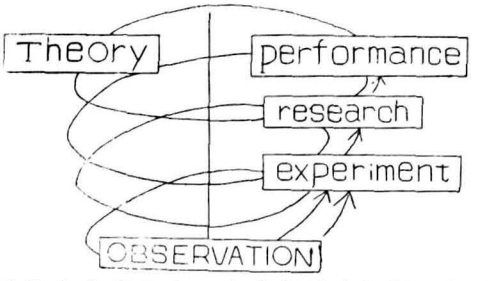
言語敎育을 爲한 Research Design

宋 堯 仁
(東 國 大 學 校)

I. 緖 言

本稿는 言語敎育에 있어서 theory와 performance를 보다 더 緊密히 連關시켜 줄 수 있는 Research Design의 問題, 그리고 이에 對한 methodological application을 可能한 限 平易한 統計學的 範圍에서 다루어 보려는 試圖이다. 여기에서 말하는 theory란 theory of learning을 말하는 것이며 이는 performance of learning에 直接 關係가 있는 것으로서 一般的인 槪念의 performance of education과는 無關한 것으로 限定시키고자 한다. 事實 言語敎育에 있어서 theory란 performance보다 너무 앞장설 수도 없는 것이며 또 앞장서도 안되는 것으로 본다. 이에 對하여 Dykstra 敎授는 performance란 observation 단계에서 experiment와 research 단계를 거쳐서 向上되고 精密化될 수 있는 것이며 이와 같은 performance는 동시에 theory에 反映되어 theory 自體의 擴大와 向上 그리고 精密化에 寄與하게 되는 말하자면 다음과 같은 Spiraling Co-consideration 關係가 成立되는 것이라고 말하고 있다[1].



오늘날 言語學界는 變形 生成 理論이 從來의 構造論을 褪色케 할 程度로 活潑히 展開되고 있고, 나아가서는 成層(Stratification)을 論하게 되었으며 言語全般에 걸친 innateness를 究明하려 드는 형편인데 不幸히도 言語敎育의 research 面에 있어서는 이렇다 할 進步를 거두지 못하고 있는것이 實情이다. 이點 John B. Carroll의 다음과 같은 1963年度의 陳述이 4年後인 오늘날 是正이 되었다고는 볼 수 없는 것이 숨김없는 實情이기도 하다.

[1] Gerald, Dykstra, "Spiraling Co-consideration," (unpublished manuscript), Teachers College, Columbia University, 1966.

「Educational research has contributed very little to foreign language teaching methodology aside from general knowledge regarding the construction of achievement tests, the role of foreign language aptitude in the language process, and the psychology of bilingualism. .....Foreign language teachers' researches have been inadequate with respect to research methodology」[2].

이를 國內의 外國語教育 舞臺로 옮겨 본다면 事情은 더 딱하다. 이에 對한 原因은 制度上의 缺陷等 여러 가지가 있겠으나 가장 큰 原因은 教師와 學生들의 授業 過負擔이라고 보는 것이 옳을 것이다. 그들이 도저히 이와 같은 教課外의 實驗 및 研究에 時間을 割愛할 수 없을 程度로 우리네의 教課編成은 緊迫感을 주게끔 짜여져 있다. 그러나 이러한 狀況下에 놓여 있다 하여 傍觀만 할 수 없는 것이 우리의 實情이므로 本稿에서는 教師들이 教課履行 過程에서 教課의 進行에 別다른 影響을 주지 않는 範圍內에서 얼마든지 實踐할 수 있는 research design을 micro-teaching과 single-class teaching 面에서 다루어 보려는 것이다. 事實 美國과 같은 나라에서도 圓滿한 發展을 보지 못한 이 部面이어늘 우리나라의 實情으로는 도저히 制度化된 어떠한 研究機關의 專門的이고 包抱的인 研究業績에 期待를 갖는다는 것은 無望한 일이다. 오직 言語教育에 臨하고 있는 우리들이 最大限의 融通性을 發揮한다면 斷片的이고 散發的이긴 하나, 긴 眼目으로 볼 적에는 至大한 貢獻을 할 수 있는 實驗과 研究를 할 수 있는 것으로 본다. 이런 意味에서 볼 적에 英語英文學誌 21號(英語學特輯)에 실린 申相淳教授의 「言語 實驗室과 言語技術 習得과의 相互關係」 報告書는 意義가 자못 큰 것으로 慶賀해 마지 않을 努力인 것이다.

言語教育에 있어서 앞으로 實驗과 研究를 거쳐야 될 分野는 荒蕪地에 가깝다 해도 過言이 아니다. 外國語教育을 시작할 適正年令의 決定과 같은 基本的인 問題로 부터 即刻的인 feedback을 提供하는 경우와 遲延된 feedback을 提供하는 경우의 相對的인 効率性과 같은 枝葉的인 問題에 이르기까지 아직껏 reliability와 validity를 充足시키는 實驗 및 研究의 惠澤을 받은 分野는 極히 드믄 것이다. 그러므로 本稿에서는 methodological application을 간추려 論한 다음, 끝으로 앞으로 우리가 손댈수 있는 research design에 適合한 分野를 phonology, syntax, vocabulary, reading, writing, 및 comprehension의 순서로 分類하여 各己 目錄을 맨들어 添加하기로 한다. 便宜上 母國語는 NL로 外國語는 FL(=TL)로 略稱하기로 한다.

---

[2] John B. Carroll, "Research on Teaching Foreign Languages", N. L. Gage, ed., *Handbook of Research on Teaching*, (Chicago, 1963), p. 1094. 이 點은 筆者가 今年여름 美國內 主要 言語教育機關 視察中 効率的이기로 定評있는 Foreign Service Institute 조차 Testing Unit는 編制上 갖추고는 있으나 現在로는 achievement와 aptitude test에 專念하고 있지 teaching methodology에 對하여는 一切 손을 못대고 있음을 보고 實感있게 느낀바 있다.

II. Micro-teaching Research Design

一個人 만을 對象으로 한 micro-teaching에 있어서는 우선 統計學的인 評價適用이 制限되어 있어서 一見 簡單한 實驗같이 보이기는 하나 methodological laxity를 犯하기 쉽고 主觀的인 判斷이 結論을 흐리게 할 우려가 있다. Research design으로서는 어떤 問題(Vocabulary, Pronunciation 其他)中에서 難易度와 異質度가 같은 두개의 問題를 選定하여 各各 Problem I 및 Problem II라 일컫는다. 다음에는 Presentation의 方法을 두가지 採擇하는바 例를 들어 Presentation A는 "Study" Presentation으로 Presentation B는 "Teacher" Presentation으로 定하고 experiment를 實施하는바, 이에 앞서 respondent의 해당 分野에 關한 事前知識을 測定하는 pre-test를 實施한다. 대개의 경우 pre-test의 成績은 0에서 10%程度에 그치는 것이 常例이다. 10%라는 score는 chance factor를 감안했기 때문에 可能한 數値인 것이다. Test를 시작한 다음에는 第一次로 부터 適切히 maturation effect를 考慮하여 short-range, mid-range, 및 long-range別 해당 期別로 必要에 따르는 回數의 test를 實施한다. 每 test가 끝난 뒤에는 敎師의 裁量으로 reinforcement 即 positive或은 negative feedback을 提供할 수 있다[3].

詳論에 앞서 簡單한 frame을 하나 例를 들어 만들어 보기로 한다. 이 frame은 vocabulary acquisition을 위한 두가지 presentation의 相對的인 効率性을 實驗하기 위한 것으로서 架空의 FL을 쓰기로 하고 vocabulary item은 nonsense syllable로 한다[4].

Frame 1

1. 問　題

英語를 NL로 使用하는 學生의 FL vocabulary 習得에 있어서 "Study" Presentation (A)과 "Teacher" Presentation (B) 敎授法中 short-range 見地에서 본 相對的인 効率性을 檢討하고자 함.

2. 節　次

Respondent는 音樂을 專攻하는 19歲의 學部一學年 學生인바 nonsense syllable을 使用하였으므로 pre-test는 省略하였음.

本 experiment에서는 다음과 같은 問題가 使用되었음.

Problem I.　　kamino (물),　　nendumo (담배),

---

[3] Maturation effect, reinforcement 및 其他 此後에 使用하는 用語의 定義는 다음을 參照할 것. C.V Good,.ed., *Dictionary of Education*, (New York, 1959) : Chester W., Harris, ed., *Encyclopedia of Educational Research*, (New York, 1960).

[4] 本稿에 紹介하는 research frame은 前記 Dykstra 敎授의 指導下에 筆者가 作成 實施해 본 것 中에서 adapt한 것임.

(PresentationA)	poreme (책)	mumosa (사과)
	tapso (바다)	samida (여자)
Problem II.	poruna (산)	mokole (감자)
(PresentationB)	podike (소년)	narima (불)
	kunibo (연필)	tahina (나무)

前記 問題를 各單語마다 카아드에 하나씩 記入하고 Presentation A 에서는 카아드 6 枚를 respondent 에게 주어서 自習하게 하고 Pressentation B 에서는 주지 않고 演習을 audio-lingual 方式에 依하여 實施하였음.

各 presentation 마다 90 秒동안 自習 또는 演習을 시키고 最初의 test 는 presentation 이 끝난 以後 30 秒가 지난 뒤에 實施하였음. 第二回 test 는 2 時間後, 第三回 test 는 24 時間後, 第四回 test 는 72 時間後, 그리고 第五回 test 는 一週日後에 實施하였음. 그리고 post-test 는 pre-test 와 같은 理由로 省略하였음. 場所는 換氣가 잘 되어있는 書齋에서 每 test 는 午前 9 時에서 10 時 사이에 實施하였으며 每回마다 positive feedback 을 提供하였음. 本 experiment 에서는 事前에 50 % null hypothesis 를 設定하였음[5].

3. 結　果

다음과 같은 percentile score 가 記錄되었음.

最　初 test	Presentation A	83
	Presentation B	67
第二回 test	Presentation A	67
	Presentation B	50
第三回 test	Presentation A	50
	Presentation B	33
第四回 test	Presentation A	50
	Presentation B	33
第五回 test	Presentation A	50
	Presentation B	33
平均成績	最初 test	75
	第二回 test	58.5
	第三回 test	41.5
	第四回 test	41.5

[5] Hypothesis 에 關하여 留意할 事項은 Donald T. Campbell, and Julian C. Stanley, "Experimental and Quasi-experimental Designs for Research on Teaching," Gage, ed., *Handbook*, pp. 205-6 에 詳述되어 있음.

第五回 test	41.5
Presentation A	60
Presentation B	43.2

4. 討議 및 結論

이 respondent는 音樂專攻이므로 Presentation B에 있어서 aural perception에는 銳敏한 反應을 보였고, 한편 Presentation A에 있어서는 普通以上의 集中的인 努力이 엿보였음. 然이나 test의 結果를 보면 第二回 까지는 score가 같은 差異를 두고 下落했으며 第三回 부터도 亦是 score間의 差異는 變化가 없었으나 第五回까지 兩 score가 固定化 되어 간 것을 알 수 있음. 兩 score의 差異는 全般的으로 同一하게 나타났으며 始終 Presentation A가 優勢한 것이 立證되었음. 故로 이 respondent에게는 vocabulary 習得을 위해서는 Presentation A가 Presentation B보다 더 効果的이라는 暫定的인 結論을 나릴수 있음. 따라서 前記 2項에서 設定한 null hypothesis는 確認되지 못하였으므로 이를 reject하기로 함.

위에서 結論을 暫定的으로 내린 것은 experiment 自體가 여기서 完結된 것이 아니고 여기에서 發見된 새로운 事實等에 依據하여 새로운 一連의 rival hypothesis가 생길 수 있는 것이며, 따라서 이러한 hypothesis도 亦是 같은 方式으로 確認節次를 겪게 된다는 뜻에서 이다.

위와 같은 experiment는 敎師들이 豫定된 敎課에 別다른 變動을 일으키지 않고 時間的으로도 이렇다 할 희생을 초래함이 없이 얼마든지 design하고 實施할 수 있는 것이다. 그러면 이와 같은 experiment에 對한 validity와 reliability는 어떠한 것인가를 보기로 한다. 勿論 어떠한 test에 있어서도 무턱대고 "이 test가 valid한가?"하고 質問한다는 것은 禁物이다. 왜냐하면 아무리 外見上 valid하지 못한 test일지라도, 어떠한 特定의 目的을 위해서는 valid할 수가 있기 때문이다. 그러므로 validity는 external validity와 internal validity로 區分하여 論하게 된다. External validity라 하면 generalizability 即 結果의 適用에 對한 範圍와 可能性을 따지는 것이며 internal validity는 experiment의 效果가 어떠한 特定의 case에 있어서 差異 乃至 變化를 招來하였는가를 따지게 되는 것이다[6]. 이러한 validity crtieria에 依據한다면 위의 experiment는 internal validity에 있어 優勢하고 external validity에 있어서는 劣勢하다.

다음으로 reliability를 보면, 亦是 두가지로 나눠 檢討할수 있는바, 그 하나는 judge reliability요, 다른 하나는 test reliability이다. Judge reliability는 다시 intra judge reliability와 inter judge reliability로 區分할 수 있는데 micro-teaching experimet의 경우는 inter judge reliability 即 self-reliability가 適用된다[7]. Self-reliability란 다시 말하면 judge가 test에 臨

[6] Validity에 關하여는 Gage, ed., *Handbook*, p. 175를 參照할 것이며 特히 external validity에 影響을 주는 四個要素와 internal validity에 影響을 끼치는 八種의 extraneous variables에 留意할 것.
[7] *Ibid.*, pp. 253-4.

하는 態度의 一貫性을 말한다. 本 experiment 의 경우 self-reliability 는 높은 것으로 假定해 둔다. 다음으로 test reliability 에 影響을 끼치는 要素는 아래의 네 가지가 있다.

첫째 test 의 期間이 길면 길수록 reliability 가 높아지고,

둘째 test 內容이 heterogeneous 하게 構成이 되면 될수록 reliability 를 增加시켜 주고,

셋째 test 間의 間隔이 짧으면 짧을수록 reliability 를 높여주고,

넷째 irregularity 即 室內의 不快한 溫度라든지 指示事項의 違反, 身體的 不安, motivation 의 不均衡 및 不正行爲等은 reliability 를 減少시켜 준다.

이러한 點에서 檢討한다면, 本 experiment 에 對한 test reliability 는 자못 높다고 할 수 있을 것이다. 여기에서 reliability 와 validity 의 相互關係를 밝혀 두는것이 妥當할 것으로 보는바, 이에 關해서는 Lyman 敎授가 明瞭하게 說明한것이 있다.

그는 reliability 를 "the consistency of measured performance" 라고 定義를 내리고 validity 는 "a test's ability to measure what we want it to"라고 풀이한 然後에 "Reasonably high reliability is necessary for good validity. Reliability does not insure validity; reliability is necessary for, but is no guarantee of, validity"라고 말한 바 있다[8].

Judge reliability 와 아울러 judge validity 도 亦是 論할수 있는바, judge validity 는 judge 即 敎師의 "situational, motivational, cognitive, and genetic factors"[9]에 依하여 左右되고 特히 그의 人品 態度 趣味 動機 그리고 專問的 知識, 敎育經歷 및 特定分野에 關한 經驗 및 技術等의 specific variables 에 依하여 影響을 받는다. 따라서 本 experiment 에 있어서의 judge validity 는 筆者自身의 이와 같은 與件을 檢討하여 볼진대 도저히 높이 評價할 수는 없다고 본다.

Micro-teaching Research Design 은 序頭에서 말한 바와 같이 一見 簡單하게 보이고, 또 事實實施하는데 있어서도 single-class teaching experiment 보다는 簡單한 편이다. 然이나 경계해야 될 事項이 위에 말한 點을 떠나서도 몇 가지 더 있다.

그 中 가장 重要한 것이 extraneous variable 의 統制問題이다. 어떠한 experiment 든지 variable 을 全廢시킨다는 것은 不可能한 것이므로, 한 test 단계에서 다음 test 단계로 넘어갈 적에는 "one variable at a time" 原則을 固守한다는 것이 重要하다. 例를 들면 第一回에 oral test 를 實施했으면 第二回에도 文字 그대로 oral test(錄音器 代用이나 written test 가 되면 이미 또 하나의 variable 이 介入된것임), 第一回에 椅子에 앉아서 test 를 받았으면 次回에도 椅子에 앉아서 test 를 받을것 萬若에 起立姿勢로 한다든지 하면, 이것 亦是 願치 않는 variable 이 介入되는 것이다[19].

---

[8] Howard B. Lyman, *Test Scores and What They Mean*, (Englewood Cliffs N.J., 1963), p. 31.

[9] Gage, ed., *Handbook*, p. 128.

[10] Variable 에 關한 解說은 George W. Snedecor, *Statistical Methods*, (Ames, Iowa, 1956), p. 56 및 Gage, ed., *Handbook*, p. 146 을 參照할 것.

다음으로 경계할 重要한 事項은 actual learning situation 을 演出하는 일이다. Experiment 를 위한 experiment 가 되어서는 reliability 가 自然 減少된다. 또한 筆者의 經驗에 依하면 까딱하면 test 를 實施하는 側이나 받는 側이 경솔해지기 쉽고 深刻性을 잃는 수가 있다. 이 亦是 reliability 를 減少시켜 주는 要素이므로 경계해야 된다. 그러므로 事前에 respondent 로 하여금 各 experiment 의 目的과 進行方法을 周知시킴으로서 最大限의 協調를 얻는 것이 上策일 것이다.

Micro-teaching experiment 에서 再三 強調해야 될것은 "one variable at a time"原則과 問題의 內容이 同等한 難易度와 異質度를 갖고 있어야 한다는 點이다. 다만 judge 와 respondent 가 必然的으로 겪게 되는 psycho-acoustic 및 時間的인 biological change 만은 experiment 에 關한 限 無視해 버리게 되는 것이다.

Ⅲ. Single-class Teaching Research Design

Single-class Teaching Research Design 에서는 問題構成만은 micro-teaching 과 同一하게 할 수 있으나 respondent 의 數가 늘어나므로 節次가 多少 複雜해지는 同時 experiment 의 結果가 利用될 수 있기 前에 統計學的인 評價를 받게 되어 있는것이 特色이다. Single-class teaching experiment 는 敎師가 擔當한 學級을 直接 利用할 수 있으므로 micro-teaching experiment 와 같이 別定時間과 場所를 마련할 必要가 없는 利點이 있는 反面, control 의 문제가 相當한 知識과 技術이 없이는 解消되기 어려운 難點도 있다. Single-class teaching research design 으로서는 Latin square 가 가장 많이 使用된다. 가장 代表的이고 簡便한 Latin square 로서 앞의 micro-teaching 문제를 適用한다면 다음과 같은 arrangement 를 擇할 수 있다.

	Presentation A	**Presentation B**
Problem Ⅰ	Group A	Group B
Problem Ⅱ	Group B	Group A

이와 같은 design 은 學級을 二等分하면 간단히 이룩될 수 있는 것인바 이보다도 더 徹底한 Latin square arrangement 를 擇한다면 四等分을 要하게 되는 counterbalanced design 을 適用할수 있다. 代表的인 counterbalanced design 을 表示하면 아래와 같다[11]. (X=exposure, O=observation)

---

[11] Counterbalanced design 을 包含한 quasi-experimental design 과 true experimental design 에 關하여는 Gage, ed., *Handbook*, pp. 181ff를 參照할 것. Campbell 과 Stanley 는 true experimental design 으로서 time series(0000×0000)外에 九種의 design 을 紹介하고 아울러 各 design 에 對한 sources of invalidity 를 圖表로 表示하고 있다.

Group	Time I	II	III	IV
A	X_1O	X_2O	X_3O	X_4O
B	X_2O	X_4O	X_1O	X_3O
C	X_3O	X_1O	X_4O	X_2O
D	X_4O	X_3O	X_2O	X_1O

Latin square의 長點은 internal replication의 一貫性을 維持할 수 있으며 history, maturation, testing, instrumentation, regression, selection 및 mortality에 있어서 他 design보다 validity가 優勢한 것이며 短點은 multiple-X interference에 있어서 validity가 劣勢한 點이다.

다음에는 single-class teaching experiment에 있어서의 population 問題를 檢討해 보기로 한다. 敎師가 自己가 擔當한 學級을 respondent group로 擇했을 경우의 sampling technique는 勿論 random이겠으나, population 內의 特殊한 segment를 sample했다고 볼 수 있다. 本來 experiment에 있어서는 事前에 偶然에 依한 equation 即 random sampling이 必要한 것이지, 故意的인 matching을 通한 equation이란 人間本然의 各自가 지니고 있는 先天的 및 後天的 素質과 習性 때문에 있을 수 없는 것이며 또 不可能한 것이다[12]. 그러므로 judge인 敎師는 自己의 class를 等分할 적에 故意的인 matching을 試圖할 것이 아니라, randomize할 것이 要求된다.

아래에 Latin square를 適用한 簡單한 frame을 만들어 보기로 한다.

Frame 2

1. 問　題

英語를 NL로 使用하는 18歲~20歲 年齡 group에 들어 있는 學部一學年 政治學 專攻學生들에 對한 vocabulary acquisition에 있어서 "Study" Presentation (A)과 "Teacher" Presentation (B) 敎授法의 short-range에서 본 相對的인 効率性을 調査하고자 함.

2. 節　次

Respondent group는 randomized sampling에 依하여 Group A와 Group B로 分割되었고, 그 結果 各 group마다 8名씩 配當이 되었음. Problem I에 있어서는 Group A가 Presentation A를 맡고, Group B가 Presentation B를 맡았으며, Problem II에 있어서는 Group B가 Presentation A를, 그리고 Group A가 Presentation B를 맡는 Latin square arrangement를 擇하였음. FL의 vocabulary는 便宜上 nonsense syllable로 代置하였으며, pre-test는 省略하였음. Presentation A에 있어서는 6個의 單語 即 nonsense syllable이

[12] Matching을 通한 equation을 犯한 代表的인 case로서 1936年에 Literary Digest 誌에서 電話加入者만을 sample하여 實施한 某種 survey 같은 有名한 case가 있다.

적혀 있는 카아드를 各 respondent 에게 주어서 各自의 能力과 素質에 適合한 自習을 90秒동안 시켰고 Presentation B 에 있어서는 같은 數의 카아드를 敎師가 使用하되 respondent 들에게는 보여 주지 않고 audio-lingual 方法에 依하여 가장 効果的이라고 믿어지는 方式으로 演習을 亦是 90 秒 동안 시켰음. 이點 micro-teaching 과는 달리 group drill 과 individual drill 을 번갈아가며 listen-repetition-recall 演習을 實施하였음. 問題는 前記 micro-teaching 의 경우와 同一한 것을 使用했으므로 省略함.

每 Presentation 이 지나면 30 秒가 經過한 後에 test 를 하고 第二回 test 는 2 時間後, 第三回 test 는 24 時間後, 第四回 test 는 72 時間後, 그리고 第五回 test 는 一週日後에 實施하였음. 場所는 敎室이며 最初 experiment 는 午前 10 時에서 11 時 사이에 實施하였음. 每 test 가 끝난 뒤에는 reinforcement 로서 positive feedback 을 提供하였음. 本 experiment 에 있어서는 事前에 50 % null hypothesis 를 設定하였음.

3. 結 果

다음과 같은 group 別 percentile score 가 算出되었음.

最初 test

	Presentation A	Presentation B
Problem Ⅰ	Group A—40	Group B—90
Problem Ⅱ	Group B—57	Group A—57

平均成績

Presentation A	48.5
Presentation B	73.5
Group A	48.5
Group B	73.5
Problem Ⅰ	65
Problem Ⅱ	57

第二回 test

	Presentation A	Presentation B
Problem Ⅰ	Group A—30	Group B—80
Problem Ⅱ	Group B—50	Group A—40

平均成績

Presentation A	40
Presentation B	60
Group A	35
Group B	65

Problem Ⅰ	55
Problem Ⅱ	45

第三回 test

	Presentation A	Presentation B
Problem Ⅰ	Group A—28	Group B—75
Problem Ⅱ	Group B—45	Group A—38

平均成績

Presentation A	36.5
Presentation B	56.5
Group A	33
Group B	60
Problem Ⅰ	51.5
Problem Ⅱ	41.5

第四回 test

	Presentation A	Presentation B
Problem Ⅰ	Group A—25	Group B—70
Problem Ⅱ	Group B—40	Group A—35

平均成績

Presentation A	32.5
Presentation B	51.5
Group A	30
Group B	55
Problem Ⅰ	47.5
Problem Ⅱ	37.5

第五回 test

	Presentation A	Presentation B
Problem Ⅰ	Group A—25	Group B—60
Problem Ⅱ	Group B—35	Group A—35

平均成績

Presentation A	30
Presentation B	47.5
Group A	30
Group B	47.5
Problem Ⅰ	42.5
Problem Ⅱ	35

4. 討議 및 結論

本 experiment에서 두드러지게 나타난 現象은 Presentation 間의 差違와 Group 間의 差違가 顯著하다는 것임. Group A나 B를 莫論하고 Presentation B에 있어서 越等히 좋은 成績을 나타내었으며 特히 Group B는 그 performance가 Group A의 倍 또는 그에 가까운 level을 維持하고 있는 것이 特色임. 그러나 이것은 事前에 randomized sampling을 했지만 遇然하게도 Group B에는 bilingual한 家庭的 背景을 가진 respondent가 둘이나 끼어 있었다는 事實과 關聯이 있다고 봄. 다음으로 Problem I과 Problem II에 있어서 problem I의 平均値가 優勢한 것은 history와 maturation effect를 考慮한다면 이상할 것이 없음. 全 experiment를 通하여 Presentation B가 斷然 優位를 차지했으므로 이 respondent group에 對하여는 FL의 vocabulary acquisition에 있어서 Teacher Presentation 敎授法이 보다더 效率的이라는 暫定的인 結論을 내리기로 함. 따라서 事前에 設定한 50% null hypothesis는 確認되지 못하였으므로 이를 reject할 수밖에 없음.

위의 frame에서는 統計學的인 評價를 別途로 다루어 보기 위해서 討議 및 結論事項內에 包含시키지 안했다. 勿論 micro-teaching experiment 外의 모든 경우에는 마땅히 이러한 評價를 討議事項에 包含시키도록 할 것이다.

計算의 便宜上 respondent Group A의 個人別 成績을 아래와 같이 假定하여 檢討하여 보기로 한다.

姓名(略字)	Presentation A	姓名	Presentation B
J.D.	48	J.D.	46
M.L.	46	N.R.	44
T.M.	44	G.G.	42
G.G.	44	T.M.	42
B.D.	42	J.B.L.	40
H.L.T.	42	M.L.	40
J.B.L.	40	H.L.T.	38
N.R.	36	B.D.	34

우선 mean(=mean score)[13]을 보면 Presentation A에 있어서는 42.8이 되고 Presentation B에 있어서는 40.8이 되며 이 respondent group에 對한 mean은 41.8이 된다. 다음으

[13] Mean은 average와 同一하나 modal class와 median과 區別하여 使用되는 것인바 이에 對한 解說과 此後에 나오는 統計學的 用語와 公式의 解說은 McCollough, Celeste, and Loche Van Atta, *Statistical Concepts: A Program for Self-instruction*, (New York, 1963)를 參照할것.

[14] 萬若에 score의 數가 많을 적에는 Jenkins Formula인 $\frac{{}^1/_6 \text{ top scores minus } {}^1/_6 \text{ bottom scores}}{{}^1/_2 \text{ no. of total scores}}$를 適用한다.

로는 standard deviation을 보기로 한다. 이를 위해서 公式 $\sigma=\sqrt{\frac{\Sigma\times i^2}{N}}$ (여기에서 Xi=各 score의 deviation의 數値, N=score의 數)을 適用하면 Presentation A나 B가 똑같이 $\sqrt{\frac{95.5}{8}}=3.3$이 된다[14]. 다시 말하면 Group A의 standard deviation은 Presentation A와 Presentation B에 있어서 共히 3.3을 나타내고 있는바, 이러한 standard deviation은 앞으로 standard error와 significance를 算出할 때에 다시 使用되는 것이다.

세째로 算出할 것은 correlation (C)이다. 이에는 다음과 같은 公式을 使用한다.

$$C=\frac{\text{兩 test에 있어서 上半成績圈에든 respondent 數}}{\text{respondent 總數}}$$

이를 Group A에 適用한다면 J.D., T.M., 및 G.G.가 兩 test에서 上半圈에 들어 있으므로 $\frac{3}{8}=0.37$ 即 37%라는 數値가 나온다. 이 數値는 다시 아래와 같은 correlation table에 適用시킨다.

%	25%	27%	29%	31%	33%	35%	37%	40%	43%	50%
C	.00	.13	.25	.37	.5	.6	.7	.8	.9	1

여기에서 1이라는 correlation은 perfect correlation이며, .00은 correlation이 完全히 random狀態라는 것을 表示한다. Group A의 경우는 37% 即 C=.7이 나오게 되는데 이것은 勿論 perfect correlation에 가까운 높은 數値인 것이다.

다음으로는 reliability coefficient (RC)를 算出해 보기로 한다. RC란 correlation of two measures를 말하는 것이며 公式 $\frac{2C}{1+C}$를 適用하게 된다[15]. Group A의 경우 $\frac{1.4}{1.7}=.8$이 되는데 이 .8이라는 RC는 reasonable RC이므로, 本 experiment의 reliability는 充足되었다고 할 수 있다. RC가 算出된 다음에는 果然 experiment가 얼마 만큼 test를 擴大해야 reasonable RC인 .8을 達成시킬 수 있는가를 算出하는 "enough testing" test를 實施한다.

이에는 Spearman-Brown 公式 $\frac{\text{Desired RC }(1-\text{Observed RC})}{\text{Observed RC }(1-\text{DesiredRC})}=$X times as much testing required를 適用한다. Group A의 경우는 이미 .8 RC가 算出되었으므로 enough testing 檢討는 不必要하게 되었으므로 計算을 위하여 同 RC가 .7이 었다고 假定하고 그에 따른 test의 所要 擴大量을 算出해 보기로 한다. 即 Observed RC가 .7이라고 假定한다면 $\frac{.8\times(1-.7)}{.7\times(1-.8)}=1.7$이 되는데 여기에서 1.7이라는 數値는 本 test의 倍加量을 나타내는 것이므로, 本 experiment에서 使用한 vocabulary item은 6개 였으므로, 이의 1.7倍인 10개를 使用하면 enough testing 即 .8이라는 reasonable RC를 充足시키게 된다고 할 수 있다.

Presentation A의 경우 mean이 42.8이었는데 이에 對한 standard error[16]를 算出하려면

---

[15] RC公式에는 alternate formula로서 Kuder-Richardson Formula 21인 $1-\frac{m(n-m)}{ns^2}$이 있다. (m=mean score, n=number of items, s=standard deviation).

[16] Standard error 및 그다음에 오는 significance는 通常 本 experiment에서 보다는 훨씬 많은 score를 對象으로 算出하는 것이지만 여기서는 그런 대로 適用해본다.

公式 $\frac{\text{standard detviaion}}{\sqrt{\text{no. of scores}}}$ 를 適用하여 $\frac{3.3}{\sqrt{8}}=1.2$ 가 된다. 이것은 score 가 42.8±1.2 即 41.6 에서 44 사이에 나타날 確率이 높다는 말이 된다. 앞에서 算出했던 두개의 mean score 의 差違는2 가 되는바, 이에 對한 standard error 를 또한 算出할 수 있다.

이에는 公式 $\sqrt{\Sigma \text{standard error of mean scores}^2}$ 을 適用하여 $\sqrt{2.9}=1.7$ 即 2±1.7 이라는 答이 나온다.

끝으로 重要한 것은 experiment 에 對한 significance 또는 non-significance 를 決定하는 問題가 남아 있다. 이에 對해서는 앞에서 算出한 兩 mean score 의 差違數値를 亦是 앞에서 算出한 差違數値에 對한 standard error 로 나눈 數値를 가지고 決定하게 된다.

Group A 의 경우 $\frac{2}{1.7}=1.2$ 가 되는데 이 數値가 2 以下가 될 경우는 해당 experiment 는 chance factor 가 너무도 크게 作用했기 때문에 significance 가 없는 것으로 간주되는 것이다. 萬若에 同數値가 2 에서 2.6 사이에 들 적에는 .05 level of confidence 에서 significant 한 것이라고 判斷하게 되는바, 이 말은 다시 말하면 5% 以下가 chance factor 에 依한 것이라는 말과 같다. 그리고 萬若에 同數値가 2.6 以上이 될 경우에는 .01 level of confidence 에서 significant 하다는 結果가 되는데 이말은 換言하면 1% 以下가 chance factor 에 依한 것이라고 할 수 있다. 그러니까 Group A 의 경우는 significance 가 없는바, 이것은 勿論 計算의 便宜上 score 를 假定하였기 때문이며 同 score 內에 chance factor 가 크게 作用한 것을 알 수 있는 것이다.

이제까지 大略 micro-teaching 과 single-class teaching 을 위한 research design 을 frame 을 짜서 檢討해 봤고 아울러 基本的인 範圍內에서 統計學的인 評價를 내려 봤는데, 이것은 어디까지나 methodological application 에 關聯되는 문제인바, 이보다 못지 않게 experiment 를 實施하는 敎師에 對한 implication 乃至 銘心해야 될 留意事項이 있다[17].

첫째 너무도 지나치게 意慾的인 experiment 는 避할 것이다. 그래야만 最少限度內의 variable 을 投入시켜 가며 experiment 를 構成하고 進行시킬 수 있기 때문이다.

둘째 Experiment 의 目標를 正確히 세우고 test 해볼 여러개의 language ability 中에서 一回에 一個의 ability 에만 局限시킬 것.

세째 Language acquisition 에 있어서 phonology, morphology, syntax, semantics, 및 cultural understanding 의 各分野를 個別的으로 고르게 다룰 것.

네째 Listening comprehension, speaking ability, reading comprehension 및 writing ability 를 고르게 test 할 것.

---

[17] Test construction 上의 技術的인 問題等은 Robert, Lado, *Language Testing* (New York, 1964) 및 Adelaida, Paterno, "Foreign Language Testing", Harold B. Allen, ed., *Teaching English as a Second Language* (New York, 1965), pp. 373-92 를 參照할것.

다섯째 各 ability를 測定함에 있어서 test 內容이 實際的 環境과 狀況下에서 使用되고 있는 것과 같은 條件을 갖출 것.

여섯째 Respondent로 하여금 混同을 하거나 짐작을 해버리거나 하는 따위의 모호한 內容의 問題를 構成하지 말것.

일곱째 Respondent의 知識과 能力을 널리 cover하는 代表的인 sample을 取하되 質問을 짧고 簡單한 答이 나오게끔 構成해서 採點時에 自己의 主觀이 作用하지 않도록 留意할것.

그리고 効果的인 experiment를 實施하기 위하여서는 거창한 大規模의 experiment는 禁物이다. 우선 control이 問題되고 莫大한 量의 資料는 그만큼 相對的으로 많은 extraneous variable을 透引시킬 可能性이 많은 것이다. 따라서 小規模이고 control이 徹底하게 되고 一時에 單하나의 variable의 効果만을 다루며 어떠한 特定의 方法에만 얽매이지 않는 FL 敎育의 總體的인 次元에서 合理性을 찾아 experiment를 實施한다는 것이 곧 効果的인 方法이라고 할 수 있다[18].

Ⅳ. Research Design에 適合한 分野別 目錄

本稿의 序頭에서 말했듯이 language teaching research는 不振狀態에 놓여 있고 따라서 그 對象分野가 廣汎할것은 말할 것도 없으며 그 目錄을 作成한다는 것도 힘겨운 일이다. 여기에 作成해 본 目錄은 보다더 細部的이고 技術的인 面에 置重해서 본 여러 문제들을 엮어 본 것이다. 이 目錄은 勿論 얼마든지 擴大될 수 있을 것이다. 이와는 對照가 되겠지만 좀더 一般的인 面에서 본 research design을 할 수 있는 分野에 關해서는 비록 目錄型式을 取하지는 안했지만 前述 Carroll의 論文 "Research on Teaching Foreign Languages"에 豐富한 解說과 더불어 실려 있다[19].

다음의 目錄은 phonology, syntax, vocabulary, reading, writing, 및 comprehension의 六個分野로 分類하여 作成하였다.

Ⅰ. Phonology

1. Production 演習에 있어서 minimal pair와 random word를 使用한 相對的인 効率性.
2. Recognition 能力에 있어서 "listen only"와 "listen and repeat" presentation을 提供한 경우의 相對的인 効率性.
3. Phonetic symbol과 通常 spelling symbol만을 使用해서 가르친 發音能力의 比較
4. Articulatory diagram의 使用과 口頭說明에만 依存한 發音能率의 比較.
5. Production 能力에 있어서 teacher correction과 錄音器 거울 等 device를 利用한 self-correction의 相對的인 効果.

---

[18] Lado, *Language Testing*, p. 378 參照.
[19] Gage, ed., *Handbook*, pp. 1061-100.

6. Closed-circuit TV 資料의 使用과 通常 放送資料의 使用의 發音能力 培養에 있어서의 相對的인 効果.

7. 敎師로서 言語學的 素養이 없는 native speaker 의 使用과 言語學的 素養이 있는 non-native speaker 의 使用이 發音 能力 培養에 있어서 차지하는 相對的인 効率性.

8. Pitch 와 juncture drill 에 있어서 contour line 의 使用과 gesture cue[20]의 使用이 차지하는 相對的인 効率性.

9. Stress drill 에 있어서 stress symbol 의 使用과 gesture cue 의 使用이 차지하는 相對的인 効率性.

10. Intonation drill 에 있어서 repetition 對 repetition plus response 의 相對的인 効果.

11. 發音敎育에 있어서 敎師의 個別的 指導와 少集團別 指導間의 相對的인 効率性

12. 錄音敎材에 있어서 男性對 女性 목소리의 使用이 차지하는 相對的인 効果.

13. Drill 時間의 native drill instructor 로서 男性 對 女性의 相對的인 効率性.

14. 學生의 NL sounds 의 transfer 와 fresh acquisition 間의 相對的인 効果.

15. 그림 또는 實際物件等을 使用한 視覺補助敎材의 使用과 發音單語의 口述說明 依存間의 相對的인 効果.

16. 即刻的인 feedback 을 提供한 repetition 과 遲延된 feedback 을 提供한 repetition 間의 相對的인 効率性.

17. Positive feedback 과 negative feedback 을 提供한 경우의 相對的인 習得能率.

18. 發音矯正에 있어서 賞罰制度의 相對的인 効果[21].

19. 發音時間에 學生의 NL 使用을 許容치 않는 경우와 FL 과 NL 을 兼用케 하는 경우의 相對的인 効率性.

20. 演習問題를 繼續的으로 부여하는 경우와 間隔을 두고 부여하는 경우와의 相對的인 効果.

21. 發音能力 培養에 있어서 아침공부와 저녁공부間의 相對的인 効果.

22. 錄音을 通하여 敎材錄音과 比較演習을 한 경우와 "listen and repeat"만 演習한 경우와의 production 能力에 있어서의 相對的인 効率性.

23. Cluster 와 non-cluster 單語間의 production 習得 能率의 比較.

24. 單音節 context 와 多音節 sentence 또는 phrase 에 있어서의 production drill 上의 相對的인 効率性.

25. 正常的인 速度로 production 을 가르치는 경우와 느린 速度로 articulation 을 가르치

---

[20] 例를 들면 턱을 내민다든지 눈섭을 추겨 올린다.

[21] 카아드 回收制라든지 簡單한 身體的인 苦役을 수반하는 罰方法을 利用할수 있을 것이다.

는 경우와의 production 能力에 있어서의 相對的인 効率性.

26. 一名의 native speaker 와 多數의 native speaker 에 접촉을 시킨 경우의 相對的인 發音能力 培養効果.

27. 學級內의 年齡層別 發音能力 培養의 比較.

28. Monolingual 한 學生과 bilingual (第三國語)한 學生間의 發音能力 比較.

29. 專攻分野가 다른 學生들의 發音習得 能率의 比較 (例 : 心理學 專攻學生 對 數學專攻 學生 等).

30. 發音이 NL 의 音構造에 비추어 어려운 部分을 먼저 가르치는 경우와 쉬운 部分을 먼저 가르치는 경우의 相對的인 効果.

Ⅱ. Syntax

1. 構文演習에 있어서 substitution drill 과 repetition drill 間의 相對的인 効率性.

2. Expansion drill 과 complete structure drill 間의 相對的인 効率性.

3. 構文演習에 있어서 transformation drill 과 repetition drill 間의 相對的인 効率性.

4. Response drill 과 repetition drill 間의 相對的인 効率性.

5. 前記 各種 drill 의 combination 型과 repetition drill 間의 相對的인 効率性.

6. 構文演習에 있어서 pantomime 과 視覺補助 敎材의 使用間의 相對的인 効率性.

7. 構文敎育에 있어서 "Teacher" 및 "Study" presentation 間의 相對的인 効率性 (이것은 oral vs. written approach 와 同等히 다룰수 있음.)

8. FL 의 structure 를 가르침에 있어서 inductive approach 와 deductive approach 와의 相對的인 能率.

9. 學生의 NL 의 structural pattern 과 對照說明을 하는 경우와 直接 FL 의 pattern 을 NL 과 關聯없이 가르치는 경우의 相對的인 能率.

10. Structure drill 에 있어서 native speaker 를 使用한 경우와 使用하지 않은 경우의 相對的인 能率.

11. 敎師가 學生의 NL 을 使用하여 說明하는 경우와 FL 을 使用하여 說明하는 경우의 相對的인 構文習得 能率.

12. 構文敎育에 있어서 Reed and Kellogg diagram, tree diagram, 및 IC cuts 等 各 diagram 을 使用하는 경우의 相對的인 能率.

13. 構文演習에 있어서 賞罰制度를 採擇했을 경우의 相對的인 効果.

14. Monolingual 한 學生과 bilingual(第三國語)한 學生의 構文習得 能力의 比較.

15. 專攻을 달리하는 學生間의 構文習得 能力의 比較 (例 : 數學專攻學生 對 政治學 專攻學生).

16. Pattern practice 에 있어서 錄音된 敎材를 반복하여 listen 만 하는 경우와 同敎材內

容에 對하여 active recall과 practice를 하는 경우의 相對的인 効率性.

17. 一個原則에 對한 例만을 提示하는 경우와 例를 섞어가며 對照的인 原則과 比較說明하는 경우의 相對的인 効率性.

18. NL의 structure에 關하여 造詣가 없는 學生間의 構文習得 能率 比較.

19. 文法的인 專門用語를 使用한 경우와 使用하지 않은 경우의 相對的인 習得 能率.

20. 各 grammatical category에 關한 몇가지의 對照的인 feature를 說明해주는 경우와 各 category의 一部面만을 그때 그때 說明해주는 경우의 相對的인 習得能率.

21. NL과는 過激하게 다른 FL의 grammatical feature를 먼저 다루는 경우와 NL과 비슷한 feature를 먼저 다루는 경우와의 相對的인 習得能率.

Ⅲ. Vocabulary

1. "Teacher" presentation 對 "Study" presentation 間의 相對的인 習得能率.

2. Flash card와 그림과 같은 補助敎材를 使用하는 경우와 aural-oral presentation만을 하는 경우의 相對的인 習得能率.

3. NL로의 번역을 수반하는 경우와 번역없이 FL을 使用하여 說明을 하는 경우의 相對的인 習得能率.

4. 繼續的인 學習과 controlled interrupted study에 依한 vocabulary 學習에 있어서의 相對的인 効率性.

5. 아침공부와 저녁공부간의 相對的인 vocabulary 習得能率.

6. Referent의 모양이나 機能을 pantomime을 써서 說明하는 경우와 口頭說明만 하는 경우의 相對的인 習得能率.

7. 年齡層別 vocabulary 習得能力의 比較.

8. 單音節 語彙와 多音節 語彙間의 相對的인 習得能率.

9. Closed-circuit TV를 利用한 資料와 通常放送을 利用한 資料의 相對的인 習得能率.

10. 相互關聯이 있는 單語群과 關聯이 없는單語群間의 相對的인 習得能率.

11. 學生의 知識과 經驗에 견주어 親近한 單語와 그렇지 못한 單語와의 相對的인 習得能率.

12. Vocabulary 學習에 있어서 賞罰制度를 採擇했을 경우의 相對的인 効果.

13. Vocabulary 習得에 있어서 etymology의 使用 與否에 따른 相對的인 能率.

14. 難易度 頻度 및 有用度에 있어서 서로 비슷한 單語에 對한 receptive 對 productive skill의 相對的인 培養能率.

15. Monolingual한 學生과 bilingual(第三國語)한 學生間의 vocabulary 習得能力 比較.

16. 專攻分野가 判異한 學生들의 vocabulary 習得 能率比較(例：化學專攻學生 對 歷史學專攻學生).

17. 教師가 prompting을 爲主로 하는 경우와 confirmation을 爲主로 하는 경우의 相對的인 vocabulary 習得 能率.

18. List를 爲主로 한 vocabulary 教育과 context를 爲主로 한 경우의 相對的인 習得 能率.

IV. Reading

1. Loud reading과 silent reading 間의 相對的인 効率性.

2. Oral reading에 있어서 intonation을 使用하는 경우와 使用하지 않는 경우 의相對的인 効率性.

3. 速度가 빠르고 반복되는 reading과 速度가 느리고 반복되지 않는 reading 間의 相對的인 効率性.

4. Reading을 해가며 note를 하는 경우와 本文에 underline을 하는 경우의 相對的인 効率性.

5. Reading에 있어서 專門的인 資料와 非專門的인 資料와의 相對的인 能率(例 : 專攻書籍 對 小說 또는 隨筆).

6. 監督下에 實施하는 reading과 監督이 없이 實施하는 reading 間의 相對的인 効率性.

7. Sentence 單位의 reading과 paragraph 單位의 reading의 相對的인 効率性.

8. Phonetic symbol 利用의 有無에 따른 reading의 相對的인 能率.

9. 資料를 NL로번역해가며 읽는 경우와 FL로 直接 理解해가며 읽는 경우의 相對的인 効率性.[22]

10. Sentence의 길이가 긴 資料와 짧은 資料의 相對的인 reading 能率.

11. Monolingual 한 學生과 bilingual (第三國語)한 學生의 reading 能率 比較.

12. 專攻分野가 判異한 學生間의 reading 能率 比較 (例:數學專攻學生 對 音樂專攻 學生)

13. 教室과 같은 零圍氣에서의 reading과 休憩室과 같은 雰圍氣에서의 reading의 相對的인 能率.

14. 아침과 저녁의 reading 能率比較.

15. 學生이 어느 程度의 audio-lingual skill을 取得한 然後에 가르치는reading과 audio-lingual skill을 取得하기 前에 가르치는 reading과의 相對的인 効率性.

16. 列間 또는 反對페이지에 번역이 되어있는 資料와 번역이 되어있지 않은 資料의 相對的인 能率.

17. 本文을 事前에 暗記하는 경우와 scanning만 하는 경우의 相對的인 reading 能率.

18. 本文을 事前에 聽取하는 경우와 直接읽는 경우의 相對的인 能率.

---

[22] 이런 實驗은 勿論 中級 以上 水準의 學生을 對象으로 한 것이다.

V. Writing

1. 學生의 NL로부터 FL로 번역을 해가며 writing하는 경우와 FL로 直接 writing하는 경우의 相對的인 効率性.

2. Sentence 單位로 進行시키는 writing과 paragraph 單位로 進行시키는 writing間의 相對的인 能率.

3. 筆記演習에 있어서 字劃單位로 演習하는 경우와 字單位로 演習하는 경우의 相對的인 効率性. (非 Latin alphabet의 경우에 해당함)

4. 作文의 矯正에 있어서 敎師가 하는 경우와 學生各自가 self-correction하는 경우의 相對的인 作文習得 能率.

5. Composition 敎育에 있어서programed approach를 使用하는 경우와 random approach를 利用하는 경우의 相對的인 効率性.

6. Writing에 있어서 bilingual dictionary를 使用하는 경우와 monolingual dictionary를 使用하는 경우의 相對的인 効率性.

7. Re-write를 시키는 경우와 spot correction만을 하는 경우의 相對的인 能率.

8. Blank-filling techique를 使用할 경우와 使用하지 않을 경우의 相對的인 能率.

9. FL의 grapheme에 對한 筆記演習에서 flowing style과 printed style의 相對的인 習得能率.

10. 客觀的인 素材와 主觀的인 素材에 對한 相對的인 writing能率.

11. Spelling 敎育에 있어서 "Teacher" presentation과 "Study" presentation間의 相對的인 習得 能率.

12. Spelling에 있어서 phonetic symbol을 使用하는 경우와 使用하지 않는 경우의 相對的인 能率.

13. Monolingual한 學生과 bilingual(第三國語)한 學生의 writing能率 比較.

14. 專攻分野가 判異한 學生間의 相對的인 writing能率 (例：物理學專攻學生 對 歷史學專攻學生).

VI. Comprehension

1. 주어진 資料에 對한 aural comprehension과 visual comprehension의 相對的인 能率.

2. 專門的인 資料와 非專門的인 資料의 相對的인 comprehension能率 (例：專攻書籍 對 小說 또는 隨筆等).

3. 圖解가 붙은 資料와 붙지 않은 資料의 相對的인 comprehension能率.

4. 錄音된 資料와 口頭資料의 相對的인 comprehension能率.

5. 口頭로 提供하는 資料에 있어서 男女音聲의 相對的인 comprehension能率.

6. 아침과 저녁의 相對的인 comprehension能率.

7. 年齡層別 comprehension 能率 比較.

8. Oral reading 과 mental reading 사이의 相對的인 comprehension 能率.

9. 敎室과 같은 雰圍氣와 休憩室과 같은 雰圍氣에서 實施하는 comprehension 의 相對的인 能率.

10. Closed circuit TV 資料와 通常放送資料에 對한 相對的인 comprehension 能率.

11. 資料를 口頭로 傳達하고 筆記로 答을 쓰게 하는 경우와 資料를 筆記로 傳達하고 口頭로 答을 하게하는 경우의 相對的인 comprehension 能率.

12. 對話體로 된 資料와 對話體가 아닌 文으로 構成된 資料에 對한 相對的인 comprehension 能率.

13. Monolingual 한 學生과 bilingual (第三國語)한 學生의 comprehension 能率 比較.

14. 專攻分野가 判異한 學生間의 comprehension 能率 比較 (例 : 音樂專攻學生 對 經濟學專攻學生 等)

15. 同一한 資料를 빠른 速度로 반복해서 읽어주는 경우와 徐徐히 반복없이 읽어주는 경우의 相對的인 comprension 能率.

前記의 目錄에는 主로 相互間 다른 approach 와 技術을 相對的으로 다룬 것이지만 勿論 이 中에는 "有無"의 對立關係에 依한 새로운 一連의 experiment 도 可能한 것이다. 例를 들면 即刻的인 feedback 對 遲延된 feedback 과의 對立을 feedback 을 提供하는 경우와 提供하지 않는 경우와 같은 "有無" 對立關係에 依한 experiment 가 또한 可能하게 되는 것이다. 다만 本稿에서는 그러한 "有無"對立關係의 項目은 簡便을 위해서 되도록 縮少시켜 놓았다. 前記目錄은 決코 包括的이고 exhaustive 한 것이 될 수 없을 것은 말할 것도 없거니와 筆者의 보잘것 없는 經驗과 perspective 를 土台로 하나의 可能性만을 提示하는 데 그쳤다고 하는 것이 옳을 것이다.

V. 結　　言

오늘날 言語敎育에 있어서는 아래와 같은 네 가지의 뚜렷한 傾向을 들 수가 있다.

1. Audio-lingual approach 의 置重.
2. Pattern practice 를 通한 overlearning.
3. Contrastive structure analysis 의 活用.
4. NL 使用의 抑制.

이와 같은 傾向은 오늘날 言語敎育에 이바지 하고 있는 者라면 누구나 그 効率性을 疑心치 않을 程度로 널리 認識되어 있다. 그러나 이것은 主로 言語實驗室이나 敎室에서 이루어진 實際的인 經驗과 intuition 에 依據하여 作成된 報告나 文獻에 依하여 그 優秀性이 認定되어 온 것이지 決코 徹底한 empirical research 에 依據해서 그 効率性이 立證된 것이라

고는 볼 수 없는 것이다[23]. Carroll의 말을 빌린다면 NL使用의 抑制만 하더라도 이를 뒷받침해 줄 何等의 experiment도 實施된 적이 없고 pattern practice에 對하여도 文法敎育에 있어서 其他方法과 對照해 봐서 이를 勸奬할 만한 이렇다 할 empirical research가 없을뿐더러 어떠한 特定의 drill에 對해서도 그 成果를 左右하는 variable이 果然 무엇인가조차도 밝혀낸 일이 없다고 말하고 있다[24].

이러한 根本的인 問題에 있어서도 實態가 이렇거늘 보다더 細部的이고 技術的인 面에서 볼 적에는 言語敎育을 위한 research라는 것은 實로 不毛狀態에 있다고 할 수 있다. 더구나 우리나라의 경우에는 무턱대고 주먹구구식의 模倣이나 일삼았지 言語敎育 分野에 對한 政策樹立 乃至는 curriculum編成에 寄與할 만한 本格的인 research가 行하여지지 못하고 있는 것이 안타까운 實情이다. 近來 國內大學에서는 language laboratory가 必須 不可缺의 要素처럼 되었고 一部에서는 closed-circuit TV까지 備置해놓고 있는데 이 亦是 Carroll의 말을 빌린다면 이러한 補助施設을 利用함으로서 言語敎育의 能率이 어떻게 달라지는 데 對해서도 何等의 validity와 reliability를 兼備한 實驗分析이 報告된 바 없는 것이다[25]. 勿論 그 동안 이 方面에 있어서의 實驗과 硏究가 없었다는 것은 아니다. 다만 그러한 實驗과 硏究들이 大部分의 경우 control문제를 제대로 다루지 못했을 뿐더러 valid한 research methodology를 適用하지 못했다는 非難을 免할 길이 없다는 것이다.

本稿에서는 言語敎育에 臨하고 있는 敎師들이 實踐할 수 있는 小規模의 experiment에 適合한 research design을 micro-teaching과 single-class teaching部面別로 frame을 짜서 紹介하였다. 그리고 이에 따른 validity와 reliability의 評價문제, 그리고 몇가지의 統計學的인 面에서의 評價를 아울러 例示하였다. 그리고 variable을 control하기 위한 留意事項을 提示하였고, 現在 要望되는 research design이 可能한 分野에 對한 目錄을 作成 添付하였다. 本稿의 趣旨는 緖言에서 말한 바와 같이 theory와 performance를 보다더 密接히 連關시켜줄 수 있는 research design의 문제를 提示하는 데 있는바 그렇다 해서 專門的인 言語敎育 research specialist를 對象으로 한 것은 아니다. 그러한 專門家가 우리나라에도 많이 活躍을 해주어야 되겠지만 分野自體가 廣汎하고 開放的이며, 또한 國內外를 莫論하고 그 活動이 不振狀態에 있으므로, 이를 敎師들 스스로가 part-time research project로서 實驗硏究함으로서 現在와 같은 資料의 窮乏을 部分的으로 나마 是正할 수 있지 않을까 하는 筆者의 心情을 밝히며 이 글을 맺는다.

---

[23] Gage, ed., *Handbook*, pp. 1062 f.
[24] *Ibid.*, pp. 1073 ff.
[25] *Ibid.*, p. 1080.

參 考 文 獻

Agard, Frederick B., and Harold B. Dunkel. *An Investigation of Second-Language Teaching.* Boston: Ginn, 1948.

Campbell, Donald T., and Julian C. Stanley. "Experimental and Quasi-experimental Designs for Research on Teaching," Gage, N.L., ed. *Handbook of Research on Teaching.* Chicago:Rand McNally Company, 1963, pp. 171—246.

Carroll, John B. "Research on Teaching Foreign Languages," Gage, ed. *Handbook.* pp. —1100.

Dykstra, Gerald. "Spiraling Co-consideration," unpublished manuscript, Teachers College, Columbia University, 1966.

Elzey, Freeman F. *A Programed Introduction to Statistics.* Belmont, California: Wadsworth Publishing Company, 1966.

Good, C.V., ed. *Dictionary of Education.* New York: McGraw-Hill, 1959.

Harris, Chester W., ed. *Encyclopedia of Educational Research,* New York: Macmillan, 1960.

Lado, Robert. *Language Testing, The Construction and Use of Foreign Language Tests.* New York: McGraw-Hill, 1964.

Language Learning: The Individual and the Process. IJAL Vol. 32. No. 1 (Jan. 1966), Part II. Lyman, Howard B. *Test Scores and What They Mean.* Englewood Cliffs, New Jersey: Prentice-Hall, 1963.

McCollough, Celeste, and Loche Van Atta. *Statistical Concepts: A Program for Self-instruction.* New York: McGraw-Hill, 1963.

Paterno, Adelaida. "Foreign Language Testing," Allen, Harold B., ed. *Teaching English as a Second Language.* New York: McGraw-Hill, 1965, pp. 373—392.

Research Design an dthe Teaching of English- Proceedings of the San Francisco Conference. NCTE, 1963.

Research on Language Teaching- An Annotated International Bibliography for 1945~61 Seattle, Washington: The University of Washington Press, 1962.

Snedecor, George W. *Statistical Methods.* Ames, Iowa: The Iowa State College Press, 1956.