정상 아동과 유창성장애 아동의 비유창성 유형과 빈도

놀이상황과 과제상황에서 나타난 정상 아동과 유창성장애아동의 비유창성 유형 및 빈도수는 <그림3>에 그래프로 제시하였다. 비유창성 유형의 평균 빈도수와 비율은 <표2>에 제시하였다. <표2>에서 제시한 바와 같이 정상 아동의 발화에서 가장 많이 관찰된 비유창성 유형은 삽입(평균 4.3회)으로, 전체 비유창성의 48.86%를 나타낸다. 유창성장애 아동에서의 가장 많이 관찰된 비유창성 유형은 음절반복(평균 5.7회)으로 전체 비유창성의 35.85%를 나타낸다. 정상 아동은 전형적 비유창성이 전체 비유창성의 86.36%였으며, 비전형적 비유창성이 13.64%로 나타났다. 반면에 유창성장애 아동은 전형적 비유창성이 전체 비유창성의 45.91%, 비전형적 비유창성이 54.09%로 비전형적 비유창성의 비율이 더 높게 나타났다.

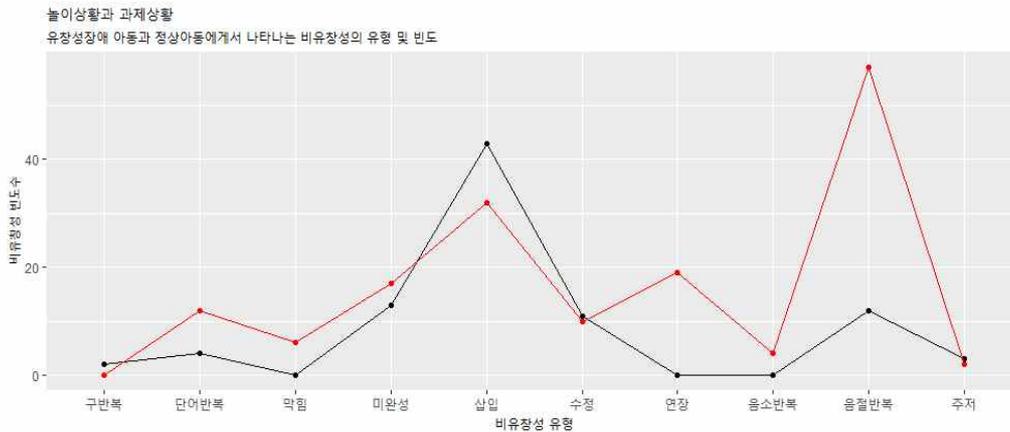
선 구분: 정상아동 · 유창성장애아동



[그림1 놀이상황에서 나타난 정상 아동과 유창성장애 아동의 비유창성 유형 및 빈도수]



[그림2 과제상황에서 나타난 정상 아동과 유창성장애 아동의 비유창성 유형 및 빈도수]



[그림3 놀이+과제상황에서 나타난 정상 아동과 유창성장애 아동의 비유창성 유형 및 빈도수]

<표2 놀이상황과 과제상황에서 나타난 정상 아동과 유창성장애 아동의 비유창성 유형별 평균 빈도수 및 비율>

		정상 아동 (n=10)			유창성장애 아동 (n=10)		
		평균	SD	비율(%)	평균	SD	비율(%)
전형적	주저	0.3	0.48	3.41	0.2	0.63	1.26
	삽입	4.3	4.08	48.86	3.2	3.22	20.12
	미완성	1.3	0.95	14.77	1.7	1.42	10.69
	수정	1.1	0.74	12.50	1.0	0.82	6.29
	구반복	0.2	0.63	2.27	0.0	0.00	0.00
	단어반복	0.4	0.70	4.55	1.2	1.62	7.55
	합계	7.6	1.55	86.36	7.3	3.13	45.91
비전형적	음절반복	1.2	0.79	13.64	5.7	4.81	35.85
	소리반복	0.0	0.00	0.00	0.4	0.52	2.52
	연장	0.0	0.00	0.00	1.9	5.04	11.95
	막힘	0.0	0.00	0.00	0.6	1.35	3.77
	합계	1.2	0.60	13.64	8.6	7.66	54.09
	총합	8.8	1.31	100.00	15.9	1.73	100.00

1-1. 정상 아동과 유창성장애 아동의 비유창성 유형과 빈도에 관한 분석 결과

연구대상과 제시상황에 따라서 유의한 차이가 있는지 확인하기 위하여 이원분산분석을 실시한다. 이원분산분석 (two-way analysis of variance)은 독립변수가 2개인 경우 집단의 평균의 차이가 있는지를 검증하는 분석방법이다.

전형적 비유창성의 평균은 정상 아동의 경우 놀이상황에서 평균 3.00회(SD=1.56), 과제상황에서 평균 4.60회(SD=3.10)였으며, 유창성장애 아동의 경우 놀이상황에서 평균 2.80회(SD=1.93), 과제상황에서 평균 4.50회(SD=2.55)였다.

- 전형적 비유창성에 대한 이원분산분석표

	제곱합	자유도	제곱평균	F	P
연구집단	0.225	1	0.225	0.040	0.842
제시상황	27.225	1	27.225	4.888	0.033
연구집단*제시상황	2.500E-02	1	2.500E-02	0.004	0.947

비전형적 비유창성의 평균은 정상 아동의 경우 놀이 상황에서 평균 0.30회(SD=0.67), 과제상황에서 평균 0.90회(SD=0.88)였으며, 유창성장애 아동의 경우 놀이상황에서 평균 4.00회(SD=3.59), 과제상황에서 평균 4.60(SD=4.81)로 나타났다.

- 비전형적 비유창성에 대한 이원분산분석표

	제곱합	자유도	제곱평균	F	P
연구집단	136.900	1	136.900	14.694	0.000
제시상황	3.600	1	3.600	0.386	0.538
연구집단*제시상황	0.000	1	0.000	0.000	1.000

2. 정상 아동과 유창성장애 아동의 말속도 분석

2-1. 전체말속도와 조음속도 분석 결과

전체말속도 평균은 정상 아동의 경우 놀이상황에서 4.94SPS(SD=0.52), 과제상황에서 3.66 SPS(SD=0.59)였으며, 유창성장애 아동의 경우 놀이상황에서 4.84SPS(SD=1.08), 과제상황에서 3.36SPS(SD=0.85)로 나타났다.

- 전체말속도에 대한 이원분산분석표

	제곱합	자유도	제곱평균	F	P
연구집단	0.390	1	0.390	0.618	0.437
제시상황	18.948	1	18.948	30.011	0.000
연구집단*제시상황	0.105	1	0.105	0.166	0.686

조음속도의 평균은 정상 아동의 경우 놀이상황에서 5.07SPS(SD=0.55), 과제상황에서 3.99SPS(SD=0.61)였으며, 유창성장애 아동의 경우 놀이상황에서 5.26SPS(SD=0.98), 과제상황에서 4.32SPS(SD=0.64)로 나타났다.

- 조음속도에 대한 이원분산분석표

	제곱합	자유도	제곱평균	F	P
연구집단	0.673	1	0.673	1.319	0.258
제시상황	10.211	1	10.211	20.001	0.000
연구집단*제시상황	4.692E-02	1	4.692E-02	0.092	0.764

III 결론

- 전형적 비유창성과 비전형적 비유창성 분석에 따른 결과

분석을 통하여 놀이상황과 과제상황에 따른 빈도에 유의한 차이가 있음을 확인할 수 있었다. ($F_{(1, 36)} = 4.888, P=0.033$) 하지만 전형적 비유창성 유형의 집단간 차이는 유의하지 않게 나왔다. 이와 같은 결과를 바탕으로 정상 아동과 유창성장애 아동 모두 상황에 따라서 전형적 비유창성에 해당하는 주저, 삽입, 미완성, 수정, 구반복, 단어반복을 발화에서 보인다는 것을 알 수 있었다. 제시되는 상황이 변화함에 따라서 연구대상 아동들이 모두 전형적 비유창성이 유의하게 변화될 수 있다는 것을 확인할 수 있었다. 놀이상황보다 과제상황에서 전형적 비유창성 빈도가 많이 보인 것을 해석하면 연구대상(아동)들이 과제상황에 부담감을 갖고 있음을 볼 수 있다.

비전형적 비유창성의 빈도에 따라 분석한 결과, 정상 아동과 유창성장애 아동 집단간에서 유의한 차이를 나타내었다. ($F_{(1, 36)} = 14.694, P<0.001$) 결과를 통하여 선행연구에 대한 근거를 확인할 수 있었다. 즉, 청자 대부분 단어내 비유창성이 있는 사람의 발화를 유창성장애라 판단, 단어간 비유창성 가지고 있는 사람을 정상이라 판단하게 될 수 있음을 볼 수 있다. 정상 아동과 유창성장애 아동은 제시상황에 따라서는 비전형적 비유창성에 유의한 차이를 나타내지는 않았다. 비전형적 비유창성이 유창성장애 아동에게만 유의하게 나타나는 단어내 비유창성이므로, 유창성장애 아동과 정상 아동의 비유창성 유형의 질적 차이를 보여준다고 해석할 수 있다.

- 전체말속도와 조음속도 분석에 따른 결과

전체말속도 데이터를 분석한 결과, 정상 아동은 제시상황별 말속도가 유의한 차이를 보였던 선행연구와 일치하게 나왔다. ($F_{(1, 36)} = 30.011, P<0.001$) 유창성장애 아동의 경우에는 과제상황별로 정상 아동과 유의한 차이가 나타나지는 않았다. 그러나 수치적으로 약간 느린 전체말속도를 나타내고 있다. 이는 유창성장애 아동이 정상 아동보다 빠른 속도로 말하지는 않는다는 연구 결과를 확인할 수 있다. 그리고 놀이상황과 과제상황에 따른 정상 아동과 유창성장애 아동의 전체말속도 변화 역시 선행연구와 일치하였다. 해당 결과는 앞서 언급한 바와 같이 과제나 상황에 대해 부담감을 가진 상태이며, 과제를 수행할 경우에는 놀이상황에 비해 과제에 대한 집중력과 언어 능력이 필요하게 되기 때문인 것이라 해석할 수 있었다.

조음속도의 경우에는 유창성장애 아동은 수치적으로 약간 빠른 조음속도를 나타내고 있었으나 유의한 통계 결과를 나타내고 있지는 않았다. 따라서 분석 결과를 토대로 정상 아동과 유창성장애 아동사이에 말속도의 유의한 차이가 없다는 선행연구결과의 근거가 될 수 있다. 조음속도의 차이는 과제상황에서 유의하게 느린 말속도를 산출하며, 전체말속도와 동일한 양상임을 확인할 수 있었다. ($F_{(1, 36)} = 20.001, P<0.001$)

분석결과를 통하여 앞서 상황에 따라 다른 양상을 보인 바와 같이 개별적 경험 등에 의한 개인차가 발생할 수 있음을 시사하고 있다고 생각한다. 따라서 개별적 음운 특성을 고려하여 치료 목표를 수립하는 데 이용하고, 앞으로 더 많은 대상자들을 포함한 연구가 이루어진다면 좀 더 유창성장애에 대한 정확한 정보를 얻을 수 있으리라 본다.