

국어 텍스트의 어휘적 응결성과 어휘의미망*

- 체계기능언어학(SFL)의 관점 -

한 정 한**

국문초록

이 글의 목적은 두 가지이다. 첫째로, 체계기능언어학(Systemic functional linguistics)의 어휘적 응결성(Lexical Cohesion)의 개념을 소개하고, 이것이 서로 다른 국어 장르(텍스트)에서 어떤 차별적 실현 양상을 갖는지 밝혀 보는 것이다. 이를 통해 어휘적 응결성과 장르의 상호 관련성을 검토해 본다. 둘째로, 어휘적 응결성의 종류와 그 특성을 잘 반영한 어휘의미망의 구축을 통해 어의 중의성 문제를 해소할 수 있는 가능성을 제시해 보는 것이다.

응결성은 어떤 종류의 언어 자원을 사용하느냐에 따라 문법 응결성과 어휘 응결성으로 나눌 수 있다. 본고는 지금까지 세월호 침몰사고 관련 ‘신문기사(장르)’ 10건과 세월호 유가족이 쓴 ‘편지글(장르)’ 7건을 대상으로 11종의 어휘적 응결성과 5종의 응결성 거리를 조사했다. 분석 결과 우리는 장르(텍스트)에 따른 차별적 실현양상을 확인할 수 있었다(4장 참조).

아울러 본 논문에서 논의한 어휘적 응결성과 그 정보를 잘 반영한 한국어 어휘의미망, 특히 연상어(Associates(AC)) 정보가 적절하게 제공되면, 한국어 텍스트의 어의 중의성 문제를 해결할 수 있는 가능성에 대해서도 언급했다. 그 예로 김민호·권혁철(2011)에서 KorLex를 활용한 어의 중의성 해소 과정을 소개했다.

그리고 특히 버클리대가 만든 영어 어휘 의미 데이터베이스인 FrameNet은 모든 단어를 개별 어휘의미단위(Lexical Unit)로 분류하고, 각각의 어휘의미단위마다, 본고에서 연상어(AC)로 정의한, 고유의 FE(Frame Element) 정보가 있음을 밝혔다.

FrameNet은 이러한 개념의미 기반 어휘의미 분석을 통해 통사론과 의미론의 연결을 형식화하고 단어 의미 중의성, 정보 추출, 기계 번역, 질의응답 등 전자사전을 활용해 중의성 문제를 해결하는 데 기여하고 있다.

[주제어] 어휘적 응결성(lexical cohesion), 장르(텍스트), 워드넷(WordNet), 연상어(Associate), 체계기능언어학(SFL)

* 이 연구는 2019학년도 단국대학교 대학연구비의 지원으로 연구되었음.

** 단국대학교 국어국문학과 교수 / jhan@dankook.ac.kr

| 목 차 |

I. 서 론	IV. 분석 결과
II. 연구 대상 텍스트 소개	V. 어휘의미망(WordNet)과 중의성 해소
III. 두 종류의 응결성	VI. 결 론

I. 서 론

이 글의 목적은 두 가지이다. 첫째로, 체계기능언어학(Systemic functional linguistics)의 어휘적 응결성(Lexical Cohesion)의 개념을 소개하고, 이것이 서로 다른 국어 장르(텍스트)에서 어떤 차별적 실현 양상을 갖는지 밝혀 보는 것이다. 이를 통해 어휘적 응결성과 장르의 상호 관련성을 검토해 본다. 둘째로, 어휘적 응결성의 종류와 그 특성을 잘 반영한 어휘의미망의 구축을 통해 어의 중의성 문제를 해소할 수 있는 가능성을 제시해 보는 것이다.

체계기능언어학(SFL) 내에서 발표된 많은 담화 텍스트 분석 연구들은 Halliday & Hasan(1976)의 Seminar Book과 그 이후의 연구 업적에 크게 영향을 받아 왔다. Baker(1992, 180), Moreno(2003, Xi 2010)에서는 그들의 업적을 “지금까지 응결성에 관해 알려진 연구 중에서 가장 상세하다.”¹⁾는 찬사를 보냈다.

여러 가지 분류 기준에 따라서 달라질 수 있지만, Halliday & Hasan(1976)에 의하면, 응결성은 우선 세 가지 종류로 나누어진다. 각각 동일지시(co-referentiality), 동일종류지시(co-classification), 그리고 동일확장지시(co-extension)가 그것이다(Halliday & Hasan 1985, 73).

먼저 동일지시는 응결 관계를 형성하는 두 개 혹은 여러 개의 요소들이 지시적 동일성(referential identity)을 갖는 경우를 말한다. 아래 예문 (1)에서 ‘a little nut tree’와 대명사 ‘it’의 관계처럼 같은 형태를 지시하는 대명사, 또는 동일 형태를 반복하는 어휘들이 여기에 속한다. 둘째로, 동일확장지시는 (1)에서 silver(은)와 golden(금) 같은 경우로, 여기서 ‘은’과 ‘금’은 모두 ‘금속’에 포함된다. 이와 같이 동일확장 지시는 두 언어 요소들 간에 상위어-하위어, 또는 하위어-상위어의 어휘 관계로 나타난다. 셋째로, 동일종류지시는 (2)에서 ‘play the cello’와 생략 조동사 ‘does’처럼 동일종류의 다른 첼로(cello)를 가리킬 때 나타난다.

(1) I had a little nut tree. Nothing would it bear. But a *silver* nutmeg. And a *golden* pear.

(2) I play the cello. My husband does, too.(Halliday & Hasan, 1985/1989: 73)

여기서 알 수 있는 것은, 개별 문장들의 응결성이 높으면, 읽는 사람은 낱낱의 문장들이 하나의 텍스트를 구성하고 있다고 인식하게 된다는 것이다. 이것을 체계기능언어학에서는 텍스트성(Textuality)이라고 한다.

1) “……the best known and most detailed model of cohesion available.”

따라서 응결성이 높은 텍스트는 전체 의미를 이해하기가 더 쉽고, 따라서 번역하기도 쉽고, 그렇게 쓰인 텍스트들은 잘 된 텍스트(well-formed texts)라는 평가를 받게 된다.

응결성을 분류하는 또 다른 기준으로는 응결성 거리(cohesion ties)라는 게 있다(Halliday & Hasan, 1976). 이것은 응결성을 가지는 두 요소 사이의 거리를 기준으로 응결력을 구분하는 것이다. 응결성 거리는 먼저 동일 문장 내부 요소들 간의 응결성과 문장과 문장 사이의 요소들 간의 거리로 나눌 수 있다. 본고에서는 전자를 WS(within-sentence), 후자를 AS(across-sentence)로 각각 분류하여 통계 처리하였다. 그리고 문장 간 거리 응결성(AS)은 다시 인접거리 유대(immediate ties), 원거리 유대(remote ties), 매개거리 유대(mediated ties)로 하위분류된다. 인접거리 유대는 아래 (3)의 ‘Alice[1]-she[2]’에서 보듯이 선행 요소와 후행 요소의 응결관계가 바로 인접한 문장에서 나타나는 것을 뜻한다. 그리고 원거리 유대는 Alice rubbed her eyes[3]-Rub as she would[기처럼 두 요소 사이의 거리가 확장되어 여러 문장에 걸치는 유대를 말한다. 마지막으로 매개거리 유대는 먼저 선행어가 나오고(Alice[3]), 한 참 떨어져서 후행 응결어(she[5]가 나오는데, 그 중간에 서로를 매개하는 매개어(she[4])가 나타나는 경우를 말한다. 후술하겠지만 이런 응결성 유대의 차이는 응결성의 종류에 따라서 상당한 편차가 있고, 또 장르에 따라서도 차이가 있다.

그리고 앞의 매개거리 유대는 다시 동일성 사슬(identity chain)과 유사성 사슬(similarity chain)로 하위 구분 된다. 동일성 사슬은 지시하는 표현과 지시대상 사이에 동일지시 관계(대명사, 반복어, 등가어 등)를 맺는 것을 말하고, 유사성 사슬은 지시하는 표현과 지시대상 개체 사이에 동일종류지시나 동일확장지시의 관계가 있는 경우를 말한다.

- (3) The last word ended in a long bleat, so like a sheep that Alice quite started [1]. She looked at the Queen, who seemed to have suddenly wrapped herself up in wool[2]. Alice rubbed her eyes, and looked again[3]. She couldn't make out what had happened at all[4]. Was she in a shop[5]? And was that really-was it really a *sheep* that was sitting on the other side of the counter[6]? Rub as she would, she could make nothing more of it[7]. (Halliday & Hasan, 1976: 330)

체계기능언어학에서 응결성(cohesion)과 일관성(coherence)의 상호 관계도 언급해 둘 필요가 있다. 가장 잘 알려진 응결성과 일관성의 차이는 응결성이 텍스트의 내적 요소들 간의 결속 관계를 다루는 분야라면, 일관성은 텍스트의 외적(또는 상황맥락적) 요소들 간의 결속 관계를 다루는 분야라는 것이다. 다시 말하면, 응결성이 문법적, 어휘적 자원들 간의 결속 관계를 통해 얻어지는 결속이라면, “일관성은 텍스트 외부에 존재하는 경험적 맥락(예: 화자, 청자), 사회문화적 맥락들 간의 결속 관계를 통해 얻어지는 결속력이다. 텍스트를 구성하는 개별 언어 표현들은 일관성을 통하여 하나의 통일된 전체로 인식하게 하는 성질을 갖는다. 다음 예를 보자.

(4) T(urn)1 : 가져가실 건가요, 드시고 가실 건가요?

T2 : 테이크아웃 잔에 주세요(이관규 외, 2021: 218).

언어 표현 자체에만 집중하면 (4)의 두 발화는 서로 전혀 상관없는 이야기를 하는 듯하다. 그러나 이 대화 속의 언어 표현들이 함께 나타날 수 있는 어떤 일관성 있는 상황 맥락(예: 커피 전문점에서 음료를 주문하는 상황)을 가정할 수 있으면 (4)는 하나의 텍스트로 볼 수 있다. 카페에서 ‘마시고 갈 것인지 가져갈 것인지’를 묻는 이유는 음료를 일회용 잔에 담아서 줄지, 다회용 잔에 줄지 결정하기 위해서다. 이런 경우 (4T1)과 (4T2)는 하나의 텍스트로 볼 수 있다.

그런데 체계기능언어학 내부에서도 응결성과 일관성은 정확히 구별되는 개념이 아니다(실제로 생략이나 대용이 많은 텍스트에서 더욱 그렇다.) 따라서 본고에서는 이것을 정도상의 의미정보(gradual semantics)로 처리하기로 하겠다.

본고에서는 다양한 종류의 어휘적 응결성을 서로 다른 장르의 텍스트를 대상으로 분석해 보려고 한다. 본고의 연구 대상 텍스트는 세월호의 ‘신문 기사문’과 세월호 희생자들의 부모들이 쓴 ‘편지글’이다. 불특정 다수를 대상으로 하는 격식적인 신문 기사문과 다분히 개인적이고 사적인 편지글에 따라서 응결성이 어떤 차이를 보이는지를 검토해 볼 계획이다. 그리고 이러한 응결성 정보를 담고 있는 잘 구축된 한국어어휘의망이 왜 필요한지도 지적하고자 한다.

II. 연구 대상 텍스트 소개

본고의 주요 분석대상은 세월호 침몰 사고²⁾의 신문 기사와 세월호 유가족이 쓴 편지글이다. ‘세월호 사건’을 공통의 주제로 하여, 불특정 다수 독자를 대상으로 쓴 신문기사와 사별한 자녀들에게 보내는 개인적인 편지글이 보이는 어휘 응결성의 차이를 밝혀보려는 것이 목적이다. 신문기사 수와 편지글 수의 문장 수를 비스듬하게 맞추기 위해서 양적 조정이 있었다. 그 내용은 다음과 같다.

〈표 1〉 세월호 침몰 사고 신문기사

번호	신문사	기사 제목	문장수
1	동아일보	여객선 세월호 침몰	16
2	서울신문	여객선 세월호 안산 단원고 학생 등 460여명 사고	9

2) 세월호 침몰 사고는 2014년 4월 16일 오전 8시 50분경 대한민국 전라남도 진도군 조도면 부근 해상에서 여객선 세월호가 전복되어 침몰한 사고이다. 이 사고로 안산시의 단원고등학교 학생들을 포함 304명이 사망하였다.

(<https://terms.naver.com/entry.naver?docId=2119309&cid=43667&categoryId=43667>).

번호	신문사	기사 제목	문장수
3	국민일보	여객선 침몰 임박 - 안산 단원고 홈페이지 마비	7
4	한겨레	진도 침몰 여객선 1명 사망 197명 구조	20
5	경향신문	여객선 진도 해상 좌초 - 정부, 대책본부 가동	13
6	조선일보	세월호 구조 작업 난항	9
7	중앙일보	진도 여객선 침몰 실종인원 290여명으로 번복 돼	17
8	한국일보	진도 여객선 침몰 중 안산단원고 학생 대부분	7
9	세계일보	침몰 세월호 수색 중단	8
10	문화일보	고등학생 등 459명 탄 여객선 침몰	15
계			121

〈표 2〉 세월호 침몰 사고 편지글

편지글	번호	제 목	문장수
1	T1	‘생각하면 눈물이 앞을 가리고’	18
2	T2	‘그곳에선 잘 지내고 있지?’	21
3	T3	‘우리 딸, 태어나면서부터’	25
4	T4	‘아기가 있는 집인지 모를 정도로’	19
5	T5	‘유치원 다닐 때도 혼자’	12
6	T6	‘학교 다니면서 공부도 잘하고’	30
계			125

Ⅲ. 두 종류의 응결성

앞장에서 응결성은 텍스트의 내적 요소들 간의 결속 관계를 다루는 분야, 그리고 일관성은 텍스트의 외적 요소(상황맥락)들 간의 결속 관계를 다루는 분야라고 소개했다. 여기서는 좀 더 구체적으로 들어가 보자. 아래 〈표 3〉은 체계기능언어학의 응결성 유형을 정리한 것이다.

〈표 3〉에 따르면, 먼저, 응결성이 작용하는 일반적인 타입(가로 층위)은 세 가지이다. ‘메시지 간 이행’(transitions between messages) 차원에서 작용하는 응결성, ‘의미하기’(meaning) 차원에서 작용하는 응결성, ‘표현하기’(in wording) 차원에서 작용하는 응결성이 그것이다. 메시지 간 이행 차원은 주로 절과 절을 이어주는 접속사(conjunction)에 의해 실현된다.³⁾ ‘의미하기’ 차원은 주로 문법적으로 ‘지시’, 어휘적으로 ‘유

의어’, ‘상위어’ 등에 의해 실현되며, ‘표현하기’ 차원은 문법적으로 ‘생략’과 ‘대체’, 어휘적으로 ‘반복’, ‘언어’ 등에 의해 실현된다.

〈표 3〉 체계기능언어학의 응결성의 종류(Halliday & Metthiessen, 2014: 608)

일반유형		문법적 영역 [문법적 단위 내부]	어휘적 영역 [어휘적 항목]
메시지 간 이행		접속사 [단위: 절]	
작용 층위	의미하기	지시 [단위: 명사군, 부가어군]	어휘적 접속 [유의어, 상하위에]
	표현하기	생략 및 대응 [단위(복합): 절, 명사군, 부사군]	[반복, 언어]

그리고 응결성의 형성에 사용되는 언어 자원의 종류에 따라서(세로 층위) 응결성을 구분하면 ‘접속, 지시, 생략 및 대응 그리고 어휘’ 이렇게 네 가지로 구분 할 수도 있다. 마지막으로 이들 네 가지는 다시 그것의 언어 자원이 ‘문법적 자원’이나 ‘어휘적 자원’이냐에 따라서 ‘문법적 응결성’과 ‘어휘적 응결성’으로 분류되기도 한다.

3.1. 문법적 응결성(grammatical cohesion)

앞장에서 설명했듯이, 응결성은 어떤 종류의 언어적 자원이 쓰였는가에 따라서 각각 문법적 응결성과 어휘적 응결성의 두 가지로 하위 구분된다. 먼저 문법적 응결성을 간단히 알아보자.⁴⁾

먼저 특정 개체가 앞에 언급된 개체를 의미적으로 가리키는 것을 ‘지시’(reference)라고 한다. 즉, 지시적 응결성은 가리키는 말과 가리켜지는 말 사이의 응결성이 문법적 성분인 인칭대명사, 지시대명사, 명사(구), 동사(구) 등에 의해서 발생하는 응결성이다. 이러한 ‘지시’는 다시 지시 대상이 지시하는 표현에 선행하는 ‘전방조응’(anaphoric)과 지시하는 표현이 지시 대상에 후행하는 ‘후방조응’(cataphoric)으로 나뉜다. 아래 (5)에서 ‘the blind mice-they’의 관계가 전방조응의 예가 되고, (6)의 ‘they-the blind mice’가 후방조응의 예가 된다.

(5) Three blind mice, three blind mice. See how they run! See how they run. (Halliday & Hasan, 1976: 31)

(6) See how they run! See how they run. Three blind mice, three blind mice.

3) 본고의 핵심 연구 주제는 어휘적 응결성이므로 접속사에 의한 응결성에 대해서는 언급하지 않기로 한다.

4) 본고의 핵심 주제는 어휘적 응결성이므로 문법적 응결성은 깊이 있게 논의되지 않을 것이다.

한편 지시가 ‘의미하기’를 기준으로 한 문법적 응결성이라면, 대용과 생략은 ‘표현하기’를 기준으로 한 문법적 응결성이라고 할 수 있다. 그리고 아래 (7)에서 명사구 ‘two poached eggs’와 대명사 ‘the same’은 대용 응결성의 예가 되고, (8)에서 조동사 ‘have’ 뒤에는 선행하는 ‘been swimming’이 생략된 것이므로 생략 응결성의 예가 된다.⁵⁾

- (7) I'll have two poached eggs on toast, please.
I'll have the same. (Halliday & Hasan 1976: 105)
- (8) Have you been swimming? Yest, I have. (Halliday & Hasan, 1976: 67)

3.2. 어휘적 응결성(lexical cohesion)

한편 어휘적 응결성은 응결에 사용되는 언어 자원들이 어휘적인 것들이다. 어휘적 응결성은 본고의 핵심 주제이므로 연구 대상 텍스트인 신문기사와 편지글을 중심으로 자세히 언급하기로 한다. 먼저 아래 <표 4>는 어휘적 응결성의 예들을 정리하여 소개한 것이다.

<표 4> 어휘적 응결의 유형 및 어휘적 관계(Gómez, 2018: 114)⁶⁾

어휘적 응결어의 유형		응결성 유대의 종류		
반복어(R)	일치 반복어(ER)	문장 내부(WT)	문장과 문장 사이(AS)	
	부분일치 반복어(IR)			굴절 반복어(IIR)
파생 반복어(DIR)				
유의어(S)	유사 유의어(NS)			인접거리 유대(IM) 매개거리 유대(MM) 원거리 유대(DM)
	명제 유의어(PS)			
반의어(O)	상보 반의어(CpO)			
	등급 반의어(AO)			
	방향 반의어(DO)			
포함어(I)	상위어(GI)			
	하위어(SI)			
연상어(AC)				

5) 응결성에 관한 국외 논문을 정리해 보면, 첫째로 응결성의 종류에 관한 연구로는 Morenno(2003), Sanders & Maat(2006), Taboada(2006), Dontcheva-Navratilova(2009)가 있고, 응결성의 표지들에 관한 연구로는(전방조용, 추상 명사, 반복어, 담화 표지 등) Ochs Keenan(1977), Tadros(1985), Francis(1986), Norrick(1987), Schmid(2000), Fraser(2005), Tannen(2007), Flowerdew(2010) 등이 있다.

6) 한국어에 맞게 일부 수정하였음을 밝힌다.

여기서는 <표 3>, H & M(2014: 608)의 분류법 대신에 <표 4>의 Gómez(2018: 114)의 분류를 사용하기도 한다. 전자는 응결성과 일관성을 구별하였지만, 후자는 응결성과 일관성을 엄격하게 구별할 필요가 없다는 입장이다.

위 <표 4>에서 보듯이 어휘적 응결은 크게 반복어(Repetition), 유의어(Synonymy), 반의어(Opposition), 포함어(Inclusion), 그리고 연상어(Associative Cohesion)로 대분류 된다. 그리고 반복어(R)는 다시 일치 반복어(ER), 부분일치 반복어(IR), 굴절 반복어(IIR), 파생 반복어(DIR)로 하위구분 된다. 일치 반복어(ER)란 형태가 동일한 반복어를 말하고, 굴절 반복어(IIR)는 조사, 어미를 제외한 나머지 부분의 반복, 파생 반복어(DIR)는 파생접사를 제외한 나머지 부분이 일치하는 반복어이다. 그리고 유의어(S)는 다시 유사-유의어(NS)와 명제-유의어(PS)로 나누어진다. 유사-유의어는 의미가 동일하거나 유사한 유의어이고, 명제-유의어는 어휘보다 큰 구나 절 표현이 낱개의 어휘와 의미적으로 유사한 관계를 맺는 경우를 말한다(cry-burst into tears).⁷⁾ 반의어(O)는 ‘남자-여자’처럼 A이면 B가 아니다가 성립하는 상보적 반의어(CpO), ‘작다-등급적 반의어(AO), ‘아래-위’처럼 방향적 반의어(DO)⁸⁾로 나누어진다. 포함어(I)는 특정한 어휘에서 일반적인 어휘로 가는 상위어(GI), 일반적인 어휘에서 특정한 어휘로 가는 하위어(SI)가 있다.

마지막으로 연상어(AC)는 특정 의미 프레임(Frame Semantics)에서 서로 연관된 어휘들 간의 관계이다. 예를 들어 아래 (9)와 같은 예문에서 인칭대명사 ‘그’가 가리키는 대상이 ‘철수’인지 ‘민수’인지 모르므로 (9)는 하나의 텍스트인지를 판단하기 어려운 중의적인 문장이다. 그러나 ‘장례식’과 ‘살해하다’를 ‘살인 사건’이라는 프레임의 틀요소(frame elements)(연상어)로 인식할 수 있으면 ‘그’가 ‘민수’를 가리키고 있다는 것을 추정할 수 있다.

- (9) S1 : 화가 난 철수가 민수를 살해하였다.
S2 : 그의 장례식은 월요일에 진행되었다.

이러한 어휘 응결성의 분류 방식을 이용하여 아래 5장에서는 세월호 침몰 사고의 신문 기사와 세월호 유가족이 쓴 편지글을 중심으로 어휘적 응결성의 특징들을 차례대로 살펴보기로 한다.

IV. 분석 결과

이 장에서는 앞장 <표 1> 세월호 침몰사고 신문기사와 <표 2> 세월호 침몰 사고 편지글 목록에서 제시한 10개 신문기사와 6개의 편지글을 대상으로 3.2.장 <표 4>에서 소개한 15개의 어휘적 응결성 유형을 비교

7) 드라마나 서정적 글쓰기 장르에서 자주 발견됨.

8) 관계적 반의어라고도 함. e.g., teacher-pupil.

검토해 보았다.

아래 <표 5>는 10개 신문 기사의 어휘 응결성 종류별 빈도를 종합한 결과이다 그리고 <표 6>은 6개 편지글의 어휘 응결성 종류별 빈도를 종합한 것이다.

<표 5> 신문 기사 어휘 응결성 종류별 빈도(종합)

신문	ER	IRR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계	카이제곱 χ^2	p-value
WS	10	1		8		2	9		4	44	41	119		
AS	185	28	30	38	8	7	3	2	3	16	77	397		
계	195	29	30	46	8	8	12	2	7	60	118	516	150.30	0.001

<표 6> 편지글 어휘 응결성 종류별 빈도(종합)

편지	ER	IRR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계	카이제곱 χ^2	p-value
WS	10	2	2	1	5		3		1	1	22	47		
AS	95	28	5	4	7	4	1	1	2	3	30	179		
계	105	30	7	5	12	4	4	1	3	3	52	226	36.84	0.00164

<표 5>와 <표 6>에서 볼 수 있듯이 전체적으로 신문 기사에서 516개, 편지글에서 226개, 총 742개의 어휘적 응결성 사례들이 분석 되었다. 이것은 동일 문장 내 어휘 응결성(WS)과 문장과 문장 사이 어휘 응결성(AS)의 사례를 모두 합산한 것이다.

여기서 제일 눈의 띄는 것은 신문 기사의 어휘 응결성 수가 편지글의 어휘 응결성 수보다 거의 2.3배 높다는 것이다. 이것은 통계적으로도 충분히 유의미한 값을 얻었다. <표 5>의 카이제곱값(chi square values)은 $\chi^2(7)^9) = 150.30$, 유의확률은 $p < .001$ 이다. 그리고 <표 6>의 카이제곱값은 $\chi^2(8) = 36.84$, 유의확률을 $p < .00164$ 이다.⁹⁾

편지글에서 어휘적 응결어들이 상대적으로 적은 것은 테너(참여자) 변수, 즉 글쓴이의 사적이고 정서적인 감정 변화가 텍스트의 정보 전달적 요소보다 더 중요하게 작용한 것으로 보인다. 또 생략 어휘들이 상대적으로 많은 것도 영향이 있어 보인다. 이에 반해 신문 기사문은 글쓴이가 직업적인 전문가들이고 다수 독자를 대상으로 쓴 계획된 글이기 때문에 이러한 결과를 얻었지 않나 생각한다.

통계분석의 두 번째 특징은 유형별 출현 빈도에 일정한 순서가 발견된다는 것이다. 먼저 신문기사에서는 아래 (10), 편지글에서는 아래 (11)의 순위 결과가 나왔다.

9) 괄호 숫자는 자유도(df).

10) p값이 3가지 유의 수준 (.001, .01, .05 모두에서 기준값보다 적으므로 유의미한 결과값이 나옴.

- (10) 신문기사 : ER > AC > SI > NS > DIR > IIR > AO > CpO > PS > GI > DO
 일치 반복어(ER, 195) > 연상어(AC, 118) > 하위어(SI, 60) > 유사 유의어(NS, 46)
 > 파생 반복어(DIR, 30) > 굴절 반복어(IIR, 29) > 등급 반의어(AO, 12) > 상보적
 반의어(CpO, 9) > 명제 유의어(PS, 8) > 상위어(GI, 7) > 방향 반의어(DO, 2)
- (11) 편지글 : ER > AC > IIR > PS > DIR > NS > AO = CpO > SI = GI > DO
 일치 반복어(ER, 105) > 연상어(AC, 52) > 굴절 반복어(IIR, 30) > 명제 유의어(PS,
 12) > 파생 반복어(DIR, 7) > 유사 유의어(NS, 5) > 등급 반의어(AO, 4)=상보 반의
 어(CpO, 4) > 하위어(SI, 3)=상위어(GI, 3) > 방향 반의어(DO, 1)

신문기사와 편지글 모두에서 반복어(R)는 가장 강력한 어휘 응결성이라는 것을 알 수가 있다. 반복어(R)는 신문 기사에서 37.7%(195/516)에 해당하고, 편지글에서도 46.4%(105/226)에 해당한다. 게다가 파생 반복어(DIR, 30)와 굴절 반복어(IIR, 29)를 일치 반복어(ER, 195)에 포함시키면 무려 신문기사는 49.2%(254/516)이고, 편지글도 파생 반복어(DIR, 7)와 굴절 반복어(IIR, 30)를 일치 반복어(ER, 105)에 합치면 무려 62.8%(142/226)가 반복어(R)가 된다. 아래 (12), (13)은 신문기사와 편지글에 나타나는 반복어의 예를 보인 것이다.

(12) 신문기사(T1) - 동아일보

- S1 : 전남 진도 해상에서 발생한 여객선 침몰사고의 사망자가 6명으로 늘었다.
 S2 : 그러나 290여 명이 실종 상태여서 사망자 수는 앞으로 더 늘 것으로 우려된다.
 S4 : 사망자 중 승무원 박지영(22·여)씨, 안산 단원고 2학년 정차웅(17)군, 권오현(17)군,
 임경빈(17)군 등 4명은 신원이 확인됐으나……

(13) 편지글(T1) - '생각하면 눈물이 앞을 가리고'

- S6 : 학교 다니면서 공부도 잘하고 혼자 다 해서 엄마, 아빠는 신경 쓸 게 없었는데……
 S9 : 엄마, 아빠가 너무 무심했어.
 S11 : 하여야, 엄마, 아빠는 잘 지내고 있어.
 S16 : 나중에 엄마, 아빠가 꼭 찾아갈게.

반복어(R)에 의한 어휘 응결이 이렇게 높게 나타나는 데는 몇 가지 이유가 있어 보인다. 우선, 반복어는 공간적으로 (멀리) 떨어져 있는 동일한 주제를 다시 가져 오는 가장 확실한 방법이다. 둘째, 반복어는 매크로 레벨 주제(예: (12)의 세월호 사건)보다 마이크로 레벨(예: (12)의 사망자)에서의 정보 처리를 집중적으로 관리해 주는 데 더 효과적이다. 셋째, 반복어는 텍스트 내에 암묵적으로 동의되고 있는 아이디어에 대한 확인 및 부인을 반복 표현하는데도 자주 나타난다. 예컨대, 아래 (14)는 '사망하다'를 반복함으로써 이 사실을 재차 확인시켜 주고 있다.

(14) 신문기사(T6) - 조선일보

S1 : 전남 진도에서 발생한 여객선 세월호 침몰 사고로 현재까지 단원고 2학년 학생을 포함, 2명이 사망한 것으로 알려졌다.

S2 : 16일 오전 8시 55분쯤 청해진해운 소속 여객선 '세월호'가 전남 진도군 관매도 인근 해상에서 침몰해 오후 3시 30분 현재 선사 여직원 박지영(22)씨와 단원고 2학년 정차용(17)군이 사망한 것으로 전해졌다.

S6 : 침몰중 구조신고를 받은 해양경찰 등이 긴급 출동해 탑승객들을 구조했으나 2명이 사망한고 293명의 생사 여부가 확인되지 않는 등 대형 참사를 피할 수 없는 상황이다.

통계 분석에서 발견되는 세 번째 특징은 어휘적 응결어의 종류별 빈도 순서가 신문기사나 편지글에서 큰 차이가 나지 않는다는 것이다.

(15) 신문기사 순서

반복어(R, 254) > 연상어(AC, 118) > 포함어(I, 67) > 유의어(S, 54) > 반의어(O, 23)

(16) 편지글 순서

반복어(R, 142) > 연상어(AC, 52) > 유의어(S, 17) > 반의어(O, 9) > 포함어(I, 6)

특히 반복어(R) > 연상어(AC)의 순서가 압도적으로 동일하고, 포함어, 유의어, 반의어는 포함어(상하위어)를 제외하면 대동소이하다. 이것은 우리가 텍스트를 분석하거나 어의 중의성을 해결할 때 반복어와 연상어가 가장 중요한 어휘적 응결어이고, 포함어, 유의어, 반의어는 텍스트의 장르적 차이를 반영하여 편차를 갖는 것으로 추정하게 해 준다.¹¹⁾

통계 분석의 네 번째 특징은 연상어(AC)의 비중이 반복어(R) 다음으로 매우 높다는 것이다. 본고에서 말하는 연상어는 기존의 언어(連語, Collocates)와 연상어(Associates)를 포함한 개념이다. 연상어의 빈도가 이렇게 높은 것은 사전이나 어휘의미망을 구축할 때 포함어(상위어, 하위어), 유의어, 반의어보다 연상어가 더 중요하고 유용한 정보라는 사실을 알려준다. 다시 말해서 연상어 자체의 사전 의미 정보도 중요하지만 연상어만의 어휘 의미망 정보도 매우 중요하다는 뜻이다.

주지하듯이 연어란 텍스트에서 우연히 발생할 확률보다 훨씬 높은 공기 가능성을 가진 어휘들의 묶음을 말한다. 이들은 대개 2개에서 6개 정도까지의 어절 사이에서 나타난다. 또 대부분 문장 내부에서 발생하며, 문장과 문장 사이에서는 거의 나타나지 않는다. 아래 (17)은 신문기사에서 발췌한 연어의 예, (18)은 편지글에서 발췌한 연어의 예들이다.

11) 특히 포함어, 즉 상위어-하위어 정보는 격식적인 글쓰기에서 많이 발견된다.

(17) 신문기사(T10) - 문화일보

S5 : 이 배는 15일 오후 9시쯤 인천여객터미널을 출항해 제주로 향하는 도중 사고를 당했다.

S10 : 또 화상 골절 등 중상을 입은 부상자 7명은 전남 목포시 목포한국병원 등으로 이송돼 치료를 받고 있다.

S14 : 김 실장은 청와대 위기관리센터에 머물며 사고 및 구조 현황을 파악하는 등 필요한 조치를 취하고 있는 것으로 전해졌다.

S15 : 특히 세월호에는 제주도로 수학여행을 가던 안산 단원고 325명이 교사들과 타고 있었는데 한때 전원 구조됐다는 소식이 들리기도 했지만 이후 사실이 아닌 것으로 밝혀져 부모들이 애를 태웠다.

(18) 편지글(T3) - 오늘도 아침 6시쯤 되니

S2 : 습관인지, 나이를 먹어서 그런지 눈이 떠진다.

S4 : 너를 학교에 보내기 위해 아침 6기면 일어나 아침 식사를 차려 주었지.

S24 : 엄마, 아빠 마음속 깊은 곳에 절대 성빈이를 잊지 않고 새겨 놓을게.

한편, 연상어(AC)는 이러한 위치 구분이 없이, 문장 내부에서도, 문장과 문장 사이에서도 나타난다. 연상어 개념의 기본적인 아이디어는 Filmore & Baker(2001)¹²⁾의 프레임 의미론(Frame Semantics)에서 제기한 것이다.¹³⁾ 그들에 의하면 어떤 개별 어휘의 의미는 외부 세계와 연결시키는 통로 역할을 하는 본질적인 세계 지식(essential world knowledge)이 없으면 이해할 수 없다는 것이다. 예를 들어, ‘팔다’라는 어휘의 의미를 상업적인 ‘상거래’의 맥락을 이해하지 못하면서 이해할 수는 없다. 즉 팔다, 구매자, 구입자, 상품, 돈 등의 연상어가 필수적으로 동원될 수밖에 없다. 이렇게 연상어는 촉발어(trigger)와 연상어(associates) 사이에 일어나는 일종의 인지적 추론 과정이라고 할 수 있다. 예를 들어 ‘세월호 사건’이라는 촉발어는 같은 경험을 공유하고 있는 사람들에게 자연스럽게 나머지 연상어들을 떠오르게 한다. 아래는 필자가 동아일보 기사(Text 1)에서 뽑아 본 연상어들의 목록이다.

(19) 동아일보(T1). 촉발어(‘세월호’) - 연상어

사망자, 해양경찰청(해경), 승객, 승무원, 탑승자, 사고, 실종자, 세월호, 사고원인, 쿵 하는 소리, 여객선, 구조, 침몰, 단원고

한편 명사가 아니라 동사도 촉발어가 될 수 있다. 아래 (20)은 촉발어가 동사인 연상어의 예를 편지글에서

12) Filmore, Charles J. & Collin F. Baker, “Frame semantics for text understanding”, *Proceedings of WordNet and Other Lexical Resources Workshop*, NAACL 2001.

13) FrameNet은 틀의미론(frame semantics)에 기반하여 International Computer Science Institute(UC Berkeley)에서 만든 어휘의미망(FrameNet lexical database)이다. FrameNet은 예를 들어, “John sold a car to Mary”라는 문장이 관점을 달리하면 “Mary bought a car from John”을 의미한다는 것을 밝히는 작업이다. 여기서 틀의미는 사건, 관계, 대상, 참여자 등(이들은 모두 개념구조 안의 성분들)을 기술하는 개념구조의 하나로 이해된다. 위키피디어 참조.
<https://en.wikipedia.org/wiki/FrameNet>

가져와 본 것이다. 편지글은 사건 자체보다는 사별한 자녀들에게 보내는 부모님들의 아픈 심리적 상태가 더 지배적인 틀(Frame)이라고 보아서, 심리적 상태를 나타내는 동사와 그 동사의 참여자들을 연상어로 처리했다. 그래서 그런지 명사 촉발어는 거의 발견되지 않았다.

(20) 편지글(T1. 촉발어(태어나다) - 연상어

태어나다, 아기, 딸, 유치원, 학교, 가족, 엄마, 아빠

그러나 본고의 연상어 연구는 아직 구체적인 방법론을 논의할 만큼 연구가 깊지 않다. 다만 예를 들어, ‘세월호 사건’을 소개하는 위키피디아의 사건 개요에 나오는 어휘들이 연상어가 될 수 있지 않을까 생각하고 있다.¹⁴⁾

통계적 분석 결과의 다섯 번째 특징은 포함어(I)이다. 포함어(I)는 신문기사에서는 세 번째 순서로 다수 출현했지만 편지글에서는 여섯 번째 순서로, 마지막 순서로 나타났다. 이것은 상위어나 하위어처럼 논리적인 어휘 관계를 나타내는 포함어가 서정적인 글쓰기에서 잘 선호되지 않았기 때문이 아닌가 생각해 본다.

실제로 거의 모든 글쓰기에서 포함어(I)가 유의어나 반의어보다 빈번하게 출현한다는 보고가 있다. 영어 텍스트를 대상으로 어휘 응결성을 다룬 Gómes(2018: 119)에서는 다자간 토론(격식)과 양자 간 전화 통화(비격식) 모두에서 포함어(I)가 유의어(S), 반의어(O)의 빈도를 앞선다고 보고하고 있다.

〈표 4〉에서 설명했듯이 포함어(I)는 상위어(GI)와 하위어(SI)를 모두 포함한다. 아래 포함어(21)은 ‘승객과 승무원’ → ‘탑승자’로의 상위어(GI) 관계를 보여준다. 그리고 (22)는 ‘잠수원’ → ‘해경’, ‘해군’, ‘경찰’로의 하위어(SI) 관계를 보여준다. 한편 (22, ②)은 포함어를 이용하여 글쓰기가 주제 전개 방식 중 파생형 주제 전개(the derived Theme pattern) 방식을 사용하고 있음을 보여 준다. 즉 ‘당국’이 ‘해경’으로 파생되었다.

(21) GI(상위어)

① T1 - 11, 12(신문기사)

여객선에는 승객과 승무원 475명이 탑승중이었다. 탑승자 가운데는 수학여행에 나선 안산 단원고 학생 325명이 포함돼 있었으며, 이날 낮 12시께 제주도 여객터미널에 도착할 예정이었다.

② T3 17, 19(편지글)

엄마, 아빠는 다른 부모님과 함께, 모든 국민이 너희를 기억하게 하여 다시는 이런 일이 일어나지 않게 하기 위해 활동하고 있다.

(22) SI(하위어)

① T1 - 5(신문기사)

현재 사고 현장에는 잠수원 178명(해경118명 · 해군 42명 · 경찰 18명), 선박 72척(해경 55

14) 물론 이것은 틀의미론(Frame Semantics)의 분석 방법론에 따라서 그 결과가 달라질 수 있다.

척·해군 17척), 항공기 18대(해경 14대·해군 4대)가 동원돼 구조 작업을 벌이고 있다.

② T8-1, 2(신문기사)

진도 해상서 학생 325명 등 460명이 탄 여객선이 침수 중이라는 조난신고가 들어와 당국이 구조에 나서고 있다.

T8 -2 해경은 경비정을 급파해 구조 중인 것으로 알려졌다.

신문기사의 경우, 총 67개의 포함어(I) 중에서 상위어(GI)가 7개(7%), 하위어(SI)가 60개(89%)로 하위어가 압도적으로 많이 쓰인다. 그리고 편지글의 경우는, 총 6개의 포함어(R) 중에서 상위어(GI)가 3개(50%), 하위어(SI)가 3개(50%)로 양쪽이 동등하게 쓰였다.

통계분석 결과의 여섯 번째 특징으로 유의어(S)를 보자. 우리는 Lyons(1977, 1981), Martin(1999)에 따라서 유의어(S)를 유사-유의어(NS)와 명제-유의어(PS)로 나누어 계산을 보았다. 그들에 의하면 이 둘은 척도적(scalar) 관계이지 이산적(either-or) 관계가 아니다. 유사-유의어는 원형적 어휘 의미와 동일하거나 점진적 척도를 갖는다. 반면에 명제-유의어(PS)는 같은 유의어지만 표현적 의미나, 문체적 의미가 다른 경우를 말한다.

아래 (23)에서 ‘눈시울’ - ‘눈물’ 등이 유사 유의어(NS)의 관계라면, 아래 (24)에서 ‘생일’ - ‘미역국을 먹다’ 등이 명제 유의어(PS)에 해당한다.

(23) 유사-유의어(NS)

① T4-1, 3(편지글)

눈시울 - 눈물

② T4-14(편지글)

분하다 - 억울하다

③ T7-1, 15(신문기사)

여객선 - 선박

④ T7-11, 13(신문기사)

탑승하고 - 타고 있다

⑤ T7-10(신문기사)

실종되다 - 생사가 확인되지 않다

⑥ T7-6, 10(신문기사)

파악하지 못했다 - 확인되지 않았다

(24) 명제 유의어(PS)

① T5-6(편지글)

생일 - 미역국을 먹다

② T5-6(편지글)

생일-배 속에 있다가 세상 밖으로 나온

③ T5-7, 10(편지글)

보고 싶다-빨리 만나서 꼬옥 안아 보고 싶고, 얼굴 비비고 싶다

④ T7-6, 8(신문기사)

숫자를 파악하다-집계하다

신문기사의 경우, 총 54개의 유의어(I) 중에서 유사-유의어(NS)가 46개(85.1%), 명제-유의어(PS)가 8개(14.8%)로 유사-유의어가 훨씬 많이 나타났다. 그리고 편지글의 경우는, 총 17개의 유의어(I) 중에서 유사-유의어(NS)가 5개(29.4%), 명제-유의어(PS)가 12개(70.5%)로 명제-유의어가 훨씬 더 많이 나타났다.

통계분석 결과의 7번째 특징으로 반의어(O)를 살펴보자. 반의어(O)는 다시 세 가지 하위구분을 갖는다. 상보 반의어(CpO)는 ‘여자-남자’처럼 이분적(either-or)인 절대값을 가진다. 따라서 한쪽을 부정하면 다른 쪽인 되는 관계이다. 그리고 등급 반의어(AO)는 ‘어렵다-쉽다’의 경우처럼 점진적이고 비양립적인 어휘 관계를 말한다. 그리고 방향 반의어(DO)는 반대 방향을 가리키거나, ‘입사-퇴사’처럼 반대 관계를 가리키는 응결성이다. 아래 (25)는 상보 반의어의 예를, (26)은 등급 반의어의 예를, (27)는 방향 반의어의 예를 세월호 텍스트에서 가져 온 것이다.

(25) 상보 반의어(CpO)

① T7-9(신문기사)

해경 관계자는 “선체가 기울며 침대 등에서 떨어져 부상을 입었거나 전기공급이 끊겨 승객들이 어둠 속에서 우왕좌왕하다 배 밖으로 탈출하지 못 했을 우려가 높다”고 밝혔다.

② T4-4(신문기사)

사망자 중 승무원 박지영(22·여)씨, 안산 단원고 2학년 정차웅(17)군·권오천(17)군·임경빈(17)군 등 4명은 신원이 확인됐으나 나머지 2명은 아직 확인되지 않았다.

(26) 등급 반의어(AO)

① T2-6(편지글)

방학도 끝나 가고 다시 새 학년, 새 학기 기다리며 지내고 있는 동생 수지와 친구들은 기쁘도, 슬픔도, 설렘도 함께(하고 있어).

② T3-5(신문기사)

사고 배에는 안산 단원고 2학년 학생 324명과 교사 10명이 탑승하고 있다.

(27) 방향 반의어(DO)

① T1-10(신문기사)

앞서 16일 오전 8시58분께 전남 진도군 조도면 병풍동 북방 1.8마일 해상에서 인천에서 출발해 제주로 향하던 6647급 여객선 세월호가 침수 중이라는 신고가 해경에 접수됐다.

② T1 - 12(신문기사)

탑승자 가운데는 수학여행에 나선 안산 단원고 학생 325명이 포함돼 있었으며, 이날 낮 12시께 제주도 여객터미널에 도착할 예정이었다.

반의어는 글쓴이가 글을 쓸 때 현재의 주제를 계속 유지시키려는 주제 전개 방식에서 많이 사용된다. 다만 같은 주제를 유지시켜 주는 방식인 '반복(R) 응결어'에 비해서 상대적으로 기피된다는 점이 다르다.

신문기사의 경우, 총 23개의 반의어(O) 중에서 상보 반의어(CpO)가 9개(39.1%), 등급 반의어(AO)가 12개(52.1%), 방향 반의어(DO)가 2개(8%)로 나타났다. 그리고 편지글의 경우는, 총 9개의 반의어(O) 중에서 상보 반의어(CpO)가 4개(44%), 등급 반의어(AO)가 4개(44%), 방향 반의어(DO)가 1개(11%)가 나왔다. 신문기사와 편지글 모두에서 방향 반의어(DO)의 빈도가 매우 낮았다.

통계분석 결과의 마지막 특징으로 신문기사와 편지글에서의 어휘 응결성 거리를 살펴보자. 응결성 거리는 두 요소 사이의 거리를 기준으로 응결력을 구분하는 것이다. 본고에서 이것은 다시 동일 문장 내부 요소들 간의 거리(WS)과 문장과 문장 사이의 요소들 간의 거리(AS)로 나누어 계산해 보았다. 후자는 다시 인접거리 유대(immediate ties), 원거리 유대(remote ties), 매개거리 유대(mediated ties)로 하위분류된다.

먼저 아래 <표 7>과 <표 8>은 신문기사의 응결성 거리와 편지글의 응결성 거리를 종합적으로 비교해 보인 것이다. <표 7>에서 보면 신문기사나 편지글을 막론하고 어휘적 응결어는 동일 문장 내부(WS)보다는 문장과 문장 사이(AS)에서 많이 나타나는 걸 확인할 수 있다. 신문기사는 WS가 119개, 23%이고, AS가 397개, 76.9%에 해당한다. 그리고 편지글에서도 WS가 47개, 20.7%이고, AS가 179개, 79.2%에 해당한다. 대부분의 어휘적 응결어들이 문장과 문장 사이에서 일어난다는 사실은 문장을 넘어 텍스트와 텍스트의 종류에 대한 정보가 얼마나 중요한 지를 잘 알려준다.

<표 7> 어휘 응결성 거리(신문기사와 편지 종합)

신문	ER	IR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
WT	10	1		8		2	9		4	44	41	119
AT	185	28	30	38	8	7	3	2	3	16	77	397
계	195	29	30	46	8	9	12	2	7	60	118	516
편지	ER	IR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
WT	10	2	2	1	5	0	3		1	1	22	47
AT	95	28	5	4	7	4	1	1	2	2	30	179
계	105	30	7	5	12	4	4	1	3	3	52	226

그리고 아래 <표 8>부터 <표 10>까지는 신문 기사를 기사별로 정리하고 어휘적 거리를 계산한 것이다.

응결성 거리(신문 기사)

<표 8> 신문 기사 별 어휘응결성 빈도

			ER	IIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
1	동아일보	WS				1		1	1			8		19
		AS	26	6	3	4	3	2		1	1	2	16	58
2	서울신문	WS	1			1						2		4
		AS	14	5	2	2					1	4	7	35
3	국민일보	WS	1						1			3		8
		AS	10	1	3	3		1				1	6	25
4	한겨레	WS	1			1						5	5	12
		AS	28	5	4	6	2			1		1	12	59
5	경향신문	WS	3	1							2	5	5	16
		AS	14	1	2	3	1		1			1	4	27
6	조선일보	WS	1			1		1	2			5	3	13
		AS	15	2	3	2			1			1	8	32
7	중앙일보	WS	1			2			2			2	5	12
		AS	20	1	6	8	1	1	1			3	7	48
8	한국일보	WS	1						1			2	3	7
		AS	11	2	1	2		1				2	5	24
9	세계일보	WS				1			1			4	3	3
		AS	14		2	2		1					6	6

			ER	IIIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
10	문화일보	WS	1			1			1		2	8	6	19
		AS	33	5	4	6	1	1			1	1	12	64

〈표 9〉 신문기사 별 동일 문장 내부 어휘 응결성 빈도(WS)

신문		ER	IIIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
동아일보	WS				1		1	1			8	8	
서울신문	WS	1			1						2		
국민일보	WS	1						1			3	3	
한겨레	WS	1			1						5	5	
경향신문	WS	3	1							2	5	5	
조선일보	WS	1			1		1	2			5	3	
중앙일보	WS	1			2			2			2	5	
한국일보	WS	1						1			2	3	
세계일보	WS				1			1			4	3	
문화일보	WS	1			1			1		2	8	6	
계		10	1		8		2	9		4	44	41	119

〈표 10〉 신문기사 별 문장과 문장 사이 어휘 응결성 빈도(AS)

신문		ER	IIIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
동아일보	AS	26	6	3	4	3	2		1	1	2	10	
서울신문	AS	14	5	2	2					1	4	7	
국민일보	AS	10	1	3	3		1				1	6	
한겨레	AS	28	5	4	6	2			1		1	12	
경향신문	AS	14	1	2	3	1		1			1	4	
조선일보	AS	15	2	3	2			1			1	8	
중앙일보	AS	20	1	6	8	1	1	1			3	7	
한국일보	AS	11	2	1	2		1				2	5	
세계일보	AS	14		2	2		1					6	
문화일보	AS	33	5	4	6	1	1			1	1	12	
계		185	28	30	38	8	7	3	2	3	16	77	397

그리고 아래 〈표 11〉부터 〈표 13〉까지는 개별 편지글에 대한 어휘적 거리를 계산한 것이다.

응결성 거리(편지글)

〈표 11〉 편지글 수별 어휘 응결성 빈도

		ER	IIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
T1	WS	1		1		1						4	9
	AS	15	5			3		1		1		4	29
<hr/>													
		ER	IIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
T2	WS	4	1	1				1				6	13
	AS	17	3	2	1	1	1				2	3	30
<hr/>													
		ER	IIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
T3	WS	3	1			1		1		1		3	10
	AS	23	7			1	1					7	39
<hr/>													
		ER	IIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
T4	WS				1			1				6	8
	AS	11	3	1	1	1			1			4	22
<hr/>													
		ER	IIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
T5	WS	1				2					1		4
	AS	7	2	1	1	1	1			1		4	18
<hr/>													
		ER	IIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
T6	WS	1				1						1	3
	AS	22	8	1	1		1					8	41

〈표 12〉 동일 문장 내 어휘 응결성 빈도(WS)

편지글		ER	IIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
T1	WS	1		1		1						6	
T2	WS	4	1	1				1				6	
T3	WS	3	1			1		1		1		3	
T4	WS				1			1				6	
T5	WS	1				2					1		
T6	WS	1				1						1	
계		10	2	2	1	5	0	3	0	1	1	22	47

〈표 13〉 문장과 문장 사이 어휘 응결성 빈도(AS)

편지글		ER	IIR	DIR	NS	PS	CpO	AO	DO	GI	SI	AC	계
T1	AS	15	5			3		1		1		4	
T2	AS	17	3	2	1	1	1				2	3	
T3	AS	23	7			1	1					7	
T4	AS	11	3	1	1	1			1			4	
T5	AS	7	2	1	1	1	1			1		4	
T6	AS	22	8	1	1		1					8	
계		95	28	5	4	7	4	1	1	2	2	30	179

끝으로 아래 〈표 14〉~〈표 16〉은 문장 과 문장 사이(WS)의 하위 구분으로 각각 인접거리 유대(IM), 매개거리 유대(MM), 원거리 유대(DM)의 빈도를 제시한 것이다. 신문기사와 편지글의 장르적 차이는 찾기 어려웠다. 다만 예상과 달리, 상대적으로 원거리 유대(DM)의 빈도가 조금 더 높다는 점이 특이하다. 〈표 14〉는 신문과 편지의 어휘적 응결어 거리 종합, 〈표 15〉는 신문기사의 어휘적 응결어 거리(종합), 〈표 16〉은 편지글의 어휘적 응결어 거리(종합) 빈도를 보여 준다.

〈표 14〉 신문기사와 편지글의 어휘적 응결어 거리(종합)

텍스트	IM	MM	DM	계
신문	126	110	161	397
편지글	48	65	66	179
합계	174	175	227	576

〈표 15〉 신문기사의 어휘적 응결어 거리(종합)

신문	IM	MM	DM	계
동아일보	8	15	35	58
서울신문	10	11	14	35
국민일보	17	6	2	25
한겨레	14	17	28	59
경향신문	4	10	13	27
조선일보	18	8	6	32
중앙일보	9	11	28	48
한국일보	11	7	6	24
세계일보	9	7	9	25

신문	IM	MM	DM	계
문화일보	26	18	20	64
합계	126	110	161	397

〈표 16〉 편지들의 어휘적 응결어 거리(종합)

편지글	IM	MM	DM	계
T1	9	10	10	29
T2	4	11	15	30
T3	12	11	16	39
T4	5	9	8	22
T5	8	5	5	18
T6	10	19	12	41
합계	48	65	66	179

V. 어휘의미망(WordNet)과 어의 중의성 해소

앞장에서 다룬 어휘적 응결성의 종류들, 즉 반복어(R), 연상어(AC), 포함어(I), 유의어(S), 반의어(O) 등은 대부분 어휘의미망(WordNet)의 관련어 성분들이다. 그러므로 본고에서 논의한 어휘적 응결성의 종류와 그 정보가 잘 반영된 한국어 어휘의미망이 구축될 수 있다면 국어 텍스트의 이해와 교육은 물론 텍스트의 기계 번역 등에서 어의 중의성 문제를 해소하는 방법으로 사용될 수가 있다. 여기서는 그 가능성 두 가지를 소개하기로 한다.

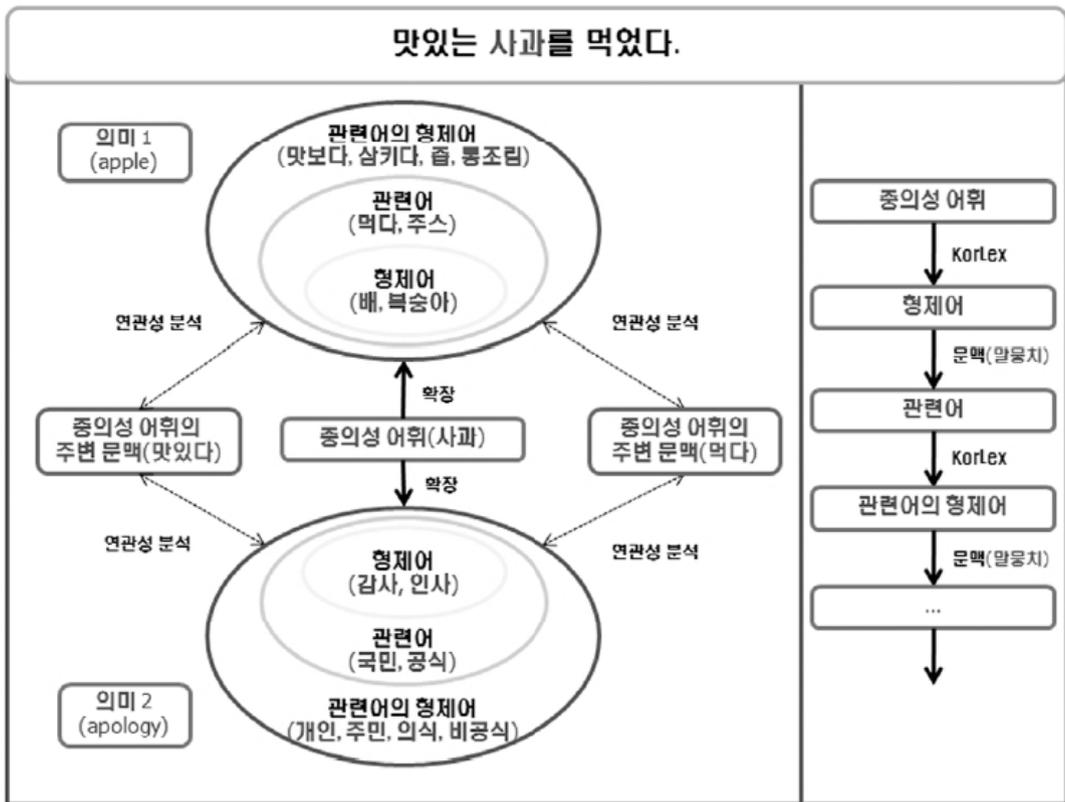
먼저 김민호·권혁철(2011)에서는 한국어 어휘의미망(KorLex)을 이용한 비감독 어의 중의성 해소 방법을 제안한다. 일반적으로 중간에 감독하는 감독 중의성 해소가 비감독 중의성 해소보다 성능이 좋게 나타나지만, 대규모 의미 부착 말뭉치가 필요하다는 단점이 있다. 그들에 의하면 중의성 어휘의 주변 문맥에 나타나는 공기(하는) 어휘들은 (이것이 본고에서 말하는 어휘적 응결어들이다.) 중의성 어휘의 의미를 판단하는 중요한 단서가 된다. 김민호·권혁철(2011: 557)에서 이렇게 대규모 의미 부착 말뭉치의 도움 없이도 중의성을 해소하는 절차는 다음과 같다.

(28) KorLex를 이용한 중의성 해소 절차

- ① 중의성 어휘의 관련어들을 한국어 어휘의미망(KorLex)에서 찾는다
- ② 중의성 어휘와 그 주변 공기 어휘 간의 공기 빈도를 대규모 말뭉치로부터 획득하여 카이스퀘어(chi square) 값을 구한다.

③ 가장 큰 카이 스퀘어 값을 가지는 공기어와 일치하는 관련어가 중의성 어휘의 의미가 된다.

아래 <그림>은 중의성 어휘 ‘사과’(apple, apology)가 들어 있는 문장에서 KorLex를 이용한 중의성 해소 절차가 어떻게 진행 되는지를 보여준다. 간단히 설명하면, (28, ①)은 KorLex에서 중의성 어휘인 ‘사과’의 형제어¹⁵⁾를 먼저 찾고, 관련어를 찾고, 마지막으로 관련어의 형제어를 찾는다. (28, ②)는 중의성 어휘(‘사과’)와 주변 문맥(‘맛있다’, ‘먹다’) 어휘와의 공기 빈도를 대규모 말뭉치로부터 카이 스퀘어 값을 구한다. (28, ③)은 가장 큰 카이 스퀘어 값을 가지는 ‘사과’(apple)이 중의성 어휘의 최종 의미가 된다.



한편, 연상어(AC) 정보를 이용하여 어의 중의성을 해소할 수 있는 방법도 있다. 프레임넷(FrameNet)¹⁶⁾은 버클리 대학교에서 만든 영어 어휘 의미 데이터베이스로서 단어 각각의 모든 언어적, 상황적 정보를 찾아 내는 기능을 위해 고안되었다. 또한 개념 기반의 어휘 분석을 통해서 통사론과 의미론의 연결을 형식화하고,

15) ‘형제어’는 김민호 · 권혁철(2011)의 용어이다.
 16) 온라인 접근은 <http://framenet.icsi.berkeley.edu/>에서 가능하다.

단어 의미의 중의성 해소, 정보추출, 기계번역, 질의응답 등 전자사전을 이용하여 자연언어 처리의 문제점을 해결하려는 데 기여하고 있다.

예를 들어, 프레임넷(FrameNet)에서는 모든 어휘들을 개별 단의(單意, sense)로 구분하고 그것을 기준으로 어휘단위(Lexical Unit, LU)를 설정한다. 이런 어휘단위들은 자기의 프레임 구조를 가지는데, 각각의 프레임은 프레임 요소(Frame Element, FE)들의 구조 차이에 의해서 구별된다. 본고의 연상어는 이런 프레임 요소들을 가리키며, 이들을 이용하여 어의 중의성을 해소할 수 있다는 것이다. 그럼 먼저 개별 프레임이 어떤 절차로 정의되는지를 어휘단위(LU), ATTACHING 프레임을 통해서 소개해 보기로 한다.

(STEP1) 아래는 말뭉치를 통해 전형적인 ATTACHING(밑줄) 의미 문장들을 뽑아 낸 것이다.

(29) They took her home and tied her to the bed with string saved from their Christmas packages.

그들은 그녀를 집으로 데려가서 크리스마스 소포에서 구한 끈으로 묶었다.

(30) Lise retrieves them to tie his hands together.

리스는 그의 두 손을 묶기 위해 그것들을 수선했다.

(31) Amelia rolled up Nina's torn sleeve and tied the tourniquet tight around her upper arm.

아메리아는 니나의 찢어진 소매를 걷어 올리고 지혈제를 그녀의 위쪽 팔에 단단히 묶었다.

(32) The calves, when only a few days old, are tethered by the neck in a stall about 60 cms wide and 150 cms long, until they are too large to turn.

태어난 지 며칠밖에 되지 않은 송아지는 너무 커서 돌릴 수 없는 때까지 가로 60cm, 세로 150cm 정도의 마구간 걸대에 묶어둔다.

(33) A garland was stapled to the underside of the mantelpiece.

화환 하나가 벽난로 밑면에 스테이플러로 꽂혀 있었다.

(STEP2) 위와 같은 말뭉치 예들을 통해서 ‘attaching’ 프레임의 느슨한 정의를 해 낼 수 있다. 즉, 누군가가 어떤 커넥터를 사용하여 어떤 것을 다른 것에 부착(또는 접착 또는 연결) 하는 상황과 관련이 있다는 것이다. 이를 통하여 이 프레임의 핵심(Core) 참여자들을 특징지을 수 있다. 이것이 바로 핵심 프레임 요소(Core frame element)이다.

(STEP3) 핵심 참여자를 차례대로 살펴보자. 먼저 행위자(AGENT) 참여자는 두 가지 물품을 부착시키는 역할을 하는 사람이다. 위의 (29)~(31)의 예에서 행위자는 *they*, *Lise*, *Amelia*로 예시되어 있다. 그리고 (32)~(33)의 예에서는 행위자가 나타나지 않는다. 둘째로, 대상(ITEM) 참여자는, 비대칭적인 경우에서, 더 큰 고정된 목표(GOAL)에 부착하는 좀 더 작은 대상을 말한다. 위 (29)~(33)에서 ‘ITEM’은 *her*, *the tourniquet*, *the calves*, *a garland*이다. 두 대상인 ITEM과 GOAL은 부착되어 떨어질 수 없다. 그러나 (30)의 *his hands*는 ITEM에 포함되지 않는다. 이것은 ATTACHING의 반대, 즉 비대칭 버전에서 다루어진다.

(29)에서 ‘크리스마스 소포에서 구한 끈으로’라는 전치사구에서 의해서 실현된 프레임 요소가 있는데, 이

Definition:

An **Agent** attaches an **Item** to a **Goal**, generally by manipulating a **Connector**, such that the **Item** is fixed to the **Goal**. Alternatively, an **Agent** uses a **Connector** to attach two **Items** to each other, where each serves as a **Goal** for the other. In either case, the **Connector**, if present, remains to bind the **Item** or **Items**. This asymmetric or symmetric relationship profiles the parts and has them remain as parts, without creating a new entity.

The robbers **TIED** Paul to his chair with a rope.

The robbers **TIED** his ankles together.

The **Goal** may be expressed solely with a preposition:

Patsy **TIED** her skirt on.

In this sentence the **Goal** is understood as being "on her body".

FEs:

Core:

The **Agent** manipulates an object. **Agent** typically refers to animate beings, but is also extended to mechanical and other inanimate objects. It is most frequently expressed as the external argument of verbs.

Agent [Agent]

The robber **TIED** Harry to the chair.

Connector [Conn]

This FE identifies the entity or substance that forms the bond that maintains the **Item** or **Items** in a fixed position; it is usually expressed in a "with"-PP.

The robber **TIED** Harry to the chair with a rope.

Goal [Goal]

Goal identifies the location to which an **Item** is attached.

The robber **TIED** Harry to the chair.

Item [Item]

This FE identifies the **Item** that the **Agent** attaches to the **Goal**.

The robber **TIED** Harry to the chair.

Items [Items]

This FE identifies the plural **Items** that are tied together, resulting in their mutual attachment.

The robber **TIED** Harry's ankles together.

<그림 1> Frame Description for Attaching

것을 커넥터(CONNECTOR)라고 부른다. 이것은 ITEM과 GOAL을 부착시키는데 사용된 물리적 대상이다. 예를 들어, glue나 staple 등이 여기에 사용될 수 있다. 마지막으로 (32)의 by the neck과 같은 것은 커넥터가 그것을 잡고 있을 수 있게 되어 있는 ITEM의 일부분이므로 핸들(HANDLE)이라고 부른다.¹⁷⁾

(STEP4) ATTACHING 프레임의 관점에서 볼 때, 이러한 핵심 프레임 요소들 외에 주변적인(peripheral) 프레임 요소들도 있다. 이들은 이 프레임의 모든 행위적 행동에 나타날 수 있기 때문이다. 여기에 해당하는 것으로 장소(PLACE), 시간(TIME), 수단(MEANS), 방법(MANNERS)이 있다. 또 결과(RESULT)도 이러한 주변 프레임 요소에 포함된다. 즉, (5)에서 together는 양손을 묶은 결과로 양손이 지금이 함께 있다는 걸 보여주기 때문이다.

이상을 정리하면, ACTOR는 의도와 동기가 부여될 수 있는 자발적 행위자이다. 그리고 ITEM은 크거나 좀 더 견고한 GOAL에 부착되는 대상으로 후자인 GOAL은 종종 언급되지 않을 수도 있다.

부착 행위는 반드시 그것이 식별되었든, 또는 동사의 의미로 융합 되었든지와 무관하게 CONNECTOR의 도움으로만 성취될 수 있다. ITEM과 GOAL은 일단 부착이 끝난 뒤에도 하나의 새로운 물체로 인식되는 않는다. 만약 그러면 FUSION 프레임이 된다. 아래는 지금까지 소개한 ATTACHING FRAME의 최종 모습을 보인 것이다.

VI. 결 론

지금까지 어휘적 응결성과 어휘의미망의 관계에 대하여 논의해 보았다. 주요 논의 내용은 두 가지였다. 첫째로, 체계기능언어학의 어휘적 응결성(Lexical cohesion)의 개념을 소개하고, 이것이 장르(텍스트)에 따라서 어떤 차이를 보이는지 살펴보았다. 둘째로, 어휘적 응결성, 그 중에서도 연상어(Associate) 정보를 갖춘 어휘의미망을 이용하여 어의 중의성 문제를 해소시킬 수 있는 가능성에 대해 지적해 보았다.

응결성은 어떤 종류의 언어 자원이 쓰였는가에 따라서 문법적 응결성과 어휘적 응결성으로 구별 된다. 본고는 지금까지 세월호 침몰 사고에 대한 10편의 신문기사와 세월호 유가족이 쓴 7편의 편지글을 대상으로 하여, 11가지 종류의 어휘적 응결성과 5가지 종류의 응결성 거리의 출현 양상을 살펴보았다. 분석 결과 다음과 같은 사실들을 알게 되었다.

첫째, 신문기사의 어휘 응결성 빈도가 편지글의 그것보다 2.3배 높게 나왔다. 이것은 직업적인 글이나, 다수 독자를 대상으로 하는 글일수록 어휘적 응결도가 높다는 것을 보여준다. 둘째, 어휘적 응결성의 유형별 출현 빈도가 신문기사와 편지글에서 큰 차이를 보이지 않았다. 그 순서는 반복어(R) > 연상어(AC) > 포함어

17) FrameNet에서 참여자의 명칭은 프레임 외부의 일반적인 광의의 의미역을 사용하지 않고, (예: AGENT, UNDERGOER, EXPERIENCER etc.) 각각의 프레임 - 특정한 방법으로 명칭을 부여하고 있다. 그들은 이러한 입장을 취하는 가장 큰 이유로, 여러 가지 이유가 있지만, 가장 중요한 것은 미리 결정된 소수의 의미역 목록에 어떻게 맞출지를 고민할 필요 없이 특정 단어 그룹을 위한 프레임 요소들에 맞는 정확한 정의를 제공할 수 있기 때문이라고 한다(p.305).

(I) >유의어(S) > 반의어(O)이다. 셋째, 신문기사와 편지글 모두에서 반복어(R)와 연상어(AC)가 가장 강력한 어휘적 응결어임을 알 수 있었다. 넷째, 연상어(AC)는 반복어(R)을 제외하면 사실상 가장 중요한 어휘의 미망 자원이다. 다섯째, 포함어(I), 즉 상하위어는 신문기사에서는 유의어(S), 반의어(O)보다 높게 나타났지만, 편지글에서는 후자보다 낮게 나왔다. 그러나 대부분의 격식적인 글쓰기에서 포함어(I)가 유의어(S), 반의어(O)보다 빈도가 높다. 여섯째, 신문기사나 편지글 모두에서 유의어(S)의 빈도가 반의어(O)를 상회하였다. 일곱째, 어휘 응결성의 거리를 분석해 본 결과, 신문기사나 편지글을 막론하고 어휘적 응결성 거리의 빈도는 동일 문장 내부(WS, with-in-sentence)보다 문장과 문장 사이(AS, across-sentence)에서 높게 나타났다. 이것은 응결성을 파악할 때 문장 내부의 어휘 정보보다 문장 밖 어휘 정보들이 더 중요하다는 것을 알려준다.

그리고 본고에서 논의한 어휘적 응결성의 종류와 그 정보가 잘 반영된, 특히 연상어(AS) 정보가 잘 반영된 한국어 어휘의미망이 구축될 수 있다면 국어 텍스트의 어의 중의성 문제를 해결 할 수 있다는 가능성을 언급하였다. 이를 위한 예시로 김민호·권혁철(2011)에서 제시한 KorLex를 이용한 중의성 해소 절차를 소개했다.

특히 버클리 대학교에서 만든 영어 어휘의미 데이터베이스인 FrameNet에서는 모든 어휘들을 개별 單意(sense)로 구분하고, (이것을 어휘단위, Lexical Unit이라고 함), 각각의 어휘단위들은 자기의 Frame Structure를 갖는데, 이것을 구성하는 Frame Element(FE)들이 본고에서 말하는 연상어(AC) 정보임을 밝혔다. FrameNet은 이러한 개념 기반 어휘 분석을 통해서 통사론과 의미론의 연결을 형식화 하고, 단어 의미의 중의성 해소, 정보추출, 기계번역, 질의응답 등 전자사전을 이용하여 자연언어 처리의 문제점을 해결하려는 데 기여하고 있다.

〈참고문헌〉

- 김민호·권혁철, 「한국어 어휘의미망의 의미 관계를 이용한 어의 중의성 해소」, 『소프트웨어 및 응용』 38: 10, 2011.
- 김서형·유혜원·이동혁·이유진·정연주(역), 「체계기능언어학의 이해」, 역학(원저: Suzanne Eggins, An Introduction to systemic functional linguistics, (2nd ed.) London: Continuum, 2004.
- 남경완·이동혁, 「틀의미론으로 분석한 ‘사다’와 ‘팔다’의 의미 분절 양상」, 『한국어학』 29: 1, 2004.
- 안경화, 「구어체 텍스트의 응결 장치 연구: 토론 텍스트를 중심으로」, 『한국어교육』 12: 2, 국제한국어교육학회, 2001.
- 유민애, 「장르-텍스트 기반 문법 교육 내용 연구: 논리적 응결 장치를 중심으로」, 『한국언어문화학』 12: 1, 한국언어문화학회, 2015.
- 이관규·김서경·노하늘·성수진·신희성·유상미·이현주·정려란·정지현·정혜현, 『체계기능언어학개관』, 사회평론아카데미, 2021.

- 정희모, 「한국어 교육과 결속성(cohesion) 및 응집성(coherence)의 문제」, 『리터러시연구』 10: 4, 한국리터러시학회, 2019.
- Baker, M., *In Other Words*. London: Routledge, 1992.
- Dontcheva-Navratilova · O. & Povolná, R., *Coherence and Cohesion in Spoken and Written Discourse*. Cambridge: Scholars Press, 2009.
- Fillmore, Ch. J. & Baker, C. F., Frame semantics for text understanding. In *Proceedings of WordNet and Other Lexical Resources Workshop*, NAACL, Pittsburgh, June. Gzipped Postscript (112 KB) PDF (189 KB), 2001.
- Flowerdew, J., Use of signalling nouns across L1 and L2 writer corpora. *International Journal of Corpus Linguistics* 15(1): 2010.
- FrameNet <https://en.wikipedia.org/wiki/FrameNet> (Wikipedia)
- FrameNet Project Official Website, <https://framenet.icsi.berkeley.edu/fndrupal/>
- Fraser, B., Towards a Theory of Discourse Markers, 2005. Retrieved from <http://people.bu.edu/bfraser/>
- Gómez González, M. Á., Lexical cohesion revisited: A combined corpus and systemic-functional analysis, *Quaderns de Filologia: Estudis Lingüístics* XXIII: 2018.
- Halliday, M. A. K. & Hasan, R., *Cohesion in English*, London: Longman, 1976.
- Halliday, M. A. K. & Hasan, R., *Language, Context, and Text: Aspect of Language in a Social-Semiotic Perspective*. Oxford: Oxford University Press, 1985.
- Halliday, M. A. K. & Matthiessen, M. I. M., *An Introduction to Functional Grammar* (4th ed.). London: Edward Arnold, 2014.
- Moreno, A., The role of cohesive devices as textual constraints on relevance: A discourse-as-process view. *International Journal of English Studies* 3: 1, 2003.
- Ochs Keenan, E., Making it last: repetition in children's discourse. In Ervin-Tripp, S. & Mitchell-Kernan, C. (ed.) *Child Discourse*. New York: Academic Press, 1977.
- Sanders, T. & Pander Maat, H. · *Cohesion and Coherence: Linguistics Approaches*. *Encyclopedia of Language and Linguistics* (2nd ed.). Elsevier: London, 2006.
- Taboada, M., *Building Coherence and Cohesion: Task-Oriented Dialogue in English and Spanish*. Amsterdam / Philadelphia: John Benjamins, 2004.
- Tadros, A., *Prediction in Text*. (Discourse Analysis Monograph 10). English Language Research, Birmingham: University of Birmingham, 1985.

* 이 논문은 2021년 11월 26일에 투고되어,
2021년 12월 22일에 심사위원을 확정하고,
2022년 1월 3일까지 심사하고,
2022년 1월 6일에 게재가 확정되었음.

Abstract

Lexical Cohesion of Korean Texts and WordNet

- Perspective on Systemic Functional Linguistics (SFL) -

Han, Jeonghan*

The purpose of this paper is twofold. First, we introduced the concept of lexical cohesion in systematic functional linguistics (SFL) and examined how it differs according to Korean genre (text). Second, we pointed out the possibility of solving the ambiguity problems by using lexical cohesion, especially the WordNet with the Associates (AS) information.

Cohesion can be divided into grammatical cohesion and lexical cohesion depending on what kind of linguistic resources are used. This paper so far examined the patterns of 11 kinds of lexical cohesion and five kinds of cohesion distances, targeting 10 'newspaper articles' about the Ferry Sewol sinking accident and 7 'personal letters' written by the Ferry Sewol bereaved family. As a result of the analysis, we were able to confirm the discriminatory realization pattern according to the genre (text) (see Chapter 4 in detail).

In addition, we mentioned the possibility of solving the ambiguity problems of Korean texts if the types of lexical cohesion discussed in this paper and the Korean wordnets reflecting its information well, especially the Associate (AC) information, can be established. As an example, of this, Kim Min-ho & Kwon Hyuk-chul(2011) introduced the procedure for resolving ambiguities using KorLex.

In particular, FrameNet, an English lexical meaning database created by Berkeley University, classified all words into individual senses (which is called lexical unit (LU)), and revealed that each lexical unit has its own Frame Element (FE) information which we define as the associate (AC) in this paper.

FrameNet is contributing to formalizing the connection between syntax and semantics through this concept-based lexical analysis and to solving the problems of ambiguity processing by using electronic dictionaries such as ambiguities of word meanings, information extraction, machine translation, and question-and-answer, etc.

[Keywords] lexical cohesion, genre (text), wordnet, associate words, Systemic Functional Linguistics (SFL).

* Dankook University, Department of Korean Language & Literature

