

# 프랑스어 기본어휘 DB의 미시구조와 기술 정보\*

고길수 · 이성현<sup>†</sup>  
서울대학교

## French Basic Vocabulary Database, Its Microstructure and Information Description

Kilsoo Ko & Seong Heon Lee<sup>†</sup>  
Seoul National University

---

### ABSTRACT

This paper shows the results of our work on two issues that are necessarily raised when making a French basic vocabulary database: first how to select basic vocabulary, and second how to represent grammatical informations in the database. For the first issue, we adopted the A-level vocabulary proposed by the *Référentiels*, but in order to revise it, we conducted frequency-based adequacy tests. 430 words are excluded because they are rarely used and thus inappropriate as basic vocabulary. Instead, we added 200 very high-frequency words that we believe should be learned at the beginner level, and we propose a basic vocabulary of 1,300 words. For the second issue, we developed a representation model with reference to the lexicon-grammar and the theory of "classe d'objet". It follows their basic framework, but it has different microstructure and representation methods so that the usage of predicate words can be understood in an integrated way.

**Keywords:** French basic vocabulary, basic vocabulary database, predicate word, grammatical information, syntactic property, argument structure

---

## 1. 서 론

본 논문은 프랑스어 교육 및 평가에 활용할 수 있는 기본어휘 DB 구축을 위한 연구의 결과를 보인다. 양질의 프랑스어 기본어휘 DB는 교육 및 평가 과정에서 중요한 준거가 될 뿐만 아니라 실제 현장에서 교수자나 학습자에게는 유용한 참고 자료로 쓰일

---

\* 이 논문은 2015년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2015S1A5A2A0 3049877).

\*\* 줄고를 읽고 논평해주신 익명의 심사위원 세 분에게 감사의 말씀을 전한다. 아울러 심사위원의 의견을 적극 반영하여 논문의 내용을 보완하였다. 하지만 여전히 개선할 부분들이 남아 있는 것으로 보인다. 이는 전적으로 우리 저자들의 몫으로 후속 연구들을 통해 지속적으로 보완해 나갈 것이다.

<sup>†</sup> Corresponding author: kilsooko@yahoo.com, lsh0717@snu.ac.kr



수 있다. 하지만 국내에는 아직 그러한 용도의 프랑스어 기본어휘 DB가 없으며, 관련 연구도 미진한 실정이다. 우리의 연구는 바로 이러한 인식에서 시작된 것이다.

기본어휘 DB를 만드는 작업은 필히 다음 2가지 문제를 동반한다. 첫째는 어떤 어휘를 기본어휘로 보아 DB에 수록할 것인가이고, 둘째는 이들 어휘와 함께 무엇을 DB에 담을 것인가이다. 즉 본 논문에서는 이 두 문제에 관한 우리의 성과를 소개하고자 한다.

우선 기본어휘 선정과 관련하여 우리는 A등급 *Référentiel*<sup>1)</sup>이 제안하는 어휘들(이하 A등급어휘)을 주된 기본어휘로 채택하되, 사용빈도를 고려하여 각 어휘를 대상으로 적정성을 검증하여 최종 선정 여부를 결정하였다. 그 결과 총 1,530여 개의 A등급어휘 중 사용빈도와 적정성이 낮다고 판단되는 430여 개를 제외시키는 한편 A등급어휘로 제시되지 않았지만 사용빈도가 높고 학습이 필요하다고 판단되는 200여 개 어휘를 추가하여, 최종적으로는 1,300개의 어휘를 기본어휘로 선정하였다.

다음으로 기본어휘와 함께 DB에 수록되는 정보와 관련하여 우리는 이들 어휘의 문법정보, 즉 이들의 형태, 통사, 의미에 관한 정보를 상세하게 기술하였다. 특히 문법정보 중 가장 중요한 것이 바로 동사를 비롯한 술어 어휘를 중심으로 문장을 만드는 정보인데, 이와 관련하여 우리는 어휘문법(Gross 1975)과 대상부류(Gross 1994) 이론의 어휘기술 방법론을 참고하여 술어 어휘의 분포 및 변형 속성과 의미 속성을 통합적으로 기술하는 모형을 개발하였다. 한편 우리는 기본어휘 DB와는 별개로 기본어휘의 사용상황을 참조할 수 있는 사용상황 유형과 유형별 대표 대화 예시 및 어휘를 담은 사용상황 온톨로지를 구축하였다.<sup>2)</sup> 여기서는 기본어휘 DB에서의 문법정보 기술에 관한 것을 살펴보기로 한다.

## 2. 프랑스어 기본어휘 선정

먼저 살펴볼 것은 기본어휘 DB에 수록할 어휘를 선정하는 문제이다. 가능한 해법으로는 기존 연구에서 제시한 기본어휘를 활용하거나 아니면 자체 기준을 마련하여 새 기본어휘를 개발하는 것이다. 우리는 첫째 해법을 따라 A등급 *Référentiel*이 제시하는 어휘, 즉 A등급어휘를 기본어휘로 채택하였다. 다만 이들 어휘를 그대로 재활용하기에는 문제가 있다고 판단하여 각 어휘를 대상으로 적정성을 검증하여 일부를 제외하고 사용빈도가 높은 어휘 중 일부를 추가하여 어휘목록을 수정하였다.

### 2.1. 기본어휘로서의 A등급어휘와 검증의 필요성

#### 2.1.1. A등급어휘의 기본어휘로서의 자격

우리가 A등급어휘를 기본어휘로 채택하는 것은 바로 이들이 국제 기준을 따라 선정된

1) Béacco et al. 2006, 2007, 2008

2) 사용상황 온톨로지는 별개의 논문을 통해 소개할 것이다.

어휘들이기 때문이다. 다시 말해서 A등급어휘는 국제 기준에서 정한 초급 수준의 프랑수어 능력을 갖추기 위해 알아야 하는 어휘인 것이다.

A등급어휘의 출처인 *Référentiel*은 <유럽공통참조기준>을 근간으로 한다. <유럽공통참조기준>은 외국어로서 유럽 언어의 학습 및 교육과 평가를 통일하고 표준화하려는 목적으로 제작된다. 이를 위해 유럽 연합의 국가들이 다년간의 공동 작업을 통해 여러 언어에 공통으로 적용되는 6개 등급의 외국어 능력<sup>3)</sup>과 각 등급의 외국어 능력에 대한 세부적인 설명을 마련하였는데, 이를 담고 있는 것이 바로 <유럽공통참조기준>이다. 이어 이를 적용한 개별 유럽 언어의 지침서들이 개발되는데, 프랑수어에 대한 지침서가 바로 *Référentiel*이다. 즉 이 책은 <유럽공통참조기준>을 따라 외국어로서 프랑수어 능력을 6개 등급으로 나누고 등급별 프랑수어 능력을 다양한 측면에서 상세하게 설명한다. 이 책은 또한 그러한 프랑수어 능력을 갖추는데 필요한 언어 자료를 함께 제공하는데, 그러한 자료 가운데 하나가 바로 어휘목록이다.

예컨대 A2등급 *Référentiel*은 “A2등급의 외국인 학습자가 <유럽공통참조기준>에서 정한 기초적인 숙달 *maîtrise dite élémentaire*을 습득했을 때 기대되는 언어능력을 설명”하고, “A2등급의 수준을 규정하는 기술항목들 *descripteurs*에 부응하는 프랑수어의 언어 자료를 확인”한다고 적고 있다(Béacco et al. 2008: préface). 즉 “프랑수어의 기초적인 숙달”의 구체적인 내용을 밝히고, 이에 필요한 프랑수어 자료를 제시하는 것이다. 이처럼 A등급어휘는 초급 수준의 프랑수어 능력을 갖추기 위해 알아야 하는 어휘들로 제시된 것이다. 그런 만큼 이들은 기본어휘에 해당한다고 볼 수 있는 것이다.

한편 *Référentiel*의 근간이 되는 <유럽공통참조기준>은 그 제정 목적에 부응하여 현재 다수의 유럽 국가에서 외국어로서 유럽 언어의 학습 및 교육, 평가를 시행하는 표준 근거로 채택되고 있다. 즉 외국어로서 유럽 언어에 대한 교육과 평가가 모두 <유럽공통참조기준>을 따라 시행된다는 것이며, 특히 주목할 것은 바로 이들 언어의 공인시험이 그러하다는 것이다. 외국인의 프랑수어 능력을 평가하는 공인시험 또한 모두 <유럽공통참조기준>을 따라 시행되고 있다. CIEP<sup>4)</sup>는 홈페이지에서 DELF와 DALF는 <유럽공통참조기준>을 따라 6개 등급으로 나누어 시행되고, 등급별 평가 항목과 내용 및 평가 방식 또한 <유럽공통참조기준>을 따라 결정, 시행된다고 적시한다.<sup>5)</sup> 결국 *Référentiel*이 제안하는 A등급어휘는 초급 수준의 프랑수어 공인시험을 통과하기 위해 알아야 하는 어휘라고 할 수 있으며<sup>6)</sup>, 그런 만큼 국내에서의 프랑수어 교육 및 평가에서도 중요하게 고려되어야 한다고 하겠다.

3) 외국어 능력은 우선 초급 A, 중급 B, 고급 C로 나뉘고, 이들은 다시 하 1와 상 2로 구분된다. 즉 우리가 관심을 갖는 초급은 초급 하 A1과 초급 상 A2로 구성된다. <유럽공통참조기준> 3장 참조.

4) Centre international d'études pédagogiques의 줄임말로, 프랑수어 공인시험을 관장하는 국가기관이다.

5) <http://www.ciep.fr/delf-tout-public/presentation-des-epreuves> 참조.

6) 다만 실제로 그러한지 확인이 불가능하다. DELF와 DALF 등 프랑수어 공인시험이 시행된 후 문항이 공개되지 않아, A등급어휘를 검증할 길이 없기 때문이다.

### 2.1.2. 적정성 검증의 필요성

하지만 A등급어휘를 있는 그대로 기본어휘로 채택하기에 아무런 문제가 없다고 하기는 어렵다.

우선 적어도 우리의 시각에서 볼 때 기본어휘로 삼기 어려운 어휘들이 적지 않게 보인다. 그러한 어휘들은 흔히 일상생활에서 접하는 다양한 사물들과 관련되는데, 뒤에서 보겠지만 이들은 대체로 사용빈도가 아주 낮은 어휘들이다. 예컨대 의복과 관련한 *baskets*(운동화), *collant*(타이즈), *slip*(팬티)이나, 의약품과 관련한 *aspirine*(아스피린), *antibiotique*(항생제), *comprimé*(알약), *sparadrap*(반창고), *sirop*(시럽) 등이 그러하다. 이들은 일상생활을 하기 위한 목적으로 프랑스어를 배우는 학습자에게는 중요하다 하겠으나, 그렇지 않은 학습자 입장에서는 이들보다는 일반적으로 널리 사용되는, 즉 사용빈도가 높은 어휘들을 익히는 것이 좀 더 나을 것이다. 그리고 우리는 적어도 초급 수준에서는 후자에 속하는 학습자가 대다수를 차지한다. 프랑스 및 프랑스어권 국가들과의 교류가 지속적으로 증가하고 있음을 감안하더라도 어느 정도 추가적인 어휘 선별 작업과 이를 위한 적정성 검증이 불가피하다고 하겠다.

한편 A등급어휘는 어휘 선정과 등급 배분에도 미심쩍은 부분이 발견된다. 아래의 표는 A등급어휘의 세부등급<sup>8)</sup>별 분포를 보인 것이다.<sup>9)</sup> 원칙상 하위등급의 어휘는 모두 상위등급에도 포함된다. 즉 A1.1의 817개 어휘는 모두 A1에 포함되고, A2는 A1의 823개 어휘를 모두 가진다. 이 경우 모든 세부등급의 어휘를 합하면 그 수는 최상위 세부등급인 A2와 같게 된다. 하지만 표에서 보듯이 그렇지가 않다. 실제로는 A2가 116개가 적은데, 이는 A1.1과 A1에서 제시된 어휘 중 116개가 A2에서 누락된 것이다.

**표 1. A등급어휘의 세부등급별 분포**

등급	A1.1	A1	A2	합계
어휘 개수	817개	823개	1,421개	1,537개

다음으로 중복 수록 내역을 보자. 총 1,537개 어휘 중 모든 세부등급에 수록된 것은 636개이고, 252개는 2개 등급에서(A1.1&A1: 1개, A1&A2: 177개, A1.1&A2: 74개),

7) *Référentiel*은 <유럽공통참조기준>을 따라 일상의 삶을 살아갈 수 있는 프랑스어 능력을 최우선 목표로 설정하고, 이에 필요한 어휘를 다수 선정, 제공한 것이다. 이러한 사실은 등급별 언어 능력을 서술하고 있는 부분에 잘 나타난다. 보다 자세한 것은 *Référentiel 2*장, <유럽공통참조기준> 3장 참조.

8) A등급은 3 개의 세부등급으로 나뉜다. 즉 A1.1, A1, A2이다. 이 가운데 A1.1은 프랑스어를 처음 배우는 입문등급으로 A1에 이르기 전 단계에 해당하며, 이 등급의 지침서가 제작된 것은 프랑스어가 유일하다. Béacco et al. (2005:3) 참조.

9) 표에 제시된 어휘는 등급별 *Référentiel*의 색인표에서 수집한 것이다. 한편 색인표는 각 어휘가 속하는 범주 *catégories*와 언어로 수행하는 기능 *fonctions* 목록을 함께 제공하는데 이들은 제외하였다. 그리고 복합어나 연어 등은 이를 구성하는 개별 단어로 나누었다. 예를 들어 *boîte aux lettres*(우편함)는 *boîte*, *aux*, *lettres* 세 개의 어휘로 분리한 것이다.

그리고 649개는 한 개 등급에서(A1.1: 106개, A1: 9개, A2: 534개) 제시되고 있다. 한편 빨간색과 파란색으로 표기된 수치에 주목하자. 전자는 하위등급에서 제시되어 있으나 상위등급에는 누락된 어휘를 나타낸다. 예컨대 A1.1에 수록된 어휘 중 106개가 A1과 A2에서 등재되지 않은 것이다. 파란색은 중간등급에서 누락된 어휘를 가리킨다. 즉 74개의 어휘가 A1.1과 A2에만 제시되고 중간등급인 A1에서는 빠져 있는 것이다.<sup>10)</sup>

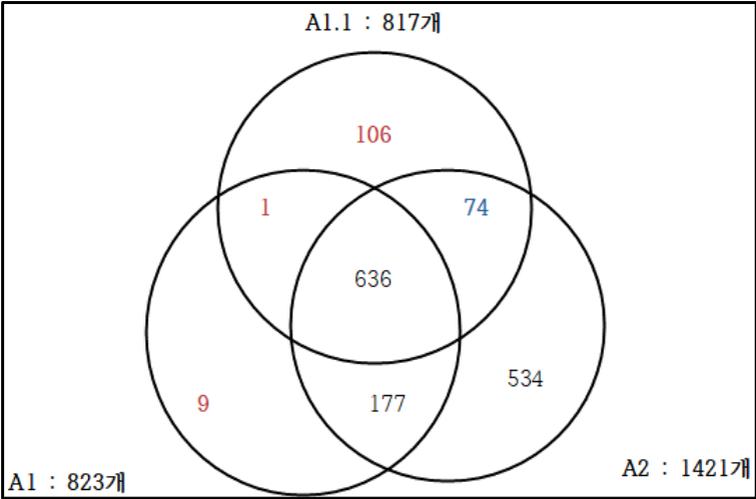


그림 1. A등급어휘의 중복 수록

이와 관련하여 *Référentiel*은 B2등급이 가장 먼저 제작되고 이후 A1.1, A1, A2, B1 등급이 만들어진다.<sup>11)</sup> 즉 낮은 등급에서 높은 등급으로 점진적으로 구성된 것이 아니다. 한편 처음 만들어진 B2등급은 어휘 선정과 관련하여 *Un niveau-seuil*(Coste et al. 1976)와 *Adaptation de "Un niveau-seuil" pour des contextes scolaires* (Porcher et al. 1979)을 따른다고 밝히고 있다(Béacco et al. 2004:9). 하지만 이후에 제작된 나머지 등급의 *Référentiel*은 <유럽공통참조기준>의 등급 기술, 학습자 및 외국어 습득에 대한 확립된 지식, 교육 목표, 교수자와 평가자의 집단적 경험, 학습자 코퍼스라는 원론적인 5 가지 기준만 언급할 뿐, 보다 구체적인 근거 자료를 제시하지 않는다.<sup>12)</sup> 즉 <유럽공통참조기준>에서 상세하게 설명하고 있는 등급 기술을 제외하고는 등급별 어휘 선정 기준과 원칙이 명확하지 않은 것이다.

10) 이 문제는 A등급어휘에서 뿐만 아니라 B등급에서도 나타나고 있다. 보다 자세한 것은 Kusseling & Lonsdale (2013), 표2 참조.  
 11) B2가 2004년, A1.1이 2006년, A1이 2007년, A2가 2008년, B1이 2011년에 발간된다. 그리고 A1.1은 A1보다 더 낮은 입문 등급에 해당한다.  
 12) 어휘를 비롯하여 언어 자료의 선정 기준은 *Référentiel* 1.2.1.에서 밝히고 있는데, 이 부분은 분량이 1쪽에 지나지 않는다.

마지막으로 A등급어휘에 대한 검증이 아직까지 이루어지지 않았다는 점 또한 문제이다. 사실 *Référentiel*이 제안하는 어휘에 대한 검증 연구는 Kusseling (2012) 및 Kusseling & Lonsdale (2013)이 유일하다고 하겠다. 이들 연구는 *Référentiel*의 어휘에 대해 객관적이지 않은 내성과 경험에 의거해서 선정되고 실제 선정 과정에 대한 상세한 설명이 없다고 비판하면서(Kusseling & Lonsdale 2013:439-440), 당시에 발간된 모든 등급 *Référentiel*의 어휘를 대상으로 검증 작업을 수행한다. 이들 연구에서 검증 기준은 사용빈도이며, 검증 단위는 단어형word-form이다. 이들이 제안하는 최종 결과물도 단어형의 목록이다. 즉 이들 연구는 중요한 참고 자료이기는 하지만 기본어휘를 마련하고자 하는 따라서 기본형lemma를 기본 단위로 하는 우리 연구와는 다소 거리가 있다고 하겠다.

이러한 이유로 우리는 A등급어휘를 기본어휘로 채택하되, 직접 적정성 검증과 추가 선정 작업을 진행하였다.

## 2.2. A등급어휘의 적정성 검증

### 2.2.1. 사용빈도 참조 적정성 검증

우리의 검증 방법은 사용빈도를 참조하여 어휘 선정이 적합한지를 살피는 것이다. 일반적으로 어휘는 사용빈도가 높을수록 우선 학습해야 하는 것으로 간주된다. 다시 말해서 사용빈도가 높으면 기본어휘이거나 적어도 기본어휘일 가능성이 높다고 하겠다. 이런 이유로 사용빈도는 여러 연구에서 교육 목표와 함께 기본어휘를 선정하는 주요 기준의 하나로 사용되고 있다(Gougenheim et al. 1964; 김언자 2006). 그런데 A등급어휘는 앞에서 보았듯이 교육 목표를 최우선 기준으로 하고 있다. 즉 이들은 <유럽공통참조기준>이 정한 초급 수준의 프랑스어 능력을 갖추기 위해 필요하다고 판단되어 선정된 것이다. 반면 사용빈도는 근거 자료의 부재로 어휘 선정 과정에서 거의 고려되지 않는다(Kusseling & Lonsdale 2013:440). 즉 우리의 검증 작업은 그러한 편중됨을 해소하고 보다 균형을 갖춘 기본어휘를 마련하기 위함이다.

사용빈도는 Routledge에서 2009년 발간된 *A Frequency Dictionary of French*(이하 빈도사전)의 것을 참조하였다. 이 사전은 2,300만 단어로 구성된 코퍼스에서 추출한 고빈도 어휘 5,000개가 등재되어 있다. 이 사전에서 우리가 주목하는 것은 바로 코퍼스인데, 그 특징은 1950년 이후 생산된 자료만을 포함하고 구어 자료와 문어 자료의 비율이 같다는 것이다(Lonsdale & Le Bras 2009; Introduction, pp. 2-3). 즉 이 사전의 사용빈도는 구어가 문어와 같은 수준으로 반영된 비교적 최신의 현대 프랑스어에서 구한 것이라고 하겠다. 사실 기존 코퍼스는 대부분이 구어 자료가 없거나 있더라도 아주 빈약하고, 20세기는 물론 18~19세기 자료도 담고 있는 경우가 적지 않았다.

한편 검증의 목적은 사용빈도가 낮은 어휘들을 수집한 뒤 이들을 제외할 것인지 살펴보는 것인데, 이와 관련하여 가용어휘vocabulaire disponible 문제가 제기될 수 있다.

가용어휘란 사용빈도는 낮지만 특정 주제가 거론되거나 이와 관련된 상황에 처하면 즉각 떠오르는 어휘를 말하는데, 이들은 여러 연구에서 기본어휘로 간주되고 있다(Gougenheim et al. 1964; 김언자 2006). 앞에서 문제라고 거론했던 *baskets*(운동화), *comprimé*(알약) 등 일상생활에서 접하는 사물들을 지칭하는 명사 어휘들이 대부분 가용어휘에 해당한다. 즉 이들은 사용빈도가 낮아도 여전히 기본어휘로 볼 수 있다는 것이다. 다만 유념할 것은 가용어휘 여부는 보충 기준이라는 점이다. 즉 교육 목표와 사용빈도를 기준으로 어휘를 선정하고 난 뒤 추가로 보충 어휘를 선정할 때 적용되는 기준이라는 것이다. 우리 연구에서도 사용빈도를 참조하여 어휘 선정이 적합한지를 점검하고 그 결과를 토대로 기본어휘를 선정하고 난 후 제외된 어휘와 고빈도 어휘를 대상으로 추가 어휘 선정 과정에서 가용어휘 문제가 함께 다루어진다.

### 2.2.2. 사용빈도 분석 결과

A등급어휘의 사용빈도를 분석한 결과는 다음과 같다.<sup>13)</sup> 표에서 보듯이 고빈도 어휘는 53%로, A등급 전체 어휘 1,537개 중 812개가 이에 해당한다. 세부등급에서도 54~62%로 비슷한 수치를 보이고 있다. 중빈도 어휘는 25~28% 사이의 비율을 차지하고 있으며, 저빈도 어휘의 경우 세부등급에서는 비율이 20% 미만이지만 A등급 전체를 대상으로 하면 20%인 것으로 나타나고 있다.

표 2. A등급어휘의 사용빈도

	고빈도	중빈도	저빈도	합계
전체	812 (53%)	420 (27%)	305 (20%)	1,537
A1.1	476 (58%)	225 (28%)	116 (14%)	817
A1	512 (62%)	205 (25%)	106 (13%)	823
A2	771 (54%)	383 (27%)	267 (19%)	1,421

즉 A등급어휘는 47%가 중·저빈도 어휘이고 특히 305개는 사용빈도 순위가 5,000등이 넘어가는 저빈도 어휘에 해당한다. 이들 저빈도 어휘는 반드시 적정성 검증이 필요하다고 할 수 있으며, 대상을 중빈도 어휘까지 확장할 경우 절반에 가까운 어휘들이 재검토되어야 한다고 하겠다.

다음은 중복 수록 여부를 나누어 사용빈도를 분석한 결과이다. 표에서 보듯이 여러 세부등급에서 중복 수록된 어휘는 고빈도 어휘의 비율이 높고, 한 개 세부등급에만 수록된 어휘는 저빈도 어휘가 상대적으로 많음을 알 수 있다. 예컨대 3개 세부등급에 모두

13) 고빈도는 빈도사전 기준 2,000등 이내를 뜻하고, 중빈도는 2,001등~5,000등 사이의 구간에 해당한다. 저빈도는 빈도사전에 등재되지 않은, 즉 5,000등이 넘어감을 의미한다.

등재된 어휘는 64%가 고빈도 어휘이지만, 한 개 등급에만 등재된 어휘는 34%와 43%만이 고빈도 어휘에 해당한다. 반대로 저빈도 어휘를 보면, 전자는 11%로 비율이 낮지만 후자는 33%와 28%로 비교적 높은 수치를 보인다.

**표 3. A등급어휘의 중복 수록 여부와 사용빈도**

		고빈도	중빈도	저빈도	합계
3개 등급 등재		407 (64%)	161 (25%)	68 (11%)	636
2개 등급 등재	A1.1 & A1	1	0	0	1
	A1.1 & A2	32 (43%)	29 (39%)	13 (18%)	74
	A1 & A2	100 (56%)	42 (24%)	35 (20%)	177
1개 등급 등재	A1.1	36 (34%)	35 (33%)	35 (33%)	106
	A1	4	2	3	9
	A2	232 (43%)	151 (28%)	151 (28%)	534

한편 위 표에서 주목할 것은 고빈도, 중빈도, 저빈도 모두에서 중복 수록된 것과 그렇지 않은 것이 혼재되어 있다는 것이다. 예컨대 저빈도 어휘 중 68개는 모든 세부등급에, 48개는 2개 세부등급에 중복해서 수록되어 있는 반면, 189개는 1개 세부등급에서만 제시된다. 이러한 결과를 감안하여, 우리는 중복 수록 여부를 적정성을 검증하는 추가 기준으로 채택한다. 즉 여러 세부등급에 중복 수록된 어휘는 비록 사용빈도가 낮아도 기본어휘일 가능성이 높다고 가정하는 것이다. 이와 관련하여 A등급어휘는 하위 등급에서 수록된 어휘가 상위 등급에서 누락되는 등 어휘 선정 및 등급 배분의 원칙과 기준이 모호하다는 점을 환기하자. 이는 원칙과 기준이 각 세부 등급에서 달랐다는 것을 의미하는 것이다. 또한 어휘가 여러 세부 등급에서 중복 수록되었다는 것은 해당 어휘가 A등급어휘를 선정하는 원칙과 기준을 여러 개를 충족함을 의미한다. 반면 하나의 세부 등급에서만 나타나는 어휘는 오직 하나의 원칙과 기준을 충족하는 것이다.

다음으로 사용빈도 분석 결과를 활용한 적정성 검증 결과를 보자.

### 2.2.3. 적정성 검증 결과

적정성 검증의 기준은 사용빈도와 중복 수록 여부이다. 다시 말해서 사용빈도가 높고 여러 세부등급에 중복 수록된 어휘는 기본어휘로서의 적정성이 높다고 보는 것이고 반면 한 개의 세부등급에서만 수록되어 있는 저빈도 어휘는 기본어휘로 보기 어렵다는 것이다. 하지만 그러한 결정이 항상 자명한 것은 아니다. 예컨대 모든 세부등급에 수록된 저빈도 어휘나 1개 세부등급에만 제시된 고빈도 어휘, 또는 2개 세부등급에 등재된 중·저빈도 어휘 등은 판단하기가 쉽지 않다.

이에 우리는 검증 기준인 사용빈도와 중복 수록 여부를 독립된 변수로 보고, 이들의 가중치를 곱한 값으로 어휘의 적정성을 산정한다. 구체적으로 우리는 이들이 각각 아래와 같은 가중치를 갖는다고 가정한다<sup>14)</sup>.

**표 4. 수록된 세부등급의 개수 기준 가중치**

중복 수록	3개	2개	1개
가중치	×3	×2	×1

**표 5. 사용빈도 순위 기준 가중치**

사용빈도 순위	~1,000등	~2,000등	~3,000등	~4,000등	~5,000등	5,000등~
가중치	×1	×0.8	×0.6	×0.5	×0.4	×0.3

따라서 어휘 적정성은 다음과 같이 0.3~3 사이의 값을 갖는다. 최대값인 3은 3개 세부등급에 수록되고 빈도순위가 1,000등 이내인 어휘의 적정성 값이고, 최소값 0.3은 1개 세부등급에만 수록되고 빈도순위가 5,000등이 넘는 저빈도 어휘가 가지는 적정성 값이다.

**표 6. 수록된 세부등급 및 사용빈도 기준 A등급어휘의 적정성 값**

	~1,000등 (×1)	~2,000등 (×0.8)	~3,000등 (×0.6)	~4,000등 (×0.5)	~5,000등 (×0.4)	5,000등~ (×0.3)
3 (×3)	3	2.4	1.8	1.5	1.2	0.9
2 (×2)	2	1.6	1.2	1	0.8	0.6
1 (×1)	1	0.8	0.6	0.5	0.4	0.3

빨간색으로 표기된 값은 두 변수의 가중치가 다르지만 이를 곱한 적정성 값이 동일한 경우이다. 예컨대 3개 세부등급에 수록된 빈도순위 4,000등~5,000등인 어휘(3×0.4)와 2개 등급에 수록된 빈도순위가 2,000등~3,000등인 어휘(2×0.6)는 모두 적정성 값이 1.2이다. 한편 2개의 적정성 값 사이에 중복 수록 여부와 사용빈도 순위가 역전되는 경우가 있다. 예를 들면, 2와 1.8이 그러한데, 전자는 2개 등급에 수록된 1,000등 이내의 고빈도 어휘인 반면 후자는 3개 등급에 수록된 2,000등~3,000등 사이의 중빈도 어휘이다.

14) 두 변수의 값은 좀 더 세분될 수 있다. 예컨대 중복 수록 여부는 어휘가 처음으로 수록되는 세부등급을 고려해 값을 더 나눌 수 있고, 사용빈도 또한 순위 구간을 500으로 나누면 값을 더 세분할 수 있을 것이다. 다만 여기서는 이런 추가 구분은 고려하지 않는다.

### 2.3. 기본어휘 확정

#### 2.3.1. 적정성 검증 결과를 활용한 최종 어휘 선정

다음은 어휘 적정성 표에 최종 선정 여부를 나타낸 것이다. 파란색 부분이 기본어휘로 최종 선정되는 영역으로, 이들은 모두 적정성 값이 1 이상이다. 즉 어휘 적정성 1이 최종 선정 기준인 것이다. 이는 2가지를 고려한 것이다. 첫째는 빈도순위 1,000등 이내의 최고 빈도 어휘를 모두 포함시키고, 둘째는 빈도순위가 5,000등이 넘는 저빈도 어휘는 가급적 제외시키기 위함이다. 전자는 최소 적정성 값이 1개 세부등급에만 수록된 어휘가 갖는 1이다. 후자는 모든 세부등급에 수록된 저빈도 어휘의 적정성 값인 0.9가 최대값이다. 즉 두 경우를 가르는 값이 1인 것이다.<sup>15)</sup>

**표 7. A등급어휘의 적정성 값 기준 기본어휘 확정**

	~1000등 (×1)	~2000등 (×0.8)	~3000등 (×0.6)	~4000등 (×0.5)	~5000등 (×0.4)	5000등~ (×0.3)
3 (×3)	3	2.4	1.8	1.5	1.2	0.9
2 (×2)	2	1.6	1.2	1	0.8	0.6
1 (×1)	1	0.8	0.6	0.5	0.4	0.3

다음은 위의 표에서 적정성 값을 어휘의 개수로 교체한 것이다. 파란색 부분에 적힌 총 921개 어휘가 기본어휘로 최종 선정되었다. 반면 하얀색 부분에 표기된 616개 어휘는 적정성 값이 1 미만으로 선정 기준을 충족시키지 못해 제외된 것이다.

**표 8. A등급어휘의 적정성 값 기준 어휘 분포**

	~1,000등 (×1)	~2,000등 (×0.8)	~3,000등 (×0.6)	~4,000등 (×0.5)	~5,000등 (×0.4)	5,000등~ (×0.3)
3 (×3)	272	135	75	47	39	68
2 (×2)	83	50	29	24	18	48
1 (×1)	167	105	94	62	32	189

15) 적정성 값 1을 최종 선정 기준으로 채택하는 것이 어쩌면 자의적으로 보일 수도 있겠다. 이는 우리가 제안하는 검증 방법론이 처음으로 시도된 것인 만큼 기준 설정과 관련하여 이론적이고 경험적인 논거를 충분히 제시하기 어려운 데서 기인한다. 이 문제는 연구 결과에 대한 검증이나 응용 등 후속 연구를 통해 보완할 수 있을 것이다.

다음으로 적정성 검증 결과 기본어휘로 채택된 것과 제외된 사례를 보자. 다음은 2개 세부등급에 수록되고 빈도순위가 3,000등~4,000등 사이에 있어 적정성 값 1로 선정된 어휘들이다. 다소 어려워 보이는 어휘가 일부 포함되어 있기는 하지만 대체로 초급 수준에서 학습이 필요하다고 여겨지는 어휘들임을 알 수 있다.

**표 9. 2개 등급 수록 3,000~4,000등 사이의 중빈도 어휘**

---

allumer(키다), ambassade(대사관), autoroute(고속도로), carré(네모난, 제곱), compagnon(동무), domestique(가정의), douzaine(12개), escalier(계단), éteindre(끄다), grand-mère(할머니), grand-père(할아버지), inscription(등록, 기입), interrogation(질문), jouet(장난감), légumes(야채), minuit(자정), panne(고장), pompier(소방관), postal(우체국의), remboursement(환불), résidence(주택, 공관), réussi(성공한), studio(원룸), tolérance(인내심)

---

반면 다음은 2개 세부등급에 수록되었으나 빈도순위가 4,000등~5,000등 사이인 중빈도 어휘로, 이들은 적정성 값이 0.8이 되어 최종 선정 과정에서 탈락한 어휘들이다. *bus*, *mairie* 등 재검토가 필요한 몇몇 어휘가 포함되어 있기는 하지만, 앞의 경우와 비교해 보면 초급 수준에서 학습할 필요성이 다소 떨어지는 어휘들임을 알 수 있다<sup>16)</sup>.

**표 10. 2개 등급 수록 4,000~5,000등 사이의 중빈도 어휘**

---

avalier(삼키다), bus(버스), chirurgien(외과의사), cochon(돼지), commerçant(상인), connecter(연결하다), coton(면), déménager(이사하다), est(동쪽), fiche(전표, 카드), humide(습한), mairie(시청), orange(폭풍우), poule(암탉), racisme(인종차별주의), stage(연수), tente(텐트), toilettes(화장실)

---

다음의 어휘들도 빈도순위가 4,000등~5,000등 사이의 중빈도 어휘이다. 하지만 이들은 세 개의 세부등급에 모두 수록되어 적정성 값이 1이 되어 기본어휘로 최종 선정된다. 실제로 앞의 경우와 비교해보면 기본어휘로 볼 가능성이 확연하게 높다고 할 수 있겠다.

**표 11. 3개 등급 수록 4,000~5,000등 사이의 중빈도 어휘**

---

auto(자동차), autobus(버스), beurre(버터), blond(금발), boulevard(대로), brouillard(안개), ceinture(벨트), commissariat(경찰서), crème(크림), dictionnaire(사전), distributeur(판매기), fièvre(열), fromage(치즈), garage(차고), gâteau(케이크), goûter(맛보다), location(임대), lunettes(안경), météo(지하철), montre(손목시계), mouton(양), nationalité(국적), oral(말의), pantalon(바지), parfum(향수), porc(돼지), portable(휴대폰, 노트북), poulet(병아리, 닭), raser(짧게 깎다), sel(소금), soif(목마름), soupe(수프), sympa(thique)(호감이 가는), taxi(택시), timbre(우표), valise(가방), vélo(자전거), visa(비자)

---

16) 이러한 판단에 대해 그 기준이 무엇이라는 질문이 제기될 수도 있겠는데, “재검토가 필요한 몇몇 어휘가 포함되어 있다는” 바로 앞에서 언급한 내용 때문에 더욱 그러하리라 생각된다. 여기서 우리가 말하고자 하는 바는 일반적인 경향이 그러하다는 것이다. 그리고 우리의 판단은 2.3.2.에서 제시하는 3 가지 추가 기준에 근거한 것이다.

### 2.3.2. 추가 어휘 선정

다음으로 보충 어휘를 추가하는 문제를 보자.

최종 선정된 921개의 어휘는 초급 수준의 프랑스어 능력을 갖추기에 충분하다고 말하기 어려운 듯하다. A등급어휘에 비해 무려 600여개가 줄었기 때문이다. 기초프랑스어 Français fondamental 역시 1,500개 어휘로 구성된다 (Gougenheim et al. 1964). 반면 Milton(2006)은 영국인 프랑스어 학습자를 대상으로 한 조사 분석을 통해 초급 수준에서는 800개~1,000개 정도의 어휘면 충분하다고 주장한다. 또한 이 연구에 따르면 영국인 학습자는 수업시간 기준 시간 당 평균 3~4개의 어휘를 습득하는데, DELF A2 자격 취득에 160~280시간<sup>17)</sup>의 학습이 필요함을 감안하면 기본어휘는 최대 1100여개로 제한된다. 보다 자세한 검토가 필요하겠지만, 우리는 우선 1,200~1,300개를 기본어휘의 적정 규모로 정하고 이에 필요한 어휘를 추가로 선정한다.

추가 어휘 선정은 2 가지 방식으로 진행되었다. 하나는 고빈도 어휘 중 기본어휘에서 누락된 것을 추가하는 것이고, 둘째는 A등급어휘 중 적정성 검토 과정에서 제외된 616개의 어휘 중에서 구제하는 것이다. 전자의 경우 검토 대상을 1,500등 이내의 고빈도 어휘로 제한하고 그 가운데 A등급어휘 670여개를 제외한 830여개의 어휘가 후보가 된다. 아울러 우리는 교육 목표, 가용어휘 여부, 학습의 수월성을 추가 선정 기준으로 채택하였다(김연자 2006). 좀 더 구체적으로 말하면 우리는 각 어휘를 3가지 관점에서 검토한 것이다: (1) 초급 수준의 한국인 학습자가 알아야 하는가?, (2) 초급 교재에서 자주 접하는 상황 및 주제와 관련하여 즉각 떠오르는 가용어휘인가?, (3) 기본어휘로 이미 선정된 어휘 또는 영어에서 배운 단어와 유사하여 쉽게 배울 수 있는가?<sup>18)</sup> 그 결과 고빈도 어휘 중에서 200개와 A등급어휘 중에서 180개 어휘가 추가로 선정되었으며, 기존의 921개 어휘와 통합하여 최종적으로는 1,300개의 기본어휘를 확정하였다.

다음은 추가로 선정된 어휘 중 일부를 보인 것이다.<sup>19)</sup>

#### 표 12. 고빈도 어휘 중 선정된 보충 기본어휘

- 
- montrer(보이다), suivre(뒤따르다), permettre(허락하다), devenir(되다), laisser(남기다), rappeler(다시 전화하다), offrir(주다), ...
  - aucun(아무), seul(혼자), meilleur(더 나은), plusieurs(여러), ainsi(그처럼), seulement(오직, 단지), ...
  - raison(이성), chose(것), groupe(그룹), problème(문제), niveau(단계), grands-parents(조부모), ...
- 

#### 표 13. A등급어휘 중 추가 선정된 보충 기본어휘

- 
- inscrire(등록하다), traverser(건너다), emporter(가져가다), rire(웃다), enlever(벗다), jeter(던지다), attacher(매다), ...
  - excellent(뛰어난), terrible(무서무서한), espagnol(스페인의), formidable(어마어마한), ...
  - policier(경찰), tableau(그림), site(사이트), acteur(배우), renseignement(정보), dos(등), spectacle(공연), ...
- 

17) 이는 알리앙스 프랑세즈 Alliance française 홈페이지의 DELF/DALF 설명에서 제시된 것이다.

18) 이 외에도 품사별 적정 배분율도 어휘 선정 기준으로 고려되어야 할 것이다. 그럼에도 본고에서 언급하지 않은 것은 본 작업에서는 기본어휘의 품사를 조사하진 하였으나, 이를 기준으로 한 통계 분석 등을 수행하지는 않았기 때문이다.

19) 첫째 그룹은 동사 어휘를, 둘째 그룹은 형용사와 부사 어휘를, 셋째 그룹은 명사 어휘를 보인 것이다.

한편 추가 선정 결과 원 출처에 따라 선정된 어휘의 범주별 분포가 차이가 난다. 고빈도 어휘에서 선정된 어휘를 보면, 명사와 동사, 형용사의 수가 크게 차이가 나지 않는다. 또한 부사도 30여개 가까이 선정되었음을 알 수 있다. 반면 A등급어휘에서 추가된 어휘의 경우 명사가 124개로 다른 범주에 비해 압도적으로 큰 비중을 차지하고 있다. 이러한 차이는 고빈도 어휘는 다양한 범주의 어휘가 혼재되어 있는 반면, A등급어휘는 적정성 검증 과정에서 명사 어휘의 비중이 컸기 때문으로 보인다.

**표 14.** 추가 선정된 기본어휘의 범주별 분포

	명사	동사	형용사	부사	기타
고빈도 어휘	54	65	46	26	9
A등급어휘	124	25	29	1	2

### 3. 기본어휘의 문법정보 및 표상

프랑스어 기본어휘 DB에는 기본어휘와 함께 이들에 관한 다양한 정보가 함께 수록된다. 여기서는 문법 정보 중 술어 어휘가 자신의 논항과 결합하여 문장을 구성하는 논항구조와 이의 표상 문제를 살펴보자.

#### 3.1. 술어의 논항구조 정보

프랑스어 교육 및 평가에 활용되기 위해서는 기본어휘 DB에 기본어휘에 관한 다양한 정보들이 함께 수록되어야 한다. 그러한 정보 중 핵심은 기본어휘의 문법 정보, 즉 이들에 어휘가 문장 내에서 쓰이기 위해 알아야 하는 형태·동사·의미에 관한 정보일 것이다. 예를 들어 다음 두 문장은 모두 동사 *aller*로 만들어진다. 하지만 두 문장에서 동사는 형태가 아주 다를 뿐만 아니라 의미해석 또한 완전히 다르다.

(1) Tu vas bien ? (너 잘 지내니?)

(2) Nous allons où ? À Paris ? (우리 어디에 갑니까? 파리에 갑니까?)

즉, 이 두 문장의 동사에 대해서 학습자가 알아야 할 것은 ㉠ 주어에 따라 다른 형태가 쓰이고, ㉡ 여러 가지 의미를 표현하며, ㉢ 의미에 따라 함께 쓰이는 요소가 다르고, ㉣ 그 결과 문장의 형태가 변한다는 것이다. 요컨대 기본어휘 DB에는 이와 같은 정보들이 모두 수록된다는 것이다.

이러한 정보들 가운데, 여기서 살펴볼 것은 논항구조에 관한 정보이다. 논항구조란

동사 등 술어 어휘가 자신의 논항과 결합하여 문장을 구성하는 것에 관한 정보이다. 조금 전 보았던 사례를 다시 언급하면, 두 문장에서 동사 *aller*가 의미에 따라 다른 요소와 결합하고 그 결과 문장의 형태가 달라지는 것이 바로 이 동사의 논항구조 정보이다. 다음은 이 동사가 쓰인 또 다른 형태의 문장들이다.

(3) Ces chaussures vont bien avec ton pantalon.

(이 신은 네 바지와 잘 어울린다.)

(4) Je vais essayer. (나는 시험 삼아 해볼 것이다.)

(5) C'est toujours lui qui va faire les courses.

(장을 보러 가는 것은 늘 그 사람이다.)

(3)에서 동사 *aller*는 주어를 비롯하여 부사어 *bien*<sup>20)</sup> 및 전치사구 *avec ton pantalon*과 함께 쓰인다. 반면 (4)와 (5)에서는 부정법 동사(구)와 결합하고 있는데, 두 문장에서 동사 *aller*가 표현하는 의미는 서로 다르다. 이러한 차이는 DB에서 정확히 그리고 명시적으로 기술되어야 한다. 즉, 동사 *aller*는 적어도 5개의 의미와 그에 따라 달라지는 논항구조 정보가 함께 기술되어야 한다.

한편 다른 논항구조가 항상 의미의 변화를 동반하는 것은 아니다. 다시 말해서 문장의 형태가 다르지만 동사가 같은 의미를 표현할 수도 있다는 것이다. 예를 들어 앞서 본 (3)은 부사(어)가 생략되고, (5)는 장소를 지칭하는 전치사구가 추가될 수 있다.<sup>21)</sup>

(3') Ces chaussures vont avec ton pantalon.

(이 신은 네 바지와 어울린다.)

(5') C'est toujours lui qui va au marché faire les courses.

(시장에 장을 보러 가는 것은 늘 그 사람이다.)

한편, 이와 반대되는 경우도 있다. 즉, 논항구조가 같아도 동사의 의미가 다른 경우로, 앞에서 본 (4)와 (5)가 이에 해당한다. 요컨대 술어 어휘의 논항구조와 의미는 서로 밀접한 관계를 갖지만 이들이 반드시 일대일의 관계를 갖는 것은 아니라고 하겠다.

논항구조와 관련한 정보에는 논항의 형태와 의미에 관한 정보도 포함된다. 술어 어휘와 그 의미에 따라 결합하는 논항의 형태와 의미가 제한되기 때문이다. 예를 들어, 앞에서 보았듯이 동사 *aller*는 전치사 *à*로 시작하는 전치사구, 부정법 동사구 등과 함께 쓰이고, (6)~(7)에서 보듯이 전치사를 생략하거나 부정법 동사구 앞에 *à*를 추가하는 것이 불가능하다.

20) 이 문장에서 부사(어) *bien*은 부가어로 분석된다.

21) 이런 이유로 (5)와 (2)는 동사가 같은 의미를 표현한다고 볼 수도 있다.

- (6) On va {à, \*E} Marseille.<sup>22)</sup>  
 (우리는 마르세이유에 갑니다.)
- (7) On va {\*à, E} entrer dans une minute.  
 (1분 후에 입장합니다.)

그리고 이와 같은 제약은 (8)~(9)에서 보인 다른 동사 어휘의 논항구조 및 논항 형태와 대비된다.

- (8) On visite {\*à, E} Marseille.  
 (우리는 마르세이유를 방문합니다.)
- (9) On commence {\*à, E} entrer.  
 (입장하기 시작합니다.)

다음은 논항의 의미가 제한되는 경우를 보인 것이다. 예에서 보듯이 동사 *aller*는 인간, 동물, 교통기관, 길 등에 속하는 의미를 가진 명사들이 주어로 나타날 수 있으나, 인간을 지칭하는 경우를 제외하면 대체로 그 쓰임이 제한적이다. 또한 식물이나 기타 무생물을 지칭하는 명사는 특별한 문맥이 주어지지 않는 한 쓰이지 못한다.

- (10) {Paul, Ce chien, Ce bus, Cette route} va à la mairie.  
 {폴, 이 개, 이 버스, 이 도로}-는 시청에 갑니다.
- (11) \*{Cet immeuble, Ce plat, Cette fleur} va à la mairie.  
 \*{이 건물, 이 음식, 이 꽃}-은 시청에 갑니다.

지금까지 동사 *aller*를 중심으로 술어 어휘의 논항구조와 관련하여 어떤 정보들이 문제가 되는지를 살펴보았다. 이를 다시 정리하면, 논항구조는 단순히 술어 어휘가 논항과 결합하여 구성하는 문장의 형태를 보이는 것이 아니라, 술어 어휘가 표현하는 의미를 비롯하여 결합할 수 있는 논항의 형태와 의미 등 구체적이고 세부적인 사항들이 통합적으로 연계되어 제시되어야 하는 정보이다. 다음으로 이들 정보의 표상 문제를 살펴보자.

### 3.2. 논항구조 정보의 표상 방안

앞에서 보았듯이 술어 어휘의 논항구조는 2가지 유형의 정보를 포함한다. 첫째는 형태에 관한 것으로, 술어 어휘가 결합하는 논항의 수와 그것의 형태가 이에 해당한다. 둘째는 의미 관련 정보로, 술어 어휘가 표현하는 의미와 각 논항에 부과되는 의미 제약이

22) 중괄호 { }와 쉼표는 교체 요소를 나열한다. E는 공백, 즉 해당 위치에 아무런 언어 요소가 실현되지 않음을 가리키고, \*는 문법적으로 잘못된 비문임을 나타낸다.

바로 그것이다. 전자는 M. Gross의 어휘문법lexique-grammaire의 이론과 방법을 참고하여 표상하고, 후자는 G. Gross의 대상부류classe d'objets 개념과 방법론을 원용하여 기술한다.

먼저 어휘문법은 가장 작은 의미 단위는 고립된 개별 단어가 아니라 하나의 단문<sup>23)</sup>이라는 공리에서 출발한다. 이러한 시각에서는 언어 분석의 기본 단위가 단문이 되고, 언어 연구의 핵심은 단문의 근간이 되는 동사, 형용사 등 술어 어휘에 대한 분석이다. 즉 어휘 문법은 술어 어휘가 갖는 통사속성, 다시 말해서 본포속성과 변형속성을 단문 층위에서 기술하고, 이를 통해서 한 언어의 어휘(부)를 구축하고자 한다<sup>24)</sup>. 다음은 어휘문법의 술어 어휘 기술 사례로, *irriter*(화나게 하다) 등 감정을 느끼는 주체를 목적보어로 가지는 심리동사들의 통사속성을 기술한 것이다.<sup>25)</sup> 표는 이들 동사가 취하는 논항의 수와 성격, 즉 논항구조로 대표되는 본포속성을 보여주고, 목적보어 생략, 대명동사 구문이나 수동 문으로의 전환 등 변형속성도 함께 기술하고 있다.

Sujet					Adjectif					Comp. direct							
N <sub>hum</sub>	N <sub>nr</sub>	le fait Ou P	V <sup>1</sup> Ω		V "concret"	N <sub>0</sub> V	s = ant	s = able	s = eux	s = [E + et] eur	N <sub>hum</sub>	N <sub>num</sub>	le fait Ou B	N <sub>0</sub> V ger Ou P	N <sub>1</sub> se V suprés de N <sub>hum</sub> de ce Ou P	N <sub>1</sub> est Vpp de ce Ou P	(passif par)
+	+	+	+	irriter	+	+	+	-	-	+	+	+	+	+	+	+	+
+	+	+	+	juguler	+	+	-	-	-	+	+	+	+	+	+	+	-
-	+	+	+	lénifier	-	+	+	-	-	+	+	+	+	+	+	+	-
+	+	+	+	léser	+	+	+	-	-	+	+	+	+	+	+	+	-
+	+	+	+	lessiver	+	+	+	-	-	+	+	+	+	+	+	+	-

그림 2. 어휘문법의 술어 어휘 기술 예 : 심리동사

대상부류 이론은 어휘문법의 기본 틀을 받아들이면서도 이 이론의 어휘기술에서 미흡하게 다루어졌던 의미 측면을 획기적으로 강화하고 이를 통해서 보다 완결된 형태의 어휘부를 구축하고자 한다. 즉 통사속성 기술 못지않게 체계적이고 형식적인 의미속성에 대한 기술을 제공함으로써 통사정보와 의미정보를 보다 균형적으로 갖춘 어휘부를

23) 단문은 하나의 술어와 그것의 고유논항들로 구성된다. cf. Gross, M.(1981).  
 24) 대표적인 연구로는 프랑스어의 경우 Gross, M. (1975), Guillet & Leclère (1992), 한국어의 경우 홍재성 (1987), 홍재성의 (1997), 고길수 (2003), 남지순 (2007) 등이 있다.  
 25) M. Gross (1975), p. 263, tableau 4.

추구하는 것이다. 그리하여 특정 의미를 공유하는 어휘들의 공통적 통사행태에 따라 어휘들의 위계적 집합, 즉 대상부류를 만드는 것이다.<sup>26)</sup> 그러한 대상부류는 다음과 같이 단순한 어휘들의 목록이 아니라 이들이 공유하는 의미속성과 통사속성을 통합적으로 기술하는 새로운 형태의 어휘 기술 모형이다.

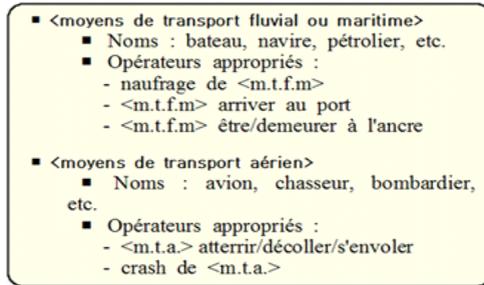


그림 3. 대상부류 이론의 기술 예: 교통수단 부류

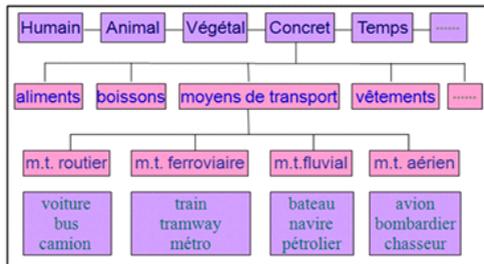


그림 4. 대상부류 체계의 기술 예: 교통수단 부류

우리는 어휘문법과 대상부류 이론을 주요 방법론으로 하여 아래와 같은 논항구조 정보<sup>27)</sup> 기술 모형을 제안한다. 우리 모형은 먼저 술어 어휘와 의미, 그리고 논항구조 형태를 보이고 이어 주어와 비릇한 각 논항의 속성을 기술한 뒤 마지막으로 변형속성을 표상한다. 앞서 보았던 어휘문법의 기술과 정보 제시 순서가 다른데, 이는 학습자가 술어 어휘가 어떤 형태의 단문을 만들고 이때 술어 어휘가 표현하는 의미가 무엇인지를 먼저 알 수 있게 하기 위함이다. 또한 이를 먼저 보임으로써 이후 제시되는 논항에 관한 세부

26) 프랑스어의 경우 Le Pesant & Mathieu-Colas (1998), Lee (2001), Gross, G. (1994, 2012), 한국어의 경우 이성현(2001a,b, 2006, 2010)을 참조.

27) 이외에도 논항구조 정보를 포함하여 어휘의 결합 정보를 기술하기 위하여 설명결합사전Dictionnaire explicatif et combinatoire의 방법론인 I. Mel'čuk의 의미-텍스트 이론 Meaning-Text theory도 적극 수 용하였다. 대표적인 연구로는 Mel'čuk (1995, 1997)이 있고, 국내에서는 홍재성, 박동호, 고길수(2001)과 이병근, 박진호(2000) 등이 있다.

정보와 변형속성 정보를 통합적으로 이해할 수 있기 때문이다. 한편 술어 어휘의 의미와 명사로 실현되는 주어와 목적보어 등 논항의 속성에 관한 정보는 대상부류 이론에서 제안하는 어휘 부류 명칭으로 기술된다.<sup>28)</sup>

			논항속성		변형속성		
어휘	의미	논항구조	주어	N0=명사	대명동사구문		
				N0=동사구			
				N0=절			
			직접목적보어	N1=명사	수동구문	동작주보어	
				N1=동사구			
				N1=절			
				생략			
			간접목적보어	Prep	...		
				N2=명사			
				N2=동사구			
				N2=절			
				생략			
			...				

그림 5. 프랑스어 기본어휘 DB의 문법정보 표상을 위한 미시구조

다음으로 앞서 보았던 동사 *aller*의 논항구조 정보를 이 모형을 따라 기술한 사례를 보자.

3.3. 논항구조 정보 표상 예 : 동사 *aller*

먼저 동사 *aller*가 만드는 단문의 형태와 의미를 중심으로 기술해야 할 논항구조 관련 정보를 환기하면 다음과 같다<sup>29)</sup>. 즉 이 동사는 최소 5 개의 의미를 표현하고, 의미에 따라 서로 다른 형태의 단문을 구성한다. (11a)는 주어 논항의 의미가 ‘인간, 동물, 교통수단, 길’을 지칭하는 명사로 제한되며, 특히 스스로 이동을 할 수 없는 무생물을 지칭하는 명사는 주어로 나타나지 못 한다. (11b)도 주어 논항에 비슷한 의미 제약이 있는데, 바로 동사가 표현하는 의미에서 알 수 있듯이 어떤 목적을 가지고 이동할 수 있는 개체를 지칭하는 명사만이 주어로 나타난다. 한편 (11c)는 부사 *bien* 또는 *mal*이 함께 쓰이고, (11d)는 전치사 *avec*이 이끄는 간접목적보어를 갖는다. (11e)는 이른바 조동사 용법으로 다른 경우와 달리 어떤 구체적인 사태를 표현하지는 않고 이어 나오는 동사가 표현하는 사태가 가까운 미래에 벌어지게 됨을 나타낸다.

28) 이는 술어 어휘의 문법정보를 표상하기 위한 모형이며, 비술어 어휘의 문법정보는 대상부류 이론의 논항 명사 부류 체계, 세종전자사전, 설명결합사전과 의미-텍스트 이론 등을 참조하여 구성된 별도의 모형을 통해 기술된다.

29) 이는 여기서 언급한 사례만을 대상으로 논항구조와 관련 정보를 정리한 것이다.

- (11) a. ~에 가다 : + 장소 표현 전치사구
- b. ~하러 가다 : + (장소 표현 전치사구+) 부정법 동사구
- c. 잘 지내다 : + bien / mal
- d. ~과 어울리다 : + (bien) avec-전치사구
- e. 근접미래 : + 부정법 동사구

다음은 위에 정리한 내용을 우리 모형을 따라 기술한 동사 *aller*의 논항구조 정보이다.<sup>30)</sup>

**표 15.** 프랑스어 기본어휘DB의 문법정보 표상 사례 : 동사 *aller*의 논항구조

aller	의미	논항구조	논항속성			
			주어	N0=명사	인간, 동물, 교통수단, 길	
간접 타동사 <sup>31)</sup>	~에 가다	N0 V Prep N1	주어	Prep	인간, 동물, 교통수단, 길	
			목적보어	N1=명사 생략	장소	-
간접 타동사	~하러 가다	N0 VVinf	주어	Vinf	인간, 동물	
			목적보어	생략	-	-
자동사	(잘) 지내다	N0 V ADV	부가어	Prep N	Prep N	à 장소
			주어	N0=명사	인간	
간접 타동사	~과 어울리다	N0 V Prep N1	기타요소	ADV	<i>bien, mal</i>	
			주어	N0=명사	+	
간접 타동사	~과 어울리다	N0 V Prep N1	주어	Prep	<i>avec</i>	
			목적보어	N1=명사 생략	+	-
조동사 <sup>32)</sup>	곧 ~할 것이다	N0 V Vinf	부가어	ADV	<i>bien, mal</i>	
			주어	N0=명사	+	

30) 이는 (11)에서 정리한 것을 우리가 제안하는 모형을 따라 표상한 것이다. 즉 이를 제외한 다른 용법과 관련 내용은 모두 정보 표상에서 제외되었다. 이는 여기서의 목적이 문법정보 표상 모형과 기술 사례를 소개하는 것이지 동사 *aller*의 완전한 기술을 추구하는 것이 아니기 때문이다.

31) 이 용법은 간접타동사로 보기 어렵다는 분석도 있다. 하지만 자동사/타동사/간접타동사 등 일반적으로 통용되는 구분에서는 하위 유형 간 경계가 모호하고 입장에 따라 다양하게 의견이 나뉘고 있는 만큼 이를 간접타동사로 보는 데는 무리가 없을 것이다. 또한 이 용법은 *tout droit*(곧장 앞으로), *à droit*(우측으로), *à gauche*(왼쪽으로) 등 방향 지칭 표현이 함께 나타날 수도 있는데 이를 기술할 정보 항목이 없는 것이 문제로 지적될 수도 있겠으나, 이를 다루지 않은 것은 여기서의 목적이 동사 *aller*의 완전한 기술이 아니라 문법정보 기술을 위한 모형과 이를 적용한 간단한 사례를 보이는 것이기 때문이다.

32) 이 용법은 준조동사이지 조동사가 아니라는 시각도 있다. 이 문제 역시 동사의 유형을 얼마나 세분할 것인가와 관련되며, 일반적인 일반동사와 조동사의 구분에서는 입장에 따라 다양한 해석이 있음을 밝힌다.

## 4. 결 론

우리는 프랑스어 기본어휘 DB 구축을 목표로 이에 수록될 기본어휘를 선정하는 과정 및 결과와 이들의 문법정보를 기술하는 표상 모형과 기술 사례를 살펴보았다.

기본어휘와 관련하여 우리는 A등급 *Référentiel*이 제안하는 어휘를 기본어휘로 채택하면서도 기본어휘로서의 적정성이 낮다고 판단되는 430여개의 어휘들을 제외하고 이를 대신하여 고빈도 어휘 중 초급 수준에서 학습할 필요가 있다고 판단되는 200개를 추가하여 총 1,300개의 어휘로 구성된 새로운 기본어휘를 작성하였다. 즉 우리는 A등급 *Référentiel*의 자료에는 한국인을 대상으로 하는 초급 프랑스어 학습에 적합하지 않은 어휘가 포함되어 있다고 보았던 것이다. 이러한 이유로 사용빈도를 기준으로 적정성 검증을 시행한 것이며, 이를 통해 사용빈도는 물론 가용성이 아주 낮은 따라서 학습의 필요성이 아주 떨어지는 어휘를 다수 확인하고 이들을 걸러낼 수 있었던 것이다. 그 과정에서 또한 우리는 A등급 *Référentiel*의 어휘가 가지는 내적인 문제점, 즉 어휘 선정과 등급 배분이 분명한 기준과 원칙 하에 이루어지지 않았음을 확인하였다.

한편 기본어휘의 문법정보 표상과 관련하여 우리는 어휘문법과 대상부류 이론을 참고하여 술어 어휘의 문법정보를 기술하는 새로운 표상 모형을 개발하였고, 이를 적용하여 동사 *aller*의 일부 의미와 용법의 논항구조 정보를 기술한 사례를 제시하였다. 우리의 표상 모형은 어휘문법과 대상부류 이론의 기본 틀을 따르되, 한국인 학습자의 입장에서 술어 어휘의 의미와 용법을 통합적으로 이해할 수 있도록 정보 항목의 구성과 표상 방식을 수정하였다.

우리 연구는 기본어휘의 선정 방식과 기본어휘 기술을 위한 DB의 구성 및 기술 방식을 체계화함으로써 다른 등급 어휘의 선정과 DB 구축에 활용될 수 있는 방법론을 정립했다는 점에서 의미가 있다고 할 수 있을 것이다.

## References

- Béacco, Jean Claude et al. (2004). *Niveau B2 pour le français. Un référentiel*, Didier.
- Béacco, Jean Claude et al. (2006). *Niveau A1.1 pour le français. Un référentiel*, Didier.
- Béacco, Jean Claude et al. (2007). *Niveau A1 pour le français. Un référentiel*, Didier.
- Béacco, Jean Claude et al. (2008). *Niveau A2 pour le français. Un référentiel*, Didier.
- Béacco, Jean Claude et al. (2011). *Niveau B1 pour le français. Un référentiel*, Didier.
- Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*, Conseil de l'Europe, Division des politiques linguistiques, Strasbourg, Didier, 2000.
- Coste, Daniel et al. (1976). *Un Niveau-seuil : Systèmes D'Apprentissage des Langues Vivantes par les Adultes*. Strasbourg : Conseil de la coopération culturelle du Conseil de l'Europe.

- Gougenheim, Georges et al. (1967). *L'élaboration du Français Fondamental, 1<sup>er</sup> Degré : étude Sur L'établissement D'un Vocabulaire et D'une Grammaire de Base*. Didier, Paris.
- Gross, Gaston. (1994). Classes d'objets et description des verbes. *Langages* 115, 15-31.
- Gross, Gaston. (2012). *Manuel D'Analyse Linguistique. Approche Sémantico-syntaxique du Lexique*. Villeneuve d'Ascq, Presses universitaires du Septentrion.
- Gross, Maurice. (1975). *Méthodes en Syntaxe*. Paris, Hermann.
- Gross, Maurice. (1981). Les bases empiriques de la notion de prédicat sémantique. *Langages*, 63, 7-52.
- Guillet, Alain and Christian Leclère. (1992). *La Structure des Phrases Simples en Français. Constructions Transitives Locatives*. Genève, Droz.
- Ko, Kilsoo. (2003). Combinatory information of non-predicate nouns in the sejong electronic dictionary. In S.-O. Lee, ed., *Korean &/or Corpus Linguistics: Proceeding of ICKL-TU Berlin International Conference on Korean/Corpus Linguistics, 21-22 July 2003*, 275-286, 경진문화사.
- Kusseling, François S. (2012). *A Corpus-Based Evaluation of the Common European Framework Vocabulary for French Teaching and Learning*. PhD Dissertation, Brigham Young University.
- Kusseling, François S. and Deryle Lonsdale. (2013). A corpus-based assessment of French CEFR lexical content. *Canadian Modern Language Review* 69(4).
- Le Pesant, Denis and Michel Mathieu-Colas. (1998). Introduction aux classes d'objets. *Langages* 131, 17-23.
- Lee, Seong Heon. (2001). *Les classes d'objets d'événements. Pour une typologie sémantique des noms prédictifs d'événement en français*. thèse de Doctorat, LLI-Université Paris 13.
- Lonsdale, Deryle and Yves Le Bras. (2009). *A Frequency Dictionary of French: Core Vocabulary for Learners*. Routledge.
- Mel'čuk, Igor. (1995). The future of the lexicon in linguistic description and the explanatory combinatorial dictionary. In I.-H. Lee, ed., *Linguistics in the Morning Calm 3* (Selected papers from SICOL-1992), Seoul, 181-270.
- Mel'čuk, Igor. (1997). *Vers une linguistique Sens-Texte. Leçon inaugurale*. Paris: Collège de France.
- Mel'čuk, Igor. et al. (1984, 1988, 1992, 1999). *Dictionnaire explicatif et combinatoire du français contemporain I, II, III, IV*, Montréal : Presses de l'Université de Montréal.
- Milton, James. (2006). Language lite? learning French vocabulary in school. *Journal of French Language Studies* 16(2), 187-205.
- Porcher, Louis. et al. (1979). *Adaptation de "Un niveau-seuil" pour des Contextes Scolaires*. Strasbourg: Conseil de l'Europe.
- 김연자. (2006). “고등학교 교육과정에서의 프랑스어 기본어휘 선정에 대하여”, 『프랑스어문교육』, 23, 7-26.
- 남지순. (2007). 『한국어 형용사 어휘분법』, 한국문화사.
- 이병근, 박진호. (2000). 결합설명 사전의 어휘 기술 방법론, 『인문논총』 43, 155-200.
- 이성현. (2001a). “서술명사 기술을 위한 대상부류 개념의 활용-불어 사건명사의 예”, 『프랑스어문교육』, 12, 129-149.
- 이성현. (2001b). “전자사전 구축을 위한 언어기술의 한 방법: 대상부류”, 『언어학』 30호, 185-206.

- 이성현. (2004). “전자사전에서의 기능동사 구문 처리 문제”, 『한국사전학』 4, 279-322.
- 이성현. (2005). “전자사전 구축과 의미부류 - 세종 명사 의미부류 체계의 예”, 『한국사전학』 5, 103-138.
- 이성현. (2006). “불어 ‘대칭’ 명사의 통사·의미론적 특성 연구 : ‘대칭’ 명사의 대상부류 구축”, 『인문논총』 56, 277-331.
- 이성현. (2010). “프랑스어 발화명사의 통사 의미 속성에 관한 연구 : 프랑스어 발화명사의 대상부류 구축”, 『프랑스어문교육』 33, 357-388.
- 홍재성. (1987). 『현대 한국어 동사구문의 연구』, 탑출판사.
- 홍재성, 박동호, 고길수. (2001). “설명결합사전의 어휘의미 기술”, 『어학연구』, 37(3), 533-562.
- 홍재성 외. (1997). 『현대 한국어 동사 구문 사전』, 두산동아.

고길수  
 선임연구원  
 언어교육원 언어능력측정센터  
 서울대학교  
 08826 서울시 관악구 관악로 1  
 전자우편 : kilsooko@yahoo.com

이성현  
 교수  
 인문대학 불어불문학과  
 서울대학교  
 08826 서울시 관악구 관악로 1  
 전자우편 : lsh0717@snu.ac.kr

접수일자 : 2018. 11. 12  
 수정본 접수 : 2018. 12. 19  
 게재결정 : 2018. 12. 20