

한국어 학습자 말뭉치 구축을 위한 기초 연구

-개인 정보 표지 체계와 오류 정보 표지 체계를 중심으로-

김정숙·김유정*

Abstract

This paper aims to present the essential elements for constructing a learners' corpus for Korean as a foreign language, with the primary purpose to propose an objective basis for the corpus. It is necessary to have an objective measuring system for the construct of the corpus. However, after constructing a learners' corpus, the analyst's subjectivity often intervenes with the interpretation of data. It is difficult to use a system with a subjective analysis in such a wide area. Furthermore, it is necessary that a schema for constructing an objective learners' corpus come first, in order for its value for implementation to increase. For this reason, I will attempt to discuss the necessity and utilization of a corpus, as well as some theoretical keys needed to construct an objective learners' corpus.

[key words] learners' corpus(학습자 말뭉치), mark-up(문서 표지), error tagging(오류 태깅)

1. 연구의 필요성과 의의

한국어 교육이 질적으로 양적으로 발전하고 있는 요즘, 새롭게 주목을 받고 있는 분야는 바로 말뭉치 구축에 관한 것이라고 할 수 있다. 한국어 교육에서 필요한 자료들과 산출된 자료들을 수집하고 데이터화하여, 이를 통해 한국어 교육의 모습을 조명하고 다시 한국어 교수 현장에 발전적으로 응용될 수 있도록 하는 다양한 연구가

* 고려대

진행되고 있음을 볼 수 있다. 그 중에서도 한국어 학습자 말뭉치 구축의 필요성, 이를 통한 학습자들의 오류 분석과 그 활용 등에 관한 연구가 다양하게 모색되고 있음을 쉽게 발견할 수 있다.¹⁾ 그러나 이러한 움직임 속에서 안타까운 것은 자료를 수집하고 자료화하는 데 필요한 객관적인 근거가 구체적으로 제시되고 있지 못하다는 것이다.²⁾

지금까지 한국어 학습자 말뭉치에 관한 연구들은 서상규·한영균(1999)를 시작으로 본격화되었다고 할 수 있다. 서상규·한영균(1999)는 말뭉치를 위한 일반적인 이론과 함께 각 나라의 말뭉치 소개, 그리고 국어 말뭉치 구축의 이론과 실체를 보여주고 있다. 이 연구 이후 김유미(2000)을 비롯하여 유석훈(2001), 성지연(2002), 유현경·서상규(2002), 서상규 외(2002) 등 한국어 학습자 말뭉치 연구로 이어지고 있다. 이 연구들은 크게 두 가지로 나뉘는데, 그 하나는 한국어 학습자 말뭉치를 구축하기 위한 기초 연구로서 구축 단계, 구축의 어려움, 말뭉치의 외국 사례 등에 대해 연구한 것이고³⁾ 다른 하나는 한국어 학습자 말뭉치를 이용한 오류 분석의 실패들을 연구한 것이다.⁴⁾

이들 연구들은 모두 한국어 학습자 말뭉치 연구의 기초 연구로서 그 의의가 크다고 할 수 있다. 그러나 서상규·한영균(1999)의 연구에도 언급된 말뭉치의 재사용성(reusability)과 기계 가독성(readability)이 제대로 실천되고 있지 못한 점 또한 현 상황이라고 할 수 있다. 이는 자료를 공개적으로 활용하지 못하는 현실의 한계도 있겠지만, 가장 중요한 것은 말뭉치를 다루는 기초 작업이 함께 공유될 정도로 공개되거나 객관화되지 못한 이유 때문이라고 할 수 있다. 다시 말하면, 그 근거가 명확하게 나타나지 못하고 단지 결과만이 제시되고 있다는 것이다.⁵⁾

1) 이와 관련한 연구는 서상규·한영균(1999), 김정숙(2001), 유석훈(2001), 김유미(2000), 성지연(2002), 서상규 외(2002) 등에서 볼 수 있다.

2) 말뭉치 구축 시에 필요한 객관적인 근거의 필요성에 대한 언급은 많으나, 구체적으로 제시된 바는 적다.

3) 유석훈(2001), 서상규 외(2002) 등을 들 수 있다.

4) 김유미(2002), 성지연(2002), 유현경·서상규(2002) 등을 들 수 있다.

5) 학습자 말뭉치를 실제로 활용한 최초의 연구인 김유미(2000)에서는 학습자 개인 정보 표지 체계와 오류 정보 표지 체계를 단지 제시할 뿐, 그 제시의 근거가 드러나고 있지 못하다.

말뭉치 구축은 계량적인 작업으로 객관화된 작업이어야 한다. 그러나 특히 학습자 말뭉치를 구축하고 그것을 실제적으로 활용하기 위해 자료를 해석하는 데 있어서는 분석자의 주관이 다소 개입하게 된다. 그러나 지나치게 주관적인 분석은 체계를 가지고 광범위한 영역에서 이용되기 어렵다. 따라서 객관화된 학습자 말뭉치 구축의 방안이 반드시 선행되어야만 학습자 말뭉치의 활용 가치를 높일 수 있을 것이다.

따라서 본 연구에서는 한국어 학습자 말뭉치의 재사용성과 기계 가독성을 최대한 보장하기 위한 기초 연구로서, 개인 정보 문서 표지 체계와 오류 정보 문서 체계에 초점을 두고 논의를 진행하고자 한다.⁶⁾

개인 정보 문서 표지 체계와 오류 정보 문서 체계는 학습자 말뭉치 구축 단계에서 초기 단계에 실행되는 작업들이며, 학습자 말뭉치를 분석하고 해석하는 데 있어서 가장 기초적인 작업이라고 할 수 있다.⁷⁾ 이것이 통일화되고 객관화될 수 있다면 말뭉치의 활용도 또한 최대한 보장받게 될 수 있을 것이다. 이러한 점에서 이 연구는 한국어 학습자 말뭉치의 기초 연구로서 그 의의를 가질 수 있을 것이다.

이를 위해 본고에서는 우선 학습자 말뭉치 구축의 필요성과 활용도, 구축 절차를 간단히 언급하고자 한다. 그리고 개인 정보 문서 표지 체계와 오류 정보 문서 체계를 표준화하고 객관화하는 이론적 기틀을 살펴 그 실례를 보이는 데 중점을 두고 논의를 진행한다.

2. 한국어 학습자 말뭉치 구축의 필요성

한국어 학습자 말뭉치 구축의 필요성은 유석훈(2001), 서상규 외(2002) 등에서 언급된 바 있다. 이들 연구들을 정리하고 누락된 사항들을 덧붙여서 언급해 보면 아래와 같다.

이후 이 연구를 바탕으로 하고 있는 논의들에서 이에 대한 적극적인 논의와 개선보다는 그냥 따르고 있다는 것이 문제라고 할 수 있다.

- 6) 물론 학습자 말뭉치 구축에 있어서는 문법 정보 주석, 형태소 분석, 구문 정보 주석, 의미 정보 주석 등이 문제가 될 수 있고, 이들에 대해서도 논의가 진행되어야 한다. 그러나 본 연구에서는 개인 정보 표지 체계와 오류 정보 표지 체계에 논의를 한정한다.
- 7) 구체적인 근거와 사례는 이후 논의에서 언급될 것이다.

첫째, 한국어 학습자의 언어 사용 자료를 수집하고 말뭉치를 구축하여 한국어 학습자들의 언어 사용 양상을 파악할 수 있다.

현재 한국어 학습자들이 의사소통 언어로 사용한 자료들을 부족하다. 또한 이들이 체계적인 절차로 수집되거나 가공되고 분석되지 못하고 있는 것 또한 사실이라고 할 수 있다. 한국어 학습자들의 언어 사용 자료를 목표를 가지고 수집·가공·분석하는 것은 한국어 학습자들이 어떻게 언어를 학습하고 사용하는지를 고찰하기 위해 반드시 필요한 작업이라고 할 수 있다.

둘째, 다양한 목적으로 사용될 수 있도록 한국어 학습자들의 문어·구어 자료를 전자 자료로 구축할 수 있다.

언어 사용 자료를 그대로 이용하는 것은 장기적이지 못하며 또한 응용하기도 어렵다. 문어·구어 자료의 특성에 맞게 전자 자료로 구축하여 연구 목적에 맞게 이용하는 것이 말뭉치의 재사용을 가능하게 할 수 있다.⁸⁾

셋째, 한국어 학습자 오류 말뭉치를 분석할 수 있는 적절한 태깅 시스템(tagging system)을 개발하여 한국어 교육의 질을 향상시킬 수 있다.

현재 한국어 학습자 말뭉치 연구는 일반적인 언어 말뭉치 연구에 기반한 것이다. 물론 말뭉치라는 점에서는 일반적인 연구의 특성도 고려되어야 한다. 그러나 한국어 학습자 말뭉치, 오류 말뭉치라는 개별적 특성이 파악되고 이에 기반한 말뭉치 구축이 필요하다. 이는 학습자 오류 말뭉치의 연구를 탄탄한 기초 위에 가능하게 하며, 지속적인 연구를 가속화하고 그 결과 한국어 교육의 질적 수준을 높이게 할 수 있다.

넷째, 한국어 학습자들의 숙달도 단계별(초급에서 고급까지) 오류 유형을 조사·분석하여 한국어 학습자들의 언어 발달 단계를 가시화 할 수 있다.

현재 한국어 학습자들의 숙달도에 관한 기술들은 다소 직관과 경험에 의존한 것들이다. 학습자 언어를 연구할 때, 그리고 학습자들을 위한 숙달도별 언어 자료와 교수법을 개발할 때 언어 발달 단계의 객관성과 신뢰성 보장은 중요한 항목이다. 따라서 숙달도 단계별 오류 유형의 분석은 학습자들의 언어 발달 단계를 객관화하고 이

8) 문어 자료를 수집하고 가공·분석하는 데 비해 구어 자료를 수집하고 가공·분석하는 작업이 어렵다는 것은 말뭉치 연구에서 여러 번 언급된 사실이다. 그러나 한국어 학습자 말뭉치 구축에 있어서도 구어 자료 수집과 분석에 관한 연구가 구체적으로 논의되어야 한다.

를 가능하게 할 수 있다.

다섯째, 한국어 학습자들의 언어권별(영어권, 일어권, 중국어권 등) 오류 유형을 조사·분석하여 학습자들의 언어 학습을 돕는 방안을 모색할 수 있다.

학습자들의 말뭉치를 보면 모국어의 영향으로 인한 학습자들의 오류 유형 발견할 수 있다. 이는 오류 수정이라는 소극적인 방법의 교육적 기여도 하게 하지만, 적극적으로는 오류 유형을 분석하여 학습자가 한국어를 사용할 때 오류를 예방하고 대처할 수 있게 한다. 이는 학습자의 언어 학습을 돕는 긍정적인 역할을 할 수 있다.

여섯째, 한국어 학습자들의 오류 유형을 조사·분석하여 실제 교실 학습에 발전적으로 수용할 수 있다.

오류 유형의 조사와 분석은 오류의 원인과 그 양상을 파악할 수 있게 하며, 이를 기반으로 하여 오류를 일으키지 않도록 하는 교수 방안을 마련하도록 한다. 이러한 작업은 한국어 교육 현장에 적극적으로 수용되어 한국어 학습을 쉽고 즐겁게 만들 수 있다.

3. 한국어 학습자 말뭉치 구축 단계

한국어 학습자 말뭉치 구축 단계를 보면 아래와 같다. 구축 단계를 살피는 것은 각 단계별로 필요한 작업이 무엇인지를 파악하게 하며, 그 중요도를 다시 한 번 인식할 수 있게 한다는 점에서 의의를 가진다.

1단계 : 예비 작업(자료 분석/이론적 분석)⁹⁾

이 단계는 '실험 말뭉치 구축 → 실험 말뭉치의 분석 → 말뭉치 설계 → 구축'으로 이루어지는데 실제 말뭉치 구축을 위한 사전 작업 단계라고 할 수 있다.

2단계 : 학습자 언어 말뭉치 자료 수집

이 단계에서는 문어 자료와 구어 자료를 어떤 비율로 어떤 방법으로 수집할지 그

9) 예비 작업의 필요성과 그 의미 및 절차에 대해서는 서상규·한영균(1999:50-54)을 참조할 수 있다.

방법이 논의되어야 한다.¹⁰⁾

3단계 : 학습자 개인 정보 문서 표지 체계 개발 및 낱말뭉치의 구축

수집된 말뭉치 자료를 활용하기 위한 첫 단계로, 학습자 말뭉치의 활용도를 높일 수 있는 학습자 개인 정보 문서 표지를 개발하고 낱말뭉치를 구축한다. 이는 학습자 말뭉치를 다양하게 활용하기 위한 기초 단계의 작업이다.

4단계 : 형태소 분석

표준 말뭉치를 위해 개발된 형태소 분석기를 통한 학습자 말뭉치의 형태소 분석하는 단계이다. 한국어 학습자 말뭉치의 형태소 분석을 위해 형태소 분석기가 개발되거나, 국어 자료들을 위한 형태소 분석기를 활용하는 방안이 모색될 수 있다.

5단계 : 학습자 말뭉치의 오류 양상 분석 및 태깅 지침 마련

학습자 말뭉치에 나타난 오류 양상에 대한 객관적인 태그셋(tag set), 즉 오류 정보 표지 체계를 개발하여야 한다. 이는 앞서서도 언급했듯이 자료의 활용도를 최대한 보장할 수 있다.

6단계 : 오류 태깅 시스템 신개발

학습자 오류 말뭉치를 자동 태깅할 수 있는 시스템의 개발하는 단계이다. 전 단계에서 오류 양상을 분석하고 오류 정보 표지 체계를 개발한 후에 이를 태깅할 수 있는 시스템을 새롭게 개발해야 한다. 이는 한국어 학습자 말뭉치의 활용을 극대화하

10) 고려대학교에서는 1급-6급 학습자를 대상으로 하여 문어 자료와 구어 자료(부분적으로) 하고 있다. 현재 수집된 자료의 양은 10만 어절 정도이다. 구어 자료는 시험 자료와 의사 소통 활동에 사용된 수업 시간 내 발표 자료들이며, 문어 자료는 시험 자료, 의사 소통 활동에 사용된 수업 시간 내 작문 자료, 숙제 작문들이 그 대상이 되고 있다. 여기에서 숙제로 학습자가 집에서 해 온 작문 과제는 많은 양을 차지하지 않는다. 이것은 학습자가 스스로 작문을 했는지 신뢰하기 어렵기 때문이다. 이는 학습자의 언어 발달 정도를 분석하는 학습자 말뭉치의 의도를 흐릴 수 있다. 따라서 학습자 스스로 한 숙제 작문 자료만을 대상으로 한다.

기 위해 반드시 필요한 작업이다.

7단계 : 오류 태깅

새로 개발된 오류 태깅 시스템을 이용하여 학습자 말뭉치에 오류를 자동으로 태깅한다. 이 단계에서는 오류 태깅의 정확성 확보가 가장 중요하다.

8단계 : 숙달도 단계별·언어권별 학습자 말뭉치 분석

태깅된 오류 말뭉치를 다양한 방법으로 분석하고, 이를 통해 교수 현장에서 학습자를 위해 어떻게 활용될 수 있을지 방안을 모색하는 단계이다.

이상에서 언급된 바와 같이 각 단계에서 중점 사항들이 잘 고려되었을 때, 한국어 학습자 말뭉치 구축과 그 활용이 그 의의를 가지게 된다.

4. 개인 정보 문서 표지 체계 개발

말뭉치의 활용도를 극대화한다는 것은 말뭉치의 재사용성(reusability)과 가독성(readability)이 확보되어야 함을 의미한다. 이를 위해 '표준 문서 표지 체계'를 만드는 것이 필요하다. 이러한 표준화 작업은 말뭉치 자료의 호환을 위해 반드시 필요한 작업이다. 따라서 우리는 여기에서 실제로 한국어 학습자 말뭉치를 구축함에 있어서 필요한 표지 체계 항목에 무엇이 있는지, 그 표지 체계를 어떻게 할 것인지 다시 한번 고려해 볼 필요가 있다. 이는 개인 정보 문서 표지의 항목이 많을수록 학습자 말뭉치의 분석을 여러 각도에서 가능하게 하는 데 영향을 미칠 수 있기 때문이다. 개인 정보 문서 표지는 학습자 정보와 문단 정보 등의 기초 정보를 부가한 것이다. 학습자 말뭉치는 표준 말뭉치와 달리 그 자체보다는 학습자의 언어 습득 발달도를 분석하고, 이를 긍정적으로 이용할 수 있도록 하는 것이 중요하다. 따라서 학습자 변인이 언어 습득에 어떻게 관련을 가지고 있는지를 읽어내는 것이 중요하다고 할 수 있다.

먼저 학습자의 개인 정보 문서 표지 체계를 세우기 위한 이론적인 바탕을 살펴보

면 다음과 같다.

모든 인간은 모국어를 습득하며, 또한 다른 언어를 습득·학습할 수 있다. 그러나 학습자가 그 언어를 습득하고자 하는 욕구와 필요성을 가질 때 더 효과적이다. 그러므로 제2언어 교사는 학습자의 특성을 고려하여, 즉 학습자를 생물학적·인지적 능력과 함께 감성적·성격적·사회적 특성을 고려한 ‘총체적 학습자’를 고려해야 한다 (신성철 역, 1996; Brown, 1980).

먼저 학습자의 생물학적 특성에는 연령과 성별, 인지적 특성으로는 지능, 적성, 학습 전략, 모국어 지식, 교육 수준, 그리고 추리력을 들 수 있고, 감성적 특성으로는 태도, 동기, 성격적 특성으로는 내·외향성, 충동·사고형, 완벽주의, 감정이입과 자존심을 들 수 있다. 마지막으로 사회적 특성으로는 타인 선호도, 사회·경제적 배경을 들 수 있다. 학습자는 이러한 요인들에 의해 영향을 받으며 언어를 습득·학습하게 되는 것이다. 따라서 학습자 말뭉치를 구축함에 있어서도 이러한 요인들이 함께 고려되어야 한다. 그러나 이러한 학습자 변인들은 완벽하게는 아니지만 객관적으로 측정될 수 있는 것도 있고, 그렇지 않은 것도 있다. 따라서 계량화 작업인 말뭉치에서는 최대한 객관성을 가질 수 있는 요인들을 그 학습자 개인 정보로 설정해야 한다. 이러한 이론을 바탕으로 객관화 할 수 있는 학습자 개인 정보를 추출해 내면 다음과 같다.

- 1) 생물학적 특성 - 성별과 나이
- 2) 인지적 특성 - 학습자 모국어, 학습자 외국어, 한국어 학습 경험, 한국어 등급
- 3) 감성적 특성 - 학습 목적(동기)
- 4) 사회적 특성 - 현재 학습 기관, 직업
- 5) 의사소통 자료 특성 - 과제 유형, 과제 장르, 시기

위에서 제시된 것 중 5)의 의사 소통 자료 특성은 학습자 말뭉치 구축이라는 목표에 있어서, 말뭉치 자료 특성이 함께 포함된 것으로 학습자 개인 정보에 함께 포함될 수 있다.

여기에서 학습자의 개인 정보 문서 표지 체계의 안을 제시하면 다음과 같다.11) 이러한 논의는 김유미(2000), 서상규 외(2002)에서 제시된 개인 정보 문서 표지 체계를 제시함에 있어서 이론적 근거를 보이지 못한 점과는 달리 이론적 근거를 제시하고 이를 통해 개인 정보 문서 표지 체계를 표준화할 수 있는 기반을 마련한다.

〈표 1. 학습자 말뭉치 개인 정보 표지 안〉

분 류		표 지
한국어 등급	등급	1. 한국어 교육 기관에서의 등급을 표시한다. : L1(1급), L2(2급), L3(3급), L4(4급), L5(5급), L6(6급),
	숙달도	1. 4등급으로 나눈다. : NO(Novice), IN(Intermediate), AD(Advanced), SU(Superior)
국가		1. 앞에서 세 자리까지 한다. KOR(한국), GER(독일), JAP(일본), USA(미국), CHI(중국), FRA(프랑스), BRA(브라질), NOR(노르웨이), AUS(호주), SWI(스위스), UZB(우즈베키스탄), KAZ(카자흐스탄), MON(몽골), PHI(필리핀), 2. 재외동포들은 뒤에 -K를 붙인다. : JAPK(재일동포), USAK(재미동포)
학습자 모국어		1. 언어명 앞에 S(Speak)를 붙인다. -> SKOR(한국어), SGER(독어), SJAP(일어), ENG(영어), SCHI(중국어), SFRA(불어),
* 학습자 외국어	언어	1. 언어명 뒤에 2(Second)를 붙인다. -> GER2(독어), JAP2(일어), ENG2(영어), CHI2(중국어), FRA2(불어), 2. 2개 이상의 언어를 학습한 학습자의 경우 이어서 표기한다. ->GER2 FRE2
	숙달도	1. 4등급으로 나누고 언어 뒤에 붙인다. : GER2NO(Novice), GER2IN(Intermediate), GER2AD(Advanced), GER2SU(Superior)

11) 김유미(2000:38-42)의 내용을 참조로 하여, 새로운 항목들을 보충하였다. 새 항목은 *로 표시한다.

* 지금까지의 한국어 학습 경험	기간	1. 학습 경험이 있을 경우, 주 당 시간은 H, 개월은 M, 연은 Y로 나타내고 앞에 수를 붙인다. : 20H, 6M (480H) 또는 2H 3Y(290H)
	기관	1. 학습한 시기의 나라 이름에 기관(INS)과 독학(PRI), 가정 학습(FAM), 개인 학습(MTM)을 붙여서 분류한다. : JAPINS, FRAFAM, KORINS....
과제 유형		1(제한된 시간 내의 자유 작문), 2(시험에서의 작문), 3(숙제로 하는 쓰기)
* 과제 장르		1(수필), 2(설명문), 3(논설문), 4(안내문), 5(광고문), 6(보고서), 7(신문 기사) ... ** 여기에서는 작문의 장르만을 제시하였다. 공식적인/개인적인 글쓰기를 나눌 수도 있다.
성별		M(남), F(여)
나이		A10(10대), A20(20대), A30(30대), A40(40대), A50(50대).....
직업		STU(DENT), TEA(ChER), WOR(KER), WIF(E), NON(E).....
* 학습 목적		ACA(Academic), TRA(Transactional), HOB(Hobby), ...
* 시기		1. 말뭉치 자료 작성 시기를 연도로 표기한다. : 1999, 2000, 2002,
* 현재 학습 기관		1. 교육 기관의 이름 앞에 I를 붙인다. : IKOR(고려대), IYON(연세대), ISEO(서울대),

여기에서 각 항목들에 대해 설명을 덧붙이면 다음과 같다.

먼저 한국어 등급은 한국어 교육 기관의 등급이나 숙달도로 표시된다.¹²⁾ 이러한 등급과 숙달도 항목에서는 반드시 각 기관의 등급별 숙달도 기준이 함께 고려되어야 한다는 것이 중요하다. 각 기관에서 목표로 하고 있는 숙달도 기준과 학습자의 언어 수준을 파악할 수 있으며, 각 단계별 언어 발달 양상을 살피는 데도 기여할 수 있다. 국가와 학습자 모국어 항목은 학습자에 대한 인지적 특성으로 언어권별 학습자 언어

12) 여기에서의 숙달도는 Ellis(1986)의 언어학적 발달의 네 가지 단계를 기반으로, ACTFL의 숙달도 지침이 크게 네 단계인 것을 참고로 제시한 것이다. 그러나 이러한 숙달도는 그 용어와 숙달도 단계별 기준이 충분히 제시된다면 다양하게 응용될 수 있을 것이다.

발달 양상을 살피는 데 도움을 줄 수 있다. 학습자 외국어는 제1언어와 제2언어 학습의 관계, 나아가서는 제2언어 학습이 제3언어 학습에 미치는 관계 등을 확인할 수 있도록 한다.¹³⁾ 덧붙여 설명하자면, 학습자는 타인 선호도의 영향을 갖는다. 즉 부모, 동료들, 지역 사회, 학교로부터 영향을 받고 타인을 의식하여 동일시하려고 한다. 따라서 학습자의 언어 양상을 살필 때 학습자의 외국어와 그 수준, 한국어 학습 경험의 유무, 학습 기간과 학습 기관이 중요하게 된다.¹⁴⁾ 특히 학습 외국어가 한국어와 가장 유사한 구조를 지닌 일본어인 경우 어떤 영향을 미치고 있는지를 판단할 수 있다. 한국어 학습 기간은 지금까지 주로 시간으로 제시된 바 있다. 그러나 이러한 시간 계산은 학습자들이 계산하기 어려워서 그 신뢰도를 확보하기가 어렵다. 따라서 주당 학습 시간과 함께 개월 수나 연수로 제시하게 하고 학습 기관과 함께 고려된다면 좀더 객관적인 학습 시간이 계산될 수 있을 것이며, 현재 학습자의 숙달도 수준과의 관련성을 파악할 수 있을 것이다. 또한 과제 장르는 학습자 말뭉치의 자료가 되는 텍스트의 과제 유형과 함께 고려되어야 하는 항목으로 학습자의 의사소통 활동이 어떤 장면에서 이루어지고 있는지를 결정하는 중요한 것이다. 이러한 과제 장르는 학습자 말뭉치의 오류를 분석할 때도 고려될 수 있다. 성별·나이·직업은 학습자의 생물학적 특성으로 한국어 학습자의 성별, 나이, 직업을 분류하고 각각의 특성을 살펴볼 수 있다. 학습 목적이 설정된 이유는 학습자의 감성적 특성을 가늠함과 동시에 학습 과제에 관련된 성취도를 파악할 수 있도록 하기 때문이다. 또한 학습 시기와 현재 학습 기간은 학습자 말뭉치 구축의 시점과 출처를 밝히는 항목이라고 할 수 있다.

이상으로 학습자의 개인 정보 표지 체계의 안을 설정하게 된 이론적인 근거와 함께 그 모델을 제시해 보았다. 물론 이는 하나의 시안이다. 그러나 그 시안을 제시함에 있어서 학습자 말뭉치의 특성과 사용 목적을 충분히 고려하고, 그 이론적 근거를

13) 이 때 숙달도 표시가 학습자의 판단에 달려 있으므로, 객관성이 떨어질 수 있다는 단점을 지닌다. 이에 대한 좀더 깊은 연구가 진행되어야 하나, 여기에서는 하나의 안으로 제시하는 데 의의를 지닌다.

14) Robert J. Fouser(1997)에서는 제3언어로서의 한국어에서 언어 전이 연구를 통해 학습자의 제2언어가 제3언어로서의 한국어 학습에 미치는 영향을 언급하고 있다.

설정하였다는 것이 중요한 의미라고 할 수 있을 것이다.

이상의 논의를 바탕으로 학습자 개인 정보 표지가 부착된 샘플을 보이면 다음과 같다.

〈표 2. 학습자 개인 정보 문서 표지 부착 샘플〉

```

<head>
<level>L1</level>
<national>JAP</national>
<mother>SJAP</mother>
<foreign>FRA2IN</foreign>
<Korperiod>20H 10M (800H) </<Korperiod>
<Lernschool>JAPPub</<Lernschool>
<task>2</task>
<genre>I</genre>
<gender>M</gender>
<age>A10</age>
<occupation>STU</occupation>
<purpose>ACA</purpose>
<time>2002</time>
<school>IKOR</school>
</head>
<text>
<body>
<p>작년 크리스마스에 많이 했스니다</p>
<p>친구를 같이 한국 음식을 먹곤했습니다</p>
<p>눈을 많이 왔습니다</p>
<p>잘 못 걸었습니다</p>
<p>집에서 사람이 많이 있었습니다</p>
<p>크리스마스 음악을 같이 들었습니다</p>
<p>날시가 주춧습니다</p>
<p>옷을 많이 입었습니다</p>
<p>나는 친구를 매원 음식 못 먹곤했습니다</p>
<p>친구를 많이 안 먹곤했습니다</p>
<p>나는 어머니 한국에 갔습니다</p>
</body>
</text>
</head>

```

5. 오류 정보 표지 체계 개발

학습자 말뭉치를 가공하는 또 다른 단계로 오류 주석이 있다. 학습자의 오류는 학습자 개인의 중간 언어를 보여 줌과 동시에 일반화된다면 한국어 학습자들의 언어 발달 양상을 가시화 할 수 있는 중요한 자료가 될 수 있다. 그러나 이러한 오류 정보 표지 체계는 분석자가 임의적으로 결정하는 경우가 많았다. 따라서 그 해석에 있어서도, 분류에 있어서도 중복되거나 자의적인 해석이 지나치게 개입되는 것을 많이 볼 수 있다.¹⁵⁾ 이러한 현상은 학습자 말뭉치가 재사용 가능해야 하고 서로 공유 가능해야 한다는 재사용성이나 가독성에 적절하지 않게 된다. 따라서 본고에서는 오류 정보 표지 체계를 개발할 때 필요한 이론적 근거를 제시하고 이를 통해 체계적이고 표준화된 지침을 마련해 보고자 한다.¹⁶⁾ 여기에서도 문어 자료를 중심으로 논의를 진행한다.

오류를 기술하기 위한 논의들 중 Ellis(1997)에서는 오류 분석의 절차로 오류의 식별과 기술, 그에 대한 설명과 평가의 단계가 있다고 하였다.¹⁷⁾

- 1) 오류의 식별은 목표어의 정상적인 또는 문법적인 문장들과 학습자가 만든 문장을 비교하여 오류를 찾아내는 단계이다.¹⁸⁾

15) 김유미(2000)에서 시도되고 있는 오류 태그 세트는 오류의 범주를 크게 언어학적 범주(음운론, 구문·형태론, 의미론, 담화), 표면 전략 범주(누락, 첨가-〈이중 표시, 과잉일반화, 단순 첨가〉, 오형태-〈과잉일반화, 대표형, 대표형 교환, 단순화〉, 오어순, 대치), 비교 분석(발달 오류, 언어내적 오류), 의사소통 효과(전국적 오류, 지역적 오류), 심리적 요인(회피, 침묵)으로 나누고 한국어 학습자들의 조사 사용 오류 유형을 50가지로 분류하고 있다. 그런데 이러한 분류가 계층을 가지고 고려되지 못하고 있음을 발견할 수 있다. 즉 대부분으로 이루어진 다섯 가지 분류가 모두 고려되고 있지 못하고, 언어학적 범주, 표면 전략에 중점을 두고 이루어지고 있다.

16) 오류의 대상에 띄어쓰기는 제외한다.

17) 각 단계는 이후 논의에서 적용되어 한국어에 어떻게 반영되는지 사례를 볼 수 있다.

18) 오류 식별 절차는 Corder(1971)에서 도표로 제시되었다(신성철 역, 1996:290 재참조). 여기에서 제시된 절차를 따라 오류는 식별 가능하기도 하고 분석 불가능하기도 하다. Corder는 여기에서 오류에는 뚜렷한 오류(overt errors)와 감추어진 오류(covert errors)가

- 2) 오류의 기술은 식별된 언어 자료를 목표어로 재구성하여 원래의 오류문과 비교한 후 그 차이를 기술하는 것이다. 여기에는 누락(omission), 오형태(misformation), 첨가(addition) 등이 있다.
- 3) 오류에 대한 설명은 오류 발생의 원인을 제공하는 것이다. 여기에는 언어간 전이와 언어내 전이, 그리고 학습의 장, 소통 전략이 그 원인이 될 수 있다.
- 4) 오류에 대한 평가로는 전반적 오류(global errors)와 국부적 오류(local errors)를 들 수 있다.

반면 오류를 기술하기 위한 몇 가지 다른 범주들이 Lennon(1991)에서 언급되었는데¹⁹⁾, 이는

- 1) 일반적인 분류 - 첨가, 생략, 대치 및 어순
- 2) 언어 층위별 분류 - 음운 혹은 철자법, 어휘, 문법, 담화²⁰⁾
- 3) 의사소통 가능성 분류 - 전반적 오류, 국부적 오류
- 4) 오류 측면의 분류 - 영역(domain), 범위(extent)²¹⁾이다.

이상을 보면 오류 분석에는 단계가 있고, 그 분류 기술에는 몇 가지가 있음을 알

있다고 하였다. 뚜렷한 오류는 문장 단위에서 비문법적인 것이고, 감추어진 오류는 문장 단위에서는 문법적으로 잘 구성되어 있으나 정상적인 의사 소통 장면에서 해석 불가능한 경우이다. 이는 전반적 오류, 국부적 오류와 같은 맥락으로 이해될 수 있다.

- 19) 이를 한국어에 적용한 구체적인 예는 김유미(2000)와 서상규 외(2002)에서 살펴볼 수 있다. 그러나 그 판단의 근거가 명확하지 않아, 여러 가지로 해석될 수 있는 경우를 찾아볼 수 있다. 또한 하나의 오류가 다양한 차원에서 해석될 가능성 또한 있다. 이에 대한 논의는 후고로 미루고, 이 논문에서는 각각의 경우를 오류 양상으로 처리하여 오류 표지 체계를 세우는 데 그 범위를 한정하고자 한다.
- 20) 물론 정확하게 어느 하나의 단계로 식별되기 어려운 경우도 있다.
- 21) 영역은 Lennon(1991)에 따르면 오류를 명백하게 만들기 위하여 문맥으로 간주해야 하는 언어 단위 즉 음소, 음절, 단어, 어절, 구, 문장, 담화의 측면이며, 범위는 문장을 수정하기 위해 생략, 대치, 보충, 혹은 어순 재배열 등을 해야 할 언어 단위의 측면이라고 한다. 영어의 경우 a scissors에서 영역은 구(句형)이며, 범위는 부정 관사 a이다.

수 있다. 그러나 지금까지의 한국어 학습자 말뭉치 연구에서 이들은 함께 혼재되어서 혹은 각각 다른 차원으로 논의된 바가 있다. 이는 분석과 해석에까지 영향을 미쳐서 연구자들이 함께 공감하기 어렵고, 미처 그 분석의 근거를 따라하기도 어려운 단점이 있다. 본고에서는 이러한 한계를 극복하기 위해 위에서 제시한 이론이 서로 유기적이고, 단계적으로 관련을 가진다는 점에 착안하여 오류 정보 표지 체계에 응용해 보고자 한다.

1) 오류 분석이 가능한지 불가능지 않은지에 대한 분석과 이에 대한 주석이 있어야 한다. 이는 Corder(1971)의 오류 식별 절차로 가능하다.

→ 오류 분석 가능(POS), 오류 분석 불가능(IMP)

예) a. 립아 걸루심에 봤어요. → IMP (또는 POS)

b. 남자 친구는 창미를 샀어요. → POS

오류 식별은 문장 차원에서만 이루어지는 것이 아니라 담화 전반적인 차원에서 이루어져야 한다. 따라서 a의 문장은 문장만을 봤을 때는 분석 불가능하지만, 담화 차원에서 보면 분석 가능하다. 물론 이는 교수 경험자일 때 훨씬 유리할 것이다.²²⁾ a를 담화 차원에서 분석한 예를 보이면 다음과 같다.

예) 나는 하가여자진그 립아에갔어요. 립아 걸루심에 봤어요. 그리거 받아근 갔어요. 그래서 버브 업었어요.

해석) 나는 한국 여자 친구와 로마에 갔어요. 로마 콜로세움을 봤어요. 그리고 바티칸에 갔어요. 그런데 교황은 없었어요.

2) 오류가 어떤 언어 층위에서 이루어지고 있는지에 대한 분석과 이에 대한 주석이 있어야 한다. 그런데 여기에서 언어 층위로 설정할 수 있는 것으로 어휘와 문법을 설정하고자 한다. 그리고 어휘와 문법 영역의 각각의 하위 층위로 형태, 의미, 사용역

22) 교수 경험자일 때 유리하다는 것은 말뭉치를 작성하게 한 교사는 언어 내용이 어디에서 나온 것인지를 파악할 수 있기 때문이다. 따라서 말뭉치 구축을 위해서는 말뭉치 연구자가 판단하기 어려운 경우 교수 경험자와 함께 분석을 하는 방안을 고려해 볼 수 있다.

을 설정하고자 한다. 그 근거는 Celce-Murcia(1988)과 Larsen-Freeman(1991)의 논의를 이어 김유정(1997)에서 언급되었듯이, 의사소통적 방법에서의 문법이란 형태, 기본 의미, 사회적 기능과 담화로서의 문법을 의미하기 때문이다. 또한 어휘 교육에 있어서 어휘란 형태와 사전적인 의미, 그리고 다른 관련 어휘와의 차이를 나타내는 사용역이 함께 다루어져야 하는 것이다.²³⁾

즉 문법에서의 형태는 철자와 발음 및 구조, 의미는 '시간+에', '장소+에'에서의 '-에'의 의미 구별을 의미하고, 사회적 기능으로서의 문법인 경어법이나 담화로서의 문법인 이, 그, 저 등은 사용역으로 묶을 수 있다.

또한 어휘의 경우 형태는 단어 형성이나 철자, 발음 그리고 의미는 사전적인 뜻, 사용역은 의사소통 장면 즉, 담화 상황과 관련된 어휘의 적절성을 나타낸다고 할 수 있다.

물론 문장 차원을 넘어서 담화의 층위에서도 오류가 분석될 수 있다. 담화의 경우는 문장 단위에서 어휘와 문법 영역을 넘어서서 분석이 이루어져야 한다. 담화의 경우 형태는 수사적 조직을, 의미의 경우는 응집성을 의미하며, 사용역은 담화의 최종 목표인 의사소통 기능의 수행 적절성을 나타낸다고 할 수 있다.²⁴⁾ 그러나 이러한 담화의 오류 분석은 앞에서 언급한 문법과 어휘 항목에 비해 그 식별과 기술이 복잡할 수 있다.²⁵⁾

이상의 논의를 종합하면 다음과 같다. 우선 문어 자료를 중심으로 기술해 보고자 한다.²⁶⁾

23) 여기에서 제시된 형태라는 의미는 형태론(morphology)과 일치하는 의미는 아니다. 단지 형태적인 결합 양상과 함께 form과 관련된 의미를 나타낸다.

24) 이에 대해서는 김유정(1999:117-120)을 참조할 수 있다. 사회언어학적 지식 항목은 한국어에서 어휘나 문법 항목으로 제시될 수 있기 때문에 담화의 영역에서는 제외한다.

25) 이러한 담화 영역의 연구는 앞으로도 지속적인 연구가 필요할 것이다.

26) 형태는 자료의 특성에 따라 문법 구조(단어 형성)와 함께 철자법, 발음으로 나눌 수 있다.

〈표 3. 오류 층위에 따른 오류 주석〉

어휘VO	형태FO
	의미ME
	사용US
문법GR	형태FO
	의미ME
	사용US

담화DI	형태FO
	의미ME
	사용US

- 예) c. 제 취미는 드라이부입니다. → VOFO²⁷⁾
 d. 한국 백화점 안 싸요. 돈이 많이 필리요(들어요). → VO(FO)ME
 e. 개가 돌아가셨어요. → VOUS
 f. 밥을 맛있어요. → GRFO
 g. 학교에 공부해요. → GRME
 h. 나는 내일부터 공부할게요.(일기) → GRUS²⁸⁾

3) 오류가 어떻게 일어나고 있는지를 분석하고 이에 대한 주석이 이루어져야 한다. 각 언어 층위에서 일어나고 있는 오류는 정상적인 한국어와 어떤 차이가 나는지를 그 양상을 기술하는 것이다.

: 누락(OMM), 첨가(ADD), 오형태(MISF), 오어순(MISO), 대치(REP), 〈회피(AVOI), 침묵(SILE)〉

물론 이러한 주석은 좀더 세분화될 수 있다. 그러나 김유미(2000)에서 볼 수 있듯이 세분화 과정에서 과잉일반화와 같은 경우는 첨가와 오형태 등에서 중복 기술되고 있다. 따라서 여기에서는 첨가는 부수적인 요소가 덧붙는 것으로, 오형태는 잘못된

27) 여기에서 제시된 자료는 고려대학교에서 구축 중인 말뭉치 문어 자료들을 이용한 것이다.

28) 이들 분석과 이후의 분석에 대한 근거는 앞에서도 언급했듯이 후고로 미룬다. 여기에서는 오류 표지 양상들을 보이는 데 그 의의를 둔다.

형태가 제시되는 것으로 정의하고 간략화 하고자 한다. 또한 심리적 요인으로 전략적으로 회피하거나 침묵으로 일관하는 오류는 말하기에서 발견될 수 있는 양상이므로, 문어 자료 말뭉치 주석에서는 제외할 수 있다.

- | | |
|---|--------|
| 예) i. 한국_____ 지금 살아요. | → OMM |
| j. 한국 음식 <u>을</u> 도 좋아해요. | → ADD |
| k. 밥을 <u>먹</u> 해요. | → NISF |
| l. 나는 <u>많</u> 이 <u>친</u> 구 마났어요. | → MISO |
| m. 나는 <u>Chicago</u> 에서 <u>Wisconsin</u> 까지 차 안에 잤어요. | → REP |

4) 오류에 대한 평가가 이루어지고 이에 대한 주석이 이루어져야 한다. 속달도별, 언어권별로 그 오류의 정도를 파악할 수 있을 것이다.

: 전반적 오류(GLO), 국부적 오류(LOC)

- 예) n. 빵가비삽는밥을삽니다. → GLO
 (빵이 비싸면 밥을 삽니다?!, 밥값이 싼 밥을 삽니다?!)
 o. 시를 써서 기분이 좋아하는데 나는 시를 쓰는 것을 좋아합니다.
 → LOC

5) 오류가 어느 영역(domain)과 범위(extent)에서 일어나고 있는지에 대한 분석이 이루어지고 이에 대한 주석이 이루어져야 한다. 이는 학습자들의 오류 범위를 알 수 있게 하고, 지속적인 연구를 통해 오류 수정기를 개발하는 데 기여를 할 수 있을 것이다.

→ 영역 : 음소(PHON), 음절(SYL), 단어(WOR), 구(PHR), 문장(SEN), 담화(TEX)

→ 범위 : 형태소 분석기를 통해 이루어진 문법 정보 주석을 이용한다.²⁹⁾ 명사

29) 김흥규·강범모(1997:13)의 분석 표지 일람을 참조할 수 있다. 본고에서는 언급하지 않지만, 이는 국어 말뭉치를 위한 문법 표지이다. 이는 한국어 교육의 모습을 보이는 말뭉치에 응용될 때 어려운 점이 있다. 따라서 한국어 교육용 말뭉치를 위한 문법 표지의 개발 또한 이루어져야 할 것이다.

〈표 4. 오류 정보 표지 체계 안〉³²⁾

1	2		3	4	5	6	7	
분석 여부	오류 층위		오류 양상	오류 평가	오류 영역	오류 범위	오류 원인	
가능 POS	어휘 VO	형태 FO	누락 OMM	전반적 GLO	음소 PHON	일반 명사 NNG	언어간 전이 INTER_TRA	
		의미 ME	첨가 ADD		음절 SYL	고유 명사 NNP		
		사용 US	오형태 MISF		단어 WOR	의존 명사 NNB		
	문법 GR	형태 FO	오어순 MISO		국부적 LOC	구 PHR	대명사 NP	언어내 전이 INTRA_TRA
		의미 ME	대치 REP			문장 SEN	수사 NR	
		사용 US	회피 AVO			담화 TEX	동사 VV	
불가능 IMP	담화 DI	형태 FO	침묵 SIL			형용사 VA	학습의 장 CON-LEAR	
		의미 ME				보조 용언 VX		
	사용 US					지정사 VC		

여기에서 오류 분석 여부와 오류 층위의 표가 나타내는 칸은 그 의미가 문법·담화에 불가능이 걸쳐 있는 것이 아니다. 어떤 오류가 분석 가능하다면 그 층위가 어휘, 문법, 담화에서 다시 분석될 수 있다. 오류 분석이 불가능한 경우에도 그 층위가 어휘인지 문법인지 담화인지를 판단해서 기재할 수 있을 것이다. 불가능한 경우에는 3의 오류 양상부터는 태깅되기 어려울 것이다. 이상의 논의를 바탕으로 오류 정보 표지를 부착한 샘플을 보이면 아래와 같다.

32) 5)의 오류 범위(error extent)는 형태소 분석을 통한 문법 정보 표지를 참조할 수 있다. 김홍규·강범모(1997:13)에서는 45개의 문법 정보 표지를 제시하고 있다.

<오류 정보 표지 부착 샘플>

예) u. <p>나는 친구를 매원 음식 못 먹곤텐니다</p>

국적	작문 내용	오류 분석	1	2	3	4	5	6	7
1 USA	나는								
1 USA	친구를	POS	GRME	REP	LOC	WORD	JK	INTE_TRA	
1 USA	매원	POS	GRFO	MISF	LOC	PHO	ETM	INTR_TRA	
1 USA	음식								
1 USA	못								
1 USA	먹곤텐니다	POS	GRFO	MISF	LOC	WORD	EP	INTR-TRA	

6. 결론

학습자 말뭉치를 구축함에 있어서 학습자 개인 정보 문서 표지 체계와 오류 정보 표지 체계의 개발은 가장 기초적인 작업이다. 그럼에도 불구하고, 표준화된 체계를 구축하는 작업은 미처 이루어지지 못한 상황이다. 따라서 본고의 논의는 한국어 학습자 말뭉치를 구축하는 데 있어서 객관성과 체계성을 확보하기 위해 이론을 바탕으로 한 하나의 시안으로서 의의를 가질 수 있다. 물론 시안이기에 때문에 실제로 말뭉치를 구축하면서 보완되고 수정되어야 할 것이다. 이러한 지속적인 연구와 작업을 통해 학습자 말뭉치를 위한 기본적 문서 표지 체계가 완성되는 데 기여할 수 있어야 할 것이다. 또한 문서 표지 체계의 표준화가 완성이 된 이후에 가장 중요한 점은 여기에서는 논의되지 않은 각 오류 양상들에 대한 판단의 근거와 기준을 분명히 제시하는 것이다. 문서 표지 체계가 표준화된 이후의 작업은 오류 분석을 위한 근거가 객관적으로 제시되어야 말뭉치 활용의 핵심인 재사용성을 보장할 수 있기 때문이다.

<참고 문헌>

강화진. 2000. "KFL 학습자들의 작문 오류분석을 통한 효율적인 작문지도." 고려대 석사학위 논문.

- 김유미, 2000. “외국어로서의 한국어 학습자 말뭉치를 이용한 오류 분석.” 연세대 교육대학원 석사학위 논문.
- 김유정, 1997. “외국어로서의 한국어 문법 교육: 문법 교육의 위치·교육 원리에 관하여.” 『한국어학』, 6. 한국어학회.
- 김유정, 1998. “외국어로서의 한국어 문법 교육 - 문법 항목 선정과 단계화를 중심으로.” 『한국어 교육』, 9권 1호, 국제 한국어 교육 학회.
- 김유정, 1999. 『한국어 능력 평가 연구』, 고려대 박사학위 논문.
- 김정수, 2001. “한국어 정보화를 활용한 한국어 교육의 방향 및 과제.” 『21세기 국어 정보화와 국어 연구』, 민연총서 어문·민속 1. 월인.
- 김정수·남기춘, 2002. “영어권 한국어 학습자의 조사 사용 오류 분석과 교육 방법.” 『한국어 교육』, 제13권 1호, 27-46. 국제한국어교육학회.
- 김흥규·강범모, 2000. 『한국어 형태소 및 어휘 사용 빈도의 분석』, 컴퓨터와 인문학 4. 고려대학교 민족문화 연구원.
- 남윤진, 2000. “한국어 교육용 말뭉치의 구상과 실제.” 제1회 한국어교육 국제워크숍 발표 자료집, 43-47. 연세대 언어정보개발연구원.
- 박경자 외 공저, 1994. 『언어교수학』, 박영사.
- 박의재 외 공역, 1999. 『제2언어 습득론』, 한신문화사.
- 서상규 편, 1999. 『언어 정보의 탐구』, 1. 연세대학교 언어정보개발연구원, 월인.
- 서상규, 2001. “외국인을 위한 한국어 학습 사전과 학습자 말뭉치.” 『제2차 한국어세계화 국제 학술대회 <외국인을 위한 한국어 교재>』 발표논문집, 145-159. 한국어세계화추진위원회.
- 서상규·유현경·남윤진, 2002. “한국어 학습자 말뭉치와 한국어 교육.” 『한국어 교육』, 제13권 1호, 127-156. 국제한국어교육학회
- 서상규·한영균, 1999. 『국어 정보학 입문』, 대학사.
- 성지연, 2002. “오류 분석을 통한 한국어 관형사형 어미 사용 연구.” 고려대학교 석사학위 논문.
- 신성철 역, 1996. 『외국어 교수·학습의 원리』, 한신문화사.
- 유석훈, 2001. “외국어로서의 한국어 학습자 말뭉치 구축의 필요성과 자료 분석.” 『한국어 교육』, 제12권 1호, 165-180. 국제한국어교육학회
- 이혜란 외 공역, 1995. 『2개언어상용과 그 이론』, 한국문화사.
- 정광 외, 1995. 『한국어 데이터베이스의 설계 및 응용을 위한 기초 연구』, 민음사.
- Alice Ommaggio Hadley, 1993. *Teaching Language in Context*. 2nd edition. Heinle & Heinle Publishers.
- Brown, H. Douglas, 1980a. *Principles of Language Learning and Teaching*. 신성철(1996) 역, 『외국어 교수, 학습의 원리』, 한신문화사.
- Corder, S. Pit, 1971. Idiosyncratic dialects and error analysis. *Internation Review of Applied Linguistics*, 9, 147-159.
- Ellis, Rod, 1986. *Understanding Second Language Acquisition in Context*. New York:

Prentice Hall.

Ellis, Rod. 1998. *Second Language Acquisition*. Oxford University Press.

Larsen-Freeman, D. 1991. Teaching grammar. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (2nd ed., pp. 279-296). New York: Newbury House.

Lennon, Paul. 1991. Error: Some problems of definition, identification, and distinction. *Applied Linguistics* 12, 180-196.

Tony McEnery and Andrew Wilson. 1996. *Corpus Linguistics*. Edinburgh University Press.

김정숙(Kim Chung-sook)

서울시 성북구 안암동 5-1 고려대학교 국제어학원 한국어교육센터

kmjane@korea.ac.kr

김유정(Kim You-jeong)

서울시 성북구 안암동 5-1 고려대학교 국제어학원 한국어교육센터

youjeong@korea.ac.kr

접수일자 2002. 9. 23

게재결정 2002. 10. 24