

## 심리학 연구에서 언어자료를 어떻게 활용할 것인가\*

정혜선†

한림대학교 심리학과

정보의 인출부터 사고과정, 내담자의 심리적 역동까지 다양한 심리과정이 말 또는 글의 형태로 드러나지만, 언어반응의 내용을 분석할 때 발생하는 해석의 주관성으로 말미암아 언어자료는 심리학 연구에 제한적으로만 사용되었다. 언어반응을 분석하는 데 따르는 문제점을 극복하고 언어자료를 과학적 심리학 연구 내에서 활용하려는 다양한 시도가 이루어져 왔는데, 대표적인 예로 인지연구에서 개발된 프로토콜 분석 및 언어자료분석을 들 수 있다. 이들 분석 기법은 글 또는 발화 자료를 체계적, 객관적으로 수집하고, 분석하는 방법으로, 인지발달, 문제해결, 전문성, 학습 등 다양한 심리학 영역에서 마음의 작동기제를 밝히는 데 공헌해 왔다. 본 논문의 목적은 언어자료분석을 중심으로 글 또는 구어의 형태로 수집되는 언어자료를 심리학 연구에서 체계적으로 사용하는 방법을 안내하는 데 있다. 이를 위하여 언어자료분석의 주요 목표, 특징과 함께 주요 연구 사례를 살펴보고, 또한 방법을 단순히 소개하는 데서 한발 더 나아가 실제 연구에 적용할 수 있도록 분석의 절차를 제시하고 고려할 점을 연구 단계별로 제시하였다. 마지막으로 언어자료를 분석하는 데 사용되는 유사 방법들과의 공통점 및 차이점을 논의하고 방법의 장단점을 살펴보았다.

주요어 : 언어자료분석, 언어자료분석, 프로토콜 분석, 언어보고, 자료, 양적분석

\* 본 논문은 2012년도 한국연구재단(NRF-2012R1A1A3015115) 및 한림대학교 교비연구비(HRF-201211-004)의 지원을 받아 연구되었음.

† 교신저자: 정혜선, 한림대학교 심리학과, 강원도 춘천시 한림대학길 1 (200-702).

E-mail: heis@hallym.ac.kr.

## 배 경

심리과정은 대부분 개인 내적으로 일어나는 과정으로 밖에서 손쉽게 관찰되지 않는다. 과학적 심리학의 도전 중의 하나는 직접적 관찰이 불가능한 심리현상에 대해서 이를 객관적으로 측정하고 연구할 수 있는 방법을 고안하는 것이었다. 현대 심리학에서 주로 사용하는 방법은 설문지를 통해서 주관적인 심리상태에 대한 질문을 하거나 또는 다양한 실험 및 자극 상황을 고안하여 배후의 심리과정과 결과물이 드러나도록 하는 것이다. 이 과정에서 척도화된 반응뿐만 아니라 반응시간, 안구운동 패턴 같은 다양한 행동지표가 종속 측정치로 사용되었다. 구어 또는 글로 표현되는 언어반응 또한 내적인 심리과정을 드러내는 중요한 행동지표이다. 언어는 정보처리와 인지과정을 이해하는 데 매우 중요한데, 기억 내용을 인출하거나 문제를 해결하는 것 같은 정보처리 과정은 언어적으로 주로 표현될 뿐만 아니라 언어적인 요인에 의해 영향을 받는다(조명한 외, 2003). 또한 언어는 복잡한 지식구조나 문제 해결의 과정을 드러내는 데 유용한 방법일 뿐만 아니라, 사회적 상호작용을 매개하는 중요한 수단으로 관계 형성, 소집단 상호작용 등의 현상을 이해하는 데 핵심적인 도구이다.

심리과정을 이해하는 데 있어 언어가 차지하는 중요성에도 불구하고 그동안 심리학 연구에서 언어자료의 활용은 제한적이었다. 언어반응이 사용되기는 했지만 대부분 구조화된 언어반응(예, “네”, “아니오” 반응, 기억 항목, 단어의 인출)이었다. 비구조화된 언어반응(예, 일기)이 사용되는 경우도 특정 단어(예, “우리” 또는 “나” 등의 대명사)의 발생 빈도에 대한

기계적인 분석이 주를 이루었다(예, Pennebaker, Mehl, & Niechofer, 2003). 비구조화된 언어반응의 내용에 대한 분석은 잘 이루어지지 않았는데, 이는 해석의 주관성으로 말미암아 객관적인 분석을 실시하는 것이 어려운 데 기인하였다. 이러한 문제점을 극복하기 위한 다양한 분석 기법이 개발되었는데, 본 논문의 목적은 이러한 분석방법 중의 하나인 언어자료분석(verbal data analysis)을 소개하고 그 사용을 안내하는 것이다. 논문의 주요 구성은 다음과 같다. 첫째, 심리학에서 언어자료가 사용된 역사 및 주요 기법들(예, 프로토콜 분석, 언어자료분석)을 살펴본다. 둘째, 언어자료분석이 사용된 주요 연구 사례를 살펴본다. 셋째, 연구단계별로 분석의 주요 절차 및 고려점을 살펴본다. 마지막으로 언어자료분석과 유사한 방법들과의 차이점을 살펴보고 언어자료 분석의 방법의 장단점을 논의하였다.

## 언어자료분석의 배경 및 특징

심리학이 과학으로 정립되던 초기 가장 먼저 사용된 연구방법 중의 하나는 연구 참여자들에게 자신의 심리과정을 내성(introspection)하게 하여 이를 언어적으로 보고하도록 하는 것이었다. 그러나 내성법은 보고내용의 주관성으로 말미암아 심각한 비판에 직면하였다. 행동주의자들은 내성으로 얻어진 언어보고가 과학적으로 타당한 자료(data)로 간주될 수 없다고 보았고, 대신 객관적으로 관찰 가능한 행동이 연구대상이 되어야 한다고 주장하였다. 내성법에 대한 비판은 언어보고 전반에 대한 무차별적인 불신으로 이어졌고, 언어보고나 언어자료는 행동주의의 득세와 함께 심리학

연구에서 한동안 모습을 감추게 되었다(이정모, 2001; Ericsson, 2006; Fox, Ericsson, & Ryan, 2011).

인지주의가 대두한 직후의 인지 연구는 지각, 기억, 언어 등에 집중되었는데 이들 연구에서 주요 사용된 종속 측정치는 반응시간, 반응정확성 등이었다. 언어반응도 종종 사용되었는데, 예를 들어 질문에 대한 답을 “네” “아니오”로 답변하거나, 단어쌍 학습에서는 기억 내용을 언어적으로 인출한 것이 종속 측정치로 사용되었다. 그러나 단순한 답변, 한 두 단어 이상을 넘어가는 긴 발화나 글, 또는 대화의 내용이 종속 측정치로 분석되는 경우는 드물었다. 단어 개수, 어절 수, 통사적 지표 등의 경우 정의가 분명하지만, 언어의 ‘내용’이 분석대상이 되면 그 속성상 어느 한쪽으로 분명하게 정하기 힘든 의미의 주관성 문제가 발생하기 때문이다. 그러나 인지 연구의 주제가 초기의 기억, 지각 문제를 넘어서서 문제 해결, 추론, 이해, 지식표상 등 상위인지과정으로 옮겨가면서 반응시간이나 정반응 같은 기존의 측정치에 의존하는 것의 제한점이 드러나게 되었다. 이 과정들은 시간적 전개가 상대적으로 더딜 뿐만 아니라 시간차이 보다는 표상의 내용, 문제 풀이 순서, 방략의 종류 등이 더 핵심적인 역할을 하기 때문이다. 문제풀이에 사용되는 해결방략(예, 수단-목표분석 방략), 표상의 내용과 구조(예, 심성모형, 지식구조)를 알아볼 수 있는 추가적인 측정치가 필요했는데, 이와 함께 언어반응을 분석하려는 시도가 증가하였다(Newell & Simon, 1972; Simon & Simon, 1978; Vosniadou & Brewer, 1992). 이 과정에서 다양한 분석기법이 개발되었는데, 심리학 내에서 발달된 대표적인 분석방법으로 인지 연구에서 사용된 프로

토콜분석(protocol analysis) 및 언어자료분석(verbal data analysis)을 들 수 있다(Chi, 1997, 2006; Ericsson, 2006; Ericsson & Simon, 1980, 1993). 두 기법은 다소 상이한 연구문제를 해결하려는 시도에서 발달되었는데, 각 분석방법의 주요 특징을 살펴보면 다음과 같다.

#### 프로토콜분석(Protocol analysis)

프로토콜 분석은 Ericsson 과 Simon(1980, 1993)에 의해서 제안되었는데, 이들은 행동주의자들의 비판과는 달리 언어반응이 수집되는 방식에 따라서 언어자료가 인지 및 기타 심리현상의 연구에 타당한, 정당한(legitimate) 자료로 사용되는 것이 가능하다고 보았다. 이들은 언어반응 수집의 두 차원을 구분하였다. 첫째, 언어반응이 과제수행이 이루어지는 동안 동시적으로 수집되었는가 아니면 과제수행이 종료된 후에 수집되었는가에 따라 동시적 보고(concurrent report) 또는 회상적 보고(retrospective report)를 구분하였다. 동시적 보고(예, 신발 끈을 매면서 그 과정을 언어적으로 보고하는 것)와는 달리, 회상적 보고(예, 신발 끈을 맨 후 또는 과거의 신발 끈 맨 경험에 대해서 회상하여 보고하는 것)에서는 언어보고가 실제 일어난 처리 보다는 기억된 처리를 반영할 가능성이 높다. 그 결과 기억 내용을 바탕으로 이차적으로 구성된 내용이 포함되거나, 개인의 지식과 믿음에 따라 보고내용이 왜곡될 가능성이 증가한다(Ericsson & Simon, 1980, 1993; Nisbett & Wilson, 1977). 둘째, 언어자료를 수집하기 위해서 이를 유도하는 지시, 질문(probe)이 종종 사용되는데, 이때 어떠한 질문 또는 지시가 사용되는가에 따라 언어보고의 내용이 영향을 받는다. 이미 참가자의 의식/단

기억에 가용한 정보이거나 주의를 기울이고 있는 정보를 언어적으로 보고하도록 하는 경우(예, 22 곱하기 33 을 암산하는 과정을 소리 내어 말하기), 언어보고의 과정은 이미 진행되고 있는 처리를 크게 변화시키지 않는다. 그러나 참가자가 주의를 기울이지 않고 있거나 생각하지 않고 있는 정보를 질문하면(예, 왜 특정한 방략을 사용했는지) 이들이 자발적으로 하지 않았을 처리를 유도하게 될 수 있다.

Ericsson 과 Simon(1980, 1993)은 언어자료의 수집 방법이 연구의 대상이 되는 정보처리에 영향을 주지 말아야 한다고 보았다. 따라서 회상적인 보고 대신 동시적인 보고를 사용할 것을, 또한 인지과제를 수행하는 동안에 참가자들의 의식에 이미 가용한 정보를 언어화할 것을 강조하였다. 이러한 자료 수집을 위해서 이들은 소리 내어 말하기(Talk Aloud) 또는 소리 내어 생각하기(Think-Aloud) 절차를 제안하였다. 이 절차는 종종 TA 로 줄여서 표기되고, TA 기법을 사용하여 수집된 언어자료를 프로토콜(protocol)이라고 부른다. 이 절차에서는 참가자가 과제를 수행하면서 자연스럽게 마음속에 떠오르는 생각을 소리 내어 말하라고 지시하는데, 예를 들어 문제풀이과정의 연구에 TA 를 사용하여 자료가 수집되는 경우, 참가자들은 문제를 푸는 동안 마음속에서 하고 있는 생각(예, “이번에는 2 에다 3 을 곱해야지”)을 소리 내어 말한다. TA 기법에서는 참가자의 의식에 이미 가용한 정보만 언어화할 것을 강조하기 때문에, 의식에 가용하지 않은 암묵적인 정보를 질문하거나(예, 하위 목표를 사용하였는가?), 추가적인 정보처리를 유도하는 질문(예, 왜 조금 전과 같은 순서로 문제를 풀었는가?) 등은 모두 바람직하지 않은 것으로 간주된다. 프로토콜 분석은 다양한 문제영역 관여

하는 인지 ‘과정(process)’을 밝히는 데 활발하게 사용되었고, 대표적인 영역으로 문제 해결(Kaplan & Simon, 1990; Newell & Simon, 1972; Simon & Simon, 1978), 의사결정(Payne, 1976), 글 읽기와 쓰기(Flower & Hayes, 1981; Pressley & Afflerback, 1995), 전문성(Ericsson, 2000, 2006; Ericsson, Delaney, Weaver, & , Mahadevan, 2004; Wineburg, 1991, 1998), 디자인(Badke-Schaub, Goldschmidt, & Meijer, 2007; Chan, Fu, Schunn, Cagan, Wood, & Kotovsky, 2011; Suwa & Tversky, 1997), 및 사용자경험연구를(Benbunan-Fich, 2001) 들 수 있다.

#### 언어자료분석(Verbal data analysis)

언어자료분석(Verbal Data Analysis) 또는 언어 분석(Verbal Analysis) 또한 인지 연구의 맥락에서 개발된 기법으로 언어자료를 주 분석 대상으로 한다. 프로토콜 분석과 많은 공통점을 지니고 있지만 몇 가지 차이점이 존재한다 (Chi, 1997, 2006). 가장 중요한 차이점은 언어 반응을 수집하는 방법에서의 차이이다. 앞에서 언급하였듯이 프로토콜 분석에서는 언어 반응, 자료가 수집되는 방법을 중시하는데, 무엇보다도 자료수집 방법이 참가자의 인지 과정에 미치는 영향이 최소화되어야 한다. 반면 언어자료분석에서는 언어반응이 이런 비지시적인 방식으로 수집될 것을 요구하지 않는다. 이는 언어자료분석이 인지발달, 전문성, 개념 학습 등 인지과정과 내용에서의 ‘변화’를 탐구하기 위한 목적으로 주로 사용된 데 기인한다. 일찍이 여러 연구자들이 언어반응을 산출하는 과정이 부정적인 방해 효과뿐만 아니라 정적인 학습 효과를 낳는 데 주목하였다 (Nisbett & Wilson, 1977). 예를 들어 자기 설명

(self-explanation) 같은 학습 방법에서는 학습자들에게 공부하면서, 문제 풀이를 하면서 내용을 설명할 것을 주문하는데(예, 왜 특정 공식이 사용되어야 하는지), 이러한 지시의 핵심은 학습자들이 자발적으로는 하지 않았을 처리를 요구하는 것이고, 이는 TA 절차에서는 바람직하지 않다고 간주하는 자료수집 방법이다. 그러나 연구의 목적이 일어나는 그대로의 정보 처리를 연구하는 것이 아니라 자기설명과 같은 지시의 효과를 알아보는 것이 목적일 때 이러한 지시가 사용될 수 있다. 자기 설명 지시는 여러 맥락에서 학습 효과를 높이는 결과를 낳았고, 관련 기제에 대한 연구를 촉발시키게 되었다(Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, de Leeuw, Chiu, & LaVancher, 1994; McNamara, 2004; Renkl, Stark, Guber, & Mandl, 1998; Rittle-Johnson, 2006). 언어자료분석은 이러한 맥락에서 수집된 언어자료를 분석하기 위해 개발되었고, 특히 지식구조화 표상의 변화를 탐색하는데 활발하게 사용되었다. 이런 면에서 언어자료분석은 보다 다양한 자료 수집 상황에 사용이 가능하도록 프로토콜 분석을 다양화하고, 확장한 것으로 볼 수 있다.

Chi(1997)의 언어자료분석법은 따라서 자료 수집보다는 분석방법에 초점을 맞추고 있다. Chi는 분석의 단계를 여덟 단계로 나누고 체계화하였는데, 분석의 핵심은 언어자료의 분석에 수반하는 주관성을 줄이고 분석 결과를 수치로 양화(quantify)하는 것이다. 문제풀이과정 연구에 주로 사용된 프로토콜분석과는 달리, 언어자료분석은 학습자가 가진 지식 표상의 내용과 구조의 특징 및 학습에 따른 변화를 분석하는 데 주로 사용되었다. 대표적인 연구 주제로 발달에 따른 아동의 인지개념의

변화(Chi & Koeske, 1983; Opfer & Siegler, 2004; Vosniadou & Brewer, 1992), 자기설명을 포함한 학습연구(Azevado & Cromley, 2004; Berthold, Nückles, and Renkl, 2007; Chi et al., 1994; McNamara, 2004; Renkl, 1997, 2002), 기타 다양한 맥락에서 일어나는 상호작용 연구에도 활발하게 사용되었다(Jeong & Chi, 2007; Okada & Simon, 1997; Dunbar, 1997; van Aalst, 2009).

## 분석 사례

앞에서 언급하였듯이 다양한 연구 영역에서 언어반응을 수집 분석하였는데, 대표적인 연구 주제로 문제해결(problem solving), 전문성(expertise), 인지발달, 및 학습 연구를 들 수 있다. 각 주제 영역에서 언어자료의 분석이 어떻게 사용되었는지를 살펴보면 다음과 같다.

### 문제해결

Newell 과 Simon(1972)은 문제풀이과정을 연구하기 위해서 “소리 내어 생각하기(Think-Aloud)” 기법을 고안하였다. 초기의 연구에 사용된 과제는 “DONALD + GERALD = ROBERT” 같은 암호산술(cryptoarithmic) 문제였는데, 암호산술문제에서 각 알파벳 낱자는 서로 다른 숫자를 나타낸다. 참가자는 위의 식을 만족시키도록 각 알파벳에 대응하는 숫자를 찾아야 하는데, 문제에 따라 힌트(예, D는 5)가 주어지기도 한다. 이러한 암호산술문제를 시작으로 하노이 탑문제, 식인종과 선교사 문제, 물리학 문제 등 다양한 문제풀이 과정에 대한 연구가 이루어졌는데, 참가자들이 이러한 문제를 얼마나 빨리, 정확하게 푸는가

뿐만 아니라 풀이의 과정, 해결 전략에 연구의 초점이 맞추어졌다(Kaplan & Simon, 1990; Simon & Simon, 1978). 예를 들어 문제해결 과정에서 참가자가 수단-목표분석(means-end analysis) 또는 언덕 오르기(hill climbing) 전략을 사용하는지 등이 연구되었는데, 수단-목표분석은 문제 해결자가 문제의 최종 목표를 염두에 두고 현재 상태와 목표상태의 차이를 체계적으로 줄여나가는 방식으로 문제를 해결하는 전략이고 언덕 오르기는 최종 목표 상태에 대한 장기적인 고려 없이 현재 상태를 조금 더 낫게 하는 방식으로 문제를 해결하는 것을 지칭한다. 참가자가 문제해결과정에서 어떠한 전략을 사용하는지를 알아보기 위해서 문제풀이하는 과정을 소리내어 말하도록 하여 프로토콜을 수집한 다음 참가자가 문제해결동안 언급한 목표상태의 유형 및 언급 시점 등을 바탕으로 사용된 문제해결 전략이 분석되었다(Newell & Simon, 1972).

최근 들어 디자인 및 설계 연구에서 언어자료분석이 활발하게 활용되고 있다(Stempfle & Badke-Schaub, 2002; Chan et al., 2011; Paletz, Schunn, Kim, 2010). 예를 들어 Chan 등(2011)은 사례에 기반한 유추(analogy)가 디자인 문제해결에 미치는 효과를 연구하였는데, 이들은 다양한 사례를 제시하고(예, 현재 문제와 비슷한 또는 상이한 사례, 글 또는 그림으로 제시되는 사례 등) 사례제시가 디자인에 미치는 영향을 알아보았다. 이 연구에서는 언어반응뿐만 아니라 참가자들이 아이디어를 생성하는 과정에서 그린 그림 또한 분석에 포함되었다. 수집된 반응을 분석하여 아이디어의 수, 질, 문제 공간을 탐색한 범위, 사례에 제시된 특징이 최종 디자인 해결책에 반영된 정도 등을 알아보았는데, 연구 결과 제시된 사례가 현재

해결해야 하는 문제와 유추거리(analogical distance)가 먼 것일수록 새로운 아이디어가 더 많이 만들어졌으나 동시에 제안된 아이디어의 수는 감소하였다. 유추거리가 먼(far-field) 사례가 새로운 아이디어 생성에 기여한 것으로 보이나 동시에 처리 부담을 야기한 것으로 보이는데, 먼 사례의 특징들을 처리하고 문제 해결에 반영하는 처리 부담으로 말미암아 아이디어의 수가 줄어든 것으로 보인다. 이들 연구에서 볼 수 있듯이, 언어반응의 분석은 단순히 문제풀이가 맞았다 틀렸다 또는 문제풀이 시간이 얼마나 걸렸는지를 아는 데서 더 나아가서 문제해결 과정 동안의 다양한 처리 방향(예, 사용된 전략, 탐색의 범위) 및 관련 변인의 효과를 보다 정확하게 측정할 수 있도록 해준다.

#### 전문성

체스 마스터, 바둑기사부터 시작해서 의사, 변호사, 운동선수에 이르기까지 경험과 훈련은 초보자에게서는 상상할 수 없는 수행을 가능하게 한다. 전문성의 획득에 영향을 주는 요인이 무엇이고 어떠한 과정을 거쳐 발달하는지 알아보기 위해서 다양한 연구가 실시되었다(Chi et al., 1989; Ericsson, 2000; Krampe & Charness, 2006; Wineburg, 1991, 1998). 초기의 연구들은 전문가와 초보자들 간의 선천적인 능력차이에 대한 가설을 검증하는 데 초점을 맞추었으나, 지능, 작업기억용량 등과 같은 선천적 차이 가설이 기각되면서(Chase & Simon, 1973), 전문가들이 가진 경험과 지식에 초점이 맞추어지게 되었다. 이 과정에서 언어적 보고가 활발하게 사용되었는데, 단순히 전문가와 초보자의 수행차이를 탐색하는 것이 아니라

지식 표상의 구조와 내용에 있어서의 차이를 알아보기 위해서 언어적인 보고에 의존하는 것이 필수적이었다(Chi, 2006; Hoffman, Shadbolt, Burton & Klein, 1995). 언어자료분석은 전문가와 초보자 간에 존재하는 다양한 차이를 밝히는 데 사용되었는데, 예를 들어 Wineburg(1991)는 역사학에서의 전문성이 무엇인가를 알아보기 위해서 역사학자들과 고등학교생들의 사료에 대한 이해를 비교하였다. 그는 참가자들에게 일련의 사료를 질문과 함께 제시하고(예, “어느 그림이 렉싱턴 그린(Lexington Green) 전투에서 일어난 일을 가장 정확하게 묘사하고 있는가?”) 답변 과정을 소리 내어 말하도록 지시하였다. 언어적 언어보고를 분석한 결과, 전문가와 초보자 집단이 역사적인 증거에 대한 추리를 다르게 하고 있었는데, 초보자들의 경우 사료의 표면적인 특징(예, 그림에 나타난 예술적인 기교)에 관심을 기울였으나, 전문가의 경우 서로 다른 사료들이 서로 일관된 해석을 지지하는지에 관심을 보였다. 또한 전문가들은 조건부의 답변을 많이 사용한 반면, 초보자들의 경우 답변에 조건이나 단서를 제시하는 경우가 거의 없었다. 즉, 전문가와 초보자는 해당 영역에 대한 지식의 양뿐만 아니라 추론 방식까지 상이함이 밝혀졌는데, 이렇듯 언어반응을 분석하여 전문가와 초보자 수행 차이의 원인을 보다 분명하게 밝히는 것이 가능하게 되었다.

#### 인지발달

아동의 인지발달 연구에도 언어반응이 종종 사용되었다. 예를 들어, Vosniadou와 Brewer(1992)는 아동의 개념형성에 관심을 기울였는데, 특히 아동이 지구의 모양에 대해서 가지

고 있는 개념이 어떻게 발달하는지를 연구하였다. 이들은 아동에게 일련의 질문에 답하고(예, 지구가 어떻게 생겼는지, 지구 위에 무엇이 있는지 등) 지구의 모습을 그리도록 한 다음, 이들의 답변을 바탕으로 아동이 가진 지구에 대한 표상, 심성모형(mental model)을 분석하였다. 그 결과 어린 아동들의 경우 부모와 교사에게서 ‘지구는 둥글다’라는 정보를 접하고 있었으나 여전히 지구의 모양이 구 모양(sphere)이라는 것을 이해하는 데 어려움을 겪고 있음을 발견하였다. 이들 아동들이 이해하고 있는 지구의 모형은 부모나 교사로부터 배운 지식과 자신들이 직관적으로 형성한 평평한 지구의 모습을 결합한 형태의 변형된 모형이었다(예, 둥근 구 안에 평평한 원판이 들어간 모습). 이들의 연구 결과는 아동의 개념 발달과정이 학교 또는 부모로부터 일방적으로 지식을 전달받는 과정이 아님을, 그보다는 아동이 자신의 일상경험을 통해 구성한 모형과 학교나 부모로부터 배우는 지식을 통합하는 과정이고, 따라서 발달과 학습에서 아동이 가지고 있는 선행 지식(pre-existing knowledge)이 적극적으로 고려되어야 함을 시사하였다.

Siegler(1995)는 아동을 대상으로 수와 관련된 보존(conservation) 개념의 획득을 연구하였는데, 특히 이 과정에서 피드백 제공의 효과를 알아 보았다. 세 실험 조건에서의 수행이 비교되었는데, 아동이 문제를 푼 다음에 피드백만 받는 조건(feedback only), 피드백과 함께 자신의 추리를 설명하는(feedback plus explain-own reasoning) 조건, 피드백과 함께 실험자의 추리를 설명하는(feedback plus explain-experimenter's reasoning) 조건이 존재하였다. 보존과제에서의 수행 정확성은 실험자의 추리를 설명하는 마지막 조건에서 가장 우수하였는데, 실험조건

간 차이의 원인을 추가적으로 탐색하는 데 언어자료분석이 사용되었다. 아동이 보존과제를 수행하면서 제공하는 설명을 분석하여 아동의 설명이 조건에 따라서 어떻게 변화함을 살펴 보았는데, 그 결과 단순한 길이 등 물리적 특징에 기반한 설명이 아니라 조작(operation)에 기반한 설명이 증가하는 것이 뚜렷하게 관찰되었다. 이들 연구 사례에서 볼 수 있듯이 아동의 인지발달을 연구하는 데도 언어자료분석이 효과적으로 사용되었는데, 아동의 언어반응을 분석하여 이들이 형성한 표상의 내용과 발달 경로를 알아보는 것이 가능하다.

## 학습

Chi 등(1989)은 학습 성공에 관여하는 요인이 무엇인지 알아보기 위해서 학생들에게 물리학 문제를 풀게하고 학습 정도에 따라 학습이 우수한 집단과 저조한 집단으로 나누어 이들의 학습방법과 문제풀이 방식을 비교하였다. 문제풀이 과정을 소리내어 말하도록 하여 이들이 어떻게 문제풀이를 하면서 학습하는지를 분석한 결과 학생들이 예시(worked-out examples)를 얼마나 잘 설명하는가에 따라 학습 효과가 달라짐을 발견하였다. 학습이 우수한 집단의 경우 학습하는 동안 상당한 양의 자기설명(self-explanation)을 생성하였으며, 문제 해결책이 적용되는 조건에 대한 이해가 뛰어났다. 반면 학습이 저조한 집단의 경우 학습하는 동안 자기설명을 적게 생성하였으며, 제시된 예에만 의존해서 문제를 푸는 경향이 존재하였다. 후속 연구에서는 자기설명효과를 실험적으로 검증하기 위해서 실험집단에게는 학습하는 동안 자기설명을 하도록 촉진(prompt)하고 통제집단과 비교하였는데, 자기설

명 조건에서 학습하는 동안에 더 많은 추론을 생성하였고(Chi et al., 1994; McNamara, 2004), 학습 전이(transfer)도 촉진되었다(Rittle-Johnes, 2006).

문제풀이, 학습과정 동안의 언어적 보고뿐만 아니라 평가의 목적으로 수집되는 언어반응(예, 질문에 대한 답변) 또는 발화가 아닌 글로 적은 답변도 분석에 사용될 수 있다. 예를 들어 Coleman, Brown, 및 Rivkin(1997)은 설명하기와 요약하기라는 학습활동의 효과를 세 조건 하에서 연구하였는데, 첫 번째 조건에서 학생들은 자신이 학습한 내용을 스스로에게 설명 또는 요약하였고, 두 번째 조건에서는 읽은 내용을 동료에게 설명 또는 요약하였고, 마지막 조건에서는 동료의 설명 또는 요약을 듣는, 총 여섯 개의 학습조건이 존재하였다. 학생들은 진화에 대한 글을 학습한 다음 그 내용을 요약 또는 설명하는 글을 작성하고 일련의 전이(transfer) 문제(예, 다윈(Darwin)과 월레스(Wallace)의 이론을 사용하여 어떻게 북극곰이 하얗게 변했는지에 답하기)에 답하였다. 요약과 설명에 언급된 아이디어의 수와 질, 그리고 아이디어가 통합된 정도를 분석한 결과 학습할 때 요약보다는 설명하는 것이, 그리고 듣기보다는 자신과 다른 사람에게 설명을 하는 것이 더 효과적이었다. 전이문제에 대한 답변에서도 조건 간에 유사한 차이가 관찰되었다. 이들 연구에서는 학습결과를 알아보기 위해서 객관식 주관식 시험에서 몇 점을 받았는가를 측정하기 보다는 학생들의 설명/요약과 질문에 대한 답변에 학습자료에서 제시된 핵심 내용이 얼마나 많이, 그리고 정확하게 반영이 되었는지를 분석하였는데, 이러한 분석을 통하여 학습에 대한 보다 정교한 측정이 가능해 지고, 수행차이가 어떠한 요인



에 기인하는지에 대한 이해가 가능하게 되었다.

### 언어자료의 수집 및 분석

앞에서 살펴보았듯이 언어자료분석은 심리학의 다양한 영역에서 심리현상과 기제에 대한 유용한 정보를 제공하는 데 공헌하였다. 이러한 유용성에도 불구하고 심리학 내에서 프로토콜 분석이나 언어자료분석에 대한 문헌은 그리 흔하지 않은데, Ericsson 과 Simon(1980, 1993)은 언어자료의 타당화에 강조를 두었고, 분석방법에 대한 실용적인 지침을 제공하지는 못했다. Chi(1997)은 분석의 단계를 8 단계로 나누는 등 보다 실용적인 지침을 제시하였으나, 프로토콜 분석과의 차별화를 강조하여, 두 연구 방법을 통합적으로 사용하기 위한 적절한 지침을 제공하는 데는 부족하였다. 본 논문에서는 양적인 방법론으로서 프로토콜 분석과 언어자료분석의 공통점에 주목하고, 이들을 통합하여 발화 및 글의 형태로 수집된 언어자료를 연구에 활용하기 위한 실용적인 지침을 제공하고자 하였다. 언어자료가 연구에 사용되는 경우에도 일반적인 심리학의 연구에서 고려되는 연구문제 설정, 자료 수집, 분석, 통계 처리 및 해석에 관여하는 모든 기준들이 마찬가지로 적용된다. 본 논문에서는 언어자료를 수집하고 분석하는 경우 추가적으로 고려하고 염두에 두어야 할 사항을 중심으로 다음의 여섯 단계로 나누어 언어분석의 절차를 기술하였다. 첫째, 연구주제 및 설계, 둘째, 언어반응의 유도, 셋째, 언어자료의 수집, 넷째, 축어록 작성 및 자료의 선택/표집, 다섯째, 분석, 여섯째, 통계분석 및 결과보고. 편의상 위

와 같이 연구의 단계를 구분하였으나, 각 단계들은 모두 유기적으로 연결되어 있고, 연구문제와 연구맥락에 따라서 각 단계가 반복, 생략, 확장될 수 있다. 기술의 편의를 위해서 언어자료분석을 중심으로 기술하였으나, 프로토콜 분석과 구분이 필요한 경우에는 따로 언급하였다.

#### 연구주제 및 설계

언어자료분석은 앞에서 예를 든 문제풀이, 전문성, 인지발달 등의 연구주제뿐만 아니라 기타 언어과정에 의해 매개되고 영향 받는 다양한 인지 관련 현상과 기제, 특히 상위인지 과정을 연구하는 데 유용한 연구방법이다. 인지 연구가 아니라도 언어반응에 그 기제가 드러나는 다양한 심리, 사회 및 정서 과정의 연구에도 사용 가능한데, 예를 들면 소집단 상호작용 연구, 인터넷 상의 활동을 연구하는데 활용될 수 있다(Okada & Simon, 1997; Paletz et al., 2011). 그렇다고 모든 연구문제에 언어자료분석을 사용하는 것이 적절하지는 않는데, 연구주제와 관련해서는 다음 두 요인이 고려되어야 한다. 첫째, 연구문제에서 다루는 심리 과정이 언어 또는 언어관련 반응에 포착될 수 있어야 한다. 예를 들어, 지각 과정, 특히 초기 지각과정은 비언어적일 뿐만 아니라 매우 빠르게 일어나서 언어적으로 보고하거나 기술하는 것이 거의 불가능하고, 시도한다고 해도 얻어진 자료를 신뢰하기 힘들다(Nisbett & Wilson, 1977). 그렇다고 언어자료분석이 반드시 의식적인, 내성이 가능한 인지과정을 연구하는 데만 적합한 것은 아니다. 의식되지 않는 정보처리라고 하더라도 언어행동을 통해서 드러날 수 있는데, 예를 들면 인지부하가 존

재할 때 말 더듬는 것이 증가하거나(Bosshardt, 2006), 말할 때 무의식적으로 사용하는 “어(uh)”나 “음(um)”이 화자의 의도에 대한 단서가 되기도 한다(Clark & Fox Tree, 2002). 둘째, 실증주의적 심리학 연구에 공통된 가정이지만, 언어반응이라는 종속측정치를 수집하는 과정이 연구 관심사인 독립변인의 작용에 영향을 주지 않아야 한다. 자기설명연구에서처럼 자료수집을 위한 지시가 연구의 독립변인이 아니라면, 언어자료를 수집하는 과정, 즉, 사고를 언어화하고, 질문에 답하는 과정이 참가자의 수행에 영향을 주는 경우에는 언어자료를 종속측정치로 사용하는 것이 적절하지 않다. 예를 들어, 시각적인 정보처리, 통찰적인 문제 해결 과정은 언어보고과정 자체에 의해서 방해받거나 왜곡될 수 있다(Schooler, Ohlsson, & Brooks, 1993; Schooler & Engstler-Schooler, 1990). 언어보고에 따른 왜곡 자체가 연구 대상이 아니라면, 연구 대상인 심리과정이 언어반응의 수집절차에 의해서 영향을 받을 가능성이 존재하는 경우 언어자료의 수집은 신중하게 고려되어야 한다.

연구 설계 측면에서 언어자료분석은 다양한 연구 설계와 병행될 수 있다. 관찰/기술(descriptive) 연구는 물론 실험연구에서도 사용될 수 있다. 유일한 차이는 반응시간 또는 다른 전통적인 실험 측정치 대신에 또는 그에 추가해서 언어반응이 수집된다는 점이다. 언어자료분석이 실험과 병행된 예로 McNamara(2004)의 연구를 들 수 있는데, 이 연구에서는 SERT(Self-Explanation Reading Training)라고 명명한 자기설명기법의 효과를 알아보기 위해서 기법을 훈련 받는 실험 집단과 통제 집단을 비교하였다. 훈련 기법의 효과를 알아보기 위해서 전통적인 글 이해 측정치(예, 사

후 검사 답변)에 더해 자기설명 프로토콜을 수집하여 두 집단이 글 읽기 동안 사용한 전략이 어떻게 다른지가 비교되었다. 복잡한 언어자료분석이 실시되는 경우 해석의 용이성을 위해서 실험설계를 단순하게 하는 것이 권장된다.

관찰/기술(descriptive) 연구에서 언어자료분석이 사용될 때는 실험 처치나 조작을 하지 않고 자연적인 맥락에서 발생하는 언어자료를 있는 그대로의 수집한다. 자료수집은 실험실 같은 통제된 상황에서 이루어질 수도 있으나(Chi et al., 1989) 실험실에서 일어나는 과학자들의 대화 또는 가정에서 가족 간의 대화 같은 자연적인 상황에서 이루어지기도 한다(Dunbar, 1997; Paletz et al., 2012). 최근 들어서는 인터넷, 블로그 등의 매체를 통해서 자기 생각을 표현하는 경우가 증가하고 가상공간에서의 상호작용 또한 증가하고 있는데, 이런 맥락에서 발생한 언어자료도 분석에 사용될 수 있다(Kimmerle, Moshaliuk, & Cress, 2011).

언어자료분석은 연구의 다양한 단계에서 사용될 수 있다. 즉, 이미 부분적으로 검증된 가설이나 변인의 효과에 대해서 확증적인 연구를 실시하는 경우뿐만 아니라 연구 초기에 가설을 탐색하는 단계에서 사용될 수도 있다. 확증적인 연구의 예로 자기설명의 효과를 새로운 과제나 학습 영역에서 검증하는 것을 들 수 있는데(Rittle-Johnson, 2006), 이 경우 자료분석은 상대적으로 간단한데, 선행연구에서 사용된 방법을 참고하여 해당 연구의 주제에 맞게 확장 또는 보완하면 된다. 언어자료분석이 주 종속 측정치에서 얻어진 결과를 보완하고 조건 간의 차이를 추가적으로 탐색하는 데 사용되기도 한다. 예를 들어 앞에서 소개한 Siegler(1995)는 5세 아동의 보존개념 수행을

비교하였는데, 조건 간의 수행차이를 알아보는 주 측정치는 보존과제에서의 반응 정확성이었으나, 조건 간의 차이를 보다 자세하게 알아보기 위해서 추가적으로 아동이 제공한 언어적 설명이 분석되었다. 언어자료분석은 또한 가설을 탐색하는 단계에서 매우 유용하다. 선행이론이 부족한 상태에서 연구를 수행하는 경우 또는 수행의 차이가 존재하나 기제가 불확실한 경우, 언어자료를 바탕으로 잠재적인 관련 변인과 기제를 탐색하고 가설을 만드는 데 활용할 수 있다. 이 경우 자료 분석은 ‘열린’ 과제가 되는데, 다양한 분석 단위와 틀을 시도하면서 그 과정에서 가설을 구체화하게 된다.

마지막으로 연구 설계 단계에서 중요한 것은 표본의 크기에 대한 결정이다. 언어자료분석이 사용되는 경우에도 몇 명의 참가자로부터 자료를 수집할 것인가의 문제는 다른 양적 연구에서와 동일하다. 특히 통계적 검증을 실시하고자 하는 경우는 충분한 수의 참여자로부터 자료를 수집해서 분석하는 것이 필요하다. 단 언어자료분석은 분석에 시간이 많이 걸린다. 다수의 참가자를 대상으로 자세한 분석을 하는 것은 상당한 시간과 자원을 필요로 하기 때문에 이 점을 염두에 두고 자료 수집과 분석을 계획하는 것이 필요하다. 언어자료의 분량이 짧거나 분석이 복잡하지 않은 경우 사례수가 많다고 분석이 반드시 오래 걸리는 것은 아니다. 하지만, 참가자 당 자료의 분량이 많은 경우(예, 참가자 당 인터뷰 길이가 2-3 시간) 또는 탐색적이거나 복잡한 분석의 경우에는 시간이 오래 걸리므로 절충안을 고려하는 것을 생각해 볼 수 있다. 예를 들어, Siegler(1995) 연구에서처럼 조건 간의 차이는 행동수행 측정치를 사용하고, 차이의 원인을

탐색하는 데 언어자료분석을 실시하거나, 또는 McNamara(2004) 연구에서처럼 각 조건에서 소수의 사례만 추출하여 분석을 실시할 수도 있다. 하지만 소수의 사례에 대해서만 분석이 실시되는 경우에는 통계적 일반화가 불가능하기 때문에, 확증적인 연구보다는 탐색적인 목적으로 연구가 이루어질 때 적합하다. 또한 모든 연구에서 마찬가지로, 자료를 여러 각도에서 분석하여 얻어진 결과의 타당성을 검증하는 것이 더욱 중요해 진다.

#### 언어반응의 유도

언어자료를 수집하는 것은 크게 자연적인 맥락에서 자료가 수집되는 경우와 그렇지 않은 경우로 나눌 수 있다. 자연스러운 과제 수행의 일부로 언어반응이 요구되거나(예, 대화나 회의 상황), 인터넷 등에 이미 공개된 반응(예, 공개된 게시판, 위키피디아)을 연구하는 경우 언어반응을 이끌어 내기 위한 특별한 지시나 질문이 필요하지 않다. 그러나 이 반응을 기록, 녹화하는 과정에서 참가자들의 반응성을 야기할 수 있는데, 자신의 활동이나 대화에 대해서 녹음/녹화가 이루어지는 것을 의식하여 평상시와 다르게 행동할 가능성이 존재한다. 이러한 이유로 자료수집을 할 때 자연스러운 언어반응이 일어날 수 있는 환경과 상황을 조성하는 것이 중요하다. 녹음기나 카메라라는 사람들의 주의를 끌지 않는 배경에 설치하고 사람들이 익숙해질 수 있도록 시간을 주는 것 등이 한 방법이다.

언어반응이 상황에서 자연스럽게 나타나는 반응이 아닌 경우, 어떻게 언어반응을 유도할지가 관건이 된다. 어떻게 언어반응을 유도하는 것이 적절한지는 연구 주제 및 가설에 따

라 달라지나, 크게 참여자의 온라인(online), 즉 현재 진행 중인 정보처리를 탐색하기 위해서 언어반응을 수집하는 경우와, 참여자가 이미 습득한 지식, 생각, 경험 등을 알아보기 위해서 언어반응을 수집하는 두 경우로 나누어 생각해 볼 수 있다. 우선 온라인 정보처리 연구가 목적일 때는 과제를 수행하면서 언어보고를 하도록 하는 동시적 언어보고가 원칙이다. 동시적 언어보고를 위해서 다양한 지시가 사용될 수 있는데, 이 경우 연구 질문, 주제에 따라서 적절한 지시가 달라진다. Ericsson 과 Simon(1980, 1993)은 과제가 수행되는 동안 동시적인 언어보고를 수집할 것과 언어보고에 대한 지시가 비지시적일 것을, 즉, 참여자가 이미 생각하고 있는 내용만 언어화할 것을 강조하였다. Ericsson 과 Simon의 경우처럼 참가자들이 문제를 풀면서 일어나는 사고과정의 규명에 관심이 있다면 참가자들이 과제를 수행하는 동안 떠오르는 생각을 소리 내어 말하라는 TA의 비지시적인 지시가 적합하다. 그러나 자기설명연구에서처럼 지시 자체가 처치인 경우도 있는데, 이 경우는 읽고 있는 글의 내용을 설명하라는 것 같은 구체적인 지시를 주고 언어자료를 수집할 수 있다.

바람직하지는 않지만, 과제를 수행하는 동안 언어자료를 수집하는 것이 여의하지 않아서 사후에 과제수행에 대한 언어보고를 수집하는 경우가 있다. 즉, 과제를 수행하고 난 다음에 그 과제를 어떻게 수행했는지를 답하도록 하는 것이다. Ericsson 과 Simon(1980, 1993)이 강조하였듯이 사후보고는 과제 수행 과정을 왜곡할 가능성이 많기 때문에 사용하지 않는 것이 바람직하다. 예외적으로 참여자가 과제 상황에서 답변을 할 수 없는 상황이거나(예, 다른 사람과 상호작용 해야 하는 상황),

어린 아동의 경우처럼 아직 언어 능력이 발달하지 못해서 과제 수행과 언어보고를 동시에 하는 것이 과도한 인지 부담(cognitive load)을 야기하는 경우에 회고적 보고가 불가피할 수 있다. 하지만 연구자가 내릴 수 있는 결론에 대한 제약이 존재함을 인식해야 하는데, 과제를 수행한 다음에 사후보고를 수집하는 경우 그 내용이 참여자의 온라인 처리를 반영하기 보다는 참여자가 기억한, 또는 해야 한다고 생각한 처리를 반영하기 때문이다.

온라인 정보처리 과정이 아니라 사람들이 이미 형성한 지식 구조, 생각을 알아보는 것이 연구의 목적일 때는 인터뷰 형식의 질문이 사용되기도 한다. 이러한 질문은 다양한 맥락에서 이루어지는데, 학습 연구에서 참가자들의 지식을 측정하기 위한 사전/사후 검사의 맥락에서 이루어질 수도 있고 또는 아동이나 전문가들의 경험과 지식을 추출(elicit)하기 위한 목적으로 이루어질 수도 있다. 이러한 목적으로 수집된 언어자료를 지식 프로토콜(knowledge protocol)이라 칭하기도 한다(Chi & Koeske, 1983). 연구 목적과 질문에 따라서 질문은 일반적일 수도 있고(예, XXX에 대해서 아는 바를 말씀해 주시기 바랍니다) 또는 매우 구체적일 수도 있다(예, XXX가 고장 났을 때 어떻게 해결하겠습니까?). 연구가설이 질문 설계에 반영될 수도 있는데, 예를 들어 자기 설명이 정보의 저장보다 추론에 영향을 준다는 가설을 가지고 있다면, 이를 알아보기 위해서 기억과 추론을 측정하는 질문을 구성할 수 있다. 반면 아동이 발달과정에서 형성하는 다양한 오 개념(misconception)의 여부를 알고 싶다면, 이를 알아 볼 수 있는 질문을 사용할 수도 있다(예, 아동이 지구가 평평하다고 믿는지 알아보기 위해서 '현재 위치에서 직선

으로 계속 걸어가면 어디에 도달하는가?’ 같은 질문을 하는 것). 단계적으로 질문을 하거나, 여러 가지 가설, 시나리오를 설정하고 답변에 따라 추가로 탐색질문(probe question)을 할 수도 있다.

다양한 질문 형식이 사용될 수 있으나, 동시적인 보고를 수집할 때와 마찬가지로 질문이 답변자의 사고과정에 미치는 영향을 고려해야 한다. 특히 학습이나 기타 과제의 수행 전에 사전 검사 형태로 주어지는 질문들은 이후 학습활동에 영향을 줄 가능성이 많기 때문에 이러한 영향이 최소화되도록 질문이 설계되고 제시되어야 한다. 동일한 질문이라고 하더라도 어떤 맥락에서 제시되는가에 따라 서로 다른 의미를 전달할 수 있다. 예를 들면 앞의 질문이 뒤의 질문에 대한 힌트를 제공할 수 있기 때문에, 개별 질문이 제시되는 순서에도 주의를 기울여야 한다. 일반적으로 비특정적인 질문을 먼저하고 그 다음에 구체적인 질문을 하는 것이 이러한 영향을 줄이는 방법이다. 질문이 구조화되는 경우에도, 단답 보다는 서술식 답변을 얻을 수 있는 질문을 하는 것이 바람직한데, 이러한 질문이 참가자의 사고의 내용과 과정을 더 있는 그대로, 풍부하게 드러내 주기 때문이다.

#### 언어자료의 수집

어떤 방식으로 언어반응을 유도할 지가 결정되고 나면 그에 따라 언어반응을 수집하는데, 이 때 몇 가지 고려할 사항이 존재한다. 우선 소리내어 생각하기 같은 동시적 언어 보고를 수집할 때 참가자들이 소리 내어 말하기에 익숙하지 않은 경우가 존재한다. 이 문제에 대한 해결책으로 동시적인 언어보고를 수

집하는 경우 자료 수집 전에 간단한 연습이 실시할 수 있다. 연습은 본 실험 또는 연구에서 사용하는 것과는 무관한 재료를 사용하여 실시하는데, 연습하는 동안에 참가자들이 소리내어 말하는 것에 익숙해지도록 하고, 필요한 경우 피드백을 제공한다. 연습을 충분히 한 다음에도 과제수행 동안 참가자가 소리 내어 말하는 것을 잊는 경우가 발생한다. 이 경우 참가자에게 말하는 것을 상기(prompting)시켜줄 수 있는데, 언제 어떻게 상기시키는가에 따라 참가자의 수행에 영향을 줄 수 있기 때문에 주의가 필요하다. 예를 들면, “XX에 대해서 생각하셨나요?” 등의 특징적인 질문을 하는 것은 질문에 포함된 내용에 대한 생각을 유도하는 효과가 있기 때문에 가능한 중성적인 방식으로 상기시키는 것이 필요하다. 일반적으로 “소리 내어 생각하는 것 잊지 마세요” 등의 프롬프트가 사용된다. 중성적인 프롬프트를 사용할 때도 언제 사용하느냐에 따라 수행에 영향을 줄 수 있기 때문에 신중하게 사용되어야 하는데, 참가자들이 프롬프트 제시 시점에 의미를 부여하여 무엇인가를 이야기하지 않으면 안 되는 지점으로 생각할 수 있기 때문이다. 이러한 오해를 제거하기 위해서는 자료 수집을 시작하기 전에 참가자들에게 이런 주의의 환기가 있을 수 있다는 것을 미리 알려주고, 나름의 규칙을 정하고 일관되게(예, 침묵 후 15 초) 상기를 하는 것이 필요하다. 또한 단순히 소리 내어 말하기가 익숙하지 않은 경우도 있지만, 여러 이유로 참여자들이 소리 내어 생각하는 지시에 소극적으로 반응할 수 있다. 예를 들어, 생각이 없거나 몰라서 아무 말도 하지 않는 것이 아니라, 정확한 발언인지 자신이 없어서 말을 하지 않는 것이다. 이러한 자기 검열을 제거하는 것이 중요한데,

확실하지 않아도, 단순한 추측이어도, 정제된 형태의 발언이 아니어도 떠오르는 생각을 가감 없이 말하는 것이 중요함을 지시에서 강조하고 허용적인 분위기를 조성해 주는 것이 필요하다.

언어자료의 수집은 주로 오디오 녹음 또는 비디오 촬영의 형태로 이루어진다. 글의 형태(written protocol)로 언어보고를 수집하는 경우도 존재하고(Berthold et al., 2007; Coleman et al., 1997), 문헌을 분석하는 경우도 존재한다. 그러나 온라인 정보처리가, 즉, 현재 진행 중인 정보처리가 즉시 연구하는 것이 연구목적일 때는, 연구 대상이 글쓰기 과정이나 쓰는 것을 수반하는 과정이 아니라면, 구두로 언어 자료를 수집하는 것이 바람직하다. 즉각적으로 작성되고 보내지는 단문의 메시지가 아니라면, 글은 온라인 정보처리를 드러내는 데 제한적이다. 과제가 수행되는 동안 이를 반영하는 언어반응을 수집하는 것이 불가능하거나 자료나 문제의 성격 상 기록된 문서에 의존할 수밖에 없는 경우에 글의 형태로 수집된 언어 자료를 분석할 수 있으나, 이에 따르는 해석의 제약을 고려해야 한다. 글쓰기는 말하기보다 자기검열이 더 많이 일어나는 과정으로 내적 사고과정에 대한 정보를 있는 그대로 드러내는 데 더 제한적이기 때문이다.

초기에는 언어보고를 수집하는 데 음성적 녹음이 가장 많이 사용되었지만, 최근 들어서서 기술의 발달로 동영상 자료를 수집하는 것이 증가하고 있다. 비디오 자료는 청각적인 정보 이외에 다양한 시각적인 정보를 전달한다는 장점이 있다. 도구의 사용 또는 여러 사람들이 모여서 상호작용하는 것을 연구하는 경우 행동, 움직임, 얼굴표정, 제스처 등 다양한 시각적인 정보가 존재하는데, 비디오는 이

러한 정보를 포착하는 데 탁월하다. 어떤 형태든 녹음이나 녹화를 할 때, 특히 현장에서 자료가 수집될 때 중요하게 고려해야 하는 것은 녹음의 질이다. 실험실처럼 조용하고 통제된 공간이 아니기 때문에 배경소음이 존재할 수 있고, 동시에 여러 사람이 말하는 경우 누가 말을 했는지 확인하는 것이 어려울 수 있다. 이런 경우 참가자별로 마이크를 착용하여 다채널로 녹음하거나, 고성능 마이크를 사용 또는 여러 대의 녹음기를 활용하는 방법을 고려하는 것이 필요하다. 동시에 여러 대의 카메라가 자료 수집에 사용되기도 하는데, 집단 활동을 할 때 여러 명의 얼굴을 각각의 카메라로 잡거나, 또는 얼굴과 보고 있는 화면을 같이 촬영하는 경우도 있다. 자료 수집에 여러 채널이 사용되는 경우 독립된 기기에서 수집된 녹음 또는 녹화 자료를 동기화하는 추가적인 작업이 필요하다.

#### 축어록 작성 및 자료의 선택

분석을 위해서 수집된 오디오/비디오 녹화 자료를 축어록으로 푸는 작업이 수행된다. 분석을 위해서 반드시 축어록을 만들어야 하는 것은 아니고, 비디오나 음성파일을 직접 분석하는 것도 가능하다(예, Meier, Spada, & Rummel, 2007). 하지만 이는 분석이 정교하지 않을 때, 언어적인 반응보다는 비언어적인 반응을 분석할 때 적절하다. 분석이 언어반응의 미세한 측면까지 고려하는 것을 요구할 때는 녹음된 것을 듣거나 시청하는 것만으로는 어렵고 녹음된 내용을 글로 풀어 축어록을 작성하고, 통상적으로 분석은 축어록에 기반하여 이루어진다.

녹화된 테이프를 축어록으로 만드는 녹취

또는 전사(transcribing)는 그 자체가 여러 판단을 요하는 분석과정의 일부이다. 축어록 작성은 많은 시간과 노력이 요구되는 작업이다. 축어록을 자세하게 만들수록 작성에 걸리는 시간이 늘어나기 때문에 전사를 시작하기 전에 축어록에 포함될 정보의 범위와 기준을 정하는 것이 필요하다. 어떤 정보까지를, 얼마나 자세하게 녹취하는가, 어떤 방식으로 녹취하는가의 문제는 연구문제, 연구분야, 분석기법에 따라 달라진다. 언어학, 특히 대화분석(Conversation Analysis)에서는 일반적으로 Jeffersonian Transcription System 이라고 불리는 녹취 체제를 따른다(Jefferson, 2004). Jeffersonian 체제는 매우 자세한 체제로 발화의 내용뿐만 아니라 대화의 중첩, 억양의 변화, 목소리 크기, 속도까지 자세하게 전사한다. 인지연구에서 이런 자세한 축어록을 사용하는 것은 드물지만, 여전히 언어반응에 대한 글자 그대로(verbatim)의 풀어쓰기가 중요하다. 사소해 보이는 행동도 기저의 심리과정을 작동을 나타내는 지표가 될 수 있기 때문에 말을 더듬거나 반복하는 것, 발음 오류뿐만 아니라 언어반응의 해석에 필요한 여러 비언어적인 정보(예, 침묵, 한숨, 몸짓) 및 행동에 대한 정보(예, 웃음, 일어나서 움직이기)도 축어록에 포함하는 것이 바람직하다. 동일한 언어반응도 맥락에 따라 해석이 달라지기 때문에 축어록에 상황과 맥락 정보를 병기하여 보다 정확한 언어 의미 파악이 이루어질 수 있도록 하기 위함이다. 특히 축어록을 푸는 사람과 분석하는 사람이 다른 경우가 많은데 이런 경우 축어록이 풍부한 맥락 정보를 전달할수록 분석의 정확성이 증가하고 분석하면서 원 자료를 다시 보아야 하는 수고를 덜 수 있다. 축어록에 맥락 정보가 포함되어도 분석을 하다보면 해석이

모호한 경우가 종종 발생하고, 이런 경우에는 원 녹음이나 녹화를 참고하는 것이 필요하다.

짧은 시간의 녹음이나 녹화라도 축어록으로 풀면 많은 분량의 자료가 발생한다. 말하는 속도에 따라 발생하는 축어록의 양은 달라지지만 1 시간 동안의 언어반응을 축어록으로 풀면 약 15-50 장 정도의 분량이 된다(Chi, 1997). 10 명 정도의 참가자들에게 각각 두 시간 정도 분량의 프로토콜을 받아 축어록으로 풀다고 하면, 적게는 삼 백장에서 많게는 천장이 넘는 자료를 분석해야 한다는 것을 의미한다. 대규모로 수행되는 프로젝트의 경우 수십, 또는 수백 명에 대해서 수십 시간의 자료가 수집되는데(예, 학교 장면에서 한 학기 동안 수업 시간을 계속 관찰하고 녹화하는 경우, 또는 장기간 인터넷 커뮤니티에서 일어난 상호작용을 연구하는 경우), 이 경우 수집된 자료의 양은 종종 어느 개인 또는 개별 연구팀에서 할 수 있는 분석의 범위를 넘어서게 되고, 이런 경우 종종 자료의 일부만 표집하여 분석에 사용한다.

자료를 표집하는 방법에는 여러 가지가 있다. 첫째, 자료를 무선적으로 표집하는 방법이 있다. 예를 들어 50 명에게서 프로토콜을 수집했으나 이를 다 분석할 수 없을 때 무선적으로 10 명의 자료를 표집하여 분석할 수 있다. 참가자 이외의 다른 단위가 표집에 사용될 수도 있는데, 예를 들어 학교나 조직에서 몇 달 동안 매일 녹화를 한 경우 그 중 일부를 무선적으로 표집하여 분석할 수 있다. 표집의 단위를 인위적으로 만들고 선택할 수도 있는데, 예를 들면 녹화된 테이프를 10 분 또는 30 분 단위로 쪼개고 그 단위들 가운데서 분석할 단위들을 무선적으로 표집할 수 있다. 어떤 단위를 사용하든 중요한 것은 연구의 대상이 되

는 활동이나 사건이 표집 단위에 유의미하게 포함될 수 있는 단위를 선택하는 것이다. 둘째, 심리학 연구에서 많이 사용되는 참가자의 특징 및 기타 영향 요인을 통제된 표집을 생각해 볼 수 있다. 예를 들어 각 실험조건에서 참여자 일부의 자료를 표집해서 분석하거나, 또는 참가자 중 남자와 여자의 성비 또는 연령대가 고르게 표집되도록 하는 것이다. 연구 주제에 따라서 자료가 수집된 시간적인 순서가 중요할 수도 있는데, 이 경우 예를 들어 학습이나 상호작용의 초기, 중기, 말미를 골고루 표집하거나, 인터넷 커뮤니티의 형성을 연구할 때 초기와 중기, 후기 등으로 나누어 표집하는 것을 생각해 볼 수 있다. 이런 표집이 이루어진 경우 조건 간, 성별, 또는 시간에 따른 추이를 비교분석하는 것도 가능해진다. 마지막으로, 언어반응의 내용을 고려하여 표집을 할 수 있다. 연구 문제에 따라서 다양한 기준을 사용하는 것이 가능한데, 예를 들어, 수집한 자료 중에서 갈등이 있었던 사례/사건만을 분석하거나, 문제해결에 성공 또는 실패한 사례만 분석하거나 하는 것이다. 이 경우 기준이 되는 사건, 개념에 대한 정의가 연구자마다 다를 수 있기 때문에, 표집 기준이 되는 사건을 명확하게 정의하는 것이 매우 중요하고 가능한 객관적인 지표에 의존할 것이 권장된다.

## 분석

설문지를 사용하거나 실험을 하는 경우, 참가자들은 평정 척도에서 자신의 심리상태나 행동을 나타내는 수치를 선택하거나, 실험과제에서 요구되는 반응을 키를 눌러 표시하거나, 경우에 따라서 ‘네, ‘아니오’ 같은 언어반

응을 한다. 이런 반응은 구조화된 형태로 수집되기 때문에, 자료분석의 핵심은 이들 반응에 대한 기술 통계치(예, 평균반응시간, 횟수)를 구하거나 조건 간의 차이를 통계적으로 검증하거나 또는 모형을 사용하여 변인 간의 관계를 탐색하는 것이 된다. 반면 언어자료분석에는 구조화 되지 않은 언어반응, 또는 기타 형태의 자료(예, 그림, 사진, 동영상)가 분석 대상이 되는데, 언어자료분석의 목적은 이러한 비구조화된 자료를 체계적으로 구조화하여, 객관적인 분석이 실시될 수 있도록 하는 것이다. 이 작업은 여러 단계를 거쳐 이루어지는데, 크게 (1) 분석 단위/대상 정하기 (2) 코딩 스킴/분석 틀 정하기 및 (3) 코딩 및 코딩 신뢰도 확인 단계로 나눌 수 있다. 편의상 분석의 단계를 나누었으나 각 단계는 독립적이고 계열적으로 진행되기보다는 병렬적으로 진행되고 순환/반복되는 과정이다.

### 분석의 단위/대상 정하기

분석을 위해서 가장 먼저 결정해야 하는 것은 분석의 단위를 무엇으로 할 것인가이다. 분석단위 결정에 가장 중요한 고려요인은 연구 질문이다. 언어반응을 문장단위에서 분석할 수도 있고, 아이디어나 주제 수준 또는 담화수준에서 분석할 수도 있지만, 어떤 분석 단위를 사용할지의 결정은 일차적으로 연구 질문 또는 가설에서 무엇을 답변하고 측정하고자 하는가에 달려있다. 예를 들어, 학생들이 습득한 어휘를 얼마나 자유자재로 사용하는가를 알아보고자 한다면, 어휘가 분석의 단위가 되어야 한다. 반면 보고서를 전반을 관통하는 논리를 알아보고자 한다면 보고서 전체가 분석의 단위가 되어야 한다. 즉, 연구에서 측정하고자 하는 개념, 가설을 가장 잘 반영하는



단위가 분석 단위가 되어야 한다.

분석의 단위는 결국 코딩, 범주화의 단위가 되기 때문에 분석 단위의 결정은 결국 무엇을 코딩 단위로 할 것인가의 결정이다. 코딩 및 해석의 용이성을 위해서는 연구 질문에서 다루는 여러 사건을 다 포괄할 수 있는 분석단위를 선택하는 것이 바람직하다. 예를 들어 학생들이 문제풀이 하는 동안 어떠한 처리/활동을 하는가를 알아보고자 하고 특히 (1) 문제 읽기 (2) 문제 풀이, (3) 모니터링 활동에 관심이 있다고 한다면, 이 세 활동이 모두 포착될 수 있는 분석단위를 정하는 것이 이상적이다. 하지만 연구의 하위 가설에서 다루는 개념이 상이하여 동일한 분석단위를 사용하기 어려운 경우에는 가설에 따라 분석 단위를 달리하기도 한다. 예를 들어 Roscoe(2013)는 개인교사(tutor)의 효과를 연구하면서 교사와 학생의 상호작용의 유형을 분석할 때는 주제 중심의 에피소드를 분석단위로 사용하였지만, 튜터의 모니터링을 분석할 때는 문장/진술문을 분석단위로 사용하였다.

경우에 따라서 동일한 개념, 사건에 대해서 여러 수준 및 유형의 분석 단위가 가능하다. 이것은 일차적으로 같은 사건이 정보처리의 여러 단계, 수준에서 일어나기 때문이다. 예를 들어 글 이해 동안에 일어나는 추론을 연구한다고 할 때, 글 이해에서 일어나는 추론의 유형은 매우 다양하고 따라서 여러 단위에서 추론을 살펴보는 것이 가능하다. “철수는 못을 박았다”는 문장을 읽고 철수가 ‘망치’를 사용했을 것이라는 도구에 대한 추론은 문장 단위에서 드러나는 사건이다. 그러나 “철수는 못을 박았다. 손에서 피가 흘렀다”라는 문장을 읽었을 때 두 사건 간의 인과 관계 또는 동기에 대한 추론을 포착하기 위해서는 여러 개의 문

장을 아우르는 단위가 필요하다. 궁극적으로는 연구 질문에 가장 부합하는 단위가 무엇인지를 고려해서(예, 소형구조 형성에 관여하는 추리 대 대형구조 형성에 관여하는 추리) 가장 적절한 단위를 선택하는 것이 필요하다. 그러나 시간적인 제약이 적다면 여러 분석 단위에 대해서 분석을 모두 실시하는 것을 고려할 수도 있다(Chi, 1997). 이 경우 코딩의 일관성을 점검할 수 있을 뿐만 아니라 추론이 분석 단위에 따라 어떻게 다르게 나타나는지를 알아볼 수 있다는 장점이 존재한다.

분석 단위는 또한 수집된 자료의 종류에 따라서 달라진다. 같은 개념, 가설에 대한 분석 단위를 정한다고 할지라도, 언어자료를 분석할 때의 단위와 시각적인 자료를 분석할 때의 단위가 달라질 수밖에 없는데 언어자료의 경우 단어, 문장 등의 언어구조와 관련된 단위들이 사용될 수 있는 반면 시각적인 자료에는 이런 단위가 적절하지 않고 자료의 성격에 맞는 대안적인 단위를 찾아야 한다. 예를 들어, 노드와 망으로 이루어진 그림이면 노드와 화살표가, 디자인 스케치이면 스케치의 하위 요소가 분석단위가 될 수 있다. 사람들의 움직임이나 활동이 포함된 동영상 자료에는 또한 언어자료에 존재하는 자연스러운 경계(단어, 문장)가 존재하지 않는데, 따라서 분석을 위해서 임의적인 경계를 만들어 분석의 단위로 사용하기도 한다. 예를 들어 동영상을 시간 단위로 쪼개거나 또는 사건을 중심으로 쪼개어(예, 새로운 인물의 등장) 분석의 단위로 활용할 수 있다.

분석단위를 결정할 때 분석의 민감성(sensitivity)이 고려되어야 하는데, 분석단위의 크기는 얻어진 결과가 해당 현상을 얼마나 자세하고 민감하게 측정하는가와 관련이 있다.

분석단위가 작을수록 연구 대상이 되는 사건을 세밀하게 포착할 수 있으나 코딩의 양은 늘어난다. 반대로 단위의 크기가 클수록 코딩해야 하는 단위 수가 줄어들지만 동시에 분석에서 얻어낼 수 있는 정보가 제한된다(Chi, 1997). 예를 들어 부부가 갈등적인 상호작용을 하는지를 알아보고자 할 때, 전체 상호작용 에피소드를 한 단위로 보고 갈등이 있었는지 여부를 코딩할 수 있지만, 이런 분석 단위는 갈등의 규모, 시간 전개 등에 대한 자세한 이해를 안내하기 힘들다. 이러한 정보를 얻어내기 위해서는 더 작은 단위(예, 대화의 순서)를 사용하여 분석하는 것이 필요하다. 어떤 분석 단위를 사용하든지, 한 분석에서 일관된 분석 단위를 사용하는 것이 중요하다. 분석단위는 코딩의 단위인데 코딩의 단위가 일정해야 얻어진 결과에 대한 해석이 분명해질 수 있기 때문이다. 자료의 성격에 따라 분석 단위의 크기가 정형화되지 않거나 일관되지 않을 수 있는데, 예를 들어서 대화에서 질문-반응 쌍을 분석한다고 할 때 상황에 따라서 질문에 대한 답이 짧을 수도 있지만 길어질 수도 있고, 경우에 따라서 답변에 후속 질문이 포함될 수도 있다. 이런 경우에는 분석 단위의 크기 및 길이 등을 추가적으로 분석하는 것이 필요하다.

분석 단위를 정하고 나면 이를 적용하여 전체 축어록을 분석 단위로 나누거나(segmentation) 또는 축어록에서 분석 단위를 확인하는 작업을 하게 된다. 이 작업을 위해서는 분석의 단위를 개념적으로 정의하는 것뿐만 아니라 조작적으로 정의하는 것이 필수적이다. 즉, 축어록 또는 기타 자료에서 분석단위 구분에 사용될 수 있는 단서, 기준을 명세하는 것이 필요한데, 여러 단서와 기준이 사용될 수 있으나 크게 자료에 존재하는 물리적

/형식적 단서를 사용하는 경우와 의미 단서를 사용하는 두 경우로 나눌 수 있다. 물리적/형식적 단서를 가지고 분석단위를 정의하는 경우는 문장, 단어 등 물리적으로 구분될 수 있는 단위를 분석단위로 사용하는 것이다. 언어의 속성이 아닌 다른 물리적 기준을 사용하여 단위를 정할 수도 있는데, 말을 하다 잠시 침묵한 지점, 축어록이 인쇄된 종이에서의 줄, 동영상의 1분, 5분 등의 시간 단위로 구분한 것을 분석단위로 정할 수도 있다. 대화를 분석하는 경우 말한 순서, 가상공간의 상호작용의 경우 게시판에 올린 글이 될 수도 있다(예, van Aalst, 2009; Jeong, 2013). 어떤 단서를 사용하든지, 물리적/형식적 단서를 사용하는 경우 자료에서 코딩의 단위를 확인하는 작업은 상대적으로 손쉬운 작업이 된다.

이와는 달리 의미 단서를 바탕으로 분석단위를 정의하기도 하는데, 명제, 아이디어, 설명, 에피소드 등이 대표적인 의미 단서에 기반한 분석단위이다. 이런 의미단서에 기초한 분석단위는 연구자의 이론적 관점이나 해석에 따라서 그 경계가 달라질 수 있는데, 예를 들어, 새로운 '생각'을 분석단위로 한다고 했을 때 연구자에 따라서 무엇을 새로운 생각으로 볼지, 아이디어의 경계를 어디로 할지가 달라질 수 있다. 의미적 단서가 가진 이러한 모호성 때문에 이 경우 각 분석단위가 무엇인지 또한 축어록 또는 기타의 자료에서 어떻게 정의되고 확인될 수 있는지 구체적인 단서가 조작적으로 분명하게 정의되지 않으면 안 된다. 예를 들어, 아이디어나 추론이 어떤 사건을 지칭하는지, 어느 수준/크기에서 정의되었는지가 예와 함께 분명하게 제시되어야 한다. 이 경우 분석단위를 구분하는 것 또한 판단이 필요한 일종의 코딩 작업이 되고, 신뢰도도 검

증되어야 한다(코딩 신뢰도에 대해서는 후속 절 참고).

### 분석 틀/코딩 스킴 정하기

코딩 스킴은 분석 단위를 어떻게 범주화하고 분류할지에 대한 체계이다. 코딩 스킴을 정하는 것은 연구 가설을 명세화 하는 과정에 비유할 수 있다. 예를 들어 회의 중에 참석자들이 하는 질문을 분석할 때 단순히 참가자들이 질문을 했다는 것을 알아보고자 하는지, 아니면 질문 유형이 열린 질문인가 닫힌 질문인가, 질문이 누구에게 향한 것인가, 질문의 순서 등을 보고자 하는 것인지 등을 정해야 한다. 코딩 스킴은 이 결정을 체계화한 것으로 종종 분석에 사용되는 코딩 범주 및 판단 규칙의 형태를 띤다. 코딩 스킴에서는 코딩 범주를 정의하고, 어떠한 언어적 발화 또는 행동이 각 범주의 증거가 되는지를 구체적으로 명시하여야 한다. 예를 들어, 질문을 열린 질문과 닫힌 질문으로 코딩한다고 하면 이들 질문의 특징이 무엇인지를 개념적으로 정의하는 것뿐만 아니라, 구체적으로 어떠한 언어적 발화를 각 범주의 사례로 볼 지를 조작적으로 정의한다(예, 질문의 형태를 취하지 않았으나 의미상 질문인 경우 어떻게 코딩할 것인가). 이 때 해당 범주에 속하는 사례의 특징을 정의하고 전형적인 예를 들어주는 것뿐만 아니라 그 범주에 속하지 않은 반대 사례가 무엇인지를 정의하고 예를 들어주는 것도 필요하다. 시간이 걸려도 가능하면 자세하게 코딩 스킴을 정의하는 것이 중요한데, 코딩 스킴이 분명하게 정의되고 기술될수록 코딩의 주관성이 줄어들고, 분석의 신뢰도가 높아진다. 또한 코딩 스킴을 문서화하는 것이 바람직한데, 신뢰도 검증 및 반복 검증을 위해서 뿐만 아니

라 논문 출간에도 필요하다.

코딩 스킴을 만드는 과정은 선행연구/이론에 의해서 주도될 수도 있고, 반대로 자료에 기반하여 이루어질 수도 있다. 코딩 스킴을 만드는 과정이 이론, 또는 자료 주도적으로 이루어지는가는 언어자료분석이 탐색적 연구의 맥락에서 사용되는가 아니면 확증적 연구에서 사용되는가와 관련된다. 선행 연구와 이론이 존재하는 경우 이들을 참고하여 코딩 스킴을 만들거나 필요한 경우 선행연구에서 사용한 코딩 스킴을 현재 연구에 맞게 수정 보완할 수 있다. 선행연구가 충분하지 않은 경우에는 수집한 자료를 바탕으로 귀납적으로 코딩 스킴을 만들어 가는 작업이 필요하다. 축어록을 읽으면서 어떤 범주가, 증거가 원하는 질문에 대한 답이 될 수 있을까를 생각하는 과정에서 자료에 존재하는 패턴, 구조가 드러나게 된다. 이 경우 수집된 자료에 대한 친숙성이 중요한데, Chi(1997)는 연구자가 축어록을 여러 번 읽으면서 자료에 대한 감을 익히고 자신의 직관을 드러낼 수 있는 코딩 스킴을 개발할 것을 제안하고 있다. 그러나 강조점만 다를 뿐 연역적인 접근과 귀납적인 접근이 코딩 스킴을 만들거나 정할 때 모두 필요하다. 연역적인 방법으로 코딩스키를 만들어도 이를 자신이 수집한 자료에서 확인하고 다듬는 과정이 필요하고, 자료에 근거해서 코딩스키를 만들더라도 관련 이론을 배제할 수 없기 때문이다.

새로운 지식 탐구가 목적인 연구의 속성상 연구자들은 대부분 자신의 연구 질문에 적합한 코딩 스킴을 새로 만든다. 하지만 연구에 따라서 다른 연구자가 개발한 코딩 스킴을 사용할 수도 있다. 선행연구에서 얻어진 결과를 새로운 상황에 확장하거나(예, 다른 지역에서

확인) 또는 연구 질문은 달라도 선행연구와 동일한 과제를 사용하는 경우 선행 연구의 코딩 스킴이 큰 수정 없이 그대로 사용될 수 있다(예, Azvedo & Cromley, 2004). 또한 측정도구처럼 다양한 상황에 사용될 목적으로 개발된 코딩 스킴을 사용할 수도 있는데, Meier 등(2007)은 화상 통화를 사용한 의사와 심리학자 간의 상호작용 질 평가를 위한 평정척도를 개발하였고, Baker, Andriessen, Lund, van Amelsvoort 및 Quignard(2007)는 인터넷 상의 논쟁(debate)을 7 개의 하위범주로 분석하는 코딩 스킴을 개발하였다. 이들 분석 스킴들은 다양한 상호작용과 논쟁의 맥락에서 사용될 목적으로 개발된 경우이나, 이러한 목적으로 개발된 도구가 아니라도 하더라도, 필요에 따라서 다른 연구에서 개발한 코딩 스킴을 사용하거나 또는 자신의 연구의 목적에 맞게 수정하여 사용할 수 있다. 그러나 척도 사용에 있어서도 마찬가지로이지만, 기존에 개발된 코딩 스킴이나 선행연구에서 사용된 코딩스킴을 사용할 때는 연구 질문을 고려할 때 적절한 스킴인가, 또한 분석하고자 하는 자료에 적용이 가능한지를 검토해야 한다. 예를 들어 동영상 분석을 위해 만들어진 코딩 스킴이 축어록 분석에도 사용될 수 있는지, 대학생들 대상으로 개발된 코딩 스킴을 아동에게 적용하고자 한다면 그대로 적용이 가능한가를 점검할 필요가 있다.

코딩은 척도화 과정이다. 어떠한 방식으로 척도화 할 것인지 정하는 것이 필요한데, 많은 경우 명목척도(nominal scale)가 사용된다. 명목척도로 코딩하는 것은 관찰된 언어 또는 기타 행동을 특정 범주로 코딩하는 것인데, 특정한 처리의 발생 여부(예, 설명을 했는가) 또는 유형(예, 사례에 기반한 설명, 원리에 기반

한 설명, 개인적인 경험에 의존한 설명)을 코딩할 수 있다. 이때 코딩의 결과물은 설명이 5 회, 또는 사례 기반 설명이 5 회, 원리에 기반한 설명이 10 회 등과 같은 빈도가 된다. 언어 자료분석에서 대부분의 코딩은 명목척도로의 코딩을 사용하나, 경우에 따라 Likert 척도가 사용되기도 한다. 사건의 강도 또는 정도를 포착하고 싶을 때 사용되는데, 예를 들면 설명의 질을 1-3 또는 1-5 척도 상에 코딩하는 것을 들 수 있는데, 답안이 설명을 포함하지 않은 경우 0, 설명을 했지만 정확하지 않은 경우 “1”, 정확하기는 하지만 정교하지 않은 경우 “2”, 정확하면서 정교한 설명은 “3”으로 구분하여 코딩할 수 있다.

이론에 기반하여 도출된 모형(model) 또는 템플릿(template)에 기반하여 분석이 진행되기도 한다. 이 방법은 특히 초기의 프로토콜 분석을 사용한 문제해결 연구에서 많이 사용되었는데, 연구 참여자의 수행이 이론이나 논리적으로 도출된 모형과 어느 정도 부합하는지를 알아보기 위한 목적으로 주로 사용된다. 이를 위해서는 인지적 과제분석(cognitive task analysis)이 필수적인데, 이는 연구 참여자가 수행해야 하는 과제를 정보처리 요구 면에서 세 말하게 분석하는 것을 지칭한다. 인지적 과제 분석을 통해서 문제풀이 동안 참가자가 사용할 수 있는 방략의 수, 문제 해결의 경로, 검색의 범위, 필요한 지식 등을 도출하거나 학습 상황에서 학생들이 학습해야 하는 지식의 범위가 무엇인지를 결정할 수 있다(Chan et al., 2011; Kaplan & Simon, 1990). 분석의 핵심은 참여자의 수행이 논리적, 이론적으로 도출된 모형과 어느 정도 부합하는지를 비교하는 것인데, 그 결과는 템플릿이나 모형에 일치/불일치한지 여부가 될 수도 있고, 또는 일치나 불

일치의 정도를 수치화 한 것이 될 수도 있다.

코딩이 단순히 정해진 범주의 사건이 몇 번이나 있었는지를 알아보는 것으로 끝날 수도 있지만, 언어자료분석의 강점이자 전통은 단순한 표면적인 사건을 코딩하고 그 발생 빈도를 알아보는 것이 아니라 배후에 있는 인지과정과 표상의 속성을 드러내는 분석을 하는 것이다(Chi, 2006). 이를 위해서는 일차적 코딩에서 멈추지 말고 범주 간의 구조, 관계, 순서를 탐색하는 추가적인 분석을 실시하는 것이 종종 필요하다. 이 추가적인 분석을 정형화하기는 쉽지 않은데, 예를 들어 문제풀이과정에서 특정 개념을 언급한 여부나 횟수를 분석하는 것에서 멈추지 않고, 배후의 심성 모형을 분석하여 문제풀이의 기제를 밝히거나, 또는 아동이 보존개념문제를 얼마나 정확하게 풀었는가의 횟수를 코딩하는 것에서 더 나아가서 정반응의 시간적인 전개를 분석할 수 있는데, 이러한 분석을 통해서 발달의 경로가 단선적인지 아니면 나선형인지 등을 알아보는 것이 가능하다. 또는 분석의 결과를 지식 망(knowledge network)으로 표상할 수도 있는데 이렇게 표상된 지식 망은 노드의 수, 클러스터링의 정도 등에서 그 특징이 분석될 수 있다(Chi & Koeske, 1983).

언어자료분석의 장점 중의 하나는 동일한 자료를 여러 번 코딩할 수 있다는 것이다(Chi, 1997). 이는 여러 맥락에서 다양한 이유로 이루어지는데, 첫째, 동일한 코딩을 여러 분석단위에서 실시하는 경우로, 결과의 일관성을 알아보기 위해서 이루어진다. 둘째, 일차적으로 실시한 코딩의 결과를 바탕으로 중요한 사건에 대해서 더 자세한 추가 분석을 실시하는 것이다. 예를 들어 처음에는 질문과 답변으로 코딩을 하였다가 이후 질문에 대해서 이를 하

위 유형별로 코딩하는 것이 그것이다. 마지막으로 서로 상반되는 또는 독립적인 코딩을 실시하는 방법도 존재한다. 대립적인 가설을 비교하기 위해서 또는 새로운 관점에서 자료를 분석하기 위해서 이루어진다. 예를 들어 상호작용 패턴을 분석한다고 할 때 한번은 학교장면에서 나타나는 상호작용 패턴과의 유사성을 분석하고 그 다음에는 지식 공유의 관점에서 분석할 수 있다. 또는 한번은 축어록에 나타난 문제해결의 인지적인 측면을 분석하고, 그 다음 번에는 정서적인 측면을 분석할 수 있다.

#### 코딩 및 코딩 신뢰도

본격적인 코딩을 하기 전에 예비 코딩, 즉, 일종의 파일럿 코딩을 해보는 것이 필요하다. 코딩 스킴의 초안이 정해지고 나면 이를 실제 자료 분석에 적용해 보는 것인데, 자료를 직접 코딩하는 과정에서 미리 생각하지 못한 코딩 스킴의 애매성이나 문제가 드러나게 되고 이를 점점 보완하는 기회가 되기 때문이다. 이런 과정을 거친 다음 자료를 본격적으로 코딩하는 작업이 이루어진다. 코딩은 의미처리, 범주화, 판단력 등을 요하는 작업이고, 적절하게 코딩작업을 수행하기 위해서는 다양한 배경 지식과 경험이 필요하다. 우선적으로 코딩 기법과 절차에 대한 지식과 경험이 필요하다. 예를 들어, 분석단위가 무엇인지, 어떻게 구분하는지, 코딩의 결과를 어떻게 기록하고 통합하는지 등 분석을 수행하는 데 필요한 지식이나 경험이 필요하다. 이런 종류의 지식은 그리 복잡하거나 습득에 오랜 시간이 걸리는 것이 아니어서, 복잡한 코딩이 아닌 경우 학부생도 훈련을 받으면 어렵지 않게 습득할 수 있다. 그 다음으로는 분석 대상이 되는 인지과정, 개념에 대한 지식이 필요하다. 예를 들

어 ‘설명’ ‘추론’ 또는 ‘심성모형’을 코딩하는 경우 설명이나 추론이 무엇인지, 심성모형(mental model)이 무엇인지를 개념적으로 이해할 수 있는 배경지식이 있어야 한다. 마지막으로 자료가 수집된 영역의 지식이 필요하다. 과제가 간단한 경우(예, 식인종과 선교사 문제 풀이) 일반 상식 수준의 지식만 지니고 있으면 되지만, 과제가 전문적인 경우 코딩의 위해서 상당한 양의 영역지식이 필요할 수 있다. 예를 들어 Paletz 등(2012)은 화성탐험(Mars Exploration Rover)에 참여하는 과학자들의 팀워크(teamwork)를 연구하면서 상호작용에서 나타나는 유추(analogy)의 사용과 갈등 간의 관계를 살펴보았는데, 과학자들의 대화를 이해하고 어떤 유추가 발생했는지를 코딩하기 위해서는 공학, 지질학, 기상학 등 여러 과학 분야에 대한 지식이 필수적이다. 분석을 담당하는 코더가 코딩에 필요한 모든 지식을 가지고 있는 것이 가장 바람직하나, 전문적인 지식이 요구되는 경우 코딩을 위해서 적절한 배경지식을 지닌 전문가를 코딩에 참여시키거나 도움을 받는 것이 필요하다.

언어 분석은 언어의 의미에 대한 판단과 범주화를 요구하는데, 이 작업에는 주관성이 개입할 수 있다. 이제까지 기술된 여러 분석 절차들은 모두 분석의 주관성을 줄이고 결과를 가능한 객관적으로 만들기 위한 장치들이다. 즉, 판단의 단위가 되는 분석단위를 분명하게 정의하고, 코딩 스킴을 조작적으로 정의하여 판단의 기준을 명세하고, 코더 훈련을 통해서 분석자 간 개념정의와 판단기준을 일치시키려는 노력은 모두 주관성이 개입할 여지를 줄이고자 하는 노력의 일환이다. 이러한 노력과 장치에도 불구하고 분석에 주관성이 개입할 가능성이 여전히 존재하는데, 때문에 얻어진

결과가 신뢰로운지, 즉, 분석의 결과가 특정한 코더의 주관적인 해석의 반영인지 아니면 다른 코더에게도 일관된 결과인지를 확인하는 과정이 필요하다. 코딩의 신뢰도는 상이한 코더의 코딩이 일치하는 정도를 알아보는 것으로 분석의 객관성을 확인하려는 시도이다. 통상적으로는 서로 다른 코더간의 코딩이 일치하는지를 보는 코더 간 신뢰도(inter-rater reliability)를 측정하지만, 경우에 따라서는 코더 내 신뢰도(intra-rater reliability)를 점검하기도 한다. 평정의 주관성, 변산성은 평정자 간에만 존재하는 것이 아니라 평정자 내에도 존재하고, 같은 사람이라도 언제 그 분석을 하는가에 따라서 평정이 달라지기 때문이다. 이상적으로는 두 신뢰도를 모두 계산하는 것이 바람직하나 신뢰도 계산에 걸리는 시간과 노력으로 인하여 코더 간 신뢰도만 사용하는 것이 대부분인데, 코더 내 신뢰도보다 결과의 객관성을 더 담보할 수 있는 방법이기 때문이다.

자료를 코딩할 때 전체 자료를 한 사람의 코더가 코딩하는 경우도 있고, 자료가 많은 경우 여러 사람이 나누어서 진행하기도 한다. 두 경우 모두 코딩 신뢰도의 계산이 필요하다. 우선, 한 사람이 전체 자료를 모두 코딩하는 경우 코딩의 신뢰도를 알아보기 위해서 다른 코더가 같은 코딩을 독립적으로 실시해서 코더 간 신뢰도(inter-rater reliability)를 점검한다. 이상적으로는 전체 자료를 두 명 또는 그 이상의 코더가 독립적으로 코딩하여 코딩 신뢰도를 점검하는 것이 바람직하다. 하지만 분석에 오랜 시간이 걸리는 점 때문에 자료의 일부를 대상으로 코딩의 신뢰도를 확인하는 경우가 대부분인데, 통상적으로 전체 분석 자료의 20%에 대해서 코딩 신뢰도를 점검한다. 전체 표본 수가 작은 경우 20% 규칙이 바람직

하지는 않은데, Neuendorf(2002)는 이 경우 최소한 50 사례 이상을 신뢰도 분석 표본에 포함시킬 것을 권장하고 있다. 코딩 분량이 많아서 여러 사람이 전체 자료를 나누어서 코딩을 진행하는 경우도 있는데, 이 경우에는 코딩을 시작하기 전에 서로 분석자들 간에 코딩스킴에 대한 이해 정보가 동일한지를 점검하는 과정이 필요하다. 이를 위해서는 코딩의 목적, 절차를 공유하고, 사례의 일부에 대해서 독립적으로 코딩을 진행하여 서로 견해가 일치하지 않는 부분을 확인하고 논의를 통해 이를 조정하는 것이 필요하다. 코더 간 신뢰도 (inter-rater reliability)가 적정 수준 이상 이루어지고 난 다음 독립적으로 코딩을 진행한다.

코딩 분석의 신뢰도는 코딩의 단계에서 의사결정과 판단이 일어나는 매 지점에 대해서 점검되어야 한다. 분석 단위를 코딩하는 것에 대한 신뢰도 점검은 물론이고, 분석단위를 나누는 데 의미 단서를 사용해서 주관성이 개입할 여지가 있다면 이에 대해서도 역시 신뢰도가 계산되어야 한다(Rattleff, 2007; Strijbos, Martens, Prins, & Jochems, 2006). 신뢰도 점검은 코딩 초기, 파일럿 코딩을 마친 다음에 실시하는 것이 좋다. 여러 명의 코더가 작업하는 경우에는 자연스럽게 이 작업이 코딩 초기에 일어나게 되나 혼자서 코딩을 하는 경우에는 코딩을 끝낸 후 신뢰도를 점검하는 경우가 종종 발생한다. 코딩이 다 이루어진 다음에 신뢰도를 확인하는 데는 위험부담이 존재하는데, 신뢰도가 낮은 경우 재코딩을 해야 할 수도 있기 때문이다. 또한 초반에 신뢰도를 확인하고 전체 코딩을 진행했다고 하더라도, 코딩의 양이 많거나 오랜 기간에 걸쳐 코딩이 이루어지는 경우, 자료의 특성이 바뀔 때 또는 코딩의 중간 및 마지막 시점에 신뢰도를 다시 확인하

는 작업이 필요하다. 이 경우 코더 간 신뢰도 뿐만 아니라 코더 내 신뢰도를 같이 확인하는 것이 바람직하다.

다양한 통계치가 코딩의 신뢰도 계산에 사용된다(de Wever, Schellens, Valcke, & Keer, 2006; Eye & Mun, 2005). 과거에는 각 분석단위 별로 코더 간 코딩의 일치도(%)가 많이 사용되었으나, 우연에 의한 일치도의 과대평가 문제로 인하여 최근에는 Cohen's Kappa( $\kappa$ )가 가장 많이 권장된다. 카파 값이 .75 또는 .80 이상일 때 신뢰로운 코딩으로 간주되는데, 최근에는 .80 이상의 값이 권장되고 있다. 코딩 신뢰도가 만족스럽지 않은 경우, 예를 들어 .80 이하의 값이 얻어진 경우, 불일치의 원인을 확인하고 신뢰도를 높이려는 시도가 필요하다. 우선 무엇 때문에 신뢰도가 낮은지를 확인해야 하는데, 낮은 신뢰도는 여러 원인에 기인한다. 첫째, 낮은 코딩 신뢰도는 코더가 코딩스킴을 잘못 이해 또는 적용한 결과로 발생할 수 있다. 예를 들어 A 라는 행동은 a 범주로 코딩해야 하는데, 코더의 실수로 다른 범주로 잘못 코딩한 경우이다. 이런 오류가 발생한 경우는 오류를 수정하고 신뢰도를 다시 계산하면 된다. 둘째, 코딩 스킴이 충분히 정의가 안 된 상태에서 코더들이 자의적으로 코딩을 하여 불일치가 발생한 경우이다. 해석의 주관성 보다는 필요한 약속이 코딩 스킴에 포함되지 않아서 불일치가 발생한 경우인데, 예를 들어 쉬는 시간에 발언한 내용을 분석하는가, 도구적 추론을 설명으로 볼 것인가, 또는 맥락 정보를 포함하는가 등에 대해서 코더가 서로 다른 기준을 사용했기 때문에 코딩 차이가 발생한 경우를 들 수 있다. 이 경우 정의가 부족한 부분을 찾아서 코딩 스킴을 보완하고 다시 코딩하여 신뢰도를 계산하면 된다. 이때

새로운 표본을 가지고 신뢰도를 알아보는 것이 일반적이지만, 불일치의 유형/원인에 따라 (예, 한 코더는 쉬는 시간의 발언을 포함하고 다른 코더는 제외한 경우) 문제되는 코딩만 수정하고 다시 신뢰도를 계산할 수도 있다. 마지막으로, 코딩 범주에 대한 해석이 달라서 신뢰도가 낮아지는 경우가 있다. 이 경우 원인이 매우 다양할 수 있는데, 논의를 통해 불일치의 원인을 손쉽게 찾을 수 있으면 다행이나, 원인이 불분명하거나 또는 원인을 알아도 자세하게 명세하거나 불일치를 제거하기 힘든 경우도 있다. 우선은 논의를 통해 불일치의 원인을 탐색하고 합의점을 찾는 것이 필요한데, 그 결과 코딩 스킴이 수정될 수 있다. 이 경우 포괄적인 코딩 스킴의 수정이 대부분이고, 따라서 추가적인 표본을 사용하여 다시 코딩 신뢰도를 점검하는 것이 필요하다. 여러 번 코딩 스킴을 구체화하는 노력을 기울였음에도 신뢰도가 낮은 경우가 있는데, 이는 그 개념에 대한 정의와 해석이 여러 분석자들 사이에서 일치시키기 어려운 것이거나 또는 무엇 때문에 이견이 발생하는지가 분명하지 않은 경우이다. 두 세 번의 시도를 거쳐도 기준 이상의 신뢰도가 얻어지지 않는다면 획기적인 다른 시도가 존재하지 않는 한 분석을 포기하는 것이 바람직하다.

#### 통계, 결과보고 및 분석보조도구

언어자료분석의 결과는 특정 범주에 대한 빈도 또는 점수로 종종 요약된다. 이것은 일차 코딩 범주의 빈도일 수도 있고, 심성모형 분석이나 지식 망 분석처럼 이차 분석의 결과로 얻어진 수치일 수도 있다(예, 노드의 수). 경우에 따라서는 코딩은 명목척도로 이루어지

더라도 그 결과는 비율척도로 표상될 수 있는데, 예를 들어 질문의 횟수를 코딩한다고 하면 개별 언어반응에 대해서 분석자가 내리는 결정은 각 분석 단위가 질문인가 아닌가의 범주화이지만, 참가자 별로 또는 집단 별로 결과가 합해질 때 최종 결과는 질문의 횟수가 된다. 얻어진 결과에 대해서 관찰 사례가 충분한 경우 통계분석을 실시하여 집단 간, 또는 사전, 사후 간 차이의 유의미성을 검증하는 작업이 이루어지는데, 이는 여타 심리학 연구에서처럼 얻어진 결과가 모집단에 일반화 가능한 결과인지에 대한 정보를 제공한다. 탐색적인 용도로 언어자료분석이 사용되었거나 또는 사례 수가 적어서 통계적 분석이 실시되지 못하는 경우에는 대안적인 방법을 사용하여 얻어진 결과의 타당성을 확인하는 것이 바람직하다. 예를 들어 문헌에서 보고된 결과와 비교하거나, 참가자 내에서 시간에 따른 수행 변화를 비교하거나, 사전 검사 점수를 사용해서 이후의 수행을 예측하는지 등 결과의 내적인 합치도와 정합성을 검증하는 분석을 수행하는 것이 필요하다.

심리학 연구 방법을 보고할 때 다른 연구자에 의한 반복검증이 가능하도록 하는 것이 목적이듯이 언어자료분석에서도 다른 분석가들에 의해서 결과가 반복검증 될 수 있도록 분석의 내용과 방법을 가능한 자세하기 기술하는 것이 필요하다. 어떠한 코딩 스킴과 결정 규칙이 사용되었는지는 물론, 다른 연구자들이 만든 코딩 스킴을 사용할 때는 반드시 출처를 밝히고 수정이 이루어진 경우 무엇이 왜 수정되었는지를 기술하는 것이 필요하다. 또한 코딩 스킴뿐만 아니라 코딩을 수행한 사람의 수, 수준(예, 학부생), 훈련의 정도/내용, 신뢰도 검증 방법과 절차 및 결과가 보고되어야



한다. 신뢰도는 각 분석 범주별로 보고한다.

언어자료분석은 대부분 의미처리를 요하기 때문에, 자료수집과 분석에 많은 시간이 걸린다. 특히 코딩은 시간과 노력이 많이 드는, 다소 또는 매우 지루한 작업이다. 이 작업을 기계나 프로그램이 완전히 대신할 수는 없으나, 자료 분석에 걸리는 시간을 단축하고 보다 효율적으로 만들기 위한 다양한 보조도구와 장치들이 개발되었다. 예를 들어, 녹음/녹화된 자료를 축어록으로 푸는 작업에 사용되는 도구(예, transcribing machine, foot pedal) 및 프로그램이 존재하며, 비디오 및 오디오 등 다양한 매체로 수집된 자료들을 통합해서 분석할 수 있게 해주는 프로그램(예, Tatiana; Dyke, Lund, & Girardot, 2009), 및 축어록, 동영상 자료의 코딩에 사용되는 프로그램(예, NVivo, Atlas-ti)이 존재한다. 최근에는 코딩 과정을 자동화하는 인공지능 프로그램도 개발되고 있는데(예, TagHelper; Rosé, 2008), 아직 기계가 온전히 모든 작업을 대신해 줄 수는 없으나, 조만간 분석의 상당부분이 자동화 될 수 있을 것으로 기대된다.

### 유사 기법과의 비교

언어분석과 프로토콜 분석은 상이한 연구문제를 설명하려는 시도 하에 발달하였고, 그 결과 자료 수집 방법, 분석 기법에서의 차이가 존재한다. 프로토콜 분석에서는 언어반응을 유도할 때 소리내어 생각하기 또는 말하기 같은 비지시적인 방법으로 동시적인 보고를 수집할 것을 강조하는 반면, 언어자료분석은 자료수집 방법의 제한을 두지 않는다. 분석 기법에서도 차이를 보이는데, 프로토콜 분석

은 인지적 과제분석에 기반한 모형과의 일치 여부가 주요 분석대상이었으나 언어자료분석에서는 자료에서 드러나는 패턴을 포착하는데 더 초점을 맞추었다. 이러한 차이점에도 불구하고 두 방법 간에는 많은 공통점이 존재한다. 무엇보다도 두 방법은 모두 언어자료로 대표되는 내용, 행동 자료를 체계적으로 분석하여 객관적 심리학의 방법으로 사용했는데, 이런 점에서 두 방법은 서로를 보완하는 관계로 볼 수 있다.

언어자료분석은 통상적으로 심리학 연구에서 사용되는 자기보고(예, 자신의 기분 상태 보고)와는 구분되어야 한다. 척도화되었던, 그렇지 않은 자기보고는 참여자의 주관적인 생각, 현상학적인 경험(예, 지금 얼마나 행복한가?) 자체를 있는 그대로 받아들이고 이를 분석하는 것을 목적으로 한다. 이와는 달리 프로토콜 분석이나 언어자료분석에서 언어반응을 분석한다는 것은 언어적으로 표현된 주관적 보고를 액면 그대로 받아들이는 것을 의미하지 않는다. 언어반응을 통해 드러나는 배후의 심리적 기제를 이해하는 것이 목적이고, 언어자료는 그러한 탐구 과정에서 사용되는 단서의 역할을 하기 때문이다. 예를 들어, CEO의 의사결정과정을 이해하기 위해서, 이들에게 자신이 어떻게 결정을 내리는지를 보고 하라고 하고 이들이 제공한 답변을 요약하는 분석을 할 수 있다. 언어자료분석에서 이러한 분석을 배제하지는 않지만, 이들의 자기보고를 요약하기 보다는 CEO들의 자신의 의사결정과정을 설명할 때 사용한 개념의 유형, 순서, 특징, 언급된 사건들 간의 관계를 분석하여 의사결정과정이 실제로 어떻게 진행되는가에 대한 객관적 증거를 찾아내는 것이 분석의 핵심이다.

자료 수집의 용이성 때문에 자기보고는 현재 심리학 연구에서 과도하게 사용되고 있는데, 자기보고 측정치가 필요한 연구뿐만 아니라 부적절한 경우에도 무작위적으로 사용되고 있다. 예를 들어서 참가자들이 어떤 행동을 실제로 했는지를 측정하는 것이 아니라 그런 행동을 했는지를 묻는 것이 비일비재한데, 이 경우 자기보고는 불완전한 기억뿐만 아니라 평가근심, 모니터링 오류 등의 다양한 요인으로 인하여 신뢰롭지 못한 측정치가 되기 쉽다. 참고로 Veenman, Prins, 및 Verheij(2003)은 학생들의 학습스타일(learning style)의 측정방법으로 설문지식 자기보고와 소리내어 생각하기 절차에서 얻어진 프로토콜을 비교하였는데, 자기보고식 학습스타일 측정치 보다 프로토콜을 사용한 측정치가 수행 결과를 더 잘 예측했으며, 두 측정치 간의 요인구조는 유사하였으나 상관은 매우 낮음을 보고하였다. 자기보고식 측정치가 필요한 경우가 존재한다. 참가자들의 주관적인 정서 상태나 지각, 경험을 이해하는 것이 목적인 때는 자기보고식 측정치를 사용하는 것이 필수적이다. 하지만 자료수집과 분석의 편의성만 고려하여 자기보고 자료에 과도하게 의존하는 것은 자제되어야 한다.

언어자료분석은 또한 언어의 속성에 대한 분석과는 구분되는데, 언어자료분석은 언어반응의 형식적, 문법적 특성(예, 수동태, 형태소수)에 대한 분석을 지칭하지 않는다. 언어자료 분석과정에서 언어적 속성에 대한 이해가 필요하고(예, 형태소, 문장), 언어가 마음에서 처리되는 방식에 대해서도 어느 정도의 이해가 필요하다(예, 표면구조, 심층구조). 언어분석에서도 언어의 구조나 형식적 특성이 분석될 수 있지만, 그 목적은 수집된 언어반응의 언어적 특성을 이해하는 것이 아니라 그 특성들을 통

해서 드러난 배후의 인지구조와 과정을 이해하는 것이다. 이와 관련하여 언어자료분석은 언어에 대한 단순한 양적분석과도 구분되어야 한다. 언어의 양적인 지표(예, 특정 단어가 사용된 횟수)를 양화하여 이를 다양한 심리현상(예, 우울증)과 관련짓는 연구가 존재하고 그 자체로서의 가치를 가지지만(Pennebaker, Booth, & Francis, 2007), 언어자료분석은 언어의 형식적인 단위를 단순히 세는 분석을 지칭하지 않는다(Chi, 1997). 물론 특정 단어의 발생 빈도가 분석에 포함될 수는 있으나, 언어자료분석은 언어자료의 내용에 대해 체계적인 분석을 실시하여 배후의 심리과정을 밝히는 것이다.

심리학 이외의 분야에서도 언어자료분석과 유사한 기법이 개발되어 왔는데, 대표적인 예로 내용분석(content analysis)을 들 수 있다(Holsti, 1969; Krippendorff, 2004; Neuendorf, 2002). 내용분석은 초기에 문서 등 기록된 매체의 내용을 분석하기 위해 개발되었으나, 최근에는 영화, TV, 인터넷 등 다양한 매체의 내용을 분석하는 용도로도 확장되고 있다. 전통적으로 언어자료분석이나 프로토콜 분석이 구어, 발화를 중심으로 이루어진 반면 내용분석은 주로 기록된 매체를 대상으로 했다는 면에서 차이가 있다. 또한 연구 주제 면에서도 달랐는데, 언어분석이 심리학 내의 발달, 인지, 사회 등의 연구 주제를 주로 다루어 온 반면, 내용분석은 주로 심리학 외의 분야에서 사회, 정치, 문화적인 연구 주제를 연구하는 데 사용되어져 왔다. 연구 기법 상에서도 차이가 존재하는데, 특히 초기의 내용분석은 특정 단어의 발생 빈도를 알아보는 등의 내용 자료의 표면을 주로 분석하였으나, 최근에 들어서는 자료의 배후의 내재적인 의미에 대한 분석도 실시하고 있다(Krippendorff, 2004; Neuendorf,

2002). 내용분석에는 양적인 전통과 질적인 전통이 공존하는데(Hsiu-Fang, H., & Shannon, 2005), 양적인 내용분석에서는 언어자료분석과 유사하게 코딩을 실시하고 신뢰도를 계산하는 등 분석 기법 상에 있어서는 경계가 모호한 경우도 종종 존재한다.

마지막으로, 언어자료분석은 질적인 분석과 구분되어야 한다. 질문에 대한 답변, 대화 등에 대한 분석이 질적인 연구에서도 이루어지는 까닭으로 언어자료분석을 질적분석 방법으로 오해하거나 언어자료를 분석하면 모두 질적인 연구라고 생각하는 오해가 종종 존재한다. 그러나 질적 연구와 양적 연구를 구분하는 것은 어떤 자료를 분석하는가에 의해서 결정되는 것이 아니고, 동일하게 언어자료가 분석된다고 하더라도, 자료에 대한 접근 및 분석방법에 따라서 질적인 분석과 양적인 분석이 구분된다. 강조하는 측면과 기법에서 차이를 보이기는 하지만 프로토콜, 언어자료분석 등의 방법은 모두 실증주의적인, 양적인 연구 전통 하에 놓여 있는 분석기법들이다. 이들 기법은 모두 현상에 대한 객관적인 탐구를 강조하고, 반복 가능하고, 일반화 가능한 결과를 지향한다. 사례 수가 충분하지 않아서 통계적 검증이 실시되지 않을 수 있지만, 가능한 경우에는 모집단에서의 일반화 가능성이 적극적으로 검증된다(Shadish, Cook, & Campbell, 2002). 그러나 질적연구의 전통은 이러한 실증주의적 전통과는 다른 입장을 취하고 있는데, 다양한 질적 분석방법이 존재하고(Gee & Green, 1998; Jordan & Henderson, 1995; Sacks, 1992; Straus & Corbin, 1980) 이들 분석 방법을 다 개관할 수는 없으나, 여러 질적인 연구 전통에서 공통된 입장 중의 하나는 양적인 연구전통에서 가정하는 객관적인 실재를 인정하지 않는 것이

다(조용환, 1999). 질적인 연구 전통에서는 진정한 객관성은 불가능하다는 인식론적 입장을 취하는데, 따라서 연구에 있어서도 일반화 가능한 결과보다는 특정 사례의 독특성과 주관성을 탐색하는 것을 목적으로 한다(Bryman, 1984). 분석기법 또한 양적인 방법에 의존하기 보다는 현상학적, 개념적 기술을 중시한다. 코딩이 실시되기도 하지만, 그 결과는 통계적 처리의 대상이 되지 않고 모집단에 일반화 가능한지는 고려 대상이 되지 않는다(질적인 분석에 대한 보다 자세한 개관 및 양적인 분석, 혼합 분석과의 차이를 위해서는 조용환, 1999 및 Johnson & Christensen, 2008 참고).

## 결 론

본 논문에서는 언어자료분석을 중심으로 심리학 연구에서 언어반응이 연구에 사용된 배경과 분석 사례를 살펴보고, 분석의 주요 절차와 고려점을 연구 단계별로 살펴보았다. 언어반응은 인지과정을 포함한 다양한 심리과정에 대한 유용한 정보를 제공하고, 심리학의 여러 영역에서 중요한 발견을 하는데 공헌하였다. 언어자료분석은 처음에 개인의 인지를 연구하려는 시도 하에서, 특히 문제풀이와 학습과정 동안에 참가자들이 사용하는 방략과 표상의 변화를 알아보려는 목적으로 개발되었다. 연구방법론은 연구 주제와 함께 진화하는데, 전문성, 디자인 등은 물론, 최근에는 상호작용과 그룹인지에 대한 관심이 증가하면서 집단의 인지과정에 대한 물음을 답하기 위해서도 적극적으로 사용되고 있다(Jeong & Chi, 2007; van Aalst, 2009). 또한 디지털 테크놀로지 사용 및 가상공간의 연구에도 적극적으로

적용되고 있다(de Wever et al., 2006; Jeong, Chen, & Looi, 2011; Schneider, Passant, & Bresslin, 2010).

언어자료분석은 전통적인 실험 연구나 설문지 연구에 익숙한 연구자들에게는 아직 낯선 방법이지만, 심리학 연구에 보다 적극적으로 사용될 필요가 있다. 언어자료분석의 가장 큰 장점 중의 하나는 연구의 지평을 넓혀 준다는 데 있다. 언어반응을 분석함으로써 설문지를 통한 척도화된 반응, 반응시간이나 정반응률로는 포착하기 어려운 다양한 심리과정에 대한 연구가 가능하다. 언어에 의해서 매개되거나 언어반응에 심리과정이 드러나는 다양한 심리현상은 물론이고, 언어가 아닌 유사한 다른 행동 자료(예, 동영상)의 분석도 가능하다(Chan et al., 2011). 또 다른 장점은 앞서서도 언급하였듯이 배후의 심리과정에 대해 보다 더 정확한 정보를 얻는 것이 가능하다는 것이다(Veenman, Prins, & Verheij, 2003). 특히 단답식 검사나 자기보고식 설문지와 비교할 때 언어보고는 종종 기저의 심리과정에 대한 더 정확한 정보를 제공하며, 그 외의 다른 경우에도 기존의 측정치를 보완하여 연구문제에 대한 더 풍부한 정보를 제공해 주는 역할을 한다. 마지막으로 언어자료분석은 탐색적 연구의 유용한 도구가 될 수 있는데, 연구 가설을 구체화할 충분한 선행연구가 존재하지 않는 경우 연구 문제를 다각도에서 점검하고 관련 변인을 탐색하는데 유용한 도구가 된다.

이러한 장점에도 불구하고 언어자료를 연구에 사용하는데 따르는 몇 가지 단점이 존재한다. 가장 큰 단점은 분석에 오랜 시간이 걸린다는 것이다. 설문지 연구에서는 한 번에 다수의 참가자로부터 자료를 수집할 수 있고, 상대적으로 짧은 시간에 분석을 완료할 수 있

지만, 언어자료분석을 사용하는 경우 자료 수집부터 시작해서 이를 축어록으로 변환하고 분석을 실시하는데 많은 시간과 노력이 요구된다. 인터넷 등에서 이미 가용한 자료를 사용하는 경우 수집에 걸리는 시간이 단축될 수 있지만, 여전히 분석에 걸리는 시간을 무시할 없다. 또한 분석을 위해 자료의 의미, 내용을 처리하지 않으면 안 되기 때문에(예, 참가자가 말한 내용이 특정 주제에 관한 것인지 아닌지) 해석의 주관성의 문제를 완전히 제거할 수는 없다. 하지만 분석의 절차를 충실하게 따른다면(예, 조작적 정의, 신뢰도) 해석의 주관성 문제는 크게 우려할 정도는 아닌 것으로 보인다. 인지발달, 전문성, 학습 등 여러 분야에서 언어자료분석을 사용하여 얻어진 연구 결과들은 후속 연구에서 반복 검증되어왔고 학계의 주요 연구결과로 자기 매김하고 있기 때문이다(예, Chi et al., 1989, 1994; Simon & Simon, 1978). 또한 이러한 단점을 보완하기 위한 다양한 시도가 존재하는데, 자동화된 분석 알고리즘과 보조 도구들의 발달로 초보적인 분석에 걸리는 시간은 머지않아 상당히 단축될 수 있을 것으로 예상된다.

모든 연구에서 마찬가지이지만, 연구방법은 연구 질문/이론과 밀접한 관계를 갖는다. 연구를 안내하는 가장 큰 축은 연구 질문/이론적 접근이고, 방법론은 연구에서 알아보고자 하는 것을 구현하는 도구 역할을 한다. 연구 질문이 없이 방법론만으로 연구를 할 수는 없듯이 언어자료분석도 마찬가지이다. 방법론에 대한 지식은 연구 질문을 대신하지도 않고, 무소불위의 도구도 아니다. 언어자료분석이 유용한 도구이기는 하지만, 연구방법론의 가치는 그 자체로 존재하는 것이 아니라 연구 질문과의 관계 속에서 존재한다. 연구방법을

정하기 전에 연구 질문과 가설에 대한 고민이 우선이고, 이를 바탕으로 연구 질문에 답하는데 가장 적절한 연구방법론을 선정하는 것이 필요하다. 본 논문은 그 동안 적절한 방법론의 부재로 연구에 어려움을 느낀 연구자들에게, 또한 연구를 시작하는 단계에서 어떠한 방법이 가능한지를 탐색하는 연구자들에게 유용한 지침을 제공하기 위해 작성되었다. 본 논문에서 소개한 언어자료분석이 다양한 분야에서 새로운 연구를 촉발시키는 역할을 하기를, 그리하여 무엇보다도 한국 심리학 연구의 지평을 확대하는 데 공헌할 수 있기를 기대한다.

### 참고문헌

- 이정모. (2001). 인지심리학: 형성사, 개념적 기초, 조망. 아카넷. 서울.
- 조명한, 이정모, 김정오, 신현정, 이광오, 도경수, 이양, 이현진, 김영진, 김소영, 고성룡, & 정혜선. (2003). 언어 심리학. 학지사. 서울.
- 조용환. (1999). 질적연구: 방법과 사례. 교육과학사. 서울.
- Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, 96(3), 523-535.
- Baker, M., Andriessen, J., Lund, K., van Amelsvoort, M., & Quignard, M. (2007). Rainbow: A framework for analyzing computer-mediated pedagogical debates. *International Journal of Computer-Supported Collaborative Learning*, 2, 315-357.
- Benbunan-Fich, R. (2001). Using protocol analysis to evaluate the usability of a commercial web site. *Information & Management*, 39, 151-163.
- Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, 17(5), 564-577.
- Bosshardt, H.-G. (2006). Cognitive processing load as a determinant of stuttering: summary of a research programme. *Clinical Linguistics & Phonetics*, 20(5), 371-85.
- Bryman, A. (1984). The debate about quantitative and qualitative research: A question of method or epistemology. *The British Journal of Sociology*, 35(1), 75-92.
- Chan, J., Fu, K., Schunn, C., Cagan, J., Wood, K., & Kotovsky, K. (2011). On the benefits and pitfalls of analogies for innovative design: Ideation performance based on analogical distance, commonness, and modality of examples. *Journal of Mechanical Design*, 133(8), 081004.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Science*, 6(3), 271-315.
- Chi, M. T. H. (2006). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. Feltovich & R. Hoffman (Eds.), *Cambridge Handbook of Expertise and Expert Performance* (pp.167-184). Cambridge: Cambridge University Press.
- Chi, M. T. H., & Koeske, R. D. (1983). Network

- representation of a child's dinosaur knowledge. *Developmental Psychology*, 19(1), 29-39.
- Chi, M. T. H., Bassok, M., Lewis, M. T., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145-182.
- Chi, M. T. H., de Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18, 439-477.
- Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84, 73-111.
- Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The effect of instructional explanations on learning from scientific texts. *The Journal of the Learning Sciences*, 6(4), 347-365.
- De Wever, B., Schellens, T., Valcke, M., & van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers and Education*, 46, 6-28.
- Dunbar, K. (1997). How Scientists Think: On-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes*. (pp.461-493). Washington: American Psychological Association.
- Dyke, G., Lund, K., & Girardot, J.-J. (2009). *Tatiana: An environment to support the CSCL analysis process*. Paper presented at the International Conference of the Learning Sciences, Rhodes, Greece.
- Ericsson, K. A. (2000). Expertise in interpreting: An expert-performance perspective. *Interpreting*, 5(2), 187-220.
- Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. Feltovich & R. Hoffman (Eds.), *Cambridge Handbook of Expertise and Expert Performance*. Cambridge: Cambridge University Press.
- Ericsson, K. A., Delaney, P. F., & Weaver, G. (2004). Uncovering the structure of a memorist's superior "basic" memory capacity. *Cognitive Psychology*, 49, 191-237.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal report as data. *Psychological Review*, 87(3), 215-251.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal report as data*. Cambridge, MA: The MIT Press.
- Eye, A. V., & Mun, E. Y. (2005). *Analyzing rater agreement*. Mahwah, NJ: Earlbaum.
- Flower, L. S., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387.
- Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, 137(2), 316-44.
- Gee, J. P., & Green, J. L. (1998). Discourse analysis, learning, and school practice: A methodological study. In *Review of Research in Education* (Vol. 23). Washington, DC:

- AERA.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of The Learning Science*, 4(1), 39-103.
- Hoffman, R. R., Shadbolt, N. R., Burton, A. M., & Klein, G. (1995). Eliciting knowledge from experts: A methodological analysis. *Organizational Behavior and Human Decision Processes*, 62, 129-158.
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Menlo Park, CA: Addison-Wesley Publishing Company.
- Hsiu-Fang, H., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp.13-31). Amsterdam: John Benjamin Publishing Company.
- Jeong, H. (2013). Verbal data analysis for understanding interactions. In C. E. Hmelo-Silver, A. M. O'Donnell, C. Chan, & C. A. Chinn (Eds.), *The international handbook of collaborative learning*. London: Taylor & Francis.
- Jeong, H., Chen, W., & Looi, C. K. (2011). Analysis of group understanding in artifact-mediated discourse. In *the Proceedings of Computer-Support for Collaborative Learning*.
- Jeong, H., & Chi, M. H. (2007). Knowledge convergence and collaborative learning. *Instructional Science*, 35, 287-315.
- Johnson, R., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches*. Thousand Oaks, CA: Sage.
- Kaplan, C. K. , & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, 22, 374-419.
- Kimmerle, J., Moskaliuk, J., & Cress, U. (2011). Using Wikis for Learning and Knowledge Building: Results of an Experimental Study. *Educational Technology & Society*, 14, 138-148.
- Krampe, R. T., & Charness, N. (2006). Aging and expertise. In K. A. Ericsson, N. Charness, P. Feltovich, & R. Hoffman (Eds.), *Cambridge Handbook of Expertise and Expert Performance* (pp.723-742). Cambridge, UK: Cambridge University Press.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage Publications.
- McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, 38(1), 1-30.
- Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning*, 2, 63-86.
- Neuendorf, K. A. (2002). *The content analysis: Guidebook*. London: Thousand Oaks.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231-259.
- Opfer, J. E., & Siegler, R. S. (2004). Revisiting preschoolers' living things concept: A microgenetic analysis of conceptual change in

- basic biology. *Cognitive Psychology*, 49, 301-332.
- Okada, T., & Simon, H. A. (1997). Collaborative discovery in a scientific domain. *Cognitive Sciences*, 21(2), 109-146.
- Paletz, S. B. F., Schunn, C. D., & Kim, K. H. (2012). The interplay of conflict and analogy in multidisciplinary teams. *Cognition*, 126, 1-19.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, 16(2), 366-387.
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Operator's manual: Linguistic inquiry and word count: LIWC2007. Austin, TX: LIWC.net.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: our words, our selves. *Annual Review of Psychology*, 54, 547-77.
- Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Erlbaum.
- Rattleff, P. (2007). The reliability of content analysis of computer conference communication. *Computers and Education*, 49, 230-242.
- Renkl, A. (1997). Learning from worked-out examples: A study on individual difference. *Cognitive Science*, 21(1), 1-29.
- Renkl, A. (2002). Worked-out examples: Instructional explanations support learning by self-explanations. *Learning and Instruction*, 12, 529-556.
- Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology*, 23(23), 90-108.
- Rittle-johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction, *Child Development*, 77(1), 1-15.
- Roscoe, R. D. (2013). Self-monitoring and knowledge-building in learning by teaching. *Instructional Science*, 42(3), 327-351.
- Rosé, C. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, 3, 237-271.
- Sacks, H. (1992). Lectures on conversation. Malden, MA: Blackwell.
- Schneider, J., Passant, A., & Breslin, J. G. (2011). Understanding and Improving Wikipedia Article Discussion Spaces. In SAC'11.
- Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 22(1), 36-71.
- Schooler, J. S., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, 122(2), 166-183.
- Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, 28, 225-273.
- Simon, D. P., & Simon, H. A. (1978). Individual differences in problem solving physics problem.



- In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum.
- Stempfle, J., & Badke-Schaub, P. (2002). Thinking in design teams - an analysis of team communication. *Design Studies*, 23(5), 473-496.
- Straus, A., & Corbin, J. (1990). Basics of qualitative research. Newbury Park: Sage publication.
- Strijbos, J., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers and Education*, 46, 29-48.
- Suwa, M., & Tversky, B. (1997). What do architects and students perceive in their design sketches? A protocol analysis. *Design Studies*, 18(4), 385-403.
- van Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourse. *International Journal of Computer-Supported Collaborative Learning*, 4, 259-287.
- Veenman, M. V. J., Prins, F. J., & Verheij, J. (2003). Learning styles: self-reports versus thinking-aloud measures. *The British Journal of Educational Psychology*, 73(Pt 3), 357-72.
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24, 535-585.
- Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83(1), 73-87.
- Wineburg, S. (1998). Reading Abraham Lincoln: An expert/expert study in the interpretation of historical texts. *Cognitive Science*, 22(3), 319-346.

1차원고접수 : 2014. 01. 20.

수정원고접수 : 2014. 06. 12.

최종게재결정 : 2014. 07. 21.

## How to Make Use of Verbal Data in Psychological Research?

**Heisawn Jeong**

Department of Psychology, Hallym University

Many of the psychological processes are revealed in language responses. Verbal reports in the form of Think-Aloud protocols, self-explanations, or responses to various questions have been successfully used to understand psychological processes that are difficult to capture by utilizing traditional measures such as reaction time and Likert-scale responses. The goal of this paper is to introduce a particular method to analyze verbal data for scientific research of psychology. Called as verbal data analysis, this method is designed to systematically analyze verbal data to produce objective and quantitative results. The paper describes core features of the method along with research examples in which verbal data analysis was successfully applied. Detailed procedures of the analysis are then described with issues that researchers need to be aware of at each stage of data collection and analysis. Lastly, verbal data analysis is compared to other similar methods. Strengths and weaknesses of the method are also discussed.

*Key words* : verbal data analysis, protocol analysis, verbal reports, data, quantitative, analysis