

읽기에서의 어휘 시소러스(thesaurus)의 응용*

조형일**

< 차례 >

- I. 디딤-읽기, 읽어내기
- II. 뺏음-시소러스(thesaurus)와 읽기
 - 1. 시소러스(thesaurus)와 어휘 해석
 - 2. 시냅스형 시소러스(Synaptic Thesaurus system)
 - 3. 읽기에서의 시냅스형 시소러스(STs)의 응용 방안
 - 4. 읽기에의 시소러스 응용의 실제
- III. 닫음

I. 디딤-읽기, 읽어내기

텍스트가 전제된 언어의 이해교육 영역 관점에서 ‘읽기’는 전통적으로 ‘독서’가 함의하는 교육적 효과는 물론 이 행위의 근원적 목표가 되는 지식의 습득과 함께 이해, 논리적 탐구, 지적 향유라는, 다시 말해서 ‘읽기 행위’를 통해서 얻을 수 있는 총체적인 결과물이 상정된 교육 행위의 통칭으로 볼 수 있다. 읽기 영역에 대해서 국어교육에서는, 읽기 활동을 통해서 얻고자 하는 것과 얻을 수 있는 것에 대한 논의로부터 시작해서, 읽기의 내용·방법·평가 항목 등에 대한 논의의 연쇄 안에서 그 위계를 파

* 이 논문은 제48회 국어교육학회 전국학술발표대회(2011. 4. 9.)에서 발표한 것을 수정·보완한 것이다.

** 가톨릭대학교 국어국문학전공, 교육전담초빙교수(jeongam@gmail.com)

악할 수 있을 것이다.

이 연구에서는 읽기의 위계에 대한 논의는 잠시 미뤄 두고, 국어교육에서의 읽기 자료의 구축과 읽기의 교육 효과 달성을 위한 교수·학습 방법에 직접적으로 적용 가능한 ‘시소러스 응용 방안’을 논의해 보았다. 이 논의가 다시 회귀적으로 읽기의 위계를 정하는 데에는 크게 기여할 것 같지는 않지만, 읽기에 대한 교육 방법을 풍성하게 만들어 줄 수 있을 것으로 기대한다.

일반적으로 어휘를 텍스트 이해의 주요 요소로 보고 이것이 텍스트 이독성(readabilities)을 결정짓는 핵심 요소의 하나로 파악하는 데에 주저하지 않는다. 하지만 이와 동시에 어휘가 과연 ‘어떻게’, 그리고 모든 상황에 ‘반드시’ 관여하는지에 대해서 의구심이 드는 것도 사실이다. 어휘를 텍스트의 난이도를 결정하는 핵심 요소로 파악하려면 필연코 어휘를 어휘 이상의 층위에서 파악해야 하는데, 문장의 연쇄적 구조 안에서 논항 간 상호 참조 되는 양상 즉, 규칙 지배(rule governed)를 받는 정도를 명확하게 예측해 내기란 쉬운 일이 아니다. 이러한 난점을 해소할 수 있는 방안으로 시소러스(thesaurus) 즉, ‘목적성 있게 가공된 분류어휘집’의 구축과 활용을 제안해 볼 수 있다. 읽기 텍스트의 해석은 물론 생산에서부터 관여하는 어휘 목록이 제공된다면, 이것은 당연히 텍스트 이독성/난이도 측정의 지표가 될 수 있다) 이 연구에서 제안하고 있는 시냅스형 시소러스(Synaptic Thesaurus system : 이하 STs)는 부족하나마 어느 정도 이를 충족시켜 줄 수 있을 것이다.

분류어휘집인 시소러스라는 개념을 읽기에 적용하기에 앞서, 읽기에 대해서 간단히 개념을 정리해 둘 필요가 있어 보인다. 주지하듯이 텍스트에 ‘목표된 지식’ 또는 ‘구축된 지식’)은 단 한 번의 읽기 행위만으로 즉

-
- 1) 그러므로 이 연구에서는 시소러스를, 읽기 행위가 지향하는 이해와 지식 습득의 앵커(anchor)이자 접점(node)으로 기능하는 것으로 본다.
 - 2) 이런 표현을 쓸 수 있을지 모르겠다. 이에서 이야기하고자 하는 ‘텍스트에 목표된 지식’이란 읽기 행위를 염두에 둔, 텍스트가 궁극적으로 제공하고자 하는 지식을 말하며, ‘텍스트에 구축된 지식’이란 텍스트 자체의 이독성에 관여하는 지식으로 보았다.

각적으로 이해되거나 행위자의 이해 영역으로 체계화되기 어렵다. 따라서 읽기는 텍스트의 구축에서부터 독자의 수준을 염두에 둘 수밖에 없다.

읽기의 중요성에 대한 논의는 다음의 천경록·이재승(1997)에서 정리된 것으로 충분해 보인다.

◎ 읽기의 중요성

- (1) 읽기를 통해 정보(지식)를 얻을 수 있다.
- (2) 읽기를 통해 문화를 전수하고 유지, 발전시키게 된다.
- (3) 읽기는 사고력을 기르는 한 방법이다. 글을 읽는 과정에서 독자는 글의 내용을 분석, 종합, 비판하는 과정을 밟게 되고 이 과정에서 독자는 사고력을 기르게 된다.
- (4) 읽기를 통해 정서를 함양할 수 있다.
- (5) 읽기는 그 자체가 목적일 뿐만 아니라 다른 학습을 하는 도구가 된다.
- (6) 읽기는 언어 발달을 가져온다. 글을 읽음으로써 많은 어휘와 지식을 습득하게 되는데, 이는 읽는 사람의 언어를 발달시킨다.
- (7) 다른 사람과 의사소통을 원만히 하는 데도 읽기는 필요하다.

이처럼 읽기는 종합적 사고와 그 과정에 대한 본질적 접근을 가능하게 해 주는 학습 영역이라고 할 수 있다. 그렇다면 읽기의 교육 목적을 실현하기 위한 교육 방법을 구안하기 위해서는 우선 읽기 텍스트의 생산 단계에서부터 적절한 교육적 안배가 필요하다고 볼 수 있다. 읽기의 교육성과를 고려하여 다양한 텍스트를 선정하고 교육적으로 활용 가능하도록 가공하는 것과 함께, 처음부터 교육적 활용 가능성을 고려하여 텍스트를 생산할 수 있어야 하는 것이다. 바로 이때 시소러스를 응용할 수 있다.³⁾⁴⁾

3) 읽기의 대상이 되는 텍스트에 대한 이해는 문식성(literacy)의 측면에서 접근되는 것이 옳다. ‘정보와 언어 처리 기능의 습득이라는 관점에서 개인의 스키마를 동원한 의미 구성 능력과 텍스트에 대한 독자의 반응이 강조되고 증시되는 것(박인기, 2002)’, ‘문자 언어 식별 기능 수준을 넘어선 매체 언어와 인간의 사고력을 설명하기 위한 개념으로, 고등 수준의 비판적 문식성이 초점화되고 있음(박수자, 2003)’ 등에서 보이듯 문식성은 교육 텍스트 또는 텍스트를 활용한 교육에서 중요한 측면이라고 볼 수 있다.

4) 이처럼 텍스트의 생산 단계에서 적용 가능한 어휘와 문형 자료를 시소러스로 구축하고 이를 응용하는 방안은 앞으로 텍스트 생산에 상당한 도움을 줄 수 있을 것이다. 따라서

II. 뽀뽀 - 시소러스(thesaurus)와 읽기

1. 시소러스(thesaurus)와 어휘 해석

시소러스란 ‘분류어휘집’을 말한다. 이는 간단히 말해서 목적성 있게 가공된 어휘 목록이라고 할 수 있다. 도서관에서 서지사항별로 책을 분류하는 것에서부터 인터넷 검색용 페이지를 데이터베이스로 구축하는 것에 이르기까지 시소러스의 개념은 매우 유용하게 활용될 수 있다. 조형일(2010)에서는 시소러스를 다음처럼 정리해 놓고 있다.

- 1) 자료의 탐색 시에 질의어(query)에 포함된 용어의 의미를 연계성을 기반으로 확대하여 사용자가 원하는 정보를 찾게 해 주는 색인으로써 어휘를 분류해 놓은 것.
- 2) 특정 분야에서 활용하기 위해 의도적·제한적으로 용어를 정리해 놓은 분류어휘집.
- 3) 체계적인 분류와 함께 분류된 대상 간의 관계에 대한 표시를 명확하게 나타낸 것.
- 4) 논리적인 분석을 통한 기준이 필요하고 이를 위해 필산규칙인 알고리즘을 고려한 설계가 필수적으로 수반되어야 하는 것.

시소러스를 구축한다는 것은 어휘 간의 정보를 결합하는 것으로 볼 수 있다. 따라서 제한된 영역에서 시소러스는 매우 유용한 수단으로 기능하게 된다. 예를 들어서 어떤 인터넷 검색기 검색창에 질의어(query)를 넣었을 때, 그 질의어에 가장 적당한 원문 텍스트를 찾아서 보여 주는 기법은 시소러스 때문에 구현 가능한 기술이다. 이때 질의어의 다양한 변수를 예측하고 시소러스를 복잡하게 구축하면 할수록 검색 결과가 좋아지는 것은 당연하다. 그러므로 시소러스를 구축할 때 4번 항목이 가장 중요한 원

이제부터 제안하고자 하는 시소러스는 단순한 어휘집 수준이 아닌 복합적 시각에서 읽기 텍스트의 해석과 생산에 직접적으로 적용 가능한 개념으로 제안될 것이다.

리가 된다. 실재(實在)하는 정보를 찾아서 연결시켜 주는 간단한 방식의 검색 엔진(retrieval engine)을 구축하기 위해서도 상당히 복잡한 기준과 알고리즘이 필요하다. 따라서 읽기 교육을 위해서 텍스트의 난이도를 판별하고, 텍스트 생산 단계에서부터 적용 가능한 시소러스를 구축하는 것은 상당히 복잡한 기준을 세우고 이에 걸맞은 알고리즘을 구축하는 고된(?) 작업이 된다.

지금까지의 시소러스는 그 유용성에 비해서 다소 제한적 영역에서 응용되어 왔다고 판단할 수 있다. 비록 주제별 분류와 함께 어휘장, 동음이의, 유의, 상하의, 반의 등의 의미 관계에 따른 어휘 분석 기준을 이용하고 있기는 하지만 그것만으로는 시소러스의 진가를 드러내기에는 부족해 보인다.⁵⁾

다음의 <표 1>은 한국어 시소러스 및 시소러스 개발 기준을 조형일(2010 : 63)에서 정리해 놓은 것이다.

<표 1> 시소러스 개발 지침과 그 내용

시소러스 개발 지침	시소러스 관계
한국자료베이스진흥센터 (2000)	대등관계 : 우선어와 비우선어의 등가어 집합 형성, 유사어, 유사 동의어 구분
	계층관계 : 상·하위어, 전체·부분 관계
	연관관계 : 동일 범주/상이 범주
ISO가이드라인(2788 : 단일어, 5964 다국어)	유사관계 : 동의어/유사 동의어
	계층관계 : 종속 관계, 전체-부분 관계
	연관관계 : 유사/계층 관계에 속하지 않는 관계
ANSI가이드라인(2003)	대등관계 : 동의어, 어휘 파생어, 이형 동의어, 유사어, 복합명사(상호 참조 가능 시)
	계층관계 : 상·하위어, 종속/전체-부분, 사례, 다층 관계
	연관관계 : 동일 범주 간/타 범주 동시 출현, 관련어, 기계가독 특정 코드 구분

5) 사실 우리가 지니고 있는 추상적 언어 체계와 능력은 개인별 시소러스 체계와 그 수준으로 바꾸어 말할 수 있다. 추상적으로 동일하다고 믿는 언어 체계와 능력은 사실 저마다 각자 방식으로 다시 체계화되어 추상의 영역에 실재하게 되는데 이를 시소러스 체계로 바꾸어 부를 수 있을 것이다.

시소러스 개발 지침	시소러스 관계
UNESCO지침(1981)	대등관계 : 우선어/비우선어, 동의어/유사 동의어
	계층관계 : 최상위어/파생 하위어, 속성 관계/전체-부분, 사례 관계
	연관관계 : 군집, 동일 범주/이 범주의 관련성 고려
	상호범주관계 : 범주의 복수성 고려, 특정 주제

이에서 보이듯 시소러스 개발을 위해 마련해 놓은 지침은 서로 대등 소이하다. 어휘의 대등성과 유사성 그리고 계층성, 연관성 및 상호 범주 간의 관계를 중심으로 크게 구분되어 있는 것이다. 이러한 지침은 일반적으로 언어학, 국어학을 전공한 사람이라면 누구나 알 수 있는 수준인데, 문제는 이들을 어떻게, 어느 정도 수준으로 판별하고 조직화할 것인가와 과연 이것만으로도 충분한가에 달려 있다. 이는 전산학에서의 성과와 국어학에서의 연구 성과가 적절한 수준에서 상호 수용되어야 하는 부분이라고 하겠다.

어휘의 의미 관계에 따른 상호 참조식의 시소러스는 어휘의 의미 속성에 전적으로 기댈 수밖에 없다. 중언부언 같지만 그러하기에 어휘의 의미 속성의 범위를 벗어나기 어렵다. 이렇게 구축된 시소러스는 제한된 상황에서만 유용할 뿐이다. 어휘를 색인용 노드(node)나 앵커(anchor) 이상으로 볼 수 없게 되는 것이다.

어휘는 내용과 속성, 기능 등의 다양한 층위에서 해석되어야 한다. 김광해(1995 : 229~234)에서는 “단어의 의미에 대한 설명을 위해서는 어떤 한 어휘소와 다른 어휘소가 의미상으로 관계를 맺는다는 것에 대한 고찰이 필요하다.”고 했다. 이러한 고찰은 다시 생각과 말과 세계가 연결된 삼각도식으로 부연 설명되는데, 이에서는 이들 간의 상호 관련성의 정도에 따라서 어휘의 의미관계에 대한 기술이 가능해질 것으로 보고 있다.

이러한 논의에 기대보면 어휘의 의미 관계에 따른 ‘단순 참조식’ 시소러스의 구축은 그 자체가 말이 안 되는 개념이 된다. 생각과 말과 세계는 각자 또는 상호 조응하면서 상대적인 상황 내에서 다양한 변인의 결과물

로 표출될 것이다. 이러한 복합적 결과물은 표면적으로는 어휘 연쇄와 문장 구조로 드러날 것이지만 필경 상황과 관계 등의 요소에 영향을 받으면서 함께 사유의 철학적 언명까지 내재하고 있게 된다. 그러므로 이에 대한 완벽한 해석은 애초부터 요원한 일일지도 모른다. 하지만 어느 정도 설명 타당한 방안을 모색하는 노력을 지속적으로 해야 하는 것도 사실이다. 그러므로 이제 어휘의 의미 관계 즉, 의미 속성에 대한 관계의 기술을 넘어서는 개념이 필요해졌다.

어휘의 의미 관계에 대한 논의를 포함하는 동시에 이를 넘어설 수 있는 적절한 개념으로는 어휘의 자릿값을 들 수 있다. 일반적으로 이는 문장 안에서 하나의 논항인 어휘가 필요충분 성분으로 요구하는 다른 논항의 개수를 말한다. 그런데 이 연구에서는 하나의 논항으로 선택되는 어휘를 출발점으로 삼고 이와 결합하는 문장형식과 다른 어휘가 제시되는 개념으로 본다. 어휘의 자릿값은 어휘가 변수가 되는 함수로 해석해야 하고 이때의 등식은 당연히 방향성을 갖게 된다. 이는 ‘단순 참조식’의 시소러스가 아닌 인지속성과 유사한 시소러스를 구축하기 위한 개념인데, 이 연구에서는 이를 시냅스형 시소러스(STs)로 제안했다.

2. 시냅스형 시소러스(Synaptic Thesaurus system)⁶⁾

시냅스형 시소러스(STs)란 졸고(2010)에서 한국어교육을 위해 고안 제시한 개념이다. 이는 교육용 어휘와 문형을 난이도별로 선정·배열하고 이를 교육적으로 활용 가능한 기준에 맞게 세분화하여 학습 단위로 상정한 것이다. 원래 이 개념은 한국어교육에서 중점적으로 이루어지고 있는 문형 중심의 교육이 갖는 한계를 어휘로 극복해 보고자 설정한 것이었다. 따라서 이는 어휘를 교육적으로 활용 가능하게 분류하여 제시한 어휘 체계라고 할 수 있다.⁷⁾ 그런데 한국어교육에서가 아닌 국어교육의 관점에서

6) 이 부분은 조형일(2010)에서의 내용을 이 논문의 취지에 맞게 요약 발췌하여 수정하였음을 밝힌다.

이를 응용하면 전술한 바와 같이 텍스트의 생성 및 난이도 판단을 할 수 있는 하나의 평가 기준을 마련할 수 있게 된다. 이에 대한 논리적 전개를 위해서 우선 시냅스형 시소러스(STs)의 배경과 원리, 특징을 제시하면 다음과 같다.

- 1) [배경] 각각의 정보가 하나의 연결점(node)으로 접점을 이루면서 참조되는 개념인 신경망 구조는 언어 표현의 다양한 가능성을 사전에 차단해 버리는 오류를 범할 수 있다. 선·후행 한 두 단어를 참조하는 것만으로 목표어와 동형어의 관계나 다의 관계에 있는 어휘를 분명하게 구분할 수 있다고 보기 어렵고 상호 참조되는 어휘 각각의 세밀한 차이를 설명해 주지도 못한다. 따라서 교육적으로 활용 가능한 시소러스의 구조는 신경망 구조보다 조금 더 복잡한 구조를 가질 수 있어야 한다.
- 2) [원리] 신경망 구조 대신 시냅스(synapse) 간극을 설정하여 어휘 각각에 부여되는 조건에 따라서 수의적으로 선택할 수 있도록 어휘 간의 정보 체계를 열어 두는 것을 생각해 볼 수 있다. 교육적으로 활용 가능한 시소러스는 어휘에 대한 기본 정보까지 고려하여 설계되어야 하는데 이때 어휘와 통사 정보를 분리하여 수의적인 결속 우위에 따라 연결시키는 방법을 제안해 볼 수 있는 것이다.
- 3) [특징] 시냅스형 시소러스의 가장 큰 특징은 출발점이 정해져 있다는 것에 있다. 시냅스형 시소러스는 대상 어휘를 선택하고 이를 중심으로 하는 어휘 목록을 선정한 후 대상 어휘와 결합하는 문형에 따라서 선정된 어휘와의 관계 유지 여부를 판별하는 원리를 갖는다. 이를 이용하여 교사는 교육적으로 설명 가능한 설명 어휘장을 확보하게 되고 제시 가능한 문형을 확인할 수 있게 된다. 이러한 구조는 통합적 관점에서 어휘 교육을 바라보는 입장에 부합한다고 할 수 있다.

시냅스형 시소러스(STs)는 앞서 설명한 것처럼, 기존 시소러스 체계가 어휘 간 의미적 속성에 따라서 종합적으로 연결되는 신경망 체계로 보고

7) 조형일(2010)에서는 시냅스형 시소러스(STs)를 “어휘의 ‘습득’을 지향하는 ‘학습’ 모델”로 보고 있다. 외국어로서의 언어 학습에서 어휘 학습은 어휘의 정확한 의미의 학습뿐만 아니라 어휘의 사용 환경에 대한 학습 즉, 문법적 기능까지 효율적으로 학습하여 적확하게 사용할 수 있도록 해 주는 것이다. 따라서 어휘는 단순한 의미 학습의 단위가 아닌 문장 구성에 관계하는 중요 요소가 된다. 그리고 이는 습득의 개념을 지향하고 있다.

있는 것에 반해서, 구축된 어휘와 문형 간의 수의적인 선택 조합에 따라서 그 결과값을 참(true)과 거짓(false)으로 보여주는 가부(可否) 결정 체계라고 할 수 있다. 이때 목록으로 구축된 어휘와 문형은 그 빈도수에 따라 교육적으로 활용 가능한 등급을 갖게 된다.⁸⁾

따라서 시냅스 개념을 포함하는 시소러스 체계에는 각각의 어휘마다 교육적으로 응용 가능한 정보 체계가 표시되어 있어야 한다. 시냅스형 시소러스(STs)의 원리는 다음의 <표 2>에서 보이듯 시소러스 구축을 위한 어휘의 선정 후에 선택 어휘를 기준으로 하여 각각의 적용 가능한 격틀 정보로 유의어들을 연결하고 이들이 그 격틀과의 결합 시에 결과값을 참(True) 또는 거짓(False)으로 표시해 주는 것으로 다시 간단하게 설명할 수 있다.

<표 2> 시냅스형 시소러스의 구축 방법

선택 어휘	결합 격틀	유의어 1	유의어 2	유의어 3
A	Composition Grp1-A	B	C	D
판정 값		True	True	False

‘격틀⁹⁾’은 기존의 ‘문형’ 개념으로 어휘와 결속 가능한 문법적 관계를 나타내는 것이다. 이는 다시 복합적인 문법 요소와의 결합(Composition Group1)과 단일한 문법 요소와의 결합(Composition Group2)으로 구분할 수 있다.¹⁰⁾ 시냅스형 시소러스(STs)는 앞서 논의한 것처럼 단순히 어휘 간의 접점 구조를 이루는 것이 아니기 때문에 격틀 정보가 매우 중요한 요소가

8) 이와 비슷한 연구로는 남지순(2007)이 있다. 이는 어휘부와 통사적인 속성들을 기반으로 전체 5,300여 개의 한국어 형용사 술어 구문을 검토하여 15개의 클래스로 나누어 기술한 연구인데, 시냅스형 시소러스(STs)가 어휘의 의미 관계와 문형(구문)을 종합적으로 고찰하여 교육적으로 응용 가능한 결과물이라면, 이는 형용사 중심의 거시적이고 종합적인 어휘문법적 기술 방안에 관한 연구로 볼 수 있다.

9) 이에서의 ‘격틀’과 ‘Composition Grp1, 2’의 용어는 권재일(2000)을 따랐다.

10) Composition Group1은 ‘N을/를 N에게 V’와 같은 세 개 이상의 논항이 관계하는 문형을 말하고 Composition Group2는 ‘N이/가 V/A’와 같이 두 개 이하의 논항이 관계하는 문형을 의미한다. 이는 다시 ‘복합 문형’과 ‘단순 문형’으로 환언할 수 있다.

된다. <표 2>는 선택 어휘 A와 이의 유의어로 일반적으로 제시되는 어휘 B, C, D가 있을 때 A와 B, C, D 간의 유의 관계가 격틀 Grp1이 A를 취할 때 결합값유의 관계를 유지할 수 있느냐 없느냐에 따라서 참(True), 거짓(False)으로 판정되는 것을 보여 주고 있다. 이때 참값을 갖는 유의어 B와 C가 결합 격틀을 기준으로 하여 A와 함께 유의어 목록으로 결정되는 것을 확인할 수 있다.

시냅스형 시소러스(STs)에서 선택 어휘와 유의어로 제시되는 어휘는 일차적으로 유의 관계에 따라서 분류되어야 하며 각각의 어휘마다 격틀 정보의 결합값을 가지고 있어야 한다. 이러한 방식을 취하게 되면 동사와 형용사 어휘의 경우 격틀과의 결합은 한 어휘의 기능적 형태를 보여 줄 수 있을 뿐만 아니라 그 어휘가 몇 개의 논항을 취할 수 있는지도 보여 줄 수 있게 된다.

이에서는 어휘 A를 출발점으로 삼고 문법 형태와의 결합 요소를 중점에 두고 등급별 유의어의 결합 관계를 보였지만 추후 이를 심화 확대하여 반의어, 상하의어, 다의어, 주제화 어휘 집단 등으로 발전시켜야 한다. 이런 단계를 거쳐 구축된 시소러스가 비로소 텍스트의 구성과 판단 등에 응용할 수 있는 충분한 기준 자료가 될 수 있다.

3. 읽기에서의 시냅스형 시소러스(STs)의 응용 방안

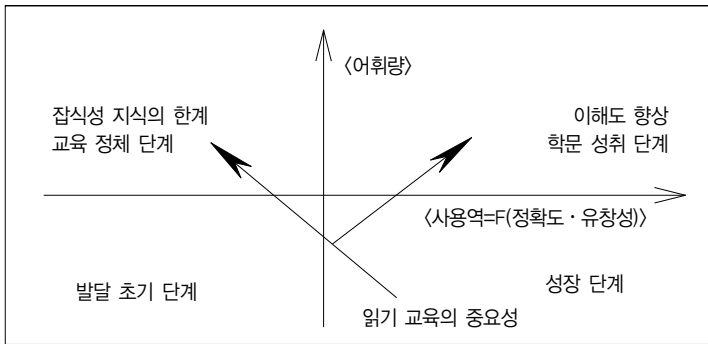
‘읽기 행위’를 통해서 학습자가 얻는 것은 다양하다. 학습자는 읽기 활동을 통해서 집중력과 함께 안정감을 기를 수 있다. 이는 학습자의 물리적 성장과 신체 정신적 밸런스를 이상적으로 끌어올리는 데에 기여할 것이다.

여러 가지 읽기의 행위 중에서 특별히 ‘독서’가 지향하는 주요 목표는 독자의 의식 향상이 될 것이다. 그러므로 다양한 텍스트 중에서 교육적 함의성이 가장 큰 텍스트를 ‘책’으로 보고, 독자인 ‘학습자’가 독서 행위를 통해서 정보 획득, 정서 함양, 재미와 여유의 치환을 얻을 수 있을 것

으로 생각할 수 있다면 독서를 ‘읽기의 복합적 활동’으로 분석할 수 있어야 한다. 따라서 읽기 행위는 독서의 일면이자 독서의 개념을 뛰어넘는 개념으로 봐야 한다. 읽기를 텍스트를 읽어내는 모든 활동으로 볼 수 있기 때문이다.

이러한 개념 아래에서 보면 읽기 활동의 측면에서 텍스트를 구성하고 있는 어휘란 텍스트의 해석을 위한 전달 매개이자 의미 그 자체로 볼 수 있다. 그러므로 독자 즉 학습자에게는 텍스트 해석에 필요한 일정 수준의 어휘량이 확보되어 있어야 한다. 그리고 텍스트 읽기를 통해서 어휘량을 증가시키는 동시에 어휘의 사용역에 대한 유기적인 확장을 수행하게 된다.

읽기 교육을 통해서 기대하는 결과는 학습자의 어휘량 증가와 사용역의 확장(정확도와 유기적인 연쇄 해석 능력) 정도로 볼 수 있다. 이를 관계도로 나타내면 다음처럼 나타낼 수 있을 것이다.



〈그림 1〉 학습자의 어휘량과 사용역 간 상관관계

이에서 보이듯 발달 초기 단계에는 학습 어휘량과 그 사용 수준이 평균 이하인 상태로 볼 수 있다. 이후 사용역에 대한 지식과 사용 능력의 변화 없이 어휘량이 많아지는 단계를 소위 잡식성 지식의 한계 상태인 ‘교육 정체 단계’로 볼 수 있다. 이와는 상대적으로 어휘량은 크게 증가하지 않고 있지만 그 사용역이 증가하는 단계를 ‘성장 단계’로, 이 두 가지 요건이 모두 상승하는 단계를 ‘학문 성취 단계’로 볼 수 있다.

이러한 발달 단계를 전제로 하는 어휘 중심의 읽기 교육이란 ‘읽기 행위를 통한 어휘 교육 방법’을 구안하는 것이 된다. 따라서 이의 목표는 어휘에 대한 이해도를 높이는 동시에 사용의 정확성/적확성, 유창성과 함께 설명적 치환 능력을 키워주는 데에 있다고 하겠다. 그리고 이러한 목표를 달성할 수 있는 방안으로 읽기에 시냅스형 시소러스(STs)를 응용하게 되면 다음과 같이 될 것이다.

- 1) 시냅스형 시소러스는 구조와 결속에 관한 것이다. 이에서는 어휘를 문장 구조를 관통하는 의미의 축으로 본다. 이를 이용하면 텍스트(책)의 독서 지수를 측정하거나 독서 지수를 고려한 읽기 텍스트 작성성이 가능해진다. 예를 들어 ‘A라는 것은 B이다.’라는 정의, 지정 또는 설명, 비유 등에 해당하는 문장은 ‘A는 B의 B"이다.’를 함의하는 것이므로 A와 B 또는 B의 구성성분이 되는 B'와 B"의 문장 내 결속 관계와 함께 각 성분들의 의미 관계를 파악하는 것으로 문장의 난이도를 계수할 수 있다.¹¹⁾
- 2) 주제 중심의 시냅스형 시소러스를 이용하여 주제 밀집 어휘의 양과 등급을 판별하는 동시에 하위 설명 개념 어휘를 추출, 측정할 수 있다. 이를 위해서 문형 검증/복합 문형의 개수/수식의 관계 등을 판단할 수 있는 알고리즘을 구축하여야 한다. 이때 “1이 아닌 양의 어떤 수를 거듭제곱하여 다른 주어진 수와 같아지는 거듭제곱 수”인 로그(log) 개념¹²⁾을 이용할 수 있다.¹³⁾ 문장과 문단의 로그값은

11) 시냅스형 시소러스(STs)는 이때 어휘 A, B, B', B" 및 이들의 결속 관계를 한정하는 ‘N₁(이)라는 것은 N₂이다’, ‘N₁은/는 N₂의 N₃이다’와 같은 문형의 난이도를 제공해 주고 있다.

12) $x=ay$ 에서, y 를 a 를 밑으로 하는 x 의 로그라 하며 $\log_a X=y$ 처럼 표시한다. 예를 들어 $100=10^2$ 을 $\log_{10}100=2$ 로 표시할 수 있다.

13) 10의 제곱은 100이 된다. 그런데 읽기는 10의 제곱이 100이라는 결과를 도출하는 것이 아니라 100을 만드는 거듭제곱 수가 10이라는 결과를 찾아내는 과정이 된다. 즉 100의 로그값을 찾는 것이다.

문단 구성 문장의 개수와 문장 내 서술어의 개수, 그리고 문장 구성에 관계하는 어휘의 등급과 함께, 문형 결합의 결속 정도, 유의어/반의어/언어 관계 등의 활용 정도를 판별해야 정확하게 얻을 수 있을 것이다.

지금까지 살펴본 시냅스형 시소러스(STs)는 1)번 항목에 집중되었다. 그런데 어휘의 의미 관계와 문형을 복합적으로 계측하는 방법과 함께 2)에서 제안하고 있는 것처럼 주제화 구성이 이루어진다면 읽기 텍스트의 판단과 생산에 직접 관여하는 객관적이고 타당한 기준을 제시해 줄 수 있을 것이다. 그리고 이렇게 구축된 시소러스는 낱말이 새로워지는 새로운 형태의 읽기 방식과 내용에 대한 지속적인 경신으로 꾸준히 응용될 수 있을 것이다.

4. 읽기에의 시소러스 응용의 실제

이 장에서는 앞서 논의된 시냅스형 시소러스의 응용 방안이 읽기 텍스트의 독서 지수 측정과 텍스트 작성에 실제적으로 어떻게 적용될 수 있는지 살펴보기로 한다. 이를 통해서 이 연구에서 제안하고 있는 시냅스형 시소러스가 읽기 텍스트의 해석과 생산에 적용 가능한 방법론이 될 수 있는지를 검증할 수 있을 것이다.

1) 텍스트 독서 지수 측정에의 응용 실제

다음 텍스트는 초등국어 4학년 1학기 교과서 읽기 영역에 실려 있는 ‘가끔씩 비 오는 날’이라는 텍스트의 일부분이다. 이 텍스트는 삶과 가치 영역으로서 용기, 창조, 개척 정신 부분의 글감으로 선정된 것이다.¹⁴⁾

14) 이 텍스트는 제2회 ‘좋은 어린이 책’ 원고 공모창작 부문 대상 수상작 ‘이가을(1998)

이삿짐이 들어왔습니다. 서랍 두 개가 딸린 책상이 하나, 일인용 침대가 하나, 낡은 책장이 두 개, 그리고 무엇이 들었는지 꽤 무거워 보이는 종이상자가 서른 개쯤 있었습니다.

‘무엇을 하는 사람일까?’

아저씨는 밤늦게까지 책상 앞에 앉아 책을 읽거나 무엇인가를 씁니다. 때때로 소리 내어 책을 읽기도 하고, 자기가 쓴 글을 읽기도 합니다. 나는 그 모든 것을 보고 듣는 것이 즐겁습니다.

이 텍스트에 쓰인 어휘와 사용 문형을 간단하게 분류해 보이면 다음과 같이 정리할 수 있다.¹⁵⁾

〈표 3〉 선정 텍스트의 등급별 어휘와 문형 분석

구분	1등급	2등급	3등급
체언/수식언	책상, 두, 개, 하나 침대, 무엇, 사람 아저씨, 밤, 앞, 책 자기, 글, 나, 그, 것 소리	서랍 책장 종이상자 꽤	이삿짐 일인용 때때로
용언	들어오다, 무겁다 있다, 하다, 늦다 읽다, 쓰다 보다, 듣다	즐겁다 보이다	떨리다 들다 내다
사용 문형 (조사와 어미)	N이/가 V A/V-아/어 보이다 N까지 N에 V N을/를 V	V-왔/였는지 -(으)는 N N쯤 V V-거나 V-기도 V A/V(으)는 것	N ₁ 이/가 V-(ㄴ)/는 N ₂ N이/가 N

기존의 텍스트 지수 계산 방식은 텍스트에 사용된 어휘량과 등급별 어휘량을 개수하는 것에서 시작했다. 하지만 <표 3>에서처럼 사용 문형

가끔씩 비 오는 날, 창작과 비평'에서 가져온 것이다.

15) 어휘 등급과 문형 등급의 객관성이 확보되어야 하나 이에서 다룰 문제가 아니므로 표에서는 임의로 조정했다.

을 추출하는 것만으로도 이 텍스트가 초등 저급학년 학습자에게는 어려운 수준이라는 것을 한눈에 파악할 수 있게 된다. 그리고 이때 파악된 사용 문형과 등급별 어휘가 결합된 시냅스형 시소러스를 비교 검사하면 각 사용 문형의 위계와 관계를 확인할 수 있게 될 것이다.¹⁶⁾

2) 텍스트 작성에의 응용 실제

읽기 텍스트는 다양한 경로로 수집·구축될 수 있다. 교육적으로 활용 가능한 텍스트를 선정하는 것과 교육적 효과를 고려하여 텍스트를 제작하는 것은 모두 ‘교육적 활용 가능성’을 중심축에 두어야 한다. 따라서 텍스트의 선정은 물론 제작에 있어서 읽기 텍스트로서 객관성과 타당성을 확보하는 것은 매우 중요한 요건이 된다.

그런데 전문가에 의해서 선정되는 텍스트는 선정 위원의 전문성에 의존하게 되고, 새롭게 구축되는 텍스트는 생산자인 작가의 전문성에 전적으로 의존하게 되는 경향이 짙다. 이러한 전문성을 다시 객관적 지표로 환언해서 제시하는 것은 쉽지 않은 일이 된다. 물론 오랜 기간 동안 텍스트를 다루어 온 전문가라면 어휘와 문장 구성, 그리고 제재와 주제화 구성, 기타 글 구성의 여러 기법들을 구분하고 분석하여 기술해 낼 수 있다. 하지만 ‘텍스트의 전문가’가 곧 ‘텍스트 내용의 전문가’를 의미하지는 않을 것이다. 이 경우에도 다양한 전문가의 저술 내용을 교육적으로 활용 가능한 수준으로 텍스트 전문가가 개작하면 되기는 한다. 하지만 일차적인 텍스트의 구축에서부터 교육적으로 활용 가능한 텍스트로 작성하거나 이의 수준을 결정하는 데에 도움을 줄 수 있다면, 읽기 텍스트에서 시냅스형 시소러스를 응용해 볼 만한 것이다.

앞의 <표 3>으로 예를 들어서 1등급 텍스트를 구축해 보이면 다음과 같은 결과물을 얻을 수 있을 것이다.

16) 이 연구는 시냅스형 시소러스의 원리와 응용 방안의 가능성을 제안하는 데에 그 목적이 있다. 이 텍스트의 읽기 지수를 완벽하게 측정해낼 수 있는 시냅스형 시소러스는 아직 구현되지 못한 상태이다.

아저씨가 이사를 왔습니다. 이삿짐에는 책상이 하나, 침대가 하나 있습니다. 무거워 보이는 종이상자도 많이 있습니다.

‘무엇을 하는 사람일까?’

아저씨는 밤에 책을 읽습니다. 자기가 쓴 글도 읽습니다. 나는 그 소리를 보고 듣습니다.

*이삿짐 : 이사할 때 옮기는 짐

*종이상자 : 종이로 만든 상자

이 텍스트를 앞서 제시한 원래의 텍스트와 비교해 보면, 문장 길이가 단순해졌을 뿐만 아니라 사용 어휘도 줄어든 것을 확인할 수 있다. 이는 <표 3>에서 제시하고 있는 1등급 수준의 어휘와 문형을 사용했기 때문이다. 이때 ‘이삿짐’과 ‘종이상자’처럼 학습자 수준보다 높은 어휘로 제시된 경우 새로 배워야 할 어휘로 제시해 줄 수 있다. 결과적으로 새로운 텍스트는 초등학교 4학년 수준이라고 하기에는 쉬운 텍스트가 되어버렸다.

이처럼 텍스트 작성을 위한 지침을 제공하게 되는 경우에 텍스트의 생산자가 보다 분명하고 직접적으로, 교육적으로 활용 가능한 텍스트를 생산할 수 있게 된다.

그런데 앞의 <표 3>은 이미 있는 텍스트를 분석해서 이를 다시 응용한 것이므로 사실 어떤 텍스트의 작성에 직접적으로 도움을 줄 수 있는 시냅스형 시소러스는 주제별로 참조 가능하거나 어휘별, 또는 문형별로 조건별 구성이 가능해야 한다. 이것은 앞서 제시한 주제 중심의 시냅스형 시소러스 개념으로 접근해야 한다. 주제별 어휘의 등급별 구축과 함께 이들 어휘의 유의어/반의어/상하의어/관용 표현 등을 보여줄 수 있어야 하고 이들이 다시 등급화된 문형과 결합되는 결합값을 확인할 수 있도록 해야 하기 때문이다.

Ⅲ. 다음

이 연구에서 주장·제안한 것을 요약하면 다음과 같다.

- 1) 단순한 의미 관계에 따른 어휘 수준과 문장의 길이를 측정하는 것만으로는 이독성과 문식성을 포함하는 읽기 텍스트 난이도 판별을 실현하기는 어렵다.
- 2) 읽기 텍스트의 난이도 판별 또는 읽기 텍스트의 작성 시에 시냅스형 시소러스(STs)를 적용해 볼 수 있다.
- 3) 시냅스형 시소러스(STs)는 단순한 어휘, 구문 간의 조합뿐만 아니라 주제화된 영역을 각각의 출발점으로부터 구축해 낼 수 있을 것이다.

지금까지 읽기에서의 어휘 시소러스 응용 방법에 관해서 논의하면서 읽기 능력 향상을 위한 방법과 함께 읽기 텍스트의 구축을 조금 더 객관적으로 계수할 수 있는 측면에서, 어휘 시소러스의 응용 가능성을 살펴보았다. 이들 중에서 어느 정도까지 호응을 받을 수 있을지 모르는 상태에서, 필자의 천착에서 비롯된 논리적인 비약을 제외한 결론을 조심스럽게 정리하면 이렇다. 앞서 제시해 보인 것처럼 읽기 교육 방안을 구안에 있어서, 교육 텍스트의 구축으로 제한하여, 시냅스형 시소러스가 충분히 응용될 수 있고, 텍스트의 읽기 난이도 선정 판별에서 시냅스형 시소러스가 이용될 수 있다. 그리고 이는 객관적이고 타당한 기준을 마련하는 데에 일조할 수 있을 것이다.*

* 본 논문은 2011. 6. 29. 투고되었으며, 2011. 7. 10. 심사가 시작되어 2011. 7. 29. 심사가 종료되었음.

▣ 참고문헌

- 권재일(2000), “국어 정보화와 용언 전자사전 구축”, 『한말연구』 제6호, 한말연구학회, pp.67-86.
- 김광해(1995), 『어휘연구의 실제와 응용』, 집문당.
- 남지순(2007), 『한국어 형용사 어휘문법』, 한국문화사.
- 박수자(2003), “21세기 문식력과 국어과교육의 과제”, 『국어교육』 Vol.110, 한국어교육학회, pp.45-65.
- 박인기(2002), “국어교육과 문화교육 : 문화적 문식성의 국어교육적 재개념화”, 『국어교육학연구』 Vol.15, 국어교육학회, pp23-54.
- 조형일(2010), “시소러스 기반 한국어 어휘 교육 연구”, 서울대학교 박사학위 논문.
- 천경록·이재승(1997), 『읽기 교육의 이해』, 우리교육.

<초록>

읽기에서의 어휘 시소러스(thesaurus)의 응용

조형일

이 연구는 국어교육에서의 읽기 자료의 구축과 읽기의 교육 효과 달성을 위한 교수·학습 방법에 직접적으로 적용 가능한 ‘시소러스 응용 방안’을 논의한 것이다. 시소러스(thesaurus) 즉, ‘목적성 있게 가공된 분류어휘집’을 구축하고 응용하여 읽기 텍스트의 해석은 물론 생산에서부터 관여하는 어휘 목록이 제공된다면, 이것은 당연히 텍스트 이독성/난이도 측정의 지표가 될 수 있다.

어휘 중심의 읽기 교육이란 ‘읽기 행위를 통한 어휘 교육 방법’을 구안하는 것이 된다. 따라서 이의 목표는 어휘에 대한 이해도를 높이는 동시에 사용의 정확성/적확성, 유창성과 함께 설명적 치환 능력을 키워주는 데에 있다. 그리고 이러한 목표를 달성할 수 있는 방안으로 읽기에 시냅스형 시소러스(STs)를 응용하게 되면 다음과 같은 결과를 가져올 수 있다.

- 1) 시냅스형 시소러스 이용하면 텍스트(책)의 독서 지수를 측정하거나 독서 지수를 고려한 읽기 텍스트 작성이 가능해진다. 문장 내 어휘 결속 관계와 함께 각 성분들의 의미 관계를 파악하는 것으로 문장의 난이도를 계수할 수 있다.
- 2) 주제 중심의 시냅스형 시소러스를 이용하여 주제 밀집 어휘의 양과 등급을 판별하는 동시에 하위 설명 개념 어휘를 추출, 측정할 수 있다.

【핵심어】 국어교육, 읽기, 어휘, 시소러스, 시냅스형 시소러스

<Abstract>

Adapting Thesauruses on Reading

ZO, Hyoung-il

This thesis suggests that a direct adaptable methods of the way of adapting thesauruses on reading for constructing the reading materials and attaining the reading educate result in the Korean education. If there is provided certain indexes which are objectively processed and branched i.e. thesauruses and it could be concerned not also a text interpretation but also text production, certainly it will be an indicator to measuring text readabilities or difficulties.

Lexicon centered reading education means that making the methods of lexicon teaching through reading activities. Therefore these goals are raising intelligibility of lexicon and at the same time promoting the ability of explaining and fluency and accuracy. And the way of accomplishing these goals, if we will adapt the Synaptic Thesaurus system on reading, it will bring about below result.

- 1) Using the Synaptic Thesaurus system make the text 'lectio quotient' measurement of the published books and text writing which is considered 'lectio quotient' possible.
- 2) Using theme centered the Synaptic Thesaurus system can discriminate the quantity of the theme centered lexicon and the grade of them and also measure the meta-language

【Key words】 Korean Education, Reading, Vocabulary, Thesaurus, Synaptic Thesaurus system