

발행인: 남승호  
편집인: 이해운  
발행처: 한국언어정보학회  
발행일: 2012년 10월 15일  
URL: <http://society.kisti.re.kr/~ksli/>

☎☎-☎☎☎

경기도 용인시 처인구 모현면  
한국외국어대학교 인문대학 언어인지과학과 이해운 교수 연구실내  
전화: (031) 330-4284  
e-mail: haeyun@hufs.ac.kr

## 1. 월례 논문 발표회

이번 10월 월례 발표회(10월 20일)에서는 손익수(경북대), 신효필(서울대), 정대호(한양대), 이한정 & 김나연(성균관대) 임동식 선생님의 발표가 계획되어 있습니다. 이번 학기의 발표 일정 및 발표요지는 아래를 참고하시고, 많은 참석과 관심을 바랍니다.

## 2. [언어와 정보] 논문 모집

- ▶ [언어와 정보]에 실을 원고를 수시로 모집하고 있습니다. 심사용 논문은 자유로운 스타일로 아래아 한글이나 MS Word로 작성하셔도 됩니다. 논문의 스타일은 학회 홈페이지 (<http://society.kisti.re.kr/~ksli/>)에 자세히 소개되어 있으니 참조하시기 바랍니다.
- ▶ 논문 심사료와 게재료는 다음과 같습니다. 논문 출판 시 라텍 편집비는 따로 받지 않습니다.

심사료: 30,000원

논문 게재료:

- 1) 연구비 수혜 논문: 270,000원 (심사료 포함 총 300,000원)
- 2) 연구비 수혜를 받지 않은 일반논문:
  - (i) 저자 중 한 사람 이상이 전임인 경우 170,000원 (심사료 포함 총 200,000원)
  - (ii) 모든 저자가 전임이 아닌 경우 20,000원 (심사료 포함 총 50,000원)

- ▶ 논문 게재에 관하여 기타 문의사항이 있으시거나, 원고 제출에 관해 의논하시려면 편집위원장이신 류병래 교수([ryu@cnu.ac.kr](mailto:ryu@cnu.ac.kr))께 연락하시면 됩니다.

## 3. 회원 가입 및 회비 납부 안내

- ▶ 우리 학회는 회원 여러분들께서 납부하신 회비로 운영됩니다. 원활한 학회 운영을 위해 연회비를 납부해 주시기 바랍니다. 회비는 월례 발표회장에서 또는 아래의 학회계좌로 납부하실 수 있습니다. (계좌번호: 우리은행 1002-839-410880 예금주: 이영주)

연회비: 정회원 40,000원, 준회원 20,000원, 기관회원 100,000원

평생회비: 400,000원

- ▶ 회원주소록은 연중 관리하고 있습니다. 주소 및 이메일, 전화번호가 바뀌신 분은 변경된 내용을 총

무이사 (haeyun@hufs.ac.kr)에게 알려주시면 많은 도움이 되겠습니다. 아울러 주변의 교수님들과 대학원생들이 신입회원으로 많이 가입할 수 있도록 적극적으로 홍보해 주시기를 부탁드립니다.

#### **4. 학술대회 안내**

▶ **The 26<sup>th</sup> Pacific Asia Conference on Language, Information, and Computation (PACLIC 26)**

- 일시: 2012년 11월 08~10일
- 장소: Bali
- 사이트: <http://pacific26.cs.ui.ac.id>

▶ **[CFP] 2012 International Conference for the Korean Association for Corpus Linguistics (KACL 2012)**

- 일시: 2012년 12월 10~11일
- 장소: 부산
- 사이트: <http://kacl.pusan.ac.kr>
- 초록마감: 2012년 10월 15일

**한국언어정보학회 2012학년도 2학기 월례회 일정표**  
**장소: 대우재단 7층 1실    시간: 9:30 - 13:00**

| 날짜           | 시간                 | 구분  | 발표자                 | 논문 제목   | 사회               |
|--------------|--------------------|-----|---------------------|---|------------------|
| 9/22<br>(토)  | 09:30<br>10:30     | 기획1 | 임동식 (CCHS-<br>CSIC) | Evidentials at semantics-pragmatics interface:<br>Perspective shifts and context-overwriting effects of<br>Korean | 류병래<br><br>(충남대) |
|              | Coffee Break (15분) |     |                     |   |                  |
|              | 10:45<br>11:45     | 특강  | 윤애선 (부산대)           | Contrastive expressions in Korean and Hungarian   |                  |
|              | Coffee Break (15분) |     |                     |   |                  |
|              | 12:00<br>12:45     | 일반  | 김지혜 (송실대)           | Tensed-S Condition (TSC) and the Determination<br>of Binding Domain of Anaphors in Korean                         |                  |
| 10/20<br>(토) | 09:30<br>10:30     | 기획  | 손익수 (경북대)           | 동의어 관계의 분포적 성질  | 양정석<br><br>(연세대) |
|              | Coffee Break (15분) |     |                     |   |                  |
|              | 10:45<br>11:45     | 특강  | 신효필 (서울대)           | Constructing Sentiment Corpus in Korean   |                  |
|              | Coffee Break (15분) |     |                     |   |                  |
|              | 12:00<br>12:30     | 일반  | 정대호 (한양대)           | Some Negative Comments on the Negative<br>Quantifier Analysis of <i>amwu(N)-to</i>                                |                  |
|              | 12:30<br>13:00     | 일반  | 이한정, 김나연<br>(성균관대)  | Usage Probability, Non-Canonical Word Order and<br>Subject-Object Asymmetry in Korean Case Ellipsis               |                  |
| 11/17<br>(토) | 09:30<br>10:30     | 기획1 | 임현서 (서울대)           | TBA   | 미정               |
|              | Coffee Break (15분) |     |                     |   |                  |
|              | 10:45<br>11:45     | 특강  | 이용훈 (충남대)           | TBA   |                  |
|              | Coffee Break (15분) |     |                     |   |                  |
|              | 12:00<br>12:30     | 일반  | 김지은 (울산대)           | TBA   |                  |
|              | 12:30<br>13:00     | 일반  | 김동성 (고려대)           | TBA   |                  |

|              |                    |     |               |                            |              |
|--------------|--------------------|-----|---------------|----------------------------|--------------|
| 12/15<br>(토) | 09:30<br>10:30     | 특강1 | 최승권 (ETRI)    | 대화시스템을 위한 사용자 의도 태깅 코퍼스 구축 | 김윤신<br>(신라대) |
|              | Coffee Break (15분) |     |               |                            |              |
|              | 10:45<br>11:45     | 특강2 | 이기용 (고려대)     | TBA                        |              |
|              | Coffee Break (15분) |     |               |                            |              |
|              | 12:00<br>12:30     | 일반  | 김지은<br>(한국외대) | TBA                        |              |
|              | 12:30<br>13:00     | 일반  | 이우승 (협성대)     | TBA                        |              |

## **한국언어정보학회 2012년 10월 월례회 발표논문 초록**

### <기획 논문 1>

#### 동의어 관계의 분포적 성질

손 익 수 (경북대)

두 단어가 동일한 혹은 유사한 의미를 가지고 있을 때 우리는 흔히 동의어 관계라고 한다. 동의어는 해당단어와 의미자질을 충분히 공유하며 의미의 대비가 거의 일어나지 않는다. 이러한 단어의 유사성에 대한 견해는 의미의 내연적 측면에 바탕을 두고 있다고 볼 수 있다. 또 하나의 의미의 유사성에 대한 접근은 분포적 측면이다. 분포적 의미는 해당 단어와 함께 공기는 단어들에 의해서 규정된다. 만약 어떤 두 단어가 공기 하는 단어들을 공유한다면 이들 간의 유사성이 존재한다고 보아야겠다.

본 연구에서는 독일어에서 나타나는 동의어 관계의 분포적 유사성을 살펴보고자 한다. 이를 위하여 게르마넷(GermaNet) 7.0 과 DISCO(extracting Distributionally related words using CO-occurrences)를 이용하였다. 동의어 관계의 분포적 성질에 대한 고찰은 두 가지 방법으로 이루어진다. 첫 번째 방법에서는 수집된 모든 동의어 관계의 유사도 분포를 대상으로 하며 반의어 관계의 유사도 분포와 비교하여 나타나는 차이점을 살펴봄으로써 동의어 관계의 분포적 성질의 파악을 시도한다. 또 하나의 방법으로써 동일한 개념을 나타내는 단어들에 한정하여 이들이 형성하는 동의어 관계의 유사도를 살펴보는 것이다. 이러한 경우 분석 단위는 게르마넷에서의 경우 동일한 신셋(synset)에 속하는 단어들로 이루어진 동의어 관계들의 집합이 된다.

분석 결과에서는 동의어 관계가 가지는 내연적 측면에서의 의미 유사성은 분포적 측면에서 크게 반영되지 않음을 나타내고 있다. 오히려 반의어 관계가 가지는 분포적 유사도의 평균과 중간 값이 동의어 관계에서 보다 더 높게 나타나고 있다. 또한 동일한 개념을 나타내더라도 동의어들의 조합에 따라 상이한 유사도를 가지지만 동의어 관계를 이루는 단어들의 뉴앙스나 사투리 등과 같은 비기

술적 의미(non-descriptive meaning)가 유사한 경우 높은 분포적 유사도가 나타난다. 이와 같은 현상은 동의어가 가지는 텍스트에서의 기능에서 그 원인을 찾을 수 있다. 동의어는 이미 언급된 개념을 다시 언급할 경우에 흔히 나타나는데 이러한 재 언급의 경우 해당 단어의 동의어의 선택에서 해당 단어와의 비기술적 측면에서의 호환이 중요하게 고려되기 때문인 것으로 추측할 수 있다.

## <특강>

### **Constructing Sentiment Corpus in Korean**

신 호 필 (서울대)

There has been much research on the automatic identification and extraction of sentiments and opinions in text. Researchers have been working on these issues by focusing mainly on subjectivity and sentiment classification either at the document or sentence level. Classifying editorials or movie reviews as positive or negative are examples of a document classification tasks while classifying individual sentences as subjective or objective would be an example of a sentence-level task.

Along with these lines of research, a need for corpora annotated with rich information about opinions and emotions has also emerged. This would allow for the development of statistical and machine learning approaches for various practical NLP applications. As such a resource, the Multiperspective Question Answering (MPQA) Opinion Corpus developed by Wiebe (2002), Wiebe et al. (2005), and Wilson et al. (2008), plays an important role in sentiment and opinion analysis. It contains the manual annotation of a 10,000 sentence-corpus of articles from the world press. Since this corpus provides a fine-grained annotation scheme, it is widely used as a source for training data in machine learning approaches and serves as the gold standard in sentiment classification tests.

We started constructing a cross-language sentiment corpus, called the Korean Sentiment Corpus. We aim to provide both a solid theoretical background for the Corpus, reflecting the characteristics of the Korean language, as well as fine-grained annotations for the 8,050 sentence-corpus of news articles. I am going to introduce annotation scheme, inter-annotator agreement tests, and annotation tool for constructing sentiment corpus in Korean.

## <일반논문 1>

### **Some Negative Comments on the Negative Quantifier Analysis of *amwu*( N)-to**

정 대 호 (한양대)

Watanabe (2004) analyzes the traditionally called negative polarity item in a language like Korean and Japanese as a negative quantifier (N-word), to account for the fact that it can be a fragmental answer to an affirmative interrogative, despite the apparent polarity mismatch between the affirmative predicate in the antecedent clause and the negative predicate in the ellipsis site. The polarity disparity resolves in his system due to Agree between the N-word and the elided negative predicate, which induces the [+neg] feature in the former to get copied into the latter and ultimately cancels out the [+neg] feature in the elided predicate as an instance of double negation. The current work, however, casts doubt on the negative quantifier analysis for the following two reasons: (i) the neg-feature-copy-followed-by-cancel-out mechanism leads to interpretation failure in some structures involving an N-word (e.g., an N-word as a short answer to a negative interrogative, an N-word in a non-negation context, and an N-word in an RNR construction); and (ii) polarity mismatch can be induced by a non-N-word (e.g. *acik* 'still, yet' as a short answer to an affirmative interrogative sentence and *selma* '(not) a chance' as a short response to an affirmative declarative sentence). As for the availability of an N-word as a fragmental response to an affirmative sentence, it is speculated in this work, conforming to Ahn and Cho (2011), that such N-word fragments involve no ellipsis and they are to be pragmatically licensed. If this is on the right track, then the semantic isomorphic condition becomes irrelevant to such fragments, and the polarity mismatch problem disappears accordingly.

## <일반논문 2>

### **On the source of subject-object asymmetries in Korean case ellipsis: The case of OSV sentences**

이 한 정 & 김 나 연 (성균관대)

It has been observed that subject case ellipsis occurs less frequently and is also less acceptable than object case ellipsis (Kim 2008; S. Lee 2009; H. Lee 2010, 2011). One particular case of this general subject-object asymmetry in case ellipsis is the dispreference for subject case ellipsis in sentences that have the non-canonical OSV word order. This has been analyzed as resulting from a violation of a structural requirement on the position of bare subject NPs (Ahn and Cho 2006a, 2006b, 2007). In this study, we present evidence from an acceptability rating experiment demonstrating that OSV sentences containing a case-ellipsed subject exhibit acceptability patterns different from ungrammatical sentences violating a core syntactic principle on case assignment and that these sentences are judged acceptable when the subject refers to expected, predictable information in context. We argue that this evidence supports the conclusion that the dispreference for subject case ellipsis in OSV sentences is due to violations of probabilistic constraints that favor case marking for rare types of subjects and such violations can be remedied by non-syntactic information.