

발행인: 강범모
편집인: 전영철
발행처: 한국언어정보학회
발행일: 2003년 10월 25일
URL: <http://society.kordic.re.kr/~ksli/>

449-791
경기도 용인시 모현면 왕산리 산89번지
한국외국어대학교 인문대학 언어인지과학과 전영철 교수 연구실내
전화: (031) 330-4239
e-mail: ycjun@hufs.ac.kr

1. [언어와 정보] 논문 모집

[언어와 정보]에 실을 원고를 수시로 모집하고 있습니다. 심사용 논문은 자유로운 스타일로 아래아 한글이나 MS Word로 작성하셔도 됩니다. 어느 경우이든 꼭 논문요약을 포함시켜 주시기 바라고, 그 외 논문의 스타일은 학회 홈페이지(<http://society.kordic.re.kr/~ksli/>)에 자세히 소개되어 있으니 참조하시기 바랍니다. 투고하실 분은 다음 주소로 보내주시기 바랍니다.

보내실 곳	우편번호 449-791 경기도 용인시 모현면 왕산리 산89번지 한국외국어대학교 인문대학 언어인지과학과 채희락 교수 (hrchae@hufs.ac.kr)
-------	--

2. 회원 가입 안내

2003년도 학회 연회비는 정회원 20,000원, 준회원 10,000원이며, 기관회원 100,000원, 평생회비는 250,000원입니다. 연회비는 발표회 모임에서 직접 내시든지 다음 은행 계좌로 송금하시면 됩니다. (예금주: 한국언어정보학회, 계좌번호 012336-01-001683, 우체국)

3. 2003년 회비 납부 현황

평생회원: 강범모(고려대), 광은주(세종대), 권희상(한밭대), 김경애(한경대), 김경학(수원대), 김량혜윤(고려대), 김미경(대덕대), 김성목(한국 IBM), 김용범(광운대), 김용환(양산대), 김종복(경희대), 김창익(호서대), 류병래(충남대), 박병수(경희대), 윤애선(부산대), 윤영은(이화여대), 윤평현(전남대), 이기용(고려대), 이남근(조선대), 이민행(연세대), 이영현(조선대), 이예식(경북대), 이익환(연세대), 이현우(인하대), 전영철(한국외대), 정대호(한양대), 정소우(성신여대), 정재현(충북대), 정찬(동서대), 채명희(조선이공대), 채희락(한국외대), 최규련(연세대), 최기선(KAIST), 최재웅(고려대), 최현숙(영남대), 한영균(울산대) 이상 36명

정회원: 김윤신(서울대), 남승호(서울대), 노용균(충남대), 박철우(서울대), 손유미(서울대), 신호필(서울대), 안희돈(건국대), 양동휘(서울대), 양정석(연세대), 염재일(홍익대), 유은정(서울대), 위혜경(한국사이버대학), 이계영(동국대), 이정민(서울대), 이혜경(건국대), 임경섭(동신대), 장석진(서울대), 전중섭(서울대), 조성은(서강대), 조준모(한동대), 채희숙(서울대), 최인철(경희대), 홍민표(명지대), 이상 23명

기관회원: 고려대, 광운대, 국회도서관, 명지대, 방송대, 세종대, 위덕대, 조선대, 한경대, 한국외대, 한양대 이상 11기관

4. 공지사항

ㄱ. 오는 12월 31일에 발간되는 [언어와 정보] 7-2호에 논문을 신고자 하시는 분은 11월 9일까지 원고를 투고하시면 가능합니다. 많은 투고 바랍니다.

ㄴ. 11월 정기 월례 모임은 다음 학술대회의 참가로 대신합니다.

한국현대언어학회, 한국언어정보학회, 한국영어학학회 공동학술대회
주제 : 언어분석과 전산처리
일시: 11월 29일 오전 10시부터
장소 : 한밭대학교(대전 유성), 유성 툅게이트에서 6-7분정도

우리 학회에서는 김종복(경희대), 조성은(서강대), 윤재학(경희대), 김애령(경남대) 네 분의 발표가 계획되어 있습니다.

ㄷ. 내년의 'PACLIC 18'이 일본의 와세다대학에서 2004년 12월 8일부터 10일까지 개최되는 것으로 결정되었음을 알려드립니다.

5. 발표논문초록

코퍼스 빈도 정보 활용을 위한 적정 통계 모형 연구

- 코퍼스 규모에 따른 타입/토큰의 함수관계 중심으로 -

박병선(고려대 국어국문학과), 양경숙(고려대 통계학과)

지금까지 코퍼스를 구축하는 데 있어서, 자료의 다양성을 고려한 자료 균형성 문제와 더불어 코퍼스 구축 규모의 문제는 매우 중요한 고려사항이다. 이런 문제는 일찍이 영어 코퍼스를 중심으로 많은 연구가 진행된 바가 있지만 한국어를 대상으로 한 엄밀한 연구는 많이 이루어지지 않았다.

이 논문의 일차적 목적은 엄밀한 통계적 방법을 적용하여 코퍼스 규모(크기)에 따른 타입과 토큰 간의 관계를 밝히고자 하는 것이다. 특히 이 연구에서는 현재까지의 국내 연구에서 다루지 않았던 규모인, 현대 한국어 코퍼스 1억여 어절을 대상으로 코퍼스 크기 증가에 따른 타입과 토큰의 상관성을 통계적 모델을 이용하여 규명하였다. 그리고 이를 위해 통계적 방법을 엄밀히 적용하여 좀더 명시적인 결과를 도출하였다.

이 논문의 연구에 의해 도출된 함수식을 활용하여 발전시키면 소규모의 자료를 이용하더라도, 대규모 자료에서의 빈도수를 예측해 낼 수 있으므로, 더욱 다양하고 정확한 통계처리의 근거를 제시할 수 있게 된다. 예를 들자면 규모가 1000만 어절인 코퍼스를 이용하여 1억어절 규모의 자료를 이용하는 것과 같은 효과를 낼 수 있다. 이 경우 1천만 어절 규모에서 저빈도인 단어를 통계처리할 때 그 결과를 신뢰하기 힘들지만 1억 어절 규모인 코퍼스에서의 빈도로 계산하여 통계처리한다면 결과의 신뢰성이 확보된다. 그리고 어느정도의 규모까지 코퍼스를 구축하면 되는지 그 방향을 제시할 수 있다.

이 연구에 사용한 자료인 현대국어 코퍼스의 총규모는 총 114,387,008어절로 이번 연구에서는 매100만 어절별로 토큰을 증가시키면서 누적된 타입 빈도를 계산하였다. 타입수의 변화가 언어 자료의 장르 특성에 영향을 받을 수 있기 때문에 분석에 사용된 코퍼스 파일 6999개를 난수를 이용하여 완전 랜덤화 하였다. 매번 완전 랜덤화된 데이터 파일로부터 토큰을 100만개씩 증가시키면서 누적된 타입수를 계산하였고 이와 같은 랜덤화 과정을 100회 반복하였다. 이와 같이 완전 랜덤화한 데이터로부터 타입수를 계산하는 과정을 100번 반복한 것은 장르특성으로부터 영향을 가능한 받지 않도록 만들면서 토큰과 타입간의 관계를 규명하기 위함이다.

결국 최종적으로 타입과 토큰간의 관계 규명을 위해 사용한 데이터는 100회 반복하여 계산된 타입수의 평균으로 큐빅모형(cubic model), 파워모형(power model) 그리고 ARIMA모형(auto-regressive integrated moving average model)에 적합시켜 그 결과를 비교하였다. 이들 3가지 모형에 대한 평가는 물론 잔차분석과 모형에 대한 설명력 등을 이용하여 이루어졌다. 이 결과들로부터

100만 어절 단위로 관측된 평균타입수(Y)에 대해 적합한 모형식은 다음과 같이 수립된다.

$$Y_t = Y_{t-1} - 0.599 Y_{t-1} - Y_{t-2} + 0.599 Y_{t-2} + Y_{t-3} - 0.599 Y_{t-3} \\ + 0.599 Y_{t-4} + a_t$$

여기서 아래첨자 t는 관측단위별 시점을 나타내는 것으로 토큰 100만 어절 단위로 표시하였다. 즉 t=10 이면 1000만어절의 토큰일 경우의 평균 타입수를 나타낸다.

본 논문에서는 이제까지 연구가 없었던 약 1억 어절이라는 대용량 데이터에 대하여 코퍼스의 크기 증가와 타입간의 상관성을 함수적 관계로 규명하였다. 더욱이 장르의 영향을 최대한 배제시키기 위해 100회의 모의실험을 통해 함수적 관계를 규명하였다.

그리고 세 모형을 통계적으로 비교하는 과정에서 오차의 가정을 벗어나는 관측값들을 살펴보면 서 통계적 처리를 위해서는 최소한 1000만 어절 이상의 규모가 필요할 것으로 판단되었다. 따라서 지금까지 '21세기 세종계획' 등에서 구축한 균형 코퍼스(balanced corpus)의 규모가 1000만 어절 규모인데 이 연구에 따르면, 그 규모는 통계적 처리에 비교적 적정 구축량을 증명할 수 있었다.

이 연구에서 밝혀진 통계모형을 코퍼스 구축에 있어서 장르별 특징과 시대별 특징을 대상으로 확대 발전시킨다면, 각 연구별 적정 코퍼스 구축량을 예측하고 가장 적절한 통계처리 연구를 할 수 있는 방법론의 근거를 제시한 의의가 있다고 하겠다. 그리고 저빈도 타입의 증가와 코퍼스 크기 증가에 따른 변화 특징 등의 연구에 응용하여, 저빈도 단어들의 적절한 통계 처리 방법을 개발해 낼 수 있다. 따라서 이 논문은 앞에서 언급한 여러 의미있는 연구를 진행하기 위한 기본적인면서 가장 중요한 검증과 방법론을 보였고 앞으로 계속 연구를 진행할 것이다.

Goal and Source: Asymmetry in their Syntax and Semantics Seungho Nam (Seoul National University)

0. This paper focuses on the syntactic and semantic asymmetry between Goal (e.g., into the store) and Source locatives (e.g., from the store). In the literature they have been treated as directional locatives, but they have not been properly analyzed from a contrastive and formal perspective. Based on their syntactic and semantic asymmetry, the paper argues that they have distinctive underlying base positions and further that they have different semantic scope/contribution in event structure. Thus we get a more explicit mapping principle between syntax and semantics of locative expressions in terms of extended VP-structure and event structure.

1. Proposal: Base Positions and semantic scopes of Locative PPs

1.1. Goal PPs (PP_G) are generated under the lower VP1, and they semantically compose a core event (result state: E2) denoted by the lower VP2.

1.2. Source PPs (PP_S) are generated under the higher VP1, and semantically modify the event (E1) denoted by the higher V1'. They do not compose the core event.

Non-directional PPs are generated as a VP1-adjunct and modify/locate the whole event denoted by VP1.

2. Goal-Source Asymmetry in Syntax: Koopman (1997) and Travis (2000) argue for the structural distinction between directional and non-directional PPs, and Kracht (2002) identifies the two in terms of "Modalizer" and "Localizer". The following observations, however, further support the asymmetry of Goal and Source.

2.1. Preposition incorporation: Goal-directional pre/post-positions get more easily incorporated into verb than Source or non-directional ones do (Baker 1990, Koopman 1997). Munro (2000) illustrates that some Chickasaw verbs can occur with more than one applicative prefixes, and Goal applicative always comes closer to the verb stem than Source or non-directional applicatives do.

2.2. Prepositional Passive: Couper-Kuhlen (1979) shows that Goal is a typical subject of prepositional passive, while Source or non-directional one is hardly found in the subject position.

2.3. Movement and Ordering: Source PPs are more easily dislocated than Goal PPs are.

2.4. Locative Alternation: Locative alternation in many languages involve Goal-type PPs rather than Source or non-directional ones. (Levin 1993; Hale & Keyser 1993, 2002) More generally, V-modifiers like Goal-type PPs allow alternations more often than VP-adjuncts like Source/Path-type or non-directional PPs do.

3. Source and Goal in Event Structure

3.1. Lexical Complex Event Structure: Extending the recent proposals on complex lexical event structure (Jackendoff 1990; Grimshaw 1990; Pustejovsky 1991, 1995), we identify the distinctive semantic contributions of Source and Goal in terms of semantic scope in event structure.

3.2. Event Structures for VPs and Locative PPs: Recent studies on adverbial modification argue for a finer-grained event structures (Eckardt 1998; Ernst 1998; Tenny 2000; Travis 2000). This paper further observes and accounts for the semantic asymmetry of Source and Goal. For instance, the following shows their asymmetry in argument orientation and ordering:

3.3. Aspectual Division: Cinque (1999) and Travis (2000) divide at least two aspectual domains in syntax - "higher/viewpoint" aspect and "middle/situation" aspect in Tenny's (2000) terms. The following data show that Source PPs scope over the situation aspect, so do not shift the aspectual character of the core event. Goal-PPs, however, composing the core event, play a crucial role in aspectual composition.

한국언어정보학회 2003학년도 2학기 월례회 일정표

장소: 대우재단 일시: 토요일 9:30-12:30

날짜	시간	발표자	논문제목	사회
9월 27일 (토)	기획논문 9:30-10:40	최인철 (U. of Texas at Austin)	Argument Structure and Case in Korean and English	유은정 (서울대)
	휴식(10분)			
	기획 논문 10:50-12:00	정덕교 (서울 시립대)	Complex Predicates under VP Hypothesis	
	일반 논문 12:00-12:30	김윤신 (서울대)	한국어 처소 이동 동사의 논항 교체 유형과 함의관계	
11월1일 (토)	기획논문 9:30-10:10	박병선, 양경숙 (고려대)	코퍼스 빈도정보 활용을 위한 적정 통계 모형 연구 -코퍼스 규모에 따른 타입/토큰의 함수관계 중심으로-	정대호 (한양대)
	일반 논문 10:10-10:50	남승호 (서울대)	Locative Arguments and their Semantics in Event Structure	
	휴식(10분)			
	논문강독 11:00-11:40	강범모 (고려대)	Discourse Representation Theory Ch. 1-2	
	논문강독 14:00-12:20	홍민표 (명지대)	Discourse Representation Theory Ch. 3	
11월29일 (토) 대전 현대 언어학회 학술대회	주제논문	김종복 (경희대)	한국어 동명사구문 분석과 전산언어학적 처리	
	일반논문	조성은 (서강대)	TBA	
	일반논문	윤재학 (경희대)	Restriction vs. Predication	
	일반논문	김애령 (경남대)	TBA	
12월20일 (토)	일반논문 9:30-10:40	채희락 (한국외대)	Nominal, Verbal and Cross categorial Particles in Korean	채명희 (조선 이공대학)
	휴식(10분)			
	일반논문 10:50-11:20	송민영 (동국대)	Tense and Choice Functions	
	논문강독 11:20-12:05	염재일 (홍익대)	Discourse Representation Theory Ch. 4	
	논문강독 12:05-12:30	정소우 (성신여대)	Discourse Representation Theory Ch. 5	

강독 자료

Kamp, Hans. 2003. *Discourse Representation Theory*. Ms.

DRAFT of article for the new edition of the *Handbook of Philosophical Logic*, jointly with Josef van Genabith and Uwe Reyle.

2003년 2학기 구구조문법연구회 워크샵

주제: Describing Semantic and Pragmatic Aspects in Constraint-Based Grammar

날짜	장소/ 시간	주제	발표자	토론자
9월 27일 (토)	대우재단/ 2:30- 5:00	Minimal Recursion Semantics	김종복 (경희대)	장석진 (서울대)
		Applications of Minimal Recursion Semantics	노용균 (충남대)	채희락 (한국외대)
11월 1일 (토)	대우재단/ 2:30- 5:00	Dative alternations in English	최혜원 (이화여대)	박병수 (경희대)
		Dative alternations in Korean	최인철 (경희대)	임경섭 (동신대)
11월 22 일 (토)	충남대/ 2:30- 5:00	Incorporating Pragmatic Information	유은정 (서울대)	차종렬 (경희대)
		Topicalization in Japanese	김용범 (광운대)	신효필 (서울대)
12월 13 일 (토)	조선대/ 2:30- 5:00	Constraint-based Semantics for Tense and Aspect	류병래 (충남대)	채명희 (조선 이공대)
		Complex Tenses in Japanese	조세연 (호남대)	정찬 (동서대)

[Reading List]:

9월 27일: Minimal Recursion Semantics and its Applications

Copestake et al. 2000. Minimal recursion semantics: an introduction. ms. Stanford University.

Egg, Markus. 1998. Wh-questions in underspecified minimal recursion semantics. *Journal of Semantics* 15.1, 37-82.

(Flickinger, D. and E. Bender. 2003. Compositional semantics in a multilingual grammar source. Paper presented in a 2003 ESSLI workshop.)

11월 1일: Dative Alternations in English/Korean

Bresnan, J. and T. Nikitina. 2003. On the Gradience of the Dative Alternation. ms. Stanford University. (available on Bresnan's website)

Levin, B. and M. Rappaport. 2002. What Alternates in the Dative Alternation? Paper presented at the 2002 Conference on Role and Reference Grammar: New Topics in Functional Linguistics: The Cognitive and Discursive Dimension of Morphology, Syntax and Semantics, Universidad de La Rioja, Logrono, Spain, July 27-28, 2002. (available on Levin's website)

Yoon, James. to appear. Non-nominative (Major) Subjects in Korean. To appear In P. Bhaskararao and K. V. Subbarao (eds.), *Non-nominative Subjects*. Berlin: Mouton de Gruyter.

11월 22일: Incorporating Pragmatic Information

Pollard & Sag. 1994. *Head-Driven Phrase Structure Grammar*, Ch.8. 332-337.

- Wilcock, Graham. 1999. Lexicalization of context. In G. Webelhuth, J.-P. Koenig, & A. Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, 373-387. CSLI Publications.
- Green, Georgia. 2000. The nature of pragmatic information. In R. Cann, C. Grover, & P. Miller (eds.), *Grammatical Interfaces in HPSG*, 113-135. CSLI Publications.
- Ginzburg & Sag. 2000. Interrogative Investigations, Ch.7.2, 255-279. (on illocutionary force)
- Fukushima, Kazuhiko. 1999. A lexical comment on a syntactic topic. In R. Levine & G. Green (eds.), *Studies in Contemporary Phrase Structure Grammar*, 199-222. Cambridge.

12월 13일: Tense and Aspect

- Van Eynde, Frank. 2000. A constraint-based semantics for tenses and temporal auxiliaries. In R. Cann, C. Grover, & P. Miller (eds.), *Grammatical Interfaces in HPSG*, 231-249. CSLI Publications.
- Bonami, Olivier. 2002. A syntax-semantics interface for tense and aspect in French. In F. Van Eynde, L. Hellan, & D. Beermann (eds.), *Proceedings of the 8th HPSG Conference*, 31-50. CSLI Publications.
- Yoshimoto, K. & Y. Mori. 2002. A compositional semantics for complex tenses. In F. Van Eynde, L. Hellan, & D. Beermann (eds.), *Proceedings of the 8th HPSG Conference*, 300-319. CSLI Publications.
- (Yoshimoto, Kei. 2000. *Tense and Aspect in Japanese and English*. Ch.4 Analyzing Japanese aspect. Peter Lang.)

자료집이 필요하신 분은 서울대 유은정 선생님께 (Email: ejyoo@snu.ac.kr) 연락 주시기 바랍니다.