

## 현대국어의 어미 및 문법 패턴의 계량언어학적 연구<sup>1)</sup>(上)

이 상 역  
(서울대학교 국어국문학과)

### 1. 서 론

#### 1.1 연구목적

본 연구는 현대국어 용언 형태론의 복잡한 교착(膠着 agglutination) 상태, 즉 선어말어미 및 연결어미, 종결어미 등의 각개 출현 및 연쇄 출현 빈도를 계량적으로 파악하고, 또 외국인이나 초학자들이 한국어를 배울 때 우선적으로 학습해야 할 문법 패턴 (또는 문법 형태)의 출현 순위를 조사하여, 학습자들이 직면하는 가장 큰 난점을 합리적으로 해결하는 방법을 마련하고, 기계번역 등에서의 절차 설정에도 통계적 바탕을 제공하려는 데에 목적을 두고 있다.

현재 세계 각처에서 한국어를 외국인에게 가르치고 있고 그 확장세가 뚜렷하지만 정작 그 교육 방법은 다른 주요 언어의 그것에 비해 아직 잘 개발되지 못한 부분이 많다. 심지어는 국내외의 초급 국어 교재에서 각 형태소의 도입 순서도 마련되어 있지 못한 것이 현실인 것이다. 본 연구의 결과는 국어사용의 현실적인 측면을 학습 장면에 도입하여 보다 실용적인 학습 장면을

---

1) 본 연구의 어미 부분의 일부원고는 1996년도 학술진흥재단의 국제협력연구과제 학술연구조성비 지원에 의해 시작한 것이었고, 또 어미 및 문법 패턴 부분의 연구는 2000년도 서울대학교 발전기금 일반 학술연구비 지원으로 이루어졌다. 이 두 연구의 모든 기초적 조사작업은 한신대 국문과 한재영 교수팀에 의해 수행되었음을 밝혀 감사한다.

구성할 수 있는 근거를 제공하게 될 것이다. 아울러 현재 사용되고 있는 교재에 대한 타당성 검토의 한 가지 기준으로 활용될 수도 있을 것이다.

본 연구의 주목표로는 다루지 않으려 하지만, 본 연구의 부차적 효과로서 최근 시도되고 있는 한영, 영한, 한일, 일한 언어간의 기계 번역에 대한 내용 확충과 효율성 증대를 위한 개선 노력에도 본 연구의 결과는 많은 도움을 제공하게 될 것이다. 번역의 질적 개선이나 외국인 한국어 교육 방법의 개선은 각기 다른 목표를 가지는 것처럼 보이나 유사한 방법으로 성취할 수 있는 내용이 많다. 두 분야는 또 연구의 필요성도, 현대 사회에서 국제적 교섭이 빈번해 지는 데 비례해서 더더욱 증대해 가고 있다.

이러한 현실에 비추어 볼 때, 본 연구가 행하려는 현대국어 선어말어미 및 연결어미, 종결어미 등의 각개 및 연쇄 출현 빈도를 계량적으로 파악하려는 작업은 다른 작업에 앞서 이루어졌어야 할 기초적인 연구라고도 할 수 있는 것이다. 따라서 본 연구에서는 종결어미의 연구와 함께 연결어미의 출현 빈도까지도 조사 집계해 볼 계획이다.

## 1.2 연구사

그 동안 이루어진 국어의 계량적 연구로는 과거 연차적 연구 계획에 의해 ‘참고문헌’에 보인 이상억(1989a, b, c, 1990a, b) 등을 들 수 있다. ‘국어 어휘 목록의 형태, 음운론적 구조 연구: 계량언어학적 표준조사’(1989a, b, c)와 ‘현대국어 음변화규칙의 기능 부담량’(1990 a) 및 ‘현대 국어의 통사적 현상에 대한 기능 부담량’(1993)이 그것으로, 그 배경에는 이상억(1986)에 소개된 것과 같이 독일 보쿰(Bochum)의 Altmann교수등이 이끄는 계량언어학(quantitative linguistics)이 자리잡고 있다. 계량언어학적 연구는 현재도 독일 트리어(Trier) 대학을 중심으로 국제학회가 구성되어 많은 연구업적을 쌓고 있으며, 이와 유사한 관점의 다른 연구 집단의 출현도 볼 수 있다.

한편 안병희(1965, 1967)는 이러한 연구들을 위한 국어의 형태론적 기술의 틀을 잘 보여 주고 있다. 그러나 종결어미 및 연결어미 출현 빈도의 조사 집계는 기이하게도 아직까지 조사가 본격적으로 수행 공개된 바가 없다. 1956

년에 실시된 문교부 주관 “우리말 찾기 조사”에서도 용언의 기본(비활용) 형태로만 조사해 놓았을 뿐 활용어미가 붙은 형태는 묵살해 버렸다. 근래에야 활용어미(종결, 연결어미 등)에 대한 조사가 관심을 끌어 사용 가능한 결과 보고(김홍규 등 1998)가 나와 부록에서와 같이 비교를 했으며, 김홍규·강범모(2000)<sup>2)</sup>는 이 원고를 탈고한 후에 나왔으나 전자와 대동소이하므로 여기에는 반영하지 않았다.

국어는 선어말 및 어말 어미의 교착/첨가로 피동 및 사동태, 시상, 준비법 및 겸양법 등을 다양하게 덧붙이는 언어로서, 굴절(inflexion)의 양상이 복잡해서 외국인이나 어린아이들이 한국어를 배워가는 초기 단계에서 큰 고비를 느끼게 한다. 가령 ‘깨-뜨리-시-었-겠-습니-다-만’ 같은 표현을 놓고 외국인이 분석하려면 어디에 형태소 경계가 있는지 쉽게 알아채기 어렵다. 보통 사전에는 각 형태소 별로 ‘깨-’ ‘뜨리-’ ‘-시-’ 등과 ‘깨뜨리다’ 한 덩어리만이 등재되어 있어서 만약 ‘깨뜨-’, ‘-리시-’ 등으로 잘못 경계를 끊었을 때는 시행착오를

2) 가장 최근에 나온 이 조사는 여러 가지 분류를 시도한 듯하나 사실은 종래에 나온 문교부(1956)와 직접 비교할 수 있도록 실질형태소만의 사용 빈도 목록이 작성되어 있지 않아 불편하다. 첨부된 CD로 작성할 수는 있겠지만, 후속 발간물이라도 김홍규·강범모(2000)를 보완하여 (문법형태소를 뺀) 실질형태소만의 빈도순 목록이 공간되었으면 한다. 그러나 김홍규 등(1998)에 보인 어미만의 빈도순 목록과 직접 비교할 수 있는 (조사와 접사를 뺀) 어미만의 목록은 게재되어 있다.

① 김홍규 등(1998): 각 항목의 전 어미 사용 중 백분율 계산은 첨부되어 있지 않음.

1. ㄴ 2. 다 3. 는 4. 고 5. 었 6. 았 7. 아 8. 어 9. 르 10. 게 11. 은 12. ㄴ다

② 김홍규·강범모(2000): 각 항목의 백분율 계산이 되어 있어 아래에 보임.

1. ㄴ 2. 다 3. 는 4. 고 5. 었 6. 어 7. 았 8. 아 9. 르 10. 은 11. 게 12. ㄴ다  
%12.6 11.3 7.8 7.2 5.7 5.1 4.7 4.6 3.9 2.7 2.7 2.3

위에 보인 짧은 대비에서 간취되듯이 각 항목의 순위는 부분적으로 순위바꿈이 있지만 국부적 선후관계의 차이일 뿐 큰 변동은 아니다. 또 각 항목의 전 어미 사용 중 백분율도 대략 계산해 보면 양자가 거의 같은 추세로 점유하고 있는 사실을 알 수 있다.

결론적으로 본 연구는 김홍규·강범모(2000)가 나오기 전에 김홍규 등(1998)에 의해 이미 다 이루어졌지만, 양자의 조사 결과가 대동소이하므로 [부록1]에 보였듯이 김홍규 등(1998)과의 대비만으로도 큰 추세를 살펴보기에는 별 손색이 없다.

겪어 많은 경험을 쌓기 전에는 그 뜻을 찾아보기 어려운 형편이다.<sup>3)</sup>

이런 난점은 ‘깨뜨-’, ‘-리시-’ 같이 잘못 자른 형태까지 사전에 등재하는 편법으로 해결할 수 있을지는 모르겠으나, 그런 연쇄형태란 그 수가 엄청나게 많을 뿐만 아니라 잘못 분석된 형태를 자꾸 보이게 된다는 점에서도 바람직한 조처는 아니라 하겠다. 그보다는 초기에 잘못 자르는 시기가 좀 있더라도 학습자가 각 형태소의 경계를 스스로 알아챌 수 있도록 점차 훈련시켜 가는 것이 문제에 대한 온당한 접근방법이라고 할 수 있다. 이를 위해서는 어떤 연쇄 형태가 가장 많이 쓰이는가를 통계적으로 조사해서, 학습과정에서 적절히 순서대로 노출시켜 주는 것이 필요하다. 예컨대 ‘-었겠-’은 가능한 연쇄 형태로서 실제 어떤 빈도로 우리가 쓰고 있는지 조사해서, 만약 그것이 자주 쓰이는 것이면 학습 초기에 도입시켜야 할 것이다. 이렇게 모든 내용을 계량적으로 파악하면 모든 형태의 적절한 도입 순서를 설정할 수 있다.

## 2. 본 론

### 2.1 연구방법

본 연구의 연구방법은 두 가지 측면에서 이야기할 수 있다. 그 하나는 대상자료에 관한 것이고, 다른 하나는 자료의 검색방법에 관한 것이다. 먼저 본 연구의 대상이 현대국어라는 점과 개별 형태의 빈도에 관심을 가지고 있다는 점에서 본다면, 현대국어의 전형적인 모습을 담고 있는 모든 자료가 검토의 대상이 되어야 할 것이나 그와 같은 작업은 불가능한 것일 뿐만 아니라 비경제적인 것이기도 한 것이다. 더구나 우리의 목적 달성에는 적절한 양의 전형적인 자료를 통한 표본 검사로 족하다 할 수 있다. 검토 대상 자료의 양은 많으면 많을수록 좋은 것이고, 앞으로 연구가 진행되어 감에 따라 검토 대상

3) 이와 더불어 외국인에게 어려운 부분은 구어투에서 대폭, 문어투에도 소폭으로 나오는 축약형에 대한 분석이다. 가령 ‘종이 치질 않는다’ 같은 문장에서 어미 ‘-지’와 결합된 조사 ‘-를’의 준말 ‘-질’을 단번에 분석해 내기가 어려울 듯하다. 사전에서 바로 ‘치질’을 그대로 찾는 초심자는 더욱 혼란스러운 해석을 하게 될 것이다.

자료의 양을 확대하여 결과에 대한 수정, 보완 및 정밀화 과정이 뒤따라야 할 것이나, 본 연구에서는 1차적인 대상 자료의 양을 약 100,000 어절 정도로 잡았다. 물론 작업의 편의 또는 수행 가능성이라는 기준 이외에 100,000이라는 수가 가지는 객관적인 근거를 제시하기는 어렵다.<sup>4)</sup> 그러나 구성 자료의 성격이 될 수 있으면 어느 한 방면에 치우쳐서는 안된다는 점과 현실성이 있는 자료로 구성이 되어야 한다는 점 그리고 그들의 내용을 어느 정도는 반영할 수 있는 하나의 단위가 될 수 있는 자료이어야 한다는 점을 자료 선택의 큰 기준으로 삼았다는 데에서 타당성을 부여할 수 있을 것이다.

<표 1> 대상 자료의 내용

자 료 명	어절수	전체 어절수에 대한 비율
신문	31,759	30.25%
성경(창세기)	24,793	23.62%
TV뉴스	4,998	4.76%
소설(깃발 없는 기수)	22,727	21.65%
드라마(짱)	5,393	5.14%
PD수첩	3,331	3.17%
시(생명의 서)	3,696	3.52%
희곡(당랑의 전설)	8,279	7.89%

위의 <표 1>에 보인 자료의 전체 어절수는 104,976개이다. 신문은 1996년 8월 17일자 ‘조선일보’와 ‘중앙일보’ 및 ‘한국일보’에서 취한 것이고, 성경은 1977년의 공동번역 가운데 ‘창세기’를 가져온 것이다. TV 뉴스는 1997년 4월 22일에 방송된 ‘MBC 뉴스 데스크’의 것이다. 소설은 선우 회의 ‘깃발 없는 기수’이며, 드라마는 1997년 5월 11일에 방송된 드라마 ‘짱 : 제129회 라이벌 죽

4) 근래에 연세대, 고려대, KAIST, 국립국어연구원, 21세기 세종계획 등의 대형 corpus 들이 구축됨에 따라 10만 어절은 상대적으로 작은 숫자이지만, 선도적 개인 연구로서 즉 대형 작업에서 아직 보고된 바 없는 연구 분야에 대한 pilot study를 통해 추세/경향을 우선 짚어 보기 위한 크기로서는 별 하자가 없는 것이다.

이기'를 취한 것이다. PD수첩은 1997년 5월에 방송된 '굶주린 북녘 두만강 접경지대를 가다'의 내용이다. 시는 1947년 행문사에서 나온 유치환의 '생명의 서'이며, 희곡은 인문평론 1940년 10월호에 실린 '螻蛄의 傳說'이다. 시사성이 있는 자료들은 현대국어의 사용어휘의 양상을 잘 반영하고, 문학작품들은 이해어휘의 내용을 보여준다. 뉴스와 드라마, PD수첩 및 희곡은 구어적인 모습을 반영하는 자료로 약 21%를 차지한다. 쉽게 생각하기로는 구어와 문어의 비율이 각각 50%인 것이 이상적인 것으로 생각할 수도 있겠으나, 구어자료의 경우에는 자료 처리상의 어려운 점 이외에 자료의 질이 덜 정제된 모습을 보인다는 점에서<sup>5)</sup> 20% 이상을 많이 넘는 것은 오히려 피해야 할 것으로 판단한 것이다. 시를 대상 자료에 포함시킨 것은 일상 생활에서는 잘 쓰이지 않으나 문학 작품 등에서 여전히 생명력을 가지고 있는 형태들을 놓치지 않으려는 배려에서이다. 1940년대의 희곡을 선택한 데에 대해서도 또 다른 이견이 있을 수 있겠으나 이는 현대국어의 사용자가 현재의 젊은이들에 국한되는 것이 아니라는 점을 기억하려 한 조처이다. 보통 1900년경부터 시작하는 현대국어의 시기를 감안하여 그 사용자들의 연령을 광범위하게 고려한 것이다. PD수첩의 경우는 여러 가지 내용 가운데 특히 '굶주린 북녘 두만강 접경지대를 가다'를 취한 것은 대화 가운데 북한의 언어 모습을 살필 수 있다는 점을 취한 것이다.<sup>6)</sup> 이러한 배려들이 결과에 반영되었는지의 여부는 오히려 2차적인 것이라 할 수 있다. 대상 자료를 선택하는 과정에서 특히 적은 량의 자료를 통하여 전형적인 모습에 다가서기 위한 작업에서는, 개별 자료가 가지는 다양한 성격을 염두에 두는 것이 온당한 태도라 할 것이다. 본 원고는 '인문논총'지에 상/하로 나누어 게재할 예정이며, '하'의 [부록 3]에 본 연구의 대상이 된 위의 기초 자료 일부를 소개하기로 한다.

- 
- 5) 자료의 정제성이라는 점에서 본다면 방송 자료는 그래도 나은 편이다. 일상적인 구어를 대상으로 삼을 경우에는 정제의 정도가 이야기로 삼기조차 어렵다 할 것이다.
- 6) 북한 언어를 개입시킨 것은 현대국어의 폭을 넓게 잡아 본 결과이나 이질적인 것은 사실이다. 그러나 이것이 전체에 대해 차지하는 비율은 1% 미만의 미미한 양이므로 어차피 조사 결과에 이질화가 크게 일어난 수준은 아니다.

이상의 자료를 통하여 개별 문법 형태의 사용 양상을 살피는 가장 정확한 방법은 하나하나 눈으로 확인하고 카드 작업화하여 정리하는 것이라 할 수 있다. 그러나 그와 같은 자료 처리 방식이 가지는 한계는 자명하다. 많은 양의 자료를 감당할 수 없다는 것이 그것이다. 그렇다고 하여 기계적인 방식으로 일률적으로 적용할 수도 없는 것이 현실이다. 아직 우리의 ‘형태소 분석기’는 문맥에 따른 다양한 가능성은 제시할 수 있되 각각의 문맥에서의 용법 판단에는 이르지 못하고 있기 때문이다. 따라서 본 연구에서는 두 가지 방식을 병용하기로 하였다. 그를 위하여 기존의 업적을 통한 문법 형태소 목록의 확인 작업으로부터 출발하였다. 다음의 <표 2>와 <표 3>이 그것이다.

<표 2>

■ 선어말어미 (23개)

거	겠	ㄴ(다)	나이(다)	느(ㄴ/냐/ㄴ지)	는(다)	더
리	사오	사옵	삼	시	았	었
으시	으오	으옵	자오	자옵	잡	오
						옵

<표 3>

■ 어말어미 (650개)

거나 --- 질 (뒤의 <표 7>에서 상세한 목록을 보임.)

위의 <표 2>와 <표 3>은 기존의 문법서들을 통하여 정리한 선어말어미와 어말어미의 목록이다. 어말어미에는 종결어미와 연결어미가 포함되어 있으며, 각각의 형태들에 대한 이형태 및 간혹 선어말어미나 조사와의 결합정도 망라 되어 있다.<sup>8)</sup>

7) 본 목록 중에 -나이-와 -로-는 선어말어미에 넣지 않는 견해도 있다. 그래서 이들은 어말어미 목록에 -나이까, -나이다, -로고, -로구나, -로구려, -로구료, -로구면, -로구면, -로군, -로다, -로되, -로라, -로세 등의 형태로 다시 나열하였다. 선어말어미 -느-, -는-, -더- 등이 앞에 복합되어 있는 형태(-느냐, -는가, -더구나 등)들과 함께 어미목록 속에 나열되어 있도록 조치한 것이다.

문법 패턴 또는 문법 형태에 대한 연구에서는 기존의 한국어 교육 현장에서 사용되고 있는 교재들을 통하여 일차적인 목록을 확보하였다. 그 과정에서 각 교재에 반영된 도입의 순서 즉 등급을 옮겨 써서 기억하기로 하였다. 이들 등급과 본 연구의 결과물이 될 빈도와의 대조는 해당 문법 형태의 도입 단계에 대한 타당성 여부를 보일 것으로 기대한다. 물론 빈도수 자체가 교육 과정에서의 도입 순서를 결정 짓는 유일한 기준이 되는 것은 아니다. 빈도가 낮다고 하더라도 해당 장면에서 필요한 문법 형태라면 당겨서 도입이 되어야 할 것이기 때문이다. 그러나 실질 의미를 가지는 어휘 빈도와는 달리 문법 형태의 빈도수는 교육 단계의 순서를 정하는 데에 중요한 기준이 된다고 할 수 있다. 실질 어휘와는 달리 문법 형태의 경우에는 발화 장면에 따른 영향이 상대적으로 덜하다고 할 수 있기 때문이다. 이는 어휘의 대체 가능성보다 문장의 대체 가능성이 상대적으로 다양하다는 데에 근거한다.

다음의 <표 4>는 기존의 한국어 교재에서 취한 문법 형태의 목록이다. 여기서는 기존 한국어 교재에서 다루고 있는 단위를 그대로 수용하기로 하였다. 복합 형태에 대해서도 별도의 형태 분석 절차를 취하지 않았음을 뜻하며, 아울러 그 교재의 급수까지 기록하였다.

<표 4> - 조사대상항목 (여기서는 각 급수의 첫 예들만 보이고, 본고 '하'의 [부록 2]에서 아래 [중략] 부분의 전 목록을 보일 예정임.)

일련번호	급수	문법형태
0001	1	(내가) 보기에는 [중략]
0180	2	-(느)냐고 (하다) [중략]
0227	3	(것) 밖에(는) [중략]
0281	4	-(기가) 예사이다 [중략]
0334	4	쭈쳐럽

- 8) 실제 분석 과정에서 이들 목록에 몇몇 형태가 보태어지고, 보태어질 형태 목록 가운데 몇몇은 작업의 편의를 위하여 목록 추가를 보류하고 관련 목록에서 함께 처리하기도 하였다. 물론 함께 처리하는 것이 나으리라는 판단이 가능한 경우에 한한다.



문장의 구조와 관계가 있는 일부 문장부사도 포함이 되어 있는, 위의 <표 4>는 기존의 한국어 교재들을 통하여 정리한 문법 형태의 목록이다. <표 4>를 작성하기 위하여 살핀 한국어 교재의 목록은 다음의 <표 5>과 같다. 상당수 중복되는 경우가 많아 해당 문법 항목의 출전을 일일이 밝히는 번거로움은 피하기로 한다.

<표 5>

Chang, Suk-in, 1982, *Modern Conversational Korean*, Seoul: Seoul Computer Press.

Park, Francis Y.T., 1984, *Speaking Korean 1·2*, Seoul: Hollym Co.

한국어 문화 연수부 편, 1989, *한국어 1-4*, 서울: 고려대학교 민족문화연구소.

재외국민교육원, 1990, *한국어 1-3: 제일 한국중학교용*, 서울: 서울대학교 재외국민교육원.

Vandesande A.V. and Yunsook Hong, 1991, *Myöngdo's Korean 1·2*, 서울: 우신사.

Park, Chang-Hai and Ki-Dawk Pak, 1992, *Korean 1·2: An Intensive Course*, 서울: 연세 대학교 출판부.

Rogers, Michael C., Clare You and Kyungnyun K. Richards, 1992, *College Korean*, Berkeley: Univ. of California Press.

서울대학교 어학연구소 편, 1993, *한국어 1-2*, 1997, *한국어 3-6*, 서울: 풍남.

서울대학교 이상억 편, 1993, *Korean through English 한국어 I·II·III*, 서울: 한림출판사, [영문판은 원제 '한국어'로 문화체육부 1992년판이 있었고, 한림출판사 1998년 개정판 에 추가항목 삽입 보완함]. [중, 일, 노, 서, 불, 독어판이 92-8년 사이에 간행됨].

Chang, Namgui and Yong-chol Kim, 1993, *Functional Korean*, Seoul: Hollym Co. 연세대학교 한국어학당 편, 1995, *한국어 1-6*, 서울: 연세대학교 출판부.

Chang, Namgui and Yong-chol Kim, 1996, *Active Korean: A Functional Approach*, Seoul: Hollym Co.

Beckers-Kim, Young Ja, 1996, *Koreanisch für Anfänger*, Seoul: Hollym Corp.

Klear-KF, 1997, Korean for English Speakers: Level I (Part A&B), Honolulu:  
Klear-KF Textbooks in Korean Language.

이들 목록 가운데에서 일차적인 검토 대상 목록을 선정하였다. 일차적인 검토 작업은 1992년 한글과 컴퓨터에서 만든 ‘문자열 검색 · 사전 검색 프로그램(HGREP)’을 가지고 행하기로 하였다. 따라서 일차적인 검토 대상 목록에서 ‘ㄴ, ㄹ, ㅁ’과 그를 포함한 형태들은 배제된다. ‘문자열 검색 · 사전 검색 프로그램(HGREP)’이 문자열에 대한 검색 프로그램이어서 음절말의 자음 ‘ㄴ, ㄹ, ㅁ’ 등을 분석하지는 못하기 때문이다.

그에 대한 보완을 위하여 취한 것은 강승식 교수(한성대)가 만든 ‘HAM(한국어 형태소분석기)’이다. 전체 대상 자료를 HAM으로 돌려 얻은 결과 가운데 ‘ㄴ, ㄹ, ㅁ’을 가지고 있는 분석 내용을 취하고 그들에 대하여 하나하나 판별하는 과정을 거친 것이다. 처음부터 HAM을 사용하지 않고 이차적인 도구로 사용한 것은 HAM이 모든 분석 가능성을 제시한다는 점에서 많은 노동량을 요구하기 때문이었다. 이와 같은 과정을 통하여 얻어진 결과는 다음 절에서 살펴하기로 한다.

## 2.2 분석의 결과

### 2.2.1 선어말어미

<표 6>

번 호	형태소목록	출현빈도	빈도순
PFE001	거	163	009
PFE002	겠	453	006
PFE003	ㄴ(다)	454	005
PFE004	나이(다)	0	014
PFE005	느(ㄴ/냐/ㄴ 지)	127	011
PFE006	는(다)	557	004

번호	형태소목록	출현빈도	빈도순
PFE007	더	196	008
PFE008	로(구나/구려)*	4	013
PFE009	리	147	010
PFE010	사오	0	014
PFE011	사읍	0	014
PFE012	샵	0	014
PFE013	시	352	007
PFE014	았	1300	003
PFE015	엇	2454	001
PFE016	였	1498	002
PFE017	오	0	014
PFE018	읍	0	014
PFE019	으시	34	012
PFE020	으오	0	014
PFE021	으읍	0	014
PFE022	자오	0	014
PFE023	자읍	0	014
PFE024	잡	0	014

\* '로(구나/구려)'는 선어말어미에서 제외할 수도 있음.

위에 보인 <표 6>는 선어말 어미의 출현 빈도와 빈도순위를 보인 것이다. 과거시제형태소인 '엇'의 출현빈도가 압도적임을 알 수 있다. 또 다른 과거시제형태소인 '였'과 '았'의 빈도를 합할 경우 선어말어미 가운데에서 과거시제 형태가 차지하는 비율은 68% 가량이다.<sup>9)</sup> 선어말어미에 대한 접근의 출발점에 과거시제가 있어야 한다는 점을 강력하게 시사하는 것이라고 이해해도 좋을 것이다. 이와 같은 결과에 대하여 우리가 취한 자료의 성격 때문이라는 지적

9) 김홍규·강범모(2000)에서도 었(50%), 았(42%) 2 가지만 합해서 92%의 높은 점유율을 보이고 있다. 이 누적백분율 속에는 아마 '-였-'이 (따로 설정된 흔적이 없으므로) 포함되어 있으리라 짐작된다.

도 가능하다. 일종의 역사자료라 할 수 있는 성경과 지난 일들에 관한 진술을 내용으로 하는 신문과 TV 뉴스 등의 비중이 크다는 점에서 그와 같은 지적은 어느 정도는 타당한 면도 가지고 있다고 할 수 있다. 하지만 다른 측면에서 본다면, 우리 언어 생활 속에서의 과거에 대한 진술이 차지하는 비중이 그만한 것이라는 이해도 가능하다고 할 것이다.

2.2.2 어말어미				043	FE022	고는	: 118
				135	FE023	고도	: 14
				346	FE024	고말고	: 0
<표 7>				066	FE025	고서	: 61
빈도	고유	어말	출현	296	FE026	고야	: 1
순위	번호	어미	빈도	235	FE027	고자	: 3
-----				143	FE028	곤	: 13
074	FE001	거나	: 54	346	FE029	관테	: 0
296	FE002	거늘	: 1	062	FE030	구나	: 70
235	FE003	거니	: 3	258	FE031	구려	: 2
258	FE004	거니와	: 2	296	FE032	구료	: 1
346	FE005	거드면	: 0	235	FE033	구면	: 3
346	FE006	거드면은	: 0	296	FE034	구면	: 1
073	FE007	거든	: 55	044	FE035	군	: 117
258	FE008	거들랑	: 2	015	FE036	기	: 531
346	FE009	거들랑은	: 0	085	FE037	기로	: 45
094	FE010	거라	: 38	346	FE038	기로서	: 0
143	FE011	건	: 13	346	FE039	기로서니	: 0
258	FE012	건대	: 2	346	FE040	기로선들	: 0
346	FE013	건넌	: 0	087	FE041	기에	: 41
346	FE014	건마는	: 0	002	FE042	ㄴ	: 5790
214	FE015	건만	: 4	053	FE043	ㄴ가	: 91
346	FE016	것다	: 0	095	FE044	ㄴ결	: 37
006	FE017	계	: 1288	113	FE045	ㄴ고	: 24
		(연결, 종결 구분없이 합산)		018	FE046	ㄴ다	: 454
346	FE018	계꿈	: 0	346	FE047	ㄴ다고	: 0
171	FE019	계나	: 8	346	FE048	ㄴ다느냐	: 0
346	FE020	게서리	: 0	055	FE049	ㄴ다는	: 85
003	FE021	고	: 3843	296	FE050	ㄴ다니까	: 1

346	FE051	ㄴ다마는	: 0	148	FE079	너라	: 12
346	FE052	ㄴ단다	: 0	092	FE080	네	: 39
296	FE053	ㄴ달	: 1	148	FE081	노니	: 12
346	FE054	ㄴ담	: 0	346	FE082	노다	: 0
235	FE651	ㄴ답니까	: 3	182	FE083	노라	: 7
346	FE055	ㄴ답니다	: 0	296	FE084	노라고	: 1
346	FE056	ㄴ답디다	: 0	346	FE085	노라니	: 0
346	FE057	ㄴ답시고	: 0	296	FE086	노라면	: 1
235	FE058	ㄴ대	: 3	258	FE659	누나	: 2
296	FE059	ㄴ대서	: 1	058	FE087	느냐	: 74
346	FE060	ㄴ대서야	: 0	135	FE088	느냐고	: 14
346	FE061	ㄴ대야	: 0	171	FE089	느뇨	: 8
036	FE062	ㄴ데	: 177	153	FE090	느니	: 11
153	FE652	ㄴ데다	: 11	346	FE091	느니라	: 0
214	FE653	ㄴ데도	: 4	182	FE092	느라고	: 7
296	FE654	ㄴ데서	: 1	004	FE093	는	: 3056
346	FE063	ㄴ들	: 0	116	FE094	는가	: 21
346	FE064	ㄴ만큼	: 0	083	FE095	는걸	: 47
182	FE065	ㄴ마	: 7	258	FE096	는고	: 2
296	FE066	ㄴ죽	: 1	171	FE097	는구나	: 8
346	FE067	ㄴ죽슨	: 0	128	FE098	는구려	: 17
059	FE068	ㄴ지	: 72	346	FE099	는구료	: 0
346	FE069	ㄴ지고	: 0	346	FE100	는구먼	: 0
346	FE070	ㄴ지라	: 0	346	FE101	는구먼	: 0
035	FE071	나	: 180	128	FE102	는군	: 17
		(연결, 종결 구분 없이 합산)		059	FE103	는다	: 72
103	FE072	나니	: 29	165	FE104	는다고	: 10
258	FE073	나마	: 2	135	FE105	는다는	: 14
346	FE074	나이까	: 0	346	FE106	는다마는	: 0
346	FE075	나이다	: 0	148	FE660	는단	: 12
040	FE076	냐	: 156	346	FE107	는단다	: 0
258	FE658	냐가	: 2	296	FE108	는달	: 1
099	FE077	냐고	: 32	346	FE109	는담	: 0
143	FE078	냐는	: 13	346	FE110	는답시고	: 0
235	FE656	냐며	: 3	346	FE111	는대	: 0
296	FE657	냐면서	: 1	346	FE112	는대서	: 0

346	FE113	논대서야	: 0	116	FE664	대	: 21
346	FE114	논대야	: 0	258	FE145	더구나	: 2
026	FE115	논데	: 312	346	FE146	더구려	: 0
296	FE661	논도다	: 1	346	FE147	더구먼	: 0
296	FE116	논마	: 1	110	FE148	더군	: 26
059	FE117	논지	: 72	153	FE149	더냐	: 11
346	FE118	논지고	: 0	051	FE150	더니	: 103
346	FE119	논지라	: 0	346	FE151	더니라	: 0
030	FE120	니	: 239	346	FE152	더니마는	: 0
057	FE121	니까	: 76	346	FE153	더니만	: 0
346	FE122	니까는	: 0	126	FE154	더라	: 18
118	FE123	니깐	: 20	143	FE155	더라도	: 13
346	FE124	니라	: 0	214	FE156	더라면	: 4
346	FE125	니만치	: 0	346	FE157	더면	: 0
346	FE126	니만큼	: 0	020	FE158	던	: 449
001	FE127	다	: 6039	153	FE159	던가	: 11
027	FE128	다가	: 284	214	FE160	던걸	: 4
192	FE129	다가는	: 6	346	FE161	던고	: 0
182	FE130	다간	: 7	153	FE162	던테	: 11
021	FE131	다고	: 440	346	FE163	던틀	: 0
258	FE662	다네	: 2	258	FE164	던지	: 2
258	FE132	다느니	: 2	346	FE165	데	: 0
028	FE133	다는	: 271	031	FE166	도	: 223
114	FE134	다니	: 22	044	FE167	도록	: 117
214	FE135	다마는	: 4	118	FE168	되	: 20
346	FE136	다마다	: 0	105	FE169	든	: 28
204	FE137	다만	: 5	296	FE666	든가	: 1
080	FE663	다며	: 49	069	FE170	든지	: 58
064	FE138	다면	: 66	039	FE171	듯	: 165
153	FE139	다면서	: 11	084	FE172	듯이	: 46
153	FE140	다시피	: 11	171	FE173	디	: 8
204	FE141	다오	: 5	112	FE174	리	: 25
235	FE142	단다	: 3	258	FE175	리거나	: 2
346	FE143	답	: 0	182	FE176	리걸	: 7
235	FE665	답니다	: 3	346	FE177	리께	: 0
346	FE144	답시고	: 0	346	FE178	리카	: 0

346	FE179	르까마는	: 0	346	FE213	르지나	: 0
346	FE180	르까말까	: 0	346	FE214	르지니	: 0
346	FE181	르까보냐	: 0	346	FE215	르지니라	: 0
346	FE182	르까보다	: 0	346	FE216	르지라	: 0
153	FE183	르께	: 11	346	FE217	르지라도	: 0
258	FE184	르꼬	: 2	346	FE218	르지어다	: 0
346	FE185	르는지	: 0	346	FE219	르지언정	: 0
346	FE186	르라	: 0	346	FE220	르진대	: 0
097	FE187	르라고	: 36	346	FE221	르진댄	: 0
346	FE188	르라치면	: 0	296	FE222	르진저	: 1
346	FE189	르래	: 0	029	FE223	라	: 256
128	FE669	르래	: 17	050	FE224	라고	: 105
346	FE190	르러니	: 0	346	FE225	라느니	: 0
346	FE191	르러라	: 0	047	FE226	라느	: 114
346	FE192	르런가	: 0	168	FE227	라니	: 9
346	FE193	르런고	: 0	182	FE228	라니까	: 7
346	FE194	르레	: 0	062	FE230	라도	: 70
346	FE195	르레라	: 0	258	FE229	라든지	: 2
346	FE196	르말로는	: 0	110	FE231	라며	: 26
346	FE197	르말로야	: 0	100	FE232	라면	: 31
258	FE198	르망정	: 2	204	FE233	라면서	: 5
258	FE199	르바	: 2	346	FE234	라서	: 0
346	FE199-1	르바에	: 0	296	FE235	라야	: 1
346	FE200	르바에야	: 0	346	FE236	라야만	: 0
214	FE201	르밖에	: 4	204	FE237	라오	: 5
296	FE202	르뿐더러	: 1	204	FE238	락	: 5
346	FE203	르새	: 0	171	FE239	란다	: 8
168	FE204	르세	: 9	346	FE240	랍	: 0
235	FE205	르세라	: 3	296	FE241	랍니다	: 1
346	FE206	르세말이지	: 0	346	FE242	랍시고	: 0
192	FE207	르수룩	: 6	171	FE668	래	: 8
346	FE208	르시	: 0	098	FE670	래도	: 35
346	FE209	르쏘냐	: 0	258	FE671	래야	: 2
346	FE210	르이만큼	: 0	148	FE243	라	: 12
346	FE211	르작성면	: 0	346	FE244	라마는	: 0
135	FE212	르지	: 14	090	FE245	러	: 40

346	FE246	러니	: 0	204	FE281	려오	: 5
346	FE247	러나라	: 0	346	FE282	련	: 0
346	FE248	러니이까	: 0	346	FE283	련다	: 0
346	FE249	러니이다	: 0	346	FE284	련마는	: 0
346	FE250	러이까	: 0	235	FE285	련만	: 3
346	FE251	러이다	: 0	346	FE286	럴	: 0
346	FE252	런가	: 0	258	FE287	럼	: 2
346	FE253	런들	: 0	235	FE288	럼니까	: 3
105	FE254	려	: 28	346	FE289	럼니다	: 0
153	FE255	려거든	: 11	346	FE290	렷다	: 0
077	FE257	려고	: 52	346	FE291	로고	: 0
346	FE256	려기에	: 0	346	FE292	로구나	: 0
258	FE258	려나	: 2	346	FE293	로구려	: 0
346	FE259	려네	: 0	346	FE294	로구료	: 0
346	FE260	려느냐	: 0	296	FE673	로구만	: 1
087	FE261	려는	: 41	346	FE295	로구먼	: 0
346	FE262	려는가	: 0	346	FE296	로구먼	: 0
296	FE263	려는고	: 1	165	FE297	로군	: 10
182	FE264	려는데	: 7	296	FE298	로다	: 1
235	FE265	려는지	: 3	346	FE299	로되	: 0
192	FE266	려니	: 6	258	FE300	로라	: 2
258	FE267	려니와	: 2	346	FE301	로세	: 0
235	FE268	려다	: 3	346	FE302	리	: 0
192	FE269	려다가	: 6	346	FE303	리까	: 0
346	FE270	려더니	: 0	258	FE304	리니	: 2
346	FE271	려더라	: 0	346	FE305	리나라	: 0
346	FE272	려던가	: 0	204	FE306	리다	: 5
346	FE273	려도	: 0	042	FE307	리라	: 134
346	FE274	려머는	: 0	346	FE308	리로다	: 0
192	FE672	려면	: 6	346	FE309	리만치	: 0
192	FE275	려면	: 6	346	FE310	리만큼	: 0
346	FE276	려면야	: 0	346	FE311	리요	: 0
214	FE277	려무나	: 4	346	FE312	리	: 0
346	FE278	려서는	: 0	346	FE313	리세	: 0
346	FE279	려서야	: 0	346	FE314	리에도	: 0
346	FE280	려야	: 0	346	FE315	리예라	: 0



296	FE316	ㅁ 직하다	: 1	346	FE348	습디까	: 0
171	FE317	마	: 8	346	FE349	습디다	: 0
192	FE674	마는	: 6	346	FE350	습딘다	: 0
296	FE318	매	: 1	346	FE351	시거늘	: 0
016	FE319	며	: 480	346	FE352	시거니와	: 0
014	FE320	면	: 562	235	FE353	시거든	: 3
025	FE321	면서	: 357	346	FE354	시어요	: 0
235	FE675	면야	: 3	192	FE695	시요	: 6
165	FE322	면은	: 10	143	FE694	시우	: 13
214	FE323	므로	: 4	008	FE355	아	: 1070
346	FE324	ㅁ 닐다	: 0	258	FE356	아다	: 2
071	FE325	ㅁ 니까	: 57	258	FE357	아다가	: 2
009	FE326	ㅁ 니다	: 870	168	FE358	아도	: 9
346	FE327	ㅁ 디까	: 0	070	FE359	아라	: 58
346	FE328	ㅁ 디다	: 0	018	FE360	아서	: 454
346	FE329	ㅁ 딪다	: 0	102	FE361	아야	: 30
346	FE330	ㅁ 쇼	: 0	296	FE362	아야만	: 1
346	FE331	ㅁ 시다	: 0	214	FE363	아야지	: 4
192	FE332	ㅁ 시오	: 6	086	FE364	아요	: 43
214	FE655	보담	: 4	346	FE365	아지이다	: 0
346	FE333	사오이다	: 0	125	FE366	왔습니다	: 19
346	FE334	사옵니까	: 0	013	FE367	야	: 656
346	FE335	사옵니다	: 0	214	FE368	야만	: 4
346	FE336	사옵디까	: 0	346	FE369	야요	: 0
346	FE337	사옵디다	: 0	087	FE370	야지	: 41
346	FE338	사외다	: 0	005	FE371	어	: 2341
153	FE339	세	: 11	296	FE372	어다	: 1
052	FE340	세요	: 99	214	FE373	어다가	: 4
068	FE341	소	: 59	077	FE374	어도	: 52
346	FE342	소서	: 0	064	FE375	어라	: 66
346	FE343	소이다	: 0	012	FE376	어서	: 674
346	FE344	쇠다	: 0	075	FE377	어야	: 53
296	FE680	쇼	: 1	346	FE378	어야만	: 0
346	FE345	습닐다	: 0	038	FE379	어요	: 170
092	FE346	습니까	: 39	346	FE380	어지이다	: 0
017	FE347	습니다	: 475	346	FE381	언마는	: 0

258	FE382	언반	: 2	135	FE417	으니깐	: 14
346	FE383	언정	: 0	346	FE418	으니라	: 0
346	FE384	엇다	: 0	346	FE419	으니만치	: 0
346	FE678	에라	: 0	346	FE420	으니만큼	: 0
090	FE385	에요	: 40	296	FE421	으되	: 1
037	FE386	여	: 174	171	FE422	으라	: 8
346	FE387	여도	: 0	182	FE423	으라고	: 7
103	FE388	여라	: 29	346	FE424	으라느니	: 0
126	FE389	여서	: 18	346	FE425	으라니	: 0
258	FE390	여야	: 2	346	FE427	으라니까	: 0
346	FE391	여야만	: 0	346	FE426	으라며	: 0
235	FE392	여요	: 3	296	FE428	으라먼	: 1
135	FE393	엿습니다	: 14	346	FE429	으라먼서	: 0
346	FE394	엿습니다	: 0	296	FE430	으락	: 1
041	FE395	오	: 154	346	FE431	으람	: 0
346	FE396	오리까	: 0	296	FE679	으래	: 1
346	FE397	오리다	: 0	192	FE432	으라	: 6
346	FE398	오이다	: 0	346	FE433	으라마는	: 0
346	FE399	올습니다	: 0	132	FE434	으러	: 15
296	FE400	올시다	: 1	171	FE435	으려	: 8
346	FE401	옵나이까	: 0	258	FE436	으려거든	: 2
346	FE402	옵나이다	: 0	214	FE438	으려고	: 4
346	FE403	옵니까	: 0	346	FE437	으러기에	: 0
346	FE404	옵니다	: 0	346	FE439	으러나	: 0
346	FE405	옵디까	: 0	346	FE440	으러네	: 0
346	FE406	옵디다	: 0	346	FE441	으러느냐	: 0
346	FE407	외다	: 0	214	FE442	으려는	: 4
011	FE408	요	: 842	346	FE443	으려는가	: 0
079	FE409	우	: 51	346	FE444	으려는고	: 0
049	FE410	으나	: 111	346	FE445	으려는데	: 0
296	FE411	으나마	: 1	346	FE446	으려는지	: 0
214	FE412	으냐	: 4	235	FE447	으러니	: 3
346	FE413	으냐고	: 0	346	FE448	으러니와	: 0
067	FE414	으니	: 60	258	FE449	으려다	: 2
135	FE415	으니까	: 14	296	FE450	으려다가	: 1
346	FE416	으니까는	: 0	346	FE451	으려더니	: 0

346	FE452	으러더라	: 0	148	FE485	으세요	: 12
346	FE453	으러던가	: 0	346	FE486	으셔요	: 0
296	FE454	으러도	: 1	346	FE487	으소서	: 0
346	FE455	으러머는	: 0	346	FE488	으시어요	: 0
346	FE456	으러면	: 0	346	FE489	으오	: 0
346	FE457	으러면야	: 0	346	FE490	으오리카	: 0
346	FE458	으러무나	: 0	346	FE492	으오이다	: 0
346	FE459	으러서는	: 0	346	FE493	으옵니까	: 0
346	FE460	으러서야	: 0	346	FE494	으옵니다	: 0
346	FE461	으러야	: 0	346	FE495	으옵디까	: 0
296	FE462	으러오	: 1	346	FE496	으옵디다	: 0
346	FE463	으런	: 0	346	FE497	으외다	: 0
346	FE464	으런다	: 0	346	FE498	으이	: 0
346	FE465	으런마는	: 0	010	FE499	은	: 858
296	FE466	으런만	: 1	192	FE500	은가	: 6
346	FE467	으럴	: 0	135	FE501	은걸	: 14
346	FE468	으럼	: 0	346	FE502	은고	: 0
346	FE469	으럼니까	: 0	346	FE571	은넌다	: 0
346	FE470	으럼니다	: 0	118	FE503	은테	: 20
346	FE471	으렀다	: 0	346	FE504	은들	: 0
346	FE491	으로리다	: 0	346	FE505	은만큼	: 0
258	FE472	으리	: 2	346	FE506	은바	: 0
346	FE473	으리카	: 0	346	FE507	은죽	: 0
235	FE474	으리니	: 3	346	FE508	은죽슨	: 0
346	FE475	으리니라	: 0	296	FE509	은지	: 1
258	FE476	으리다	: 2	346	FE510	은치고	: 0
118	FE477	으리라	: 20	346	FE511	은지라	: 0
346	FE478	으리로다	: 0	024	FE512	을	: 359
346	FE479	으리만치	: 0	346	FE513	을거나	: 0
346	FE480	으리만큼	: 0	204	FE514	을걸	: 5
346	FE481	으리요	: 0	105	FE515	을까	: 28
296	FE482	으마	: 1	346	FE516	을까마는	: 0
296	FE483	으매	: 1	296	FE517	을까말까	: 1
048	FE484	으며	: 112	346	FE518	을까보나	: 0
046	FE681	으면	: 115	346	FE519	을까보다	: 0
153	FE676	으므로	: 11	346	FE520	을께	: 0

346	FE521	을꼬	: 0	346	FE557	을지나	: 0
296	FE522	을느지	: 1	346	FE558	을지니	: 0
235	FE523	을라	: 3	346	FE559	을지니라	: 0
204	FE524	을라고	: 5	346	FE560	을지라	: 0
346	FE525	을라치면	: 0	346	FE561	을지라도	: 0
346	FE526	을락	: 0	346	FE562	을지어다	: 0
346	FE527	을랑	: 0	346	FE563	을지언정	: 0
346	FE528	을랑은	: 0	346	FE564	을진대	: 0
346	FE529	을래	: 0	346	FE565	을진댄	: 0
346	FE530	을러니	: 0	346	FE566	을진저	: 0
346	FE531	을러라	: 0	034	FE567	읍	: 193
346	FE532	을런가	: 0	346	FE568	읍세	: 0
346	FE533	을런고	: 0	235	FE569	읍에도	: 3
346	FE534	을레	: 0	346	FE570	읍에라	: 0
346	FE535	을레라	: 0	118	FE572	읍니까	: 20
346	FE536	을말로는	: 0	132	FE573	읍니다	: 15
346	FE537	을말로야	: 0	346	FE574	읍디까	: 0
346	FE538	을망정	: 0	346	FE575	읍디다	: 0
346	FE539	을바에	: 0	346	FE576	읍딘다	: 0
346	FE540	을바에야	: 0	346	FE577	읍쇼	: 0
346	FE541	을밖에	: 0	346	FE578	읍시다	: 0
346	FE542	을뿐더러	: 0	346	FE579	읍시오	: 0
346	FE543	을새	: 0	346	FE580	의	: 0
346	FE544	을세	: 0	296	FE581	이	: 1
346	FE545	을세라	: 0	214	FE582	이거나	: 4
346	FE546	을세말이지	: 0	346	FE583	이거늘	: 0
346	FE547	을소냐	: 0	296	FE684	이거니	: 1
346	FE551	을소냐	: 0	346	FE584	이거니와	: 0
346	FE552	을수룩	: 0	214	FE585	이거든	: 4
346	FE548	을수룩	: 0	346	FE586	이거들랑	: 0
346	FE549	을쏘냐	: 0	171	FE683	이건	: 8
346	FE553	을쏘냐	: 0	296	FE587	이기로	: 1
346	FE554	을이만큼	: 0	214	FE588	이기에	: 4
346	FE550	을이만큼	: 0	258	FE685	이긴	: 2
346	FE555	을작시면	: 0	296	FE686	이길래	: 1
204	FE556	을지	: 5	075	FE589	이나	: 53

235	FE590	이네	: 3	258	FE624	이로군	: 2
258	FE687	이노	: 2	235	FE625	이로다	: 3
105	FE591	이니	: 28	258	FE626	이로되	: 2
192	FE688	이니까	: 6	346	FE627	이로라	: 0
214	FE592	이니깐	: 4	095	FE628	이며	: 37
022	FE593	이다	: 411	105	FE682	이면	: 28
346	FE594	이더니	: 0	296	FE629	이므로	: 1
346	FE595	이든가	: 0	258	FE677	이므로	: 2
346	FE596	이든걸	: 0	132	FE691	이어서	: 15
346	FE597	이든고	: 0	131	FE630	이오	: 16
346	FE598	이든데	: 0	346	FE631	이올시다	: 0
072	FE600	이라	: 66	346	FE632	이옵나이다	: 0
346	FE601	이라니	: 0	082	FE690	이요	: 48
346	FE602	이라니까	: 0	054	FE633	이지	: 86
214	FE689	이라든가	: 4	346	FE634	이지마는	: 0
346	FE603	이라든지	: 0	118	FE692	이지만	: 20
153	FE604	이라면	: 11	056	FE635	인테	: 82
258	FE605	이라야	: 2	100	FE636	인지	: 31
346	FE606	이라야만	: 0	346	FE637	인지라	: 0
346	FE607	이라지	: 0	023	FE638	자	: 361
346	FE608	이람	: 0	346	FE639	자꾸나	: 0
346	FE609	이라	: 0	214	FE693	자는	: 4
346	FE610	이러니	: 0	182	FE640	자마자	: 7
346	FE611	이러니라	: 0	346	FE641	잔	: 0
346	FE612	이러니까	: 0	235	FE642	잘	: 3
346	FE613	이러니이다	: 0	032	FE696	쪼	: 218
346	FE614	이러이까	: 0	007	FE643	지	: 1109
346	FE615	이런가	: 0	114	FE644	지는	: 22
346	FE616	이런들	: 0	182	FE645	지를	: 7
296	FE617	이려니와	: 1	296	FE646	지마는	: 1
346	FE618	이런마는	: 0	033	FE647	지만	: 209
346	FE619	이런만	: 0	080	FE648	지요	: 49
346	FE620	이로고	: 0	171	FE649	진	: 8
258	FE621	이로구나	: 2	118	FE650	질	: 20
346	FE622	이로구료	: 0				
346	FE623	이로구면	: 0				

위의 <표 7>은 어말어미의 빈도를 보인 것이다. 종결어미와 연결어미를 가리지 않았다는 점에서 위의 <표 7>은 다시 둘로 나누어질 수도 있을 것이다. 그럴 경우에는 연결어미와 종결어미를 구분하지 않은 ‘고’라든가 ‘게’ 또는 ‘나’와 같은 형태에 대한 재분석이 이루어져야 할 것이다. 하지만 본 연구가 목적하고 있는 한국어 교육에서의 활용이라는 측면에서 본다면 그들에 대한 별도의 분석 작업은 보류되어도 일단은 무방할 것으로 보인다. 교육의 장면에서는 하나의 형태가 제공될 경우 같은 자리에서 다양한 의미를 제공할 필요도 있을 것으로 보이기 때문이다. 혹 피교육자의 혼동 방지와 교육 효과를 높이기 위하여 기본적인 의미나 기능을 나누어 제공한다고 하더라도 해당 형태의 비중을 파악하기 위한 빈도 조사에서는 같은 자리에서 파악하는 것이 온당한 것이라고도 할 수 있기 때문이다. 아무튼 <표 7>에 반영된 결과는 어말어미인 ‘다’와 ‘ㄴ’의 빈도가 압도적인 것임을 보여 주고 있다. 그러나 위의 <표 6>나 <표 7>에서와 같이 개별 형태의 출현 빈도수는 표면적인 것이라 할 수 있다. 교육적인 측면에서의 활용도를 높이기 위해서는 2차 해석이 필요하기 때문이다. 형태 목록을 문법 범주별로 재분류하는 것이 그것이다.

아울러 위에서 따로이 살펴본 선어말어미와 어말어미의 출현 빈도는 함께 묶어 보는 것도 유용할 것이다. 문법 형태 측면에서의 학습 진행 과정을 구성하는 데에는 함께 살펴 순서를 정하는 것이 필요할 것이기 때문이다. 본고에서는 선어말어미와 어말어미를 아우른 전체 빈도 양상을 [부록 1]로 보이기로 한다. (‘문법 패턴’에 관한 남은 원고는 다음 ‘인문논총’의 ‘하’부분에 계속)

## 참 고 문 헌

- 김영석, 이상억(1993) 현대형태론, 학연사.  
 김홍규, 성광수, 홍종선(1998) 대규모 한국어 데이터 베이스의 다원적 통계분석 연구. 한국어전산학 2. pp.1-169.  
 김홍규, 강범모(2000) 한국어 형태소 및 어휘 사용 빈도의 분석 1, 고려대 민족문화연구원.

- 남기심, 고영근(1985) 표준 국어문법론, 탑출판사.
- 문교부 주관(1956) 우리말 짓기 조사, 우리말에 쓰힌 말수의 짓기 조사. 문교부.
- 서상규(1998) 연세 말뭉치 1-9를 대상으로 한 ‘현대 한국어의 어휘 빈도’-빈도 7 이상-(상), 연세대 언어정보개발연구원.
- 안병희(1965) 문법론, 국어학개론, 수도출판사.
- 안병희(1967) 문법사, 한국어 발달사, 한국문화사대계 V, 고대 민족문화연구소. pp. 167-261.
- 이상억(1986) 계량언어학(어휘론부분): 어휘부의 음운, 형태, 의미론적 구조. 언어 11.2. pp.323-339.
- 이상억(1989a) 국어 어휘 목록의 형태, 음운론적 구조 연구: 계량언어학적 표준조사. 어학연구 25.1. pp.111-128.
- 이상억(1989b) 음절구조의 변화 원인에 대한 이런 설명은 어떨까요? 정연찬 선생 회갑기념논총, 태학사. pp.203-210.
- 이상억(1989c) A Glottometrical Study of Korean Lexicon. Harvard Studies in Korean Linguistics III, Hanshin Publishing Co. pp.159-166.
- 이상억(1990a) 현대국어 음변화규칙의 기능 부담량. 어학연구 26.3. pp.441-467.
- 이상억(1990b) 국어 어휘부의 계량언어학적 연구. 강신항 선생 회갑기념논총, 태학사. pp.331-346.
- 이상억(1993) 현대 국어의 통사적 현상에 대한 기능 부담량. 안병희 선생 회갑 기념논총, 신구문화사. pp.475-487.
- 이상억(1997) 현대국어 용언 형태의 계량언어학적 연구. 음성 음운 형태론 연구 3. pp.203-222.

[부록1] <선어말·어말 어미 빈도순>

다음 표에서 제1칸 ‘고대 순위’는 고려대 김홍규 등(1998)의 어미 빈도순 목록을 인용하여 대조시킨 것이다. 이중 ( )로 표시된 것들은 고대 쪽에 나타나지 않은 형태이며, 제1칸과 제2칸의 2종류의 순위들 사이에 >, <의 부등호가

쓰인 이유는 대조된 순위의 격차가 10위 이상 많이 다른 경우들을 보이려는 것이다.

제2칸은 <선어말·어말 어미 빈도순>으로 001은 빈도수 순위 FE(어말어미)127번 -다 : 6069의 출현회수를 나타낸다.

제3칸의 '연대 순위'는 서상규(1998)의 빈도순 7 이상의 어휘 중에서 어미(씨끝)만 찾아서 대조시킨 것이다. 이중 ( )로 표시된 것들은 연대 쪽에 나타나지 않은 형태이거나 상위 50위 이내에서 보이지 않는 형태로 그 밑 순위는 여기서 대비하지 않고 생략하였다. 부등호의 사용은 위 제1칸에서와 같이 10위 이상의 차를 보이는 쪽으로 붙였다.

제4칸의 Cf.(참조) 부분은 제2칸으로부터 '어미의 이형태를 한 데 묶은 출현빈도 순위'다.

고대 순위	<선어말·어말 어미 출현빈도순>	연대 순위	Cf. 옆 < > 표시 칸 어미의 이형태를 한 데 묶은 출현빈도 순위
2	001 FE127 다 ta : 6039	3	1. -ㄴ/-은 -n/-un(v.p., a.pr.) 6648
1	002 FE042 ㄴ n : 5790 < 25		2. -다 -ta 6039
4	003 FE021 고 ko : 3843	5	3. -왔/왔/였- -ass/ess/yess- 5252
3	004 FE093 는 nun : 3056	6	4. -고 -ko 3843
5	005 PFE015 었 ess : 2454	2	5. -는 -nun 3056
8	006 FE371 어 e : 2341	4	6. -아/-어 -a/-e 3411
( )	007 PFE016 었 yess : 1498 ( )	( )	7. -ㅁ /습니다 -p/supnita 1345
6	008 PFE014 았 ass : 1300 ( )	( )	8. -게 -key 1288
10	009 FE017 게 key : 1288	8	9. -아서/-어서 -ase/-ese 1128
(연결·종결어미가 합산된 - 게)			10. -지 -ci 1109
13	010 FE643 지 ci : 1109	9	11. -ㄴ 다/-는다 -nta/-nunta 1083
7	011 FE355 아 a : 1070 ( )	( )	12. -요 -yo 842
25 >	012 FE326 ㅁ 니다pnita: 870 < 45		13. -야 -ya 656
11	013 FE499 은 un : 858	1	14. -면 -myen 562
58 >	014 FE408 요 yo : 842 ( )	( )	15. -기 -ki 531
26 >	015 FE376 어서 ese : 674	11	16. -며 -mye 480
56 >	016 FE367 야 ya : 656 < 42		17. -겠- -keyss- 453
15	017 FE320 먼 myen : 562	14	18. -던 -ten 449



46 > 018	PFI006	는(다)nun(ta): 557	12	19. -다고 -tako	440
14 019	FE036	기 ki : 531	10	20. -이다 -ita	441
20 020	FE319	며 mye : 480	15	21. -라/-을 -l/-ul	384
19 021	FE347	습니다supnita: 475	13	22. -자 -ca	361
12 < 022	FE046	ㄴ다 nta : 454	34	23. -면서 -myense	357
18 023	FE360	아서 ase : 454	( )	24. -시- -si-	352
28 024	PFI002	졌 keyss : 453	22	25. -는데 -nuntey	312
16 < 025	FE158	던 ten : 449	16	이하는 왼쪽 033순위로 계속	
21 026	FE131	다고tako : 440	27		
( ) 027	FE593	이다 ita : 411	( )		
39 > 028	FE638	자 ca : 361	29		
17 < 029	FE512	을 ul : 359	> 7		
30 030	FE321	면서myense: 357	26		
36 031	PFE013	시 si : 352	30		
29 032	FE115	는데nuntey : 312	28		
48 > 033	FE128	다가taka : 284	35		
23 < 034	FE133	다는tanun : 271	> 23		
32 035	FE223	라 la : 256	31		
41 036	FE120	니 ni : 239	> 25		
437 > 037	FE166	도 to : 223	( )		
60 > 038	FE696	쪼 cyo : 218	( )		
27 < 039	FE647	지만ciman : 209	> 20		
( ) 040	PFE007	더 te : 196	( )		
51 041	FE567	음 um : 193	> 18		
42 042	FE071	나 na : 180	> 17		
52 043	FE062	ㄴ데ntcy : 177	( )		
230 > 044	FE386	여 ye : 174	47		
40 045	FE379	어요 eyo : 170	> 21		
91 > 046	FE171	듯 tus : 165	( )		
( ) 047	PFE001	거 ke : 163	( )		
72 > 048	FE076	냐 nya : 156	( )		
88 > 049	FE395	오 o : 154	39		
234 > 050	PFE009	리 li : 147	( )		
122 > 051	FE307	리라 : 134			
( ) 052	PFI005	ㄴ(ㄴ/냐/ㄴ지): 127			
66 > 053	FE022	고는 : 118			

144 > 054	FE035	군	: 117	64 < 090	FE648	지요	: 49
43 < 055	FE167	도록	: 117	106 > 090	FE663	다며	: 49
44 < 056	FE681	으면	: 115	( ) 092	FE690	이요	: 48
24 < 057	FE226	라논	: 114	345 > 093	FE095	논걸	: 47
37 < 058	FE484	으며	: 112	92 094	FE172	듯이	: 46
47 < 059	FE410	으나	: 111	1285 > 095	FE037	기로	: 45
22 < 060	FE224	라고	: 105	71 < 096	FE364	아요	: 43
78 > 061	FE150	디니	: 103	( ) 097	FE041	기에	: 41
84 > 062	FE340	세요	: 99	69 < 097	FE261	려논	: 41
35 < 063	FE043	ㄴ가	: 91	1042 > 097	FE370	야지	: 41
( ) 064	FE633	이지	: 86	96 100	FE245	러	: 40
38 < 065	FE049	ㄴ다는	: 85	81 < 100	FE385	에요	: 40
1285 > 066	FE635	인테	: 82	97 102	FE080	네	: 39
65 067	FE121	니까	: 76	87 < 102	FE346	습니까	: 39
82 > 068	FE087	ㄴ나	: 74	220 > 104	FE010	거라	: 38
55 < 069	FE068	ㄴ지	: 72	374 > 105	FE044	ㄴ걸	: 37
46 < 069	FE103	논다	: 72	1285 > 105	FE628	이며	: 37
54 < 069	FE117	논지	: 72	( ) 107	FE187	ㄴ라고	: 36
111 > 072	FE030	구나	: 70	623 > 108	FE670	래도	: 35
79 072	FE230	라도	: 70	109 109	PII:019	으시	: 34
73 074	FE138	다면	: 66	172 > 110	FE077	냐고	: 32
113 > 074	FE375	어라	: 66	86 < 111	FE232	라면	: 31
128 > 076	FE025	고서	: 61	1285 > 111	FE636	인지	: 31
75 077	FE414	으니	: 60	33 < 113	FE361	아야	: 30
67 < 078	FE341	소	: 59	529 > 114	FE072	나니	: 29
118 > 079	FE170	든지	: 58	( ) 114	FE388	여라	: 29
98 > 079	FE359	아라	: 58	101 < 116	FE169	든	: 28
80 081	FE325	바니까	: 57	84 < 116	FE254	려	: 28
1285 > 082	FE600	이라	: 56	74 < 116	FE515	을까	: 28
139 > 083	FE007	거든	: 55	810 > 116	FE591	이니	: 28
49 < 084	FE001	거나	: 54	1285 > 116	FE682	이면	: 28
34 < 085	FE377	어야	: 53	331 > 121	FE148	더군	: 26
1285 > 085	FE589	이나	: 53	117 121	FE231	라며	: 26
70 < 087	FE257	려고	: 52	9 < 123	FE174	ㄴ	: 25
63 < 087	FE374	어도	: 52	1042 > 124	FE045	ㄴ고	: 24
164 > 089	FE409	우	: 51	132 125	FE134	다니	: 22

695 > 125	FE644	지는	: 22	236 > 159	FE485	으세요	: 12
50 < 127	FE094	는가	: 21	472 > 159	FE660	는단	: 12
144 > 127	FE664	대	: 21	185 > 164	FE090	느니	: 11
248 > 129	FE123	니깐	: 20	185 > 164	FE139	다면서	: 11
133 > 129	FE168	되	: 20	159 > 164	FE140	다시피	: 11
167 > 129	FE477	으리라	: 20	374 > 164	FE149	더냐	: 11
116 < 129	FE503	은데	: 20	? > 164	FE159	던가	: 11
482 > 129	FE572	읍니까	: 20	275 > 164	FE162	던데	: 11
216 > 129	FE650	질	: 20	695 > 164	FE183	르께	: 11
( ) 129	FE692	이지만	: 20	695 > 164	FE255	려거든	: 11
1285 > 136	FE366	았습니다	: 19	215 > 164	FE339	세	: 11
203 > 137	FE154	더라	: 18	1285 > 164	FE604	이라면	: 11
( ) 137	FE389	여서	: 18	( ) 164	FE652	ㄴ데다	: 11
495 > 139	FE098	는구려	: 17	103 < 169	FE676	으므로	: 11
345 > 139	FE102	는군	: 17	127 < 180	FE104	는다고	: 10
271 > 139	FE669	르래	: 17	「112 < 180	FE159	던가	: 10」
( ) 142	FE630	이오	: 16	176	FE297	로군	: 10
1285 > 143	FE434	으러	: 15	176	FE322	면은	: 10
192 > 143	FE573	읍니다	: 15	179	FE204	르세	: 9
1285 > 143	FE691	이어서	: 15	179	FE227	라니	: 9
104 < 146	FE023	고도	: 14	76 < 184	FE358	아도	: 9
143 > 146	FE088	느냐고	: 14	182	FE018	게나	: 8
108 < 146	FE105	는다는	: 14	182	FE089	느노	: 8
83 < 146	FE212	르지	: 14	182	FE097	는구나	: 8
( ) 146	FE393	였습니다	: 14	182	FE173	디	: 8
107 < 146	FE415	으니까	: 14	182	FE239	란다	: 8
( ) 146	FE417	으니깐	: 14	182	FE317	마	: 8
( ) 146	FE501	은걸	: 14	182	FE422	으라	: 8
115 < 154	FE011	건	: 13	182	FE435	으러	: 8
99 < 154	FE028	곤	: 13	182	FE649	진	: 8
( ) 154	FE078	냐는	: 13	182	FE668	래	: 8
77 < 154	FE155	더라도	: 13	182	FE683	이건	: 8
( ) 154	FE694	시우	: 13	193	FE065	ㄴ마	: 7
316 > 159	FE079	너라	: 12	193	FE083	노라	: 7
657 > 159	FE081	노니	: 12	193	FE092	느라고	: 7
743 > 159	FE243	랴	: 12	193	FE130	다간	: 7

193	FE176	르결	: 7	225	FE323	므로	: 4
193	FE228	라니까	: 7	225	FE363	아야지	: 4
193	FE264	러는데	: 7	225	FE368	야만	: 4
193	FE423	으라고	: 7	225	FE373	어다가	: 4
193	FE640	자마자	: 7	225	FE412	으나	: 4
193	FE645	지를	: 7	150 < 231	FE438	으려고	: 4
203	FE129	다가는	: 6	225	FE442	으려는	: 4
203	FE207	르수록	: 6	225	FE582	이거나	: 4
203	FE266	러니	: 6	225	FE585	이거든	: 4
203	FE269	러다가	: 6	225	FE588	이기에	: 4
105 < 208	FE275	려면	: 6	225	FE592	이니깐	: 4
141 < 208	FE332	비시오	: 6	225	FE653	네데도	: 4
203	FE432	으랴	: 6	225	FE655	보담	: 4
203	FE500	은가	: 6	225	FE689	이라든가	: 4
203	FE672	려면	: 6	110 < 231	FE693	자는	: 4
203	FE674	마는	: 6	225	PII:008	료(구나/구려)	: 4
203	FE688	이니까	: 6	247	FE003	거니	: 3
203	FE695	시요	: 6	89 < 253	FE027	교자	: 3
215	FE137	다만	: 5	247	FE033	구먼	: 3
215	FE141	다오	: 5	247	FE058	네내	: 3
215	FE233	라면서	: 5	247	FE142	단다	: 3
215	FE237	라오	: 5	247	FE205	르세라	: 3
215	FE238	락	: 5	247	FE265	러는지	: 3
215	FE281	려오	: 5	247	FE268	려다	: 3
215	FE306	리다	: 5	247	FE285	런만	: 3
215	FE514	을걸	: 5	247	FE288	럼니까	: 3
215	FE524	으라고	: 5	247	FE353	시거든	: 3
119 < 221	FE556	을지	: 5	247	FE392	여요	: 3
225	FE015	건만	: 4	247	FE447	으러니	: 3
225	FE135	다마는	: 4	247	FE474	으리니	: 3
225	FE156	더라면	: 4	247	FE523	을라	: 3
225	FE160	던걸	: 4	247	FE569	음에도	: 3
225	FE201	르밖에	: 4	247	FE590	이네	: 3
225	FE277	려무나	: 4	247	FE625	이로다	: 3
95 < 231	FE323	므로	: 4	247	FE642	잘	: 3
142 < 231	FE363	아야지	: 4	247	FE651	네답니까	: 3

247	FE656	나며	: 3	270	FE662	다네	: 2
247	FE665	답니다	: 3	270	FE671	래야	: 2
247	FE675	먼야	: 3	270	FE677	이므로	: 2
270	FE004	거니와	: 2	270	FE685	이긴	: 2
270	FE008	거들랑	: 2	270	FE687	이노	: 2
270	FE012	건내	: 2	308	FE002	거늘	: 1
270	FE031	구려	: 2	308	FE026	고야	: 1
270	FE073	나파	: 2	308	FE032	구료	: 1
270	FE096	논고	: 2	308	FE034	구면	: 1
270	FE132	다느니	: 2	308	FE050	ㄴ 다니까	: 1
270	FE145	더구나	: 2	308	FE053	ㄴ 달	: 1
270	FE164	던지	: 2	308	FE059	ㄴ 대서	: 1
270	FE175	ㄴ 거나	: 2	308	FE066	ㄴ 즉	: 1
270	FE184	ㄴ 꼬	: 2	308	FE084	노라고	: 1
270	FE198	ㄴ 망정	: 2	308	FE086	노라면	: 1
270	FE199	ㄴ 마	: 2	308	FE108	논달	: 1
270	FE229	라든지	: 2	308	FE116	논마	: 1
270	FE258	려나	: 2	308	FE202	ㄴ 뿐더러	: 1
270	FE267	려니와	: 2	308	FE222	ㄴ 진저	: 1
270	FE287	럼	: 2	308	FE235	라야	: 1
270	FE300	로라	: 2	308	FE241	랍니다	: 1
270	FE304	리니	: 2	308	FE263	려는고	: 1
270	FE356	아다	: 2	308	FE298	로다	: 1
270	FE357	아다가	: 2	308	FE316	ㅁ 직하다	: 1
270	FE382	언만	: 2	308	FE318	매	: 1
270	FE390	여야	: 2	308	FE362	아야만	: 1
270	FE436	으려거든	: 2	120 < 314	FE372	어다	: 1
270	FE449	으려다	: 2	308	FE400	올시다	: 1
270	FE472	으리	: 2	308	FE411	으나마	: 1
270	FE476	으리다	: 2	308	FE421	으되	: 1
270	FE605	이라야	: 2	308	FE428	으라면	: 1
270	FE621	이로구나	: 2	308	FE430	으락	: 1
270	FE624	이로군	: 2	308	FE450	으려다가	: 1
270	FE626	이로되	: 2	308	FE454	으려도	: 1
270	FE658	냐가	: 2	308	FE462	으려오	: 1
270	FE659	누나	: 2	308	FE466	으련만	: 1

308	FE482	으파	: 1
308	FE483	으메	: 1
308	FE509	은지	: 1
308	FE517	을까말까	: 1
308	FE522	을는지	: 1
308	FE581	이	: 1
308	FE587	이기로	: 1
308	FE617	이려니와	: 1
308	FE629	이므로	: 1
308	FE646	지마는	: 1
308	FE654	ㄴ 데서	: 1
308	FE657	냐면서	: 1
308	FE661	는도다	: 1
308	FE666	듣가	: 1
308	FE673	로구만	: 1
308	FE679	으래	: 1
308	FE680	쇼	: 1
308	FE684	이거니	: 1
308	FE686	이길래	: 1
358	FE005	거드면	: 0
		(중간의 365 항목은 생략하고 근간 출처 “계량국어학”에 수록)	
358	PFI:024	잡	: 0

## Quantitative Research on Suffixes and Grammatical Patterns in Present-day Korean

Lee, Sang-Oak

In this paper, I will present a quantitative analysis of the frequency and ranking of the following items in modern Korean: (1) prefinal and final endings (the latter are used as conclusive morphemes or non-conclusive morphemes), and (2) the grammatical patterns and adverbial phrases that are connected with these sentential adverbs. Oddly, there is no research to date on the frequency of use of both suffixes and grammatical patterns. In the 1956 survey conducted by the Ministry of Education entitled “A Survey of Word Frequency in Korean” only basic form of verbal endings were surveyed. In 1997, I conducted the first survey of agglutinating suffixes(e.g., conclusive morphemes and non-conclusive morphemes), entitled “Quantitative Research on Verbal Morphemes in Modern Korean” (Lee 1997). Since then I have not been able to expand this research to a large corpus but this first pilot study already showed a valuable index such as order of frequency. A couple of others' studies on the frequency of suffixes have been reported based on their larger corpus but those of grammatical patterns not yet done by others.

### <Frequency of Prefinal and Final Verbal Suffixes>

001	FE127	다	: 6039	002 FE042	ㄴ	: 5790
003	FE021	고	: 3843	004 FE093	는	: 3056
005	PFE015	았	: 2454	006 FE371	어	: 2341
007	PFE016	였	: 1498	008 PFE014	았	: 1300
009	FE017	계	: 1288	010 FE643	지	: 1109

## &lt;Frequency of Grammatical Patterns&gt;

1	grade 1	-어(아) 가다	741
2	grade 2	-고 있다	687
3	grade 1	-지(요)	539
4	grade 1	-어 주다	468
5	grade 1	-지 않다(-지도 않다)	468
6	grade 1	-어(아) 오다	428
7	grade 1	-어 있다	392
8	grade 4	-(이)야	330
9	grade 1	-어 보다	310
10	grade 1	-어 지다	293

The above findings are useful in the following ways: First, they are instructive in developing a rational basis for pedagogical syllabi and materials in teaching Korean as a foreign or a second language. They are also useful in analyzing the order of presentation of suffixes and structures in existing materials for learners of Korean. This information is useful in improving these materials and in developing new materials. Second, the findings can also be applied to analyzing the order of presentation of morphemes in elementary Korean materials for children in Korea. As with materials for foreign learners of Korean, this information is useful in developing a natural and pedagogically effective ordering of agglutinating verbal morphemes. Third, the findings are relevant to machine translation (particularly English-Korean and Japanese-Korean) because they provide information on which Korean expressions are more preferable in translation. Machine translation, a field that has emerged from computer engineering, needs accurate linguistic information to succeed and the linguistic information, which is based on research methods from natural science, in this paper can help build a bridge from the humanities to engineering.