

‘산유화(山有花)’의 국어정보학적 연구1)  
The Study on <Sanyuhwa> from the viewpoint of Korean Informatics

이태영2)  
Lee, tae-yeong(chonbuk univ.)

전북대학교 인문대학 국어국문학과

<국문초록>

이 연구는 시를 이해하고 분석하는 데 도움이 되는 국어정보학적인 연구방법을 제시하고자 한다. 이 논문에서는 <산유화>를 대상으로 일단 문법형태소를 중심으로 이 시를 해석해 보고, 음소와 음절, 어절의 빈도를 이용하여 이 시가 갖는 여러 특징을 밝혀 보고자 한다.

문법형태소의 중요성은 이 시에서도 분명하게 나타난다. <산유화>에 나타나는 ‘산에는, 산에서, 산에’ 등은 문법형태소에 따라서 ‘산’의 의미가 달라진다. ‘꽃, 꽃이’의 경우도 마찬가지다. ‘피네’의 ‘-네’가 보여주는 기능은 시 전체를 해석하는 데 결정적인 영향을 끼친다.

시를 좀더 분석적으로 연구하기 위하여 음소, 음절, 어절, 어휘의 빈도를 자세히 내어서 그 빈도가 보여주는 특성을 통하여 정보학적인 시각으로 살펴보고자 한다. 음소 빈도 결과에서 초성과 중성, 중성에 쓰인 음소의 빈도를 정확히 파악할 수 있다. 이 시에서 사용된 음소의 특징을 살펴보면 시인이 생각하는 음소의 기능을 파악할 수 있을 것이다. 음절 빈도를 내보면 종결어미 ‘-네’, 주격조사 ‘-이’, 처격조사 ‘-에’, 특수조사 ‘-은/는’, 명사 ‘꽃’과 ‘산’, 동사 ‘피-’ 등이 가장 많이 쓰이는 핵심 음절임을 알 수 있다. 어절 빈도를 통하여 그 시의 구체적인 어휘들의 특성을 살펴보는 일은 매우 중요한 작업이 될 것이다.

<산유화>에서 우리는 시인이 음소를 배열하는 리듬을 발견하게 된다. 어절의 첫 음소의 배열 순서를 살펴보면 김소월이 <산유화>를 지으면서 인상적인 특성을 염두에 두고 음소의 배열을 고려한 모습을 살필 수 있게 될 것이다.

핵심어 : 문법형태소, 빈도, 음소, 음절, 어절, 음소배열

## 1. 서론

이 연구의 목적은 국어정보학적인 방법을 이용하여 시와 소설을 이해하는 데 도움이 되고자 하는 연구 방법을 제시하고자 하는 것이다.

현대시나 소설을 연구할 때, 기초적으로 국어학적인 탐색이 요망되지만, 국어학적 지식이 부족하여 기초적인 탐색조차도 이루어지지 않고 있다. 국어학계에서는 시와 소설의 언어 연구에 대해 국어 문체론적 접근을 주로 시도하고 있으나 국어학의 연구 자료로 적극적으로 활용하지 않고 있다.

그러나 최근 현대문학 자료에 대한 말뭉치가 대대적으로 구축되었고, 이미 많은 문학작품 말뭉치를 가지고 정보 처리를 하고 있기 때문에 국어학적인 지식과 아울러 정

1) 이 논문은 국제고려학회가 주최하여 2005년 9월 21일부터 23일까지 중국 심양에서 열린 제7차 코리아학국제 학술토론회에서 발표한 바 있다.

2) 전북대학교 인문대학 국어국문학과 교수, <겨레말큰사전> 남북공동편찬위원, 21세기세종계획 한민족언어정보화 분과 연구책임자.

보 처리 능력이 필요하게 되었다.

국어를 분석하는 프로그램이 개발되면서 음소와 음절, 어절의 빈도를 정확하게 산출하게 되었고, 음소 검색을 통하여 음운현상을 확인하고 형태소를 정밀하게 검색하면서 다양한 어휘와 형태소를 확인하게 되었다. 프로그램이 용례사전을 간단히 작성하여 주기 때문에 필요한 어휘에 대해 전체적인 예문을 한 눈에 확인할 수 있게 되었다.

따라서 하나의 작품은 물론, 작가의 작품 전체를 한꺼번에 정보처리하는 경향이 높아졌으며, 시대별 작품의 해석, 경향별 작품의 해석 등이 가능하게 되었다. 예를 들어 노천명, 서정주, 김영랑 등 각각의 시 전체를 대상으로 정보학적으로 분석하면 보다 빠르고 쉽게 작가의 시세계를 엿볼 수 있다.

본고는 김소월의 시 가운데 <산유화>를 대상으로 일단 문법 형태소를 중심으로 이 시를 해석해 보고, 음소와 음절, 어절의 빈도를 이용하여 이 시가 갖는 몇 가지 특징을 밝혀 보고자 한다.

## 2. <산유화>의 국어학적 해석

다음에 제시하는 시는 김소월의 <산유화>이다. 일단 기초적으로 국어학적인 입장에서 형태소를 중심으로 이 시를 분석해 보고자 한다.

<산유화>

- |          |        |           |           |
|----------|--------|-----------|-----------|
| ① 산에는    | ② 꽃피네  | ⑦ 산에서     | ⑧ 우는 적은새요 |
| ③ 꽃치피네   |        | ⑨ 꽃치조와    |           |
| 갈 봄 너름엽시 |        | 산에서       |           |
| 꽃치피네     |        | ⑩ 사노라네    |           |
| ④ 산에     |        | ⑪ 산에는 꽃지네 |           |
| 산에       |        | ⑫ 꽃치지네    |           |
| ⑤ 피는꽃춘   |        | 갈 봄 너름엽시  |           |
| 저만치 혼자서  | ⑥ 피여있네 | 꽃치지네      |           |

첫째, 1연과 4연의 ‘① 산에는’, 2연의 ‘④ 산에’, 3연의 ‘⑦ 산에서’는 비록 동일한 어휘인 ‘山’이지만 실제로는 의미가 다른 산이다. 의미가 다르다는 것을 파악할 수 있는 방법은 ‘山’에 연결된 조사의 기능과 의미를 확인하면 된다.

즉, ‘산에’는 정적(靜的)인 산을 말하고 있고, ‘산에서’는 동적(動的)인 산을 말하고 있다. 2연에서는 ‘꽃이 피어 있는 산’이고, 3연은 ‘새가 울고 사는 산’이다. 이것은 조사 ‘-에’와 ‘-에서’가 차이를 보이기 때문에 바로 알 수 있다. 같은 처소를 나타내는 조사이지만 ‘-에’는 상태의 처소(또는 도착점)를 나타내고, ‘-에서’는 행위의 처소(또는 출발점)를 나타내는 점에서 차이가 있다.

한편 ‘산에는’은 막연한 산을 가리키는 것이 아니다. ‘산에는’의 조사 ‘-는’은 ‘주제’ 또는 ‘대조’의 의미를 가지는 특수조사인데 여기서 ‘산’은 ‘들, 정원, 밭이 아닌 산’을 가

리키는 것이다. 달리 말하면 ‘산에’와 ‘산에서’와는 달리 ‘산에는’은 들과 정원과 밭이 있다는 사실이 전제되면서 그 중에서 ‘산’을 선택한 것이다.<sup>3)</sup>

<산유화>에서 보여주는 ‘산’은 자연과 동물이 함께 하는 장소로 묘사하고 있으며 그 속에서 상태와 동작을 반복하고 있음을 보여준다.

이러한 사실은 이미 잘 알려진 다음과 같은 구문을 통해서 쉽게 이해할 수 있다.

- (1) 철수가 우리반 반장이다.
- (2) 철수는 우리반 반장이다.

(1)의 예문에서 ‘철수’는 한 반의 다른 사람들이 전제되지 않은 상태이다. 대체로 한 반에 40명의 학생이 있는데 다른 학생과의 관련성은 전혀 없다. 단순히 철수를 지칭한 것이다. 그러나 (2)의 예문에서 ‘철수’는 다른 학생들이 전제가 되고 있다. 예를 들면 ‘영희는 미화 부장이고, 영수는 기율 부장이고, 상태는 부반장인데 비하여 철수는 반장이다’는 뜻이다. 따라서 조사 ‘-가’와 ‘-는’의 차이로 말미암아 문장의 의미가 달라지는 것을 알 수 있다.

둘째, 1연과 4연의 ‘② 꽃피네(지네)’와 ‘③ 꽃치피네(지네)’ 그리고 2연의 ‘⑤ 피는꽃춘’ 과 3연의 ‘⑨ 꽃치조와’에 나타나는 ‘꽃’의 이미지는 과연 같은 것인가? 필자의 견해로는 아주 다르다고 말할 수 있다.

1연과 4연에 나타나는 ‘② 꽃피네(지네)’는 ‘산에는 꽃피네(지네)’라는 일반적인 명제를 이야기한 것으로 보인다. 따라서 일반성은 있을 뿐, 작가가 가리키는 구체적인 꽃을 이야기하는 것은 아니다.

1연과 4연에 나타나는 ‘③ 꽃치피네(지네)’는 행을 달리하여 쓰고 있다. 따라서 ‘산에는 꽃피네 꽃이 피네’와 같이 계속 연결된 명제로 보기가 어렵다. ‘산에는 꽃피네’와 ‘꽃이 피네’는 연결된 것이 아니라 따로 행을 달리한 명제로 보인다. 이렇게 볼 때, ‘③ 꽃치피네(지네)’에서 ‘꽃이’는 구체적인 꽃으로 이해되고 ‘꽃이피네’는 구체적인 진술이라고 해석할 수 있다.

‘꽃피네’와 ‘꽃이피네’의 경우, ‘꽃피네’는 시인의 입장에서는 ‘꽃(이) 피네’와 같은 통사론적 구성이 아니라 이미 굳어진 하나의 어휘로 해석할 수 있다. 예를 들면 ‘속시끄럽다, 속좁다, 속터지다’와 같은 단어들은 ‘속이 시끄럽다, 속이 좁다, 속이 터지다’와 같은 통사론적 구성에서 왔으나 현재는 하나의 단어로 쓰이고 있다. ‘꽃피네’의 경우도 하나의 단어로 등재된 것은 아니지만 시인의 경우 ‘산에는 꽃피네’에서 ‘꽃피네’를 하나의 어휘로 사용하고 있는 것으로 해석된다.

2연의 ‘⑤ 피는꽃춘’에 나타나는 꽃은 ‘대조’의 의미를 가지면서 들이나 정원에 피는 꽃이 아닌 ‘산에 피는 꽃’을 의미한다.

셋째, 용언을 통해서도 이 시의 구조를 이해할 수 있다. 1연의 ‘피네’와 2연의 ‘피여있네’와 3연의 ‘올다, 살다’, 4연의 ‘지네’의 용언은 다음과 같은 시의 구조를 보여준다.

3) 김소월의 시에 나타나는 산은 유개념어인 일반적인 산을 보여주는 산도 있고, 중개념어로서 구체적인 산을 보여주는 산도 나타난다. 이것은 시를 짓는 일반적인 작법과 관련되기 때문에 매우 주의를 요한다. 일반적으로 시의 첫 연과 마지막 연에서는 유개념 어휘를 사용하고, 중간 연에서는 중개념 어휘를 사용하고 있다. 즉 기-승-전-결의 구조에서 기와 결에서는 유개념 어휘를 쓰고, 승과 전에서는 중개념 어휘를 쓰는 것이 일반적이다. 따라서 시를 분석할 때도 해당되는 연을 중심으로 자료를 처리하여 해석하는 것이 필요하다. (홍윤표 교수 개인 강의자료 ‘어휘 의미론’ 참고)

‘피다’와 ‘지다’는 자동사로 상태를 나타내는 동사이다. ‘피어 있다’는 상태의 지속을 보여주는 구문이다. ‘울다’와 ‘살다’는 동족목적어를 수반하는 자동사로 동작을 나타내는 동사이다. 따라서 이 시의 1연은 ‘산에 꽃이 핀’ 정적인 상태를 표현하고, 2연에서는 ‘산에 꽃이 핀’ 정적인 상태가 지속되고 있음을 보여주고, 3연에서는 동적인 산의 모습을 보여주며, 4연에서는 다시 ‘꽃이 지는’ 정적인 상태를 보여주고 있다. 1연에서 4연까지의 시적 상징은 태어나서 움직이고 또 지는 생태계의 순환적인 모습을 보여주고 있다. 이러한 모습은 산과 꽃이라는 제재를 통하여 인간의 모습을 노래하려는 모습이 아닐까 생각한다.

이 시에서 우리는 어미 ‘-네’를 주목할 필요가 있다. 어미 ‘-네’의 용법을 표준국어대사전에서 찾아보면 다음과 같다.

1. ‘이다’ 의 어간, 용언의 어간 또는 어미 ‘-으시-’, ‘-었-’, ‘-겠-’ 뒤에 붙어 하게 할 자리에 쓰여, 단순한 서술의 뜻을 나타내는 종결 어미. (자네 차례네/집이 참 넓네/나 지금 가네/여긴 눈이 많이 왔네./그러다 병나겠네.)
2. 해할 자리나 혼잣말에 쓰여, 지금 깨달은 일을 서술하는 데 쓰이는 종결 어미. 흔히 감탄의 뜻이 드러난다. (우리 아이 노래도 잘 부르네!/집이 참 깨끗하네.)

이 용법 중 2번의 용법이 이 시의 종결어미로 쓰이는 ‘-네’의 용법이다. 따라서 시인은 깨달은 어떤 사실을 감탄하면서 노래하고 있는 것으로 해석된다. 여기서 ‘깨달음’이라는 말에 주목할 필요가 있다. 한길(1991:81-2)에 의하면 서술법의 ‘-네’는 반말체로서 ‘말할이가 자신의 생각이나 느낌을 들을이에게 확인 서술함.’이라는 의미를 가지기 때문에 말할이 자신의 ‘깨달음’의 의미를 추출할 수 있다고 말하고 있다.

이러한 언어학적인 형태소의 기능 분석에 의하면 ‘꽃치피네, 꽃치지네’의 경우에서 화자의 깨달음(곧 피고 지고 다시 피는 생성의 법칙)이 있음을 알 수 있을 것이다. 곧 화자는 이러한 인생과 자연의 생성의 법칙을 깨달으면서 이를 감탄하며 노래하는 것이다.

감탄의 의미는 어미 ‘-노라네’에서도 발견할 수 있다. ‘-노라네’는 ‘-노라 + -네’의 복합된 어미 구조인데 여기서 ‘-노라’는 동사 어간이나 어미 ‘-으시-’, ‘-었-’, ‘-겠-’ 뒤에 붙어 예스러운 표현으로 ‘해라할 자리나 간접 인용절에 쓰여, 자기의 동작을 장중하게 선언하거나 감동의 느낌을 나타내는 종결 어미’이다.(표준국어대사전 참조) 따라서 3연의 ‘꽃치조와 山에서 사노라네’는 화자가 자기가 감동한 느낌을 깨달으면서 노래하고 있는 것이다. 감탄이 중복된 셈이다.

### 3. <산유화>의 정보학적 분석

한 작가의 시 전체의 음소, 음절, 어절의 특징을 눈으로 파악하는 일은 쉽지 않다. 그러나 통계를 산출하는 프로그램을 이용하여 이들의 특징을 파악하면 자세히 이해할 수 있다. 여기서는 <산유화>라는 한 편의 시만을 대상으로 분석하기로 한다.

#### 3.1. 음소를 이용한 분석

음소(phoneme)는 어휘 의미의 변별을 초래하는 소리를 말한다. 예를 들면 ‘달’과 ‘딸’은 ‘ㄷ’과 ‘ㅌ’ 때문에 두 낱말의 어휘 의미가 변별된다. 국어에는 반모음을 포함하여 ‘/ㅂ, ㅃ, ㅍ/, /ㄷ, ㅌ, ㅌ/, /ㄱ, ㅋ, ㆁ/, /ㅈ, ㅉ, ㅊ/, /ㅅ, ㅆ, ㅎ/, /ㅇ, ㄴ, ㄹ/, /ㅈ, ㅉ, ㅊ/, /ㅅ, ㅆ, ㅎ/’ 등과 같이 22개의 자음 음소가 있다. 국어에는 ‘/ㅣ, ㅏ, ㅑ, ㅓ/, /ㅕ, ㅖ, ㅗ, /ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅡ/, /ㅛ, ㅜ, ㅠ, ㅡ, ㅣ/’ 등과 같이 10개의 기본 모음 음소가 있다. 이 음소는 지역별, 세대별로 약간 다르게 존재한다. 국어의 이중 모음은 ‘/ㅟ, ㅠ, ㅡ, ㅢ, ㅣ, ㅤ/, /ㅥ, ㅦ, ㅧ, ㅨ, ㅩ/’ 등과 같이 10개가 있다. 이처럼 다양한 음소에 따라서 어휘 의미가 변별되기 때문에 음소를 이해하는 일은 매우 중요하다.

프로그램을 이용하여 음소 빈도를 내면 초성자와 중성자 및 종성자의 음소의 빈도를 내주기 때문에 시인이 선호하는 음소를 통하여 시의 특징을 이해할 수 있을 것이다. 이들 빈도는 국어자료통합처리기인 ‘깜짝새’(SynKDP)를 이용하여 검색할 수 있다. 이제 <산유화>를 좀더 분석적으로 살펴보기 위하여 현대어로 고쳐서 음소 분석을 해보기로 한다.

<산유화>

산에는 꽃피네	산에서 우는 작은 새요
꽃이 피네	꽃이 좋아
갈 봄 여름 없이	산에서
꽃이 피네	사노라네
산에	산에는 꽃지네
산에	꽃이 지네
피는 꽃은	갈 봄 여름 없이
저만치 혼자서 피어 있네	꽃이 지네

이 시를 ‘깜짝새’(synKDP)<sup>4)</sup>를 이용하여 음소 빈도를 산출하면 다음과 같다.

4) 이 프로그램은 전주대학교 소강춘 교수와 김진규, 박진양 학생이 만든 것으로 ‘국어자료통합처리기’(synKDP)라고 부른다.

초성 ㄱ: 2 (1.05820%)	중성 ㄴ: 14 (7.40741%)	중성 ㄷ: 8 (4.23280%)
초성 ㄴ: 13 (6.87831%)	중성 ㄷ: 2 (1.05820%)	중성 ㄹ: 1 (0.52910%)
초성 ㄷ: 3 (1.58730%)	중성 ㄹ: 13 (6.87831%)	전체 음소수 : 189
초성 ㄹ: 1 (0.52910%)	중성 ㅁ: 1 (0.52910%)	음소 종류수 : 30
초성 ㅁ: 2 (1.05820%)	중성 ㅂ: 1 (0.52910%)	초성 음소수 : 78
초성 ㅂ: 11 (5.82011%)	중성 ㅅ: 8 (4.23280%)	(41.26984%)
초성 ㅅ: 24 (12.69841%)	중성 ㅇ: 17 (8.99471%)	초성 종류수 : 12
초성 ㅇ: 7 (3.70370%)	중성 ㅈ: 1 (0.52910%)	중성 음소수 : 78
초성 ㅈ: 1 (0.52910%)	중성 ㅊ: 14 (7.40741%)	(41.26984%)
초성 ㅊ: 1 (0.52910%)	중성 ㅋ: 2 (1.05820%)	중성 종류수 : 10
중성 ㅊ: 14 (7.40741%)	중성 ㆁ: 4 (2.11640%)	중성 음소수 : 33
중성 ㆁ: 1 (0.52910%)	중성 ㅅ: 2 (1.05820%)	(17.46032%)
중성 ㅅ: 7 (3.70370%)	중성 ㅇ: 1 (0.52910%)	중성 종류수 : 8

음소 빈도 결과에서 초성과 중성, 중성에 쓰인 음소의 빈도를 정확히 파악할 수 있다. 이 시에서 사용한 음소의 특징을 살펴보면 시인이 생각하는 음소의 기능을 파악할 수 있을 것이다.

이 시에는 유성음인 ‘ㄴ’이 가장 많이 쓰이고 있다. ‘산, 산에는, 저만치, 혼자’ 등에 쓰이고 있다. 이처럼 유성음 중 ‘ㄴ’을 많이 사용하는 것은 전통적인 시에서는 일반적인 현상이라 할 수 있지만 운율적인 효과를 염두에 두고 사용하고 있는 것으로 해석할 수 있을 것이다.

초성 ㄴ: 13 (6.87831%) 중성 ㄴ: 14 (7.40741%)

시인은 무성 마찰음인 ‘ㅅ’을 많이 사용하고, 파찰음인 ‘ㅈ, ㅊ’을 많이 사용하여 인상적인 어휘음을 사용하고 있음을 볼 수 있다.

초성 ㅅ: 11 (5.82011%) 초성 ㅈ: 7 (3.70370%) 중성 ㅈ: 8 (4.23280%)

또한 ‘꽃, 저만치, 혼자, 지다, 피다’에서 양순파열음 ‘ㄲ’, 성문마찰음 ‘ㅎ’, 파찰음 ‘ㅈ’ 등이 쓰이고 있다. 어두의 된소리로 된 ‘꽃’을 사용하고 있다. 파열음과 마찰음, 파찰음을 써서 비록 소리가 울리지 않는 무성음이지만 강렬한 인상을 주려는 강조의 효과를 노리고 있다고 볼 수 있을 것이다. 된소리도 파열음이므로 마찬가지다.

초성 ㄲ: 8 (4.23280%) 초성 ㅈ: 5 (2.64550%)

한편 모음을 살펴보면 양성모음인 ‘ㅏ, ㅑ’의 사용 빈도가 월등히 높고, 음성모음은 훨씬 적게 쓰이고 있다. 이런 현상은 양성모음을 통하여 ‘산’과 ‘꽃’의 밝은 이미지를 활용하는 것으로 이해된다.

중성 ㅏ: 14 (7.40741%) 중성 ㅑ: 13 (6.87831%)  
 중성 ㅓ: 8 (4.23280%) 중성 ㅕ: 1 (0.52910%) 중성 ㅗ: 7 (3.70370%)  
 중성 ㅣ: 17 (8.99471%)

### 3.2. 음절을 이용한 분석

음절 빈도를 산출하면 빈도가 많은 형태소의 유형을 짐작할 수 있다. 또한 특정 형태소의 빈도수를 확인할 수가 있다. 이 시를 KDP를 이용하여 음절 빈도를 산출하면 다음과 같다.

저 : 00000001 (1.28205%) [85.89744%]	갈 : 00000002 (2.56410%) [2.56410%]
흔 : 00000001 (1.28205%) [100.00000%]	봄 : 00000002 (2.56410%) [37.17949%]
치 : 00000001 (1.28205%) [92.30769%]	은 : 00000002 (2.56410%) [71.79487%]
노 : 00000001 (1.28205%) [24.35897%]	름 : 00000002 (2.56410%) [33.33333%]
종 : 00000001 (1.28205%) [87.17949%]	없 : 00000002 (2.56410%) [56.41026%]
라 : 00000001 (1.28205%) [30.76923%]	지 : 00000003 (3.84615%) [91.02564%]
아 : 00000001 (1.28205%) [52.56410%]	서 : 00000003 (3.84615%) [51.28205%]
만 : 00000001 (1.28205%) [34.61538%]	는 : 00000004 (5.12821%) [29.48718%]
어 : 00000001 (1.28205%) [53.84615%]	피 : 00000005 (6.41026%) [98.71795%]
사 : 00000001 (1.28205%) [38.46154%]	에 : 00000006 (7.69231%) [64.10256%]
있 : 00000001 (1.28205%) [82.05128%]	산 : 00000006 (7.69231%) [46.15385%]
새 : 00000001 (1.28205%) [47.43590%]	이 : 00000007 (8.97436%) [80.76923%]
작 : 00000001 (1.28205%) [84.61538%]	꽃 : 00000008 (10.25641%) [12.82051%]
요 : 00000001 (1.28205%) [67.94872%]	네 : 00000008 (10.25641%) [23.07692%]
자 : 00000001 (1.28205%) [83.33333%]	전체 음절수 : 78
우 : 00000001 (1.28205%) [69.23077%]	음절 종류수 : 31
여 : 00000002 (2.56410%) [66.66667%]	

음절 빈도를 내보면 다음에서 보는 것처럼 종결어미 ‘-네’, 주격조사 ‘-이’, 처격조사 ‘-에’, 특수조사 ‘은/는’, 명사 ‘꽃’과 ‘산’, 동사 ‘피-’가 음절 빈도가 가장 많이 쓰이는 핵심 음절임을 알 수 있다.

지 : 00000003 (3.84615%) [91.02564%]	산 : 00000006 (7.69231%) [46.15385%]
서 : 00000003 (3.84615%) [51.28205%]	이 : 00000007 (8.97436%) [80.76923%]
는 : 00000004 (5.12821%) [29.48718%]	꽃 : 00000008 (10.25641%) [12.82051%]
피 : 00000005 (6.41026%) [98.71795%]	네 : 00000008 (10.25641%) [23.07692%]
에 : 00000006 (7.69231%) [64.10256%]	

우리는 이 음절 빈도의 결과를 가지고 이 시에서 중요한 몇 가지 사항을 추출할 수 있다.

첫째, 명사 ‘꽃’과 ‘산’이 이 시의 중심 어휘이고 동사 ‘피-, 지-’가 중심어휘인 것을 바로 알 수 있다.

둘째, 종결어미와 조사의 쓰임을 통하여 이 시인이 가지고 있는 언어 표현의 속성을 찾아 볼 수 있다. 예를 들면 주격조사 ‘-이’는 ‘꽃이 피네, 꽃이 지네’의 구문에서 5회 쓰였다. 주격조사의 특성으로 보면 이 용례는 절대성, 불변성을 말하고 있는 것이다. 따라서 ‘꽃이 피네, 꽃이 지네’의 문장은 시인에게 있어서 진리이고 명제이다.<sup>5)</sup>

이처럼 음절 빈도를 내어 시적 특징을 검토하는 일이 시를 분석할 때 기본적이고 우선적인 일이 되어야 할 것이다.

### 3.3. 어절을 이용한 분석

어절 빈도를 산출한 결과를 ‘한글’의 소트 기능을 이용하여 가나다순으로 정렬하거나, 빈도수로 정렬을 해보면 여러 현상을 알 수 있다.

이 시를 프로그램 ‘깜짝새’를 이용하여 어절 빈도를 산출하면 다음과 같다.<sup>6)</sup> 어휘 빈도를 내는 것이 가장 이상적이나 주석 말뭉치(Tagged corpus)가 구축이 되어 있지 않으면 어휘 빈도를 낼 수 없다. 따라서 이 자료는 원시 말뭉치(Raw corpus)이기 때문에 어절 빈도만을 산출할 수 있다.

---

5) ‘타고 남은 재가 기름이 됩니다.’라는 구절은 절대 불변의 진리이다. 이 문장이 ‘타고 남은 재는 기름이 됩니다.’로 쓰일 경우는 진리가 될 수 없다. 국어에서 주격조사가 보여주는 중요한 특징이다.

6) 이 프로그램을 이용하면 어절의 음절수를 확인할 수 있다. 빈도를 내는 항목에 음절어의 빈도를 확인하는 항목이 마련되어 있다. 한 작가의 작품이 어떠한 음절어의 빈도를 갖는지 확인할 수 있는 중요한 기능이다.

갈 : 000002 (5.26316%) [5.26316%]	우는 : 000001 (2.63158%) [65.78947%]
꽃 : 000002 (5.26316%) [10.52632%]	있네 : 000001 (2.63158%) [68.42105%]
꽃은 : 000001 (2.63158%) [13.15789%]	작은 : 000001 (2.63158%) [71.05263%]
꽃이 : 000005 (13.15789%) [26.31579%]	저만치 : 000001 (2.63158%) [73.68421%]
봄 : 000002 (5.26316%) [31.57895%]	좋아 : 000001 (2.63158%) [76.31579%]
사노라네 : 000001 (2.63158%) [34.21053%]	지네 : 000003 (7.89474%) [84.21053%]
산에 : 000002 (5.26316%) [39.47368%]	피네 : 000003 (7.89474%) [92.10526%]
산에는 : 000002 (5.26316%) [44.73684%]	피는 : 000001 (2.63158%) [94.73684%]
산에서 : 000002 (5.26316%) [50.00000%]	피어 : 000001 (2.63158%) [97.36842%]
새요 : 000001 (2.63158%) [52.63158%]	혼자서 : 000001 (2.63158%) [100.00000%]
없이 : 000002 (5.26316%) [57.89474%]	전체 단어수 : 38
여름 : 000002 (5.26316%) [63.15789%]	단어 종류수 : 22

어절 빈도를 낸 결과는 위와 같다. 일단 빈도가 높은 다음 예들을 검토해 보기로 한다.

꽃 : 000002 (5.26316%) [10.52632%]  
꽃은 : 000001 (2.63158%) [13.15789%]  
꽃이 : 000005 (13.15789%) [26.31579%]

산에 : 000002 (5.26316%) [39.47368%]  
산에는 : 000002 (5.26316%) [44.73684%]  
산에서 : 000002 (5.26316%) [50.00000%]

피네 : 000003 (7.89474%) [92.10526%]  
피는 : 000001 (2.63158%) [94.73684%]  
피어 : 000001 (2.63158%) [97.36842%]

이 시에서 ‘꽃’은 격조사가 생략된 ‘꽃’과 특수조사가 붙은 ‘꽃은’, 주격조사가 붙은 ‘꽃이’가 연결되어 ‘꽃’이 가지는 의미와 기능이 다름을 알 수 있다. 주격조사가 생략될 때 우리는 이미 알고 있는 정보인 구정보(old information)에서 생략이 가능함을 알 수 있다. 그러나 우리가 모르는 신정보(new information)일 경우는 생략이 불가능하다. ‘꽃이 피네’의 경우도 역시 시인만이 알고 있는 꽃을 이야기하고 있는 것으로 해석할 수 있다. ‘산’도 역시 ‘산에, 산에는, 산에서’가 쓰이고 있는데, 이때의 ‘산’도 역시 다 그 의미와 기능이 다름을 알 수 있다.

따라서 어절 빈도를 통하여 그 시의 구체적인 어휘들의 특성을 살피는 일은 매우 중요한 작업이 될 것이다. 한편 ‘깜짝새’를 이용하여 용례사전을 만들면 어절들의 쓰임을 한 눈에 살필 수 있어서 예문을 확인하면서 이해하는 데 큰 도움을 받을 수 있다. <산유화>의 용례사전을 제시하면 다음과 같다.<sup>7)</sup>

7) 김소월의 시의 원본은 띄어쓰기가 되어 있지 않고, 표기가 옛한글의 표기로 되어 있어서 정보 처리하기가 곤

갈 : 갈 봄 여름 없이 <山有花1,3>  
 갈 : 갈 봄 여름 없이 <山有花4,3>  
 꽃 : 산에는 꽃 피네 <山有花1,1>  
 꽃 : 산에는 꽃 지네 <山有花4,1>  
 꽃은 : 피는 꽃은 <山有花2,3>  
 꽃이 : 꽃이 피네 <山有花1,2>  
 꽃이 : 꽃이 피네 <山有花1,4>  
 꽃이 : 꽃이 좋아 <山有花3,2>  
 꽃이 : 꽃이 지네 <山有花4,2>  
 꽃이 : 꽃이 지네 <山有花4,4>  
 봄 : 갈 봄 여름 없이 <山有花1,3>  
 봄 : 갈 봄 여름 없이 <山有花4,3>  
 사노라네 : 사노라네 <山有花3,4>  
 산에 : 산에 <山有花2,1>  
 산에 : 산에 <山有花2,2>  
 산에는 : 산에는 꽃 피네 <山有花1,1>  
 산에는 : 산에는 꽃 지네 <山有花4,1>  
 산에서 : 산에서 우는 작은 새요 <山有花3,1>  
 산에서 : 산에서 <山有花3,3>  
 새요 : 산에서 우는 작은 새요 <山有花3,1>  
 없이 : 갈 봄 여름 없이 <山有花1,3>  
 없이 : 갈 봄 여름 없이 <山有花4,3>  
 여름 : 갈 봄 여름 없이 <山有花1,3>  
 여름 : 갈 봄 여름 없이 <山有花4,3>  
 우는 : 산에서 우는 작은 새요 <山有花3,1>  
 있네 : 저만치 혼자서 피어 있네 <山有花2,4>  
 작은 : 산에서 우는 작은 새요 <山有花3,1>  
 저만치 : 저만치 혼자서 피어 있네 <山有花2,4>  
 좋아 : 꽃이 좋아 <山有花3,2>  
 지네 : 산에는 꽃 지네 <山有花4,1>  
 지네 : 꽃이 지네 <山有花4,2>  
 피네 : 산에는 꽃 피네 <山有花1,1>  
 피네 : 꽃이 피네 <山有花1,2>  
 피네 : 꽃이 피네 <山有花1,4>  
 피는 : 피는 꽃은 <山有花2,3>  
 피어 : 저만치 혼자서 피어 있네 <山有花2,4>  
 혼자서 : 저만치 혼자서 피어 있네 <山有花2,4>

---

란하다. 필자는 노천명의 시 전체를 가지고 어절 빈도와 용례 사건을 만들어 빈도가 높은 어절의 용례를 확인하는 방법으로 노천명의 시 세계를 검토한 바 있다. 예를 들어 빈도가 높은 어휘 ‘거리’를 용례사건에서 구체적인 그 쓰임과 의미를 확인할 수 있다.

<산유화>란 한 편의 시를 이용하여 분석하기 때문에 특별히 관심을 끄는 어휘가 많지 않다. 그러나 한 작가의 전체적인 시를 가지고 분석하면 어휘의 특징을 다양하게 확인할 수 있다. 몇 시인의 예를 간략히 들어보기로 한다.

정지용 시인의 시집을 검토하면 ‘오오’, ‘아아’ 감탄사를 많이 사용하고 있다. 감탄사 뒤에는 기쁨이나 놀람의 내용이 나오는 것이 일반적인데 세상에 대한 탄식이 나오고 있다. 시인이 바라본 세상에 대한 생각과 느낌을 알 수 있다.

오오, 異國種강아지야 내발을 빨어다오. <정지용, 카·페。프랑스>  
오오, 개인 날세야, 사랑과 같은 어질머리야, <정지용, 슬픈汽車>  
오오, 나는 차보다 더 날려 가라지는 <정지용, 슬픈汽車>  
오오, 사철나무 옆에 숨은 <정지용, 해바라기씨>  
형제여, 오오, 이 꼬리긴 英雄이야! <정지용, 말2>  
오오, 나의 幸福은 나의 聖母마리아! <정지용, 또하나다른太陽>  
悲哀! 오오 나의 新婦! 너를 위하여 <정지용, 幸福>  
오오 悲哀! 너의 不死鳥 나의 <정지용, 幸福>  
오오 알맞는 位置! 좋은 우아래 <정지용, 나무>  
오오! 新約의太陽을 한아름 안다. <정지용, 나무>  
오오 좌기스대에 목을 높히 달니우고 <정지용, 勝利者金안드리아>  
오오 그들은 악한 권세로 죽인 <정지용, 勝利者金안드리아>

부사 ‘아직’은 이전과 같은 상태로 아직 때가 되지 않은 상태를 말한다. 이상화의 시에서도 ‘그러나, 애달픔, 추운’과 같은 부정적 이미지의 시어들과 함께 쓰이고 있다. 암울한 시대상황이 계속되고 있으며, 시인이 바라는 때에 이르지 못하고 있음을 표현하고 있다.

진저리나는 절름발이 노릇을 아직도 하는 것이다 <이상화, 시인에게>  
그 웃음을 아직도 <이상화, 어머니의 웃음>  
아직도 사람은 수확(收穫)을 바라고 있다. <이상화, 폭풍우를 기다리는 마음>  
숙명(宿命)이 주는 자족(自足)이 아직도 있다. <이상화, 폭풍우를 기다리는 마음>  
애인아, 내 몸이 어제같이 보이고 네 몸도 아직 살아서 내 곁에 앉았느냐. <이상화, 이별을 하느니>  
아직도 남은 애달픔이 있으려니 <이상화, 무제>

이육사의 작품 속에서 ‘소리’는 ‘물체의 진동에 의하여 생긴 음파가 귀청을 울리어 귀에 들리는 것’을 의미하는 명사로서의 ‘소리’뿐만 아니라 대체로 ‘창조의 소리’이면서 생명의 기운이 움트고 있는 긍정적 이미지를 표현한다. 이육사 시에서 ‘소리’는 ‘광명(光明)’을 뜻하며 해방과 희망을 상징한다.

우는 소리 들렸으랴 <이육사, 광야>  
돌쪽마다 소리 석양(夕陽)을 <이육사, 소년(少年)에게>  
가시네와 종달새소리에 반해 <이육사, 초가(草家)>  
별들에게도 종소리 저문 <이육사, 황혼(黃昏)>

시냇물 소리 같아서 <이육사, 황혼(黃昏)>  
 기쁨의 소리! 해조는 <이육사, 노정기>  
 해조의 소리! <이육사, 노정기>  
 자무락질치는 소리 약간 <이육사, 소공원>  
 닭소리나 들리면 <이육사, 독백>  
 첫대 소리에다 옷끈을 <이육사, 아미>  
 오 소리는 장엄(莊嚴)한 <이육사, 해조사>  
 소리! 고이한 <이육사, 해조사>  
 고이한 소리! 지축(地軸)이 <이육사, 해조사>  
 것봄의 소리! <이육사, 해조사>  
 해조(海潮)의 소리! <이육사, 해조사>

노천명의 시에서 ‘거리’라는 어휘는 주로 ‘적정한 거리, 낮은 거리, 충충한 거리, 슬픈 거리, 죽은 거리’ 등 아주 부정적인 거리로 묘사되고 있음을 볼 수 있다. 대체로 거리는 인간의 다양한 삶의 모습을 보여주는 제재인데 노천명의 시에서는 시인의 마음을 읽을 수 있는 소외된 거리로 묘사되고 있다.

거리. 낮 익은 얼굴들이 없어 낮은 거리.<노천명, 적정한거리2.1.>  
 거리. 친구들은 가고 적정한 거리.<노천명, 적정한거리1.1.>  
 거리. 흐터졌든 겨레들 모여든 거리.<노천명, 오월1.4>  
 거리거리 거리거리 골목골목.<노천명, 삼월의노래3.4>  
 거리 거리 거리에.<노천명, 약속된날있거니6.3>  
 거리거리엔 都市와 村落 거리거리엔 사슬이 풀린 사람들.<노천명, 불덩어리되어3.2>  
 거리거리엔 싱가포르의 거리거리엔 祖國의 旗사말이.<노천명, 노래하자이날을2.2>  
 거리낌에도 조그마한 거리낌에도 .<노천명, 자화상1.11>  
 거리들. 搖鈴을 흔들며 조용이 지나는데 낮은 거리들.<노천명, 만가2.1>  
 거리로 善良한 親舊들이 거리로 거리로 쏟아집니다.<노천명, 봄의서곡1.5>  
 거리로 善良한 親舊들이 거리로 거리로 쏟아집니다.<노천명, 봄의서곡1.5>  
 거리로 이 거리로 오늘은 喪輿가 한채 지나갑니다.<노천명, 만가1.3>  
 거리들. 어둠과 불안이 충충한 거리들.<노천명, 송년부4.1>  
 거리서 生의 孤寂한 거리서 그대 날 불렀건만.<노천명, 말안코그저가려오2.5>  
 거리에. 거리 거리에.<노천명, 약속된날있거니6.3>

이상 프로그램으로 간단히 살펴볼 수 있는 몇 가지의 특징만을 살펴보았다. 현대시의 연구에서도 빈도, 해당 용례 추출을 통한 특정한 어휘의 특징 연구, 의미론적인 고찰 등이 필요할 것이다.

#### 4. <산유화>의 음소 배열과 리듬

시에 쓰인 음소는 리듬과 관련된다. 유성음, 무성음, 파열음, 파찰음 등 음소의 특징을 시인이 이용하는 경우가 많다. <산유화>에서 우리는 시인이 음소를 배열하는 리

들을 발견하게 된다. 어절의 첫 음소의 배열 순서를 보면 다음과 같다.

(1연)	(2연)	(3연)	(4연)
ㅅ-ㄱ-ㅍ	ㅅ	ㅅ-ㅇ-ㅈ-ㅅ	ㅅ-ㄱ-ㅈ
ㄱ-(ㅈ)-ㅍ	ㅅ	ㄱ-(ㅈ)-ㅈ	ㄱ-(ㅈ)-ㅈ
ㄱ-ㅂ-ㅇ-ㅇ-(ㅅ)	ㅍ-ㄴ-ㄱ-(ㅈ)	ㅅ	ㄱ-ㅂ-ㅇ-ㅇ-(ㅅ)
ㄱ-(ㅈ)-ㅍ	ㅈ-ㅎ-ㅍ-ㅇ	ㅅ	ㄱ-(ㅈ)-ㅈ

어절의 첫 음소의 음의 유형을 살펴보면 다음과 같이 나눌 수 있다. 괄호 안의 자음은 받침이지만 연철되어 첫소리로 나는 소리이다.

- 비음 - ㅇ
- 마찰음 - ㅅ(연음), ㅎ(연음)
- 파찰음 - ㅈ(무기 연음), ㅊ(유기 경음)
- 파열음 - ㅂ(무성 무기 연음), ㅍ(무성 유기 경음), ㄱ(무성 무기 연음), ㄱ(무성 무기 경음)

연음은 조음기관의 긴장을 수반하지 않는 소리이며, 경음은 조음 기관의 긴장을 수반하는 소리이다. 파열음은 두 조음 기관을 접촉시켜 구강 안에 막음을 형성했다가 개방하면서 조음하는 소리이고, 마찰음은 두 조음 기관을 바짝 접근시켜 좁혀진 틈 사이로 기류를 통과시켜 조음하는 소리이다. 파찰음은 두 조음 기관을 접촉시켜 막음을 형성했다가 완전히 개방하지 않고 조금만 개방해서 두 조음 기관의 좁은 틈 사이로 기류를 통과시켜 조음하는 소리이다. 막음의 개방 후 ‘ㅎ’과 비슷한 마찰 소음을 일으키면서 분출되는 강한 기류를 기(aspiration)라고 하는데 기를 수반하는 소리를 유기음, 기를 수반하지 않는 소리를 무기음이라 한다.(이호영,1996:41-101참조)

따라서 연음보다는 경음이, 무기음보다는 유기음이 훨씬 인상적인 음이고, 마찰음보다는 파찰음이, 파찰음보다는 파열음이 훨씬 인상적인 음이라 할 수 있다.

‘산유화’의 첫 음소 배열이 보여주는 특징을 제시하면 다음과 같다.

첫째, 1연과 4연에서 각 행이 보여주는 음소 배열은 똑 같다. 1행에서 2행으로 가면서 된소리 ‘ㄱ’, 파찰음인 ‘ㅊ’, 양순파열음 ‘ㅍ’ 음을 사용하면서 2행의 음소의 인상을 강화하고, 3행에서 4행으로 가면서 마찬가지로 4행의 음소의 인상을 강화하고 있다.

‘ㅅ-ㄱ-ㅍ, ㄱ-(ㅈ)-ㅍ’에서 보면 연음에서 무기음으로 다시 유기음으로 발음이 되면서 인상이 강화되고 있다.

둘째, 2연과 3연을 보면 시인의 음소에 대한 이해가 매우 돋보인다. 2연의 1행과 2행은 부드러운 음인 무성 마찰음 ‘ㅅ’으로 시작하고 3행과 4행에서 강화된 음을 사용하고 있다. 즉 3행 ‘ㅍ-ㄴ-ㄱ-(ㅈ)’의 음소 배열은 유기 경음과 무기 경음으로 배열되어 음이 매우 강하다. 4행 ‘ㅈ-ㅎ-ㅍ-ㅇ’의 음소 배열은 연음으로 시작하여 경음으로 나가는 배열로 음이 강화되고 있다.

반면에 3연에서는 거꾸로 1연에서 2연으로 갈 때 인상이 강화되지만, 1연 ‘ㅅ-ㅇ-ㅈ-ㅅ’의 음소 배열은 마찰음에서 파찰음으로 갔다가 다시 마찰음으로 돌아오면서 음이 부드러워지고, 2연 ‘ㄱ-(ㅈ)-ㅈ’은 경음에서 연음으로 돌아오면서 부드러워지는 음소

구조를 보인다. 3행과 4행에서는 부드러운 무성 마찰음 ‘ㄱ’이 사용되고 있다. 2연과 3연의 행이 보여주는 음소 배열은 정반대의 구조를 갖는 독특한 음소 배열을 보여준다.

셋째 1연과 4연의 각 행의 음소 배열을 비교해 보면 ‘피다’와 ‘지다’의 첫 음소가 각각 유기음과 연음으로 이루어져 중요한 어휘의 음의 인상이 대조되는 효과를 보여주고 있다.

이러한 음소를 이용한 인상적인 특징을 이용한 이 시의 특징은 이미 시인이 이를 염두에 두고 창작한 것으로 이해된다.

## 5. 결론

필자는 <산유화>를 중심으로 정보를 처리하여 국어학적인 입장에서 작품을 분석하였다. 시와 소설을 연구할 때, 원래의 텍스트를 입력하여 정보를 처리하거나 정보를 검색하는 일은 매우 의미가 있는 일이다.

시의 경우, 국어 자음과 모음이 어우러져 음을 만들고 음절이 리듬을 만들기 때문에 음소와 음절, 어절의 특징을 이해하여 시의 특성을 파악하는 일은 아주 기초적인 일이다. 특히 국어 자음과 모음의 음성학적 특징을 이해하는 일이 매우 중요하다. 시인들은 자기 시의 운율적인 특성을 위해 국어의 음소를 적절히 활용하기 때문이다. 여기서는 한 편의 시를 가지고 분석을 했지만 김소월 시 전체를 가지고 분석한다면 아주 특징적인 현상을 발견할 수 있을 것이다. 또한 한 작가의 작품 전체에 나타나는 어휘를 한꺼번에 검색할 수 있어서 의미를 이해하는 데 도움을 받을 수 있다.

소설의 경우, 등장인물의 언어를 가지고 그 성격을 파악할 수 있고, 등장인물의 빈도와 발화량을 가지고 그 중요성을 가늠할 수도 있다. 등장인물이 사용하는 방언의 특징을 가지고 여러 가지로 해석이 가능하다.<sup>8)</sup>

따라서 앞으로 문학 작품을 분석하는 일에 국어정보학적인 방법이 우선적으로 적용되어 가장 기초적인 작업이 이루어져야 할 것이다.

## <참고문헌>

- 소강춘(2002), 정보처리 프로그램에 대하여 -SynKDP를 중심으로-, 한국어와 정보화, 태학사, 598-604.
- 이태영(2004ㄱ), <혼불>에 쓰인 방언의 기능과 등장인물의 성격, 혼불의 언어세계, 혼불학술총서 2, 321-336쪽.
- 이태영(2004나), 문학 작품과 방언 연구, 한국어학 25호, 102-107쪽.
- 이호영(1996), 국어음성학, 태학사, 41-101쪽.
- 한길(1991), 국어 종결어미 연구, 강원대 출판부, 81-82쪽.

## Abstract

### The Study on <Sanyuhwa> from the viewpoint of Korean Informatics

8) 이태영(2004ㄱ)에서는 최명희의 <혼불>에 나타나는 등장인물인 ‘옹구네’의 성격을 그 발화를 가지고 고찰한 바 있다.

Lee, tae-yeong(chonbuk univ.)

This study is aimed to present the method of study which is helpful for understanding and analyzing poetry. I intend to explicate <Sanyuhwa> focusing on grammatical morphemes and clarify a variety of characteristics shown in this poem using frequency of phonemes, syllables and words.

The importance of grammatical morphemes are definitely revealed in this poem as well. The meaning of 'san' in 'san-e-neun, san-e-seo and san-e' shown in <Sanyuhwa> becomes different according to the grammatical morphemes. 'kkot, kkot-i' is the same case. The function which '-ne' shows in 'pi-ne' has a dominant influence in explicating the whole poem.

In the pursuit of more analytic study on this poem, I will figure frequency of phonemes, syllables, words and vocabularies and look into the characteristics of frequency in the viewpoint of informatics. Accurate frequency used in a initial sound, the vowels (and semi-vowels) of a Korean syllable and a final consonant will be obtained through the consequence of phonemes frequency. According to syllable frequency, it turns out that a final-ending '-ne', a subject particle '-i', a locative '-e', a postposition '-eun/neun', nouns 'kkot', 'san' and a verb 'pi-' are the most frequently used key syllables. It will be very a important task to look into the attributes of the specific vocabularies expressed in a poem through frequency of words.

We can easily find out there is the well-arranged rhythm intended by its author in <Sanyuhwa>. When we pay attention to the arrangement order of first phonemes in words, we can figure out that Kim So-weol wrote <Sanyuhwa> keeping the impressive attributes of the arrangement in his mind.

Key words : grammatical morphemes, frequency, phonemes, syllables, words, phoneme arrangement