

# 언어의 과학

주하진

주의. 다음 글에는 아래 작품들에 대한 스포일러가 포함되어 있습니다.

<프로메테우스> (2012, 리들리 스콧)

<컨택트> (2016, 드니 빌뇌브)

<패션 오브 크라이스트> (2004, 멜 깁슨)

<아포칼립토> (2006, 멜 깁슨)

<반지의 제왕> 시리즈 (1954 ~ 1955, J.R.R. 톨킨)

<실마릴리온> (1977, J.R.R. 톨킨)

## 여는 글 - <프로메테우스>와 인도유럽조어

영화 <프로메테우스>는 ‘에이리언’ 시리즈의 첫 번째 프리퀄 영화로 <마션>, <글래디에이터>, <에이리언>을 연출한 리들리 스콧 감독이 다시 메가폰을 잡았지만, 그 방향성이 기존 시리즈와 좀 달라서 호불호가 굉장히 많이 갈리는 작품이다. 리들리 스콧이 제작하는 영화는 전부 하나같이 시각적인 임팩트가 엄청나기 때문에 비주얼리스트 감독이라는 별칭으로도 불린다.

이 영화의 처음에는 마이클 패스벤더가 연기한 ‘데이빗’이라는 인

조인간이 우주선 안에서 여가 활동을 즐기는 장면을 볼 수 있다. 자전거를 타며 농구를 하거나 음식을 먹으며 스크린에 나오는 AI와 함께 언어 연습을 하는 등 실제 인간이 수행하기 어려운 행동을 보여주는 연출로, 작중 데이빗이 인간을 뛰어넘는 신체 능력과 지능을 가졌다는 정보를 제공한다.

내가 이 중 관심을 가진 장면은 데이빗이 음식을 먹으면서 스크린에 나오는 AI와 함께 언어 연습을 하는 장면이다.



AI는 다음과 같은 대사를 읊는다.

이런 발음법은 인도유럽어족에선 준언어에 속하지만 고대 언어에선

소리 단위에 불과했어요.

그럼 고대어로 해보죠. 따라해보세요.

그 후에 AI가 아래의 문장을 구사하고 데이빗이 따라한다.

*\*hjewis jasmə hwælnə nahəst ak<sup>w</sup>unsəz dadʔkta.*<sup>1)</sup>

나는 **언어학**에 상당히 관심이 많아서 어떤 언어적 요소가 등장하면 엄청난 호기심을 가지고 보는 편이다. 영화에서 나온 대사로만 보면 이 고대 언어와 인도유럽어족을 별개로 언급했기 때문에, 데이빗이 구사한 고대 언어가 재구된 **원시 인도유럽어 Proto-Indo-European Language**, PIE일지 아닐지는 아직 모르는 것이었다. 그래서 이 언어가 과연 어떠한 언어일까 검색을 해봤다. 데이빗이 구사한 언어는 원시 인도유럽어가 맞았다.

내가 언어학을 좋아하는 데에는 그것이 꼭 진화생물학을 연상하게 있는데, 뒤에서 더 자세히 말하게 될 것이다. 하지만 살짝만 맛보기 위해, 잠시 원시 인도유럽어에 대해 간략히 설명하겠다.

굴절어라든가 교착어라든가 하는 복잡한 형태론적 성질들은 제쳐두고 얘기해보자. 원시 인도유럽어는 ‘인도유럽조어’로도 불리며, 이름에서 나와 있듯이 인도유럽어족의 조어이다. 인도유럽어족이라는 개념 자체가 매우 생소할 수 있지만, 일단 이 어족의 언어를 쓰는 인구가 전 세계에서 25억 명 정도라는 것을 알아두자. 우리에게 매우 익숙한 언어들이 이 어족에 속한다. 이름대로 유럽의 대부분의 언어와 중앙아시아의 대부분의 언어가 이곳에 속한다. 힌디어, 산스크리트어, 페르시아어, 구자라트어 등 서아시아-남아시아 언어부터

---

<sup>1)</sup> 국제음성기호 International Phonetic Alphabet, IPA로 써 있다. 대충 ‘헤위스 야스머 헬너 나허스트 아크운서즈 다드룩타.’ 정도로 발음한다고 보면 된다.

독일어, 영어, 스웨덴어, 노르웨이어, 프랑스어, 이탈리아어, 스페인어, 포르투갈어, 러시아어, 우크라이나어, 벨라루스어, 폴란드어, 슬로바키아어, 그리스어, 아르메니아어 등의 유럽 언어들까지.

이들 전부가 태초에는 한 언어, 인도유럽조어로부터 분기된 것이다. 따라서 이 사멸된 언어가 어떤 형태였을지는 같은 어족임이 확인된 수많은 후손 언어들로부터 추측할 수 있다. 인도유럽조어는 진화생물학의 화석과도 같은 역할을 하는 문자가 없기 때문에 이와 같은 방식으로만 연구될 수 밖에 없는데, 그래도 공통조상이 되는 언어들 중 가장 많이 재구성된 언어이다. 학자마다 구사되는 형태가 다르더라도, 어느 정도는 학계에서 정설로 여겨지는 의견이 있다는 의미다.

<프로메테우스>에서 읊어진 문장은 재구성된 인도유럽조어를 가지고 처음으로 만든 이야기<sup>2)</sup>의 첫 문장이다. 대략 이런 뜻을 가지고 있다.

<i>hjewis</i>	<i>jasmə</i>	<i>hwælnə</i>	<i>nahəst</i>	<i>ak<sup>w</sup>unsəz</i>	<i>dad<sup>r</sup>kta</i>
<i>sheep</i>	<i>on which</i>	<i>wool</i>	<i>not-was</i>	<i>horses</i>	<i>saw</i>
<i>A sheep that had no wool saw horses.</i> 털 없는 양이 말들을 보았다.					

---

2) 슐라이허의 우화 Schleicher's fable라고도 불린다.

## 조어의 연구와 인공지능

앞서 말했다시피, 사람들은 인도유럽조어가 다른 조어祖語들에 비해 훨씬 많이 재구축되어 있다고 말한다. 하지만 문자로 남아있지 않은 언어의 한계로 그것이 실제로 당시에 어떻게 발화되었을지는 아무도 모르고, 그래서 실은 학자마다 주장하는 형태가 다르다. 영화에서 나온 문장을 학자들마다 구사한 것들을 보면 다음과 같다.

h <sub>2</sub> ówis, (H)jésmin h <sub>2</sub> wlh <sub>2</sub> néh <sub>2</sub> ne éh <sub>1</sub> est, dedorke (h <sub>1</sub> )ékwons. - Lühr (2008)
owis, jāi wlnā ne eest, dedorke ékwons. - Voyles and Barrack (2009)
h <sub>2</sub> áwej h <sub>1</sub> josmėj h <sub>2</sub> wlh <sub>1</sub> náh <sub>2</sub> né h <sub>1</sub> ést, só h <sub>1</sub> ékwons derkt. - Melchert (2009, revisited 2014)
출처: Wikipedia

영화에 쓰인 버전은 Lühr (2008)의 버전을 바탕으로 발음을 단순화한 것으로 추측된다. 위 세 문장들을 보면 문장에 쓰인 어휘들이 어느 정도 일관성을 띠긴 하지만 그조차도 완전히 일치하는 것을 찾기 힘들며, 어순도 학자마다 다 다르다. 사실 비전문가인 내가 보기로는, 이건 단순히 다른 정도가 아니라 상이한 걸로 보인다. 거의 이탈리아어 - 프랑스어 수준으로 다른 언어들을 모아놓은 것 같다.

그럼 사람들은 도대체 어느 부분을 보고 인도유럽조어의 재구가 잘 이루어졌다고 하는 것일까? 일단 인도유럽조어에서 가장 잘 재구

되어 있는 부분은 음운이다. 가져온 예문들만 보더라도 재구된 인도유럽조어를 작성하는 데 쓰인 음운들은 따로 차이가 없다.  $k$ 나  $\acute{e}$ ,  $h_1$ ,  $h_2$ ,  $j$ , .... 보통 다른 조어들은 음운이 어떻게 재구되었을 것인가에서부터 논쟁이 있는 경우가 대부분이라고 하며, 인도유럽조어는 음운  $a$ 의 존재 정도를 제외하면 대부분의 음운의 존재성이 합의되어 있다.

어휘, 통사(문법) 같은 경우는 대부분 제대로 합의가 안 되어 있다. 그러니까 조어 연구에선 이 정도의 합의만 이루어졌어도 ‘충분한 재구가 이루어졌다’고 보는 것이다.

왜냐하면 조어 연구는 그 자료가 매우 제한되어있기 때문이다. **언어학은 언어를 과학적 방법으로 다루는 학문이다.** 언어학자들은 거대한 언어 데이터를 가지고서 어떻게 음운이 변화해왔고 어휘, 통사가 언제 어떻게 갈라졌는지 알기 위해 데이터를 분석하고 그것을 나름의 방법대로 해석한다. 언어학도 다른 과학들과 마찬가지로 그러한 과정을 통해 이론이 형성된다. 인도유럽조어에 대한 주장이 하나로 정리되지 못하고 있는 이유는 결국 인도유럽조어가 이론이 아닌 가설이기 때문이다. 이를테면 위에서 말한 음운  $a$ 의 존재성 같은 것도 누군가는 “ $a$ 라는 음운이 따로 존재했다”라고 하고, 누군가는 “지금의  $a$ 는  $h_2e$ 로부터 온 것이다”라고 한다. 이들 모두가 가설이다. 가설끼리 부딪히고 통일이 되지 않는 이유는 그것이 검증되지 않았기 때문이다. 더군다나 이러한 조어처럼 검증될 수 없는 것들에 대해선 논쟁이 계속될 수 밖에. 지금 당장 우주론만 보더라도 암흑 물질의 정체에 대한 논쟁, 암흑 에너지의 실재성에 대한 논쟁 등 아직

입증되기 어려운 것들에 대한 가설들이 얼마나 팽배하는가.

그런데 그런 생각을 해 봤다. 우리 인류는 현재 원하는 대상이 어떻게 작동하는지 모르는 영역에 대하여, 그것을 인공지능에게 맡겨서 스스로 규칙을 만들도록 한다. 그 규칙은 인간이 이해할 수 없는 것일 수도 있지만, 결국 원하는 것을 해내도록 한다. 데이터에 근거한 것이기 때문이다. 그리고 데이터가 매우 많은 곳에서 항상 이용되는 인공지능이, 통시언어학에서 이용되지 않으리란 법은 없다. 자연어<sup>3)</sup> 처리는 이미 인공지능을 통해 고도로 발전하였고, 지구에 존재하는 자연어<sup>4)</sup>에 대해선 이미 음운, 어휘, 통사가 많이 연구되었다. 따라서 서로 이미 계통이 밝혀져 있는 언어들 사이의 음운, 어휘, 통사 데이터를 줘서 그 변화를 인공지능에게 학습시킬 수 있으리라. 과거에서부터 현재에 이르기까지.

이미 음운, 어휘, 통사의 변동에는 일정한 규칙이 있음이 밝혀져 있고 그것을 통해 이론들이 정립된 것이지만, 이렇게 만든 인공지능은 그 이론들을 전혀 모르기도 오히려 현존하는 이론들보다도 변동을 정확히 찾아낼 수 있을지도 모른다. 그리고 이 인공지능은 우리의 목표를 위해 사용될 수도 있다. 즉, 공통 어족인 언어에 대해 인공지능이 학습한 언어의 변화 규칙을 적용시켜서 그 조어를 이끌어

---

3) 공학적인 의미에서의 자연어는 인간의 언어를 의미한다.

4) 언어학적인 의미에서의 자연어는 인간에 의해 자연적으로 발생하고 진화한 언어를 말한다.

낼 수도 있을 거란 것이다.

혹자는 ‘설마 언어학자들이 그런 시도를 안 해봤을까?’라는 생각을 할텐데, 아무리 찾아봐도 그런 기사가 나오질 않았다. 분명히 Language 같은 대형 저널에 이런 연구가 시도되고 있다거나 시도되었다는 소식이 올라왔다면 틀림없이 언론에서 대서특필할 것인데, 기사가 없다는 건 저널에도 그런 소식이 없다는 거다.

도대체 언어학자들은 지금 뭘 하고 있는 걸까? 언어학자는 결국 과학자다. 과학자들은 끊임없이 자신을 의심하며, 연구 결과를 의심하며, 이론을 의심하며 그것을 교정하려고 한다. 다른 과학자들의 자신들의 이론에 대한 끊임없는 개선을 위해 힘쓰며 인공지능을 사용하는 것처럼, 언어학에서도 인공지능은 이미 활발하게 쓰이고 있을 것이다. 아마도 그러한 연구가 진행되고 있는데 그런 정교한 알고리즘을 구성하는 데에 시간이 예상보다 너무 오래 걸리는 것이거나, 앞서 말한 긍정적인 성과가 아직 없어서 저널에 올라오지 않은 것이거나, 아니면 이제 와서 막 시도해보고 있는 걸 수도 있다. ... 최악의 경우 이런 고대 언어의 재구에 인공지능을 이용할 생각을 못 했거나.

## 영화와 인간의 언어

다시 ‘영화’로 돌아가 보자. 이런 식으로 사멸된 언어, 즉 사어가 영화 내에서 구사되는 것은 찾아보기 쉽지 않다. 아니, 언급조차 안 되는 경우가 많으리라. 그나마 떠오르는 영화는 <컨택트>인데, 그것



은 외계 언어와 사피어-워프 가설 Sapir-Whorf hypothesis를 주제로 한다. 사이퍼-워프 가설은 언어가 사고를 지배한다는 가설이다. 그래서 그나마 언어적 요소가 부각되지만, 실제 언어의 특성보다는 사피어-워프 가설에 초점을 맞춘지라 세세한 언어적 요소의 언급을 많이 줄인 편이다. 이 영화의 초반부에선 주인공인 언어학자 ‘루이즈’와 ‘웨버 대령’의 대화 중에서 산스크리트어 문헌인 ‘전쟁’이 짧게 언급되기도 한다. 영화 자체에서도 외계 문자가 원형으로 나오거나, 외계 언어가 그들의 사고를 따라 비선형적인 시간대를 나타내는 등 흥미로운 설정은 많이 등장한다.

<프로메테우스>나 <컨택트>처럼 사어는 매우 짧게 언급되는 영화도 찾기 힘든데, 사어로 만들어지는 영화는 더더욱 찾기 힘들 것이다. 하지만 **멜 깁슨** 감독은 그 어려운 것을 해냈다. 멜 깁슨은 그의 역작인 <브레이브 하트>(1995)에서 감독이자 주연 배우로 활약하여 수많은 상을 휩쓴 엄청난 인물이다. 그는 무려 당시 역사적 상황을 재현해내겠다는 일념 하나만으로 사어를 가지고 두 편의 영화를 제작했다. 두 편이나!

그 중 하나는 영화 <패션 오브 크라이스트>(2004)이다. 예수가 십자가형을 받기까지의 마지막 12시간을 그려낸 영화로, 잔혹함에 가감이 없어 그의 수난을 가장 사실적으로 표현했다고 평가받는다. 영화 <데드풀>이 나오기 전까지는 R등급 (한국으로 치면 청소년 관람불가 등급에 해당한다) 영화 중 가장 흥행한 작품이기도 했다. 이 영화가 놀라운 것은 영화에서 쓰인 언어가 당대 실제로 사용된 **아랍어**이기 때문

이다. 정확히는 아랍어화된 히브리어이며, 이것은 현재 라틴어처럼 전례로만 사용될 뿐 현실에선 사어이다. 현대어와 비교해본다면 시리아어와 아랍어, 히브리어와 가장 가까우며<sup>5)</sup>, 독자적인 문자 또한 아랍 문자, 히브리 문자와 같은 아브자드<sup>6)</sup>의 형태이다.

저는 선생님께서 가시는 곳이면 어디든지 따라가겠습니다. 감옥이  
라도, 죽음까지라도!

*Ana'atid m'hak ba'athrak al-di th'zil tamma, adoni, al bet-s øhar.  
af al mot.*

- 작중 베드로의 대사. 예수가 자신의 죽음을 예견할 때.

감독은 실제 1세기 예루살렘의 현장을 재현하고자 한 것이다. 사실 영화 <패션 오브 크라이스트>는 자막조차 없을 뻔 했다. 아예 자막까지 없애면서까지 당시의 현장을 관객들에게 체험시켜주고자 했던 것이다. 의도는 이해되지만, 실제로 그랬더라면 관객들은 영화를 보며 아무것도 알아듣지 못했을 것이다. 아무도 쓰지 않는 언어이기 때문이다. 성경에 실제로 수록되어 있는 “엘리 엘리 라마 사박다니” 정도나 알아들었으려나.

---

<sup>5)</sup> 아랍어, 아랍어, 시리아어, 히브리어는 모두 **셈어군**에 속한다.

<sup>6)</sup> 자음만을 표기하는 문자 체계를 말한다. 아랍어의 경우 ktb라는 아랍어 단어를 'kataba'라고 읽으면 '~를 쓰다'라는 동사가, 'kitaab'라고 읽으면 '책'(단수형)이 되는 식이다.

감독은 2004년 이 영화를 만들고, 과거를 다루는 영화에 그 당시의 언어를 사용하는 것에 재미가 들렸는지 2006년, 이번에는 <아포칼립토>라는 영화를 제작한다. 이 영화에선 16세기 마야 문명을 다루었다. 영화는 마야 문명의 몰락 직전 한 부족의 전사의 이야기이고, 마지막은 쿵키스타도르들의 침략으로 끝난다. 시작부터 끝까지 **마야어**로 이야기가 진행되며, 다만 여기서 사용된 언어는 사멸된 언어는 아니고 실제 사용되고 있는 마야어 중 하나를 사용하였다고 알려져 있다.

## 언어에 대한 인간의 선택과 자연 선택

잠시 내가 언어학을 좋아하는 연유를 밝히는 것이 좋을 것 같다. 앞에서 나온 언어의 복잡하면서도 간단한 구성 요소들 - 음운, 어휘, 통사의 다양성으로부터 느낀 매력도 있지만, 또다른 이유가 있다.

어렸을 때부터 우리 집은 기본적으로 세상이 7일 만에 창조되었다는 성경의 이야기를 그대로 믿고, 노아의 홍수로 지구 전체가 물에 잠겼다가 드러났다는 얘기를 그대로 믿어 왔다. 예수가 사흘 만에 부활했다는 것에 대해선 사실이라고 하자, 그것이 핵심 신앙이고 반박할 증거도 없으니. 하지만 지구의 탄생이나 우주의 탄생 같은 물리적 현상에 대해선 성경에서 말하는 창세기의 이야기들이 틀렸다는 증거가 수없이 많이 발견되고 연구됐다. 우리는 성경의 창세기에 수록된 신화적인 이야기에 대해선 신이 세상을 창조했다는 정도만 받아들이고, 구체적인 이야기들은 신화로써 받아들이는게 맞는 것이다.

하지만 적어도 우리 집에서는 창조과학이라는 유사과학이 신앙의 일부라고 봤고, 그런 의미에서 진화론은 신앙 그 자체에 반기를 드는 행위나 마찬가지였다.

어쩌면 그런 반기를 드는 행위로써 내가 언어학을 좋아하게 되었을지도 모른다. 언어의 진화는 진화라는 것이 생물 외의 다른 분야에서조차 적용될 수 있는 보편 원리라는 것의 대표적인 예시이다. **로망스어**라는 어군이 있는데, 이것은 지금은 사멸한 라틴어 Latina Vulgata, 민중 라틴어로부터 파생한 언어들에 말한다. 스페인어, 이탈리아어, 포르투갈어, 프랑스어 등의 언어들이 그러하다. 각 언어로 ‘로망스어’를 어떻게 쓰는지 정리해보면 다음과 같다.

라틴어	Linguae Romanicae
이탈리아어	Lingue romanze
스페인어	Lenguas romances
포르투갈어	Línguas românicas
프랑스어	Langues romanes

유사성이 보이는가? 위 단어들은 모두 ‘로마 계통의 언어’를 뜻하고, 모두 ‘로마 계통의’를 의미하는 단어가 앞쪽의 ‘언어’를 수식하는 순서로 배치되어 있다. 또한 ‘언어’를 뜻하는 단어가 모두 같은 어원을 공유한다는 것이 보일 것이다. 이것은 라틴어라는 한 언어로부터 이탈리아어, 스페인어, 포르투갈어, 프랑스어 등이 갈라져나와 현재

쓰이고 있음을 의미한다. 라틴어는 현재 사멸하고 없고, 그로부터 분리된 언어들이 살아남았다. 이렇게 유사한 언어들은 의사소통을 유추할 수 있을 정도로 가까울지언정 정확한 대화는 불가능하다. 서로 나뉘어진 그 시점에서부터 ‘다른 언어’로 독자적인 변화를 이어왔기 때문이다.

마치 진화론에서 이전의 종으로부터 다른 종들이 갈라져나오고, 이전의 종은 멸종하며, 다른 종들은 서로 독자적인 진화를 이어가는 모습을 보는 듯 하다.

다윈이 주장한 자연 선택은 우리 인간들이 원하는 종만 선택하여 교배하는 과정을 통해 최종적으로 원하는 형태의 종을 만들어낸 것에서부터 시작됐다. 인간의 선택 능력이 이렇게 종을 변화시키는데, 자연의 선택 능력은 얼마나 위대할까? 얼마나 많은 종과 다양성을 탄생시킬 수 있을까? 그것이 자연 선택설의 시작, 진화론의 시작이었다.

언어가 사멸하고 선택되는 과정은 인간의 명확한 의도에 의한 것이라기보다는 인간의 무의식에 더 가깝다. 그래서 난 언어의 진화사는 인간의 선택적 교배보다는 자연 선택에 가까운 모습이라고 생각한다. 물론 인간의 선택이 자연 선택은 아니다. 이를테면 자연 선택에서는 무수히 많은 돌연변이 중에서 환경에 적합한 매우 소수만이 살아남지만, 언어에 대한 인간의 선택은 매 순간순간 일어나며 그것은 언어의 수정으로서 일어난다. 그래서 자연 선택이 눈에 보이는 변화를 보이기까지는 매우매우 오랜 시간을 보이지만, 언어의 대한

인간의 선택은 인간의 시간을 따라가기 때문에, 세대에 따라 아주 빠른 변화를 보인다. 1970년대의 우리말과 지금의 우리말을 생각해 보고, 겨우 600년 전인 훈민정음 언해본—'나랏말싸미'로 대표되는 그 책—에 수록된 중세 국어와 지금의 우리말을 생각해 보라.

그래서 나는 진화를 직접적으로 느낄 수 있는 창구가 먼 곳에 있지 않다고 본다. 생각보다 가까운 곳에서 우리는 진화를 느낄 수 있다. 우리가 사용하는 언어 말이다! 겨우 우리의 조부모 세대, 부모 세대에서 쓰는 언어의 차이와 시간에 따른 언어의 전체적인 변화는 우리 모두가 몸으로 느낄 수 있다. 언어의 아름다움도 거기에 있다. 그것은 진정 사회 구성원들의 선택과 시간이 빚어낸 변화의 장엄함이다. 그 장엄함을 자연에 비할 수 있을까? 현재 존재하는 모든 생명체들은 그 긴 시간 동안, 자연의 선택과 시간이 빚어낸 예술이다. 어찌 보면 가장 확실하게 진화의 실재성과 진화론의 타당성을 입증하는 것이 언어인 것이다.

아, 언어! 그것은 인간이 자연의 어떤 동물들과도 구분됨을 말해줌과 동시에 인간이 결국 자연의 일부로써 움직인다는 점을 말해주고 있다. 언어는 결국 인간에 관한 전혀 다른 관점을 동시에 제시한다.

## 인공 언어의 극치, 톨킨의 레젠다리움

그럼 내가 생각해왔던 **인공어**에 관한 여러 생각들을 풀어내며 글을 마치고자 한다. 언어라는 것은 기본적으로 의사소통 체계이고 그것은 인간의 사고에 따라 꽤나 정형화되어 있다. 따라서 언어학적

지식이 충분하고 어휘, 통사를 구상할 수만 있다면 사람은 어느 정도 체계를 갖춘 자신만의 언어를 만들 수 있다. 물론 이렇게 ‘설계된’ 인공어에겐 자신만의 자연적인 변천사가 없다.

어떠한 인공어들은 언어 간 격차를 줄이는 것을 목표로 하여 어려운 발음을 최대한 줄이고, 자연어의 불규칙한 부분들이 없도록 설계되어 있다. 이러한 언어들은 과거의 프랑스어, 현재의 영어 대신 국제적으로 사용하는 것을 목적으로 만들어졌으며 **국제어**라고 부른다. 에스페란토 Esperanto를 들어보았는가? 네이버 제공 번역 서비스인 파파고 Papago는 에스페란토로 ‘앵무새’라는 뜻이다. 에스페란토는 세계적으로 가장 널리 쓰이는 국제어로, 세계 화자도 수백만에서 수천만으로 많고 심지어 모국어 화자도 3000명 정도일 것으로 추측된다.

물론 작품들에서만 등장하고, 실생활에서는 사용되지 않는 언어들도 존재한다! 이러한 언어들은 실용적인 목적이 아닌 소설 속 세계관에서 사용되는, 설정상으로는 ‘자연어’인 녀석들이다. 따라서 국제어와는 달리 음운이나 어휘 활용 등이 자연어 수준과 비슷하다. 이를 **예술어**라고 부른다.

나는 항상 ‘반지의 제왕 시리즈’를 최고의 판타지 소설로 꼽는다. 다른 SF, 판타지 작품들도 많은데도 굳이 ‘반지의 제왕’을 고르는 이유는, **J.R.R. 톨킨**이 만들어낸 그 어떤 소설과 영화도 따라올 수 없는 정교한 세계관에 있다. 그의 세계관은 얼마나 방대한지, 그 세계관을 집대성한 <실마릴리온>이라는 책에서 반지의 제왕의 역사가 차지하는 분량은 전체 분량 600쪽 중 2쪽에 불과하다. 그리고 치밀한

세계관 설정의 가운데에 언어가 있다. 그 어떤 치밀한 세계관이든— 그것을 판타지 이외의 장르로 뻗친다고 할지라도 그 세계관의 언어의 역사 마저 독자적으로 구상하는 일은 없다고 봐도 무방하다. 하지만 톨킨은 소설가이기 이전에 언어학자였다. 그는 그런 정교한 언어를 만든 것도 모자라 그걸 여러 개나 만들어냈고, 심지어 그 언어 하나하나에 역사에 따른 변천사와 방언까지 구현해냈다. 나는 그가 진정으로 한 세계를 창조해냈다고 생각한다.

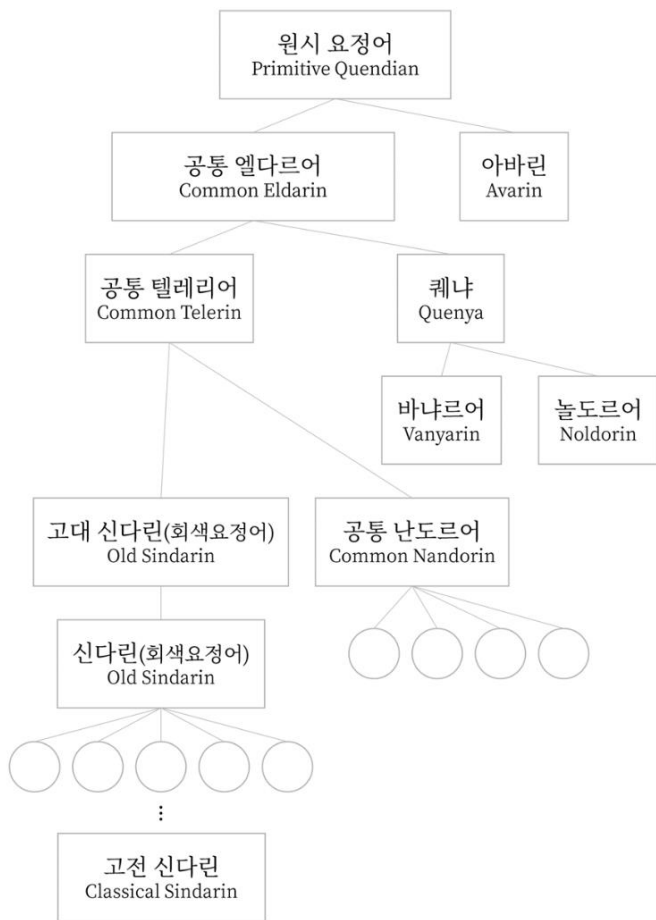
그래서 그런지 요정어만 보더라도 그 변천사가 그림처럼 방대하다. 각 분화는 대개 민족의 분리를 따라가고, 시간의 흐름에 따라 점차 변화하기도 한다. 요정의 언어는 설정상 인간의 언어에 아주 지대한 영향을 끼쳤고 이는 이후 반지의 제왕 시대에는 고전 신다린<sup>7)</sup>이 보편적인 요정어처럼, 퀘냐(놀도르어)가 귀족층의 요정어처럼 여겨진다는 설정이다.

그 뿐이 아니다. 톨킨은 이 세계관을 위한 새로운 문자체계를 세 가지 정도 고안해냈는데, 그 중 가장 우아한 것은 단언 텡과르 Tengwar다. 톨킨은 요정어를 표기하기 위해 이 문자를 만들었고, 세계관 내에선 마치 세종대왕이 한글을 창제한 것처럼 한 위대한 놀도

---

<sup>7)</sup> 먼 옛날 신다린에는 여러 방언들이 있었고, 당시 번영했던 요정들의 왕국이 무너지며 이후 <반지의 제왕> 시기에는 신다린의 지역색이 사라지고 다시 통합되는 일이 일어난다. 이렇게 뭉뚱그려진 신다린을 고전 신다린이라고 칭한다.





### ▲ 요정어의 변천사

î ṽṽṽṽ ṽṽṽṽṽṽ

▲ 신다린으로 지어진 시의 한 구절.  
‘아, 엘베레스! 길소니엘!’이라고 써 있음

르 요정이 창제한 것이라는 설정이다. 그래서 그런지 한글처럼 음운의 성질이 문자에 그대로 반영된 모습을 띠고 있다.

이 장대한 세계관의 근거한 톨킨의 작품들을 통틀어 ‘레젠다리움 Legendarium’이라고 칭한다. 이것은 <실마릴리온>으로 요약된 정식 세계관을 기반으로 한 작품부터 시작해서 톨킨의 사후 출간된, 세계관이 완전히 정립되지 않은 시절 썼던 여러 시들과 설정집들을 통틀어 말한다. 즉 톨킨의 작품집을 통틀어 말한다. 그의 작품에는 항상 그가 만든 언어가 중심이 되어 왔고, 빠지지 않는다. 그리고 그 언어들로 지어진 그의 세계에 전해지는 오래된 이야기들과 시를 통해 덕분에 우리는 그의 세계 아르다 Arda의 시간을 느끼고, 살아있는 것처럼 여기게 된다. 분명 가상의 세계라는 것을 알고 있음에도 말이다. <반지의 제왕> 시리즈가 어떤 판타지 소설들도 넘을 수 없는 벽이라 불리는 데에는 다 이유가 있다.

## 마치는 글

사실 내가 언어에 관해 제대로 썰을 푼 것은 이번이 처음, 완전히 처음이다. 그리고 이렇게 글을 길게 쓴 것도 처음이다. 나 또한 언어학의 아주 얇은 부분만을 알 뿐이지만, 당신에게 이 글이 읽기 쉬웠을지 어려웠을지 감조차 오지 않는다. 글을 쓰면서도 정말 신나서, 원래는 토막글로 쓰려고 했지만 점점 길어져서 결국 지금에 이르게 됐다.

언어는 과학적으로 굉장히 흥미롭다. 내가 이 글에서 말한 언어의

진화적인 면모는 그저 언어의 매력이라는 빙산의 극히 일부분에 불과하다. 절대 이 글에서 나온 것이 언어 그리고 언어학의 전부라고 생각하지 말라. 시간에 따른 언어의 변화를 연구하는 통시언어학은 어떤 한 순간의 언어 상태를 연구하는 공시언어학과 함께 언어학의 연구 분야를 나누는 가장 큰 구분 중 하나일 뿐이다. 세세하게 보면 음운론부터 시작해서 문자론, 형태론과 통사론<sup>8)</sup>, 의미론<sup>9)</sup>, 화용론<sup>10)</sup>, 음성학, 기호학 등 그 분야는 매우매우 다양하다. 나는 그 중 문자론을 먼저 둘러봤었다. 탕과르를 처음 만나게 된 것도, 인공어라는 개념을 처음 접하게 된 것도 이 때였다. 페니키아 문자에서부터 현재의 라틴 문자까지 변화하는 모습은 내가 언어에 본격적인 관심을 갖게 된 계기가 되었다. 그러다가 국제음성기호 International Phonetic Alphabet, IPA에 빠지게 되면서 음운론에 빠지게 되고, 그렇게 난 언어라는 것의 매력에 흠뻑 젖어들었다.

또 언어학의 범위를 보면 언어라는 한 요소가 얼마나 넓게 다루어질 수 있는지를 알 수 있다. 찾아보니 놴 촘스키 Noam Chomsky라는 언어학자가 언어를 형식과학과 엮으면서 ‘생성문법’이라는 아주 추상적인 학문 분야를 개창했다고. 이 생성문법이라는 것은 인문학보다

---

<sup>8)</sup> 형태론은 단어의 내부 구조, 통사론은 문장의 내부 구조에 대한 학문이며 우리가 아는 ‘문법’은 형태론과 통사론을 통틀어서 말한다고 봐도 된다.

<sup>9)</sup> 의미론에선 단어의 의미, 단어의 조합에 따른 의미 변화를 연구한다.

<sup>10)</sup> 화용론에선 상황과 맥락에 따른 의미를 연구한다.

는 수학과 논리학에 가깝고, 언어를 자연과학적인 범위 내에서 연구한다. 위에서 얘기한 통시언어학이나 형태론, 통사론과는 결이 아주 많이 다르다.

아, 언어에 대해 얘기하고 싶은 것이 너무 많다. 또 이러한 글을 쓸 기회가 있다면, 제발 있다면!

한 번은 문자론에 대해 얘기를 해 보고 싶다. 아랍 문자와 라틴 문자가 사실은 친척 관계라는 사실을 알고 있었는가? 그리스 문자와 라틴 문자, 키릴 문자는 왜 그리도 닮았을까? 한글이 창제한 인물과 창제 시기가 정확히 알려진 유일한 문자가 아니라는 사실을 알고 있었는가? 또 한 번은 형태론에 대해 얘기를 해 보고 싶다. 왜 우리나라 사람은 외국어를 배우기 어려워하고, 동시에 왜 외국 사람은 한국어이나 일본어를 배우기 어려워할까? 왜 영어는 문장 성분에 ‘서술어’가 아닌 ‘동사’가 있을까? 자연과학을 배우면서 일상의 여러 상황들이 설명되는 것과 마찬가지로 언어학을 파면 그냥 알파벳들 모임으로 보이던 여러 문장들이 설명될 것이다. 그리고 자연과학과 마찬가지로, 그런 부분에서 약간의 희열을 느낄지도 모른다.

언어학의 유일하고도 진정한 대상은 언어인데, 언어는 그 자체로서,  
그것만을 위하여 고찰되어야 한다.

- 언어학자 소쉬르 *Ferdinand de Saussure*