

편재(偏在)의 편재(遍在)*

L자형 편중 현상이 수많은 분야에서 보편적

인문대학 국어국문학과 명예교수 이상억

1. 언어도 자본주의적이다

필자는 국어학을 전공하는 언어학 박사일 뿐이다. 언어에 관한 특성을 탐구하다가 특이한 점을 발견하게 되어 그 궁극증을 인문학 이외의 사회·자연과학, 예술 등 다른 분야들까지 확산하면서 이 글의 내용과 같이 큰 그림을 그릴 수 있는 현상을 발견하여 모르는 사이에 ‘유레카’를 외치게 되었다.

우선 “언어도 자본주의적이다”라는 소제목이 달린 이유를 밝혀야겠다. 이렇게 대답한 질문에 들어가게 된 첫 단계는 보통 해 오던 연구들과 같이, 언어의 소리, 단어, 문장 등 각 계층별로 존재하는 규칙을 망라하여 그 각각이 언어활동에서 담당하는 기능적 부담량(functional load)을 통계적으로 조사하였던 것이다. 그런데 단어들의 규칙은 잠시 건너뛰고, 우선 쉽게 인용할 수 있는 소리와 문장에 관한 규칙을 출현 빈도수 순위로 15 내지 20위 정도만 보아도 이미 98% 정도의 급격한 상위 독과점 체제가 보인다.

1) 음성적·음운적 현상들의 기능부담량: 계량적 조사 cf. 이상억(1993)

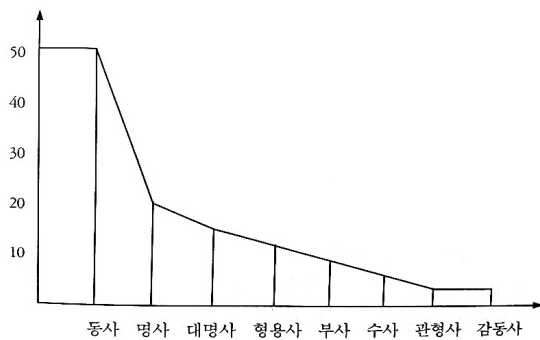
순위	명칭	발생빈도	백분율
1.	음절말 유성자음 불파화	4,947	31.84%
2.	유성음화	3,009	19.37%
	(1에서 2까지의 중간누적합계)		51.21%
	(1에서 10까지의 합계)		92.41%
	(1에서 15까지의 합계)		96.23%
	(1에서 20까지의 누적합계)		98.95%

* 편재(偏在)의 뜻: 한 곳에 치우쳐 있음. 한자가 다른, 편재(遍在)의 뜻: 두루 퍼져 있음. 즉 편중 현상이 많은 분야에서 보편적으로 보이는 사실에 대해 주목하려는 것이다.

2) 통사규칙들의 기능 부담량: 계량적 조사

1. 명사구 보문화	18.93%
2. 종속접속화	17.25%
(1에서 2까지의 중간합계)	36.18%
3. 관계화	15.58%
(1에서 3까지의 중간합계)	51.76%
(1에서 10까지의 합계)	91.60%
(1에서 15까지의 합계)	98.41%

3) 품사들의 기능 부담량: 계량적 조사



위의 독과점 내지 자본주의적 분포의 결과가 사회 배경으로 인한 것이 아니라는 점을 명백히 하기 위해, 단어 층위의 '품사'란 범주에 대한 통계를 북한학자 문영호(2001)의 '조선어 어휘 통계학'(54~55면)에서 다음과 같은 결과를 뽑아 인용하겠다. 여기도 동사의 독과점이 뚜렷하다.

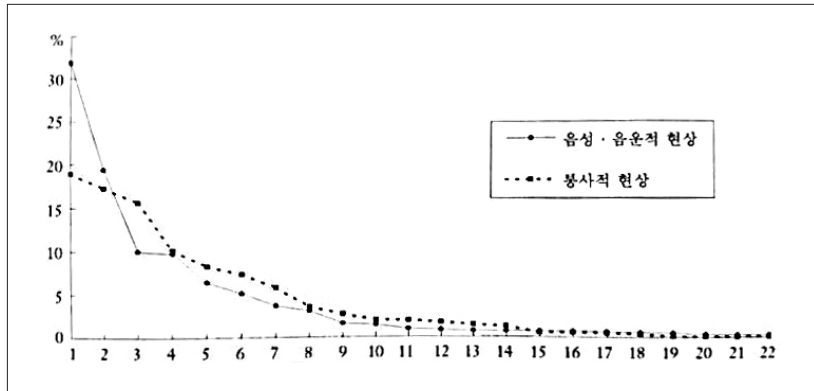
cf: 문영호(2001)

<그림 1> 8품사의 단어량

4) 소결론: 언어도 자본주의적이다

소결론을 얻기 위해 필자는 소리 단어 문장 현상 간의 병행적인 비교를 시도하고자 한다. 매우 흥미롭게도 이 현상들 사이에는 유사성이 있어 가장 많이 발생하는 두 개나 세 개의 규칙들이 50% 이상을 차지한다. 그리고 10번까지가 약 90% 이상을 차지하고 15번까지의 규칙들은 약 95% 이상을 포함한다. 품사의 분포도 비슷한 곡선으로 나타난다. 이러한 기능적 부담량을 통계적으로 조사해 보려는 본래 취지는 최신 음성공학, 전통 언어교육 등 관계되는 여러 분야에 합리적으로 활용하도록 하자는 것이었다.

그러나 위와 같은 독과점 현상이 보임에 따라 인문학 자료의 바탕에서 관찰되는 사회과학적 현상 내지 경향에 대해 지적해 두고 싶다.



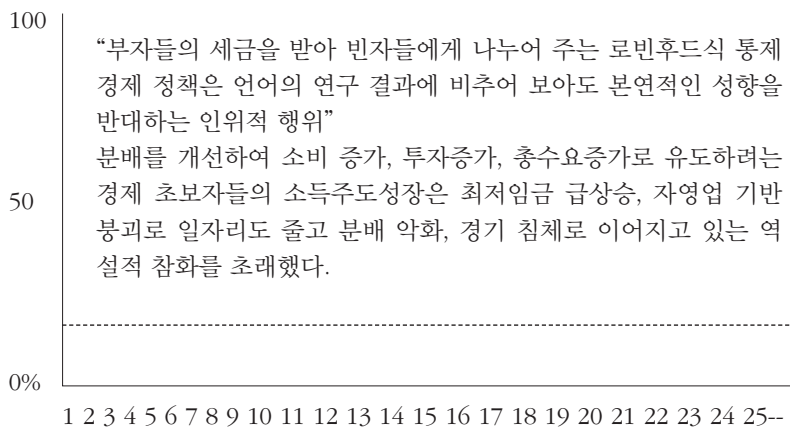
cf. <그림 2>의 출처: 이상익 (1993)
<그림 2>

각 언어 단위에 적용되는 규칙들의 기능 부담량을 관찰하면 전형적 L 자형 도표(<그림 2> 참조)로 집약된다.

상위 몇 항에 그야말로 부(富)가 집중되어 있고 10위 이하는 전체의 10% 정도를 미미하게 나누어 가지는 형태다. 이는 예기치 않게 자본주의 시장경제 구조와 흡사한 양상이다.

여기서 직감되는 사실은 부육약식(富肉弱食)의 이념으로 부자들의 세금을 받아 빈자들에게 나누어 주는 로빈후드식 통제경제 정책은 언어의 연구 결과에 비추어 보아도 만물의 본연적인 성향을 거스르는 인위적 억지임을 알 수 있다.

자본주의를 인정하고 이 본성을 활용하여 많은 기업이 부를 창출하여 흘러내리도록 하는 순리를 따라야지 이 본성을 인위적으로 뒤엎으려는 ‘소주성’ 알코올 정책은 잘못된 것임이 언어에서도 방증되었다. 돈이나 말처럼 많이 유통되는 것들은 같은 기본 원리를 가지고 규칙이나 제약을 쌓아 나가는 특성이나 경향이 있다고 추론할 수 있겠다. 이 특성을 무시하고 짜는 언어정책은 언어를 인위적으로 바꿀 수 없으니 아예 (북한 같은 비정상 국가 이외에서는)



“부자들의 세금을 받아 빈자들에게 나누어 주는 로빈후드식 통제 경제 정책은 언어의 연구 결과에 비추어 보아도 본연적인 성향을 반대하는 인위적 행위”
분배를 개선하여 소비 증가, 투자증가, 총수요증가로 유도하려는 경제 초보자들의 소득주도성장은 최저임금 급상승, 자영업 기반 붕괴로 일자리도 줄고 분배 악화, 경기 침체로 이어지고 있는 역설적 참화를 초래했다.

<그림 3>
사회주의 경제의 균분 공빈(共貧) 상황을 보이는 점선

존재할 수도 없고, 그런 식으로 설계된 경제정책은 실패할 수밖에 없다. 작은 결론을 맺자면 언어도 자본주의적이다.

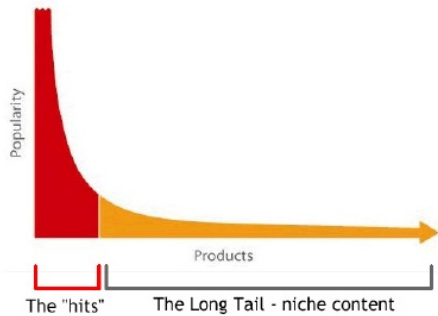
2. 파레토 법칙¹⁾: 사회·경제 영역의 경향

파레토 법칙(Pareto principle, law of the vital few, principle of factor sparsity) 또는 80 대 20 법칙(80-20 rule)은 '전체 결과의 80%가 전체 원인의 20%에서 일어나는 현상'을 가리킨다. 예를 들어, 20%의 고객이 백화점 전체 매출의 80%에 해당하는 만큼 쇼핑하는 현상을 설명할 때 이 용어를 사용한다. 2 대 8 법칙이라고도 한다. '이탈리아 인구의 20%가 이탈리아 전체 부의 80%를 가지고 있다'고 주장한 이탈리아의 경제학자 빌프레도 파레토의 이름에서 따왔다. 파레토 분포는 자연/사회/경제/웹 등에서 발견되는 재미있는 현상을 설명할 수 있는 강력한 프레임이라고 보면 될 것 같다. 이제 여기에 인문/예술 등 세상만사의 여러 분야도 추가해야 한다는 사실을 밝히려는 것이다.

파레토 법칙의 적용사례 <80/20의 법칙>

1. 기업이익의 80%는 20%의 고객으로부터 나온다.
2. 고객불만은 전체고객의 20%로부터 나온다.
3. 기업매출은 전체매출의 20%로부터 나온다./전체 판매사원의 20%로부터 나온다.
4. 통화한 사람 중 20%와의 통화시간이 총 통화시간의 80%를 차지한다.
5. 즐겨 입는 옷의 80%는 옷장에 걸린 옷의 20%에 불과하다.
6. 전체 추가상품의 80%는 상승기간의 20%의 기간에서 발생한다.
7. 20%의 운전자가 전체 교통위반의 80% 정도를 차지한다.
8. 20%의 범죄자가 80%의 범죄를 저지른다.
9. 성과의 80%는 근무시간 중 집중력을 발휘한 20%의 시간에 이루어진다.
10. 두뇌의 20%가 문제의 80%를 푼다. 우수한 20%의 인재가 80%의 문제를 해결한다.
혹은 뇌의 20%만 사용하여 문제 해결에 필요한 80%를 해결한다
11. 운동선수 중 20%가 전체 상금의 80%를 싹쓸이한다.
12. 인터넷 유저의 20%가 80%의 양질의 정보를 생산한다.

1) 파레토 법칙의 설명과 도표들은 위키백과를 주로 이용하였다. 뒤에 나올 지구, 지하자원 등에 관한 정보도 마찬가지다.



〈그림 4〉 파레토 법칙에 의한 80:20의 집중현상을 나타내는 그래프에서는 발생확률 혹은 발생량이 상대적으로 적은 부분이 무시되는 경향이 있었다. 그러나 인터넷과 새로운 물류기술의 발달로 인해 이 부분도 경제적으로 의미가 있을 수 있게 되었는데 이를 롱테일이라고 한다.

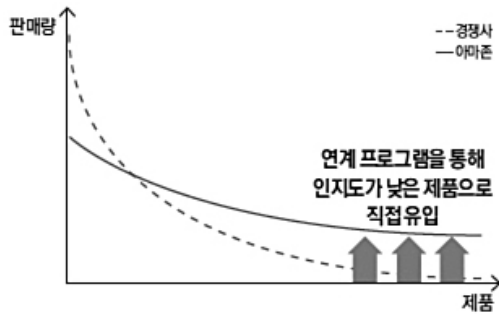
파레토 효율(Pareto efficiency) 또는 파레토 최적(Pareto optimality)이란 게임이론과 엔지니어링 및 기타 다양한 사회과학 분야에서 쓰이는 경제학적 개념이다. 경제적 효율성과 수입의 분배에 대한 연구에서 이 개념을 사용하였다. 파레토 효율성이란 하나의 자원배분 상태에서 다른 사람에게 손해가 가도록 하지 않고서는 어떤 한 사람에게 이득이 되는 변화를 만들어 내는 것이 불가능할 때 이 배분 상태를 '파레토 효율적'이라고 한다. 이상화된 조건 하에서, 자유시장 시스템은 파레토 효율적 결과로 이끈다는 것이 드러난다. 이것이 후생경제학의 기본 정리다. 반면에 파레토 비효율은 파레토 개선(Pareto improvement)이 가능한 상태를 말한다.

위의 〈그림 4〉의 아마존의 롱테일 효과에서 볼 수 있듯이, 모바일기기로 연결된 소비자, 커뮤니티 활동, 소통 관계 형성, 소비자 주도 '사용자 혁신', 브랜드/유저 커뮤니티의 지혜를 빌려 커뮤니티 기반의 혁신, 기업의 기회 등이 가능하며, 1. 선도적 소비자가 모이게 하고 2. 소비자가 머물러 교류하고 3. 소비자가 제품을 발전시키게 하며 4. 소비자의 역할이 무한계로 5. 소비자 아이디어가 실행으로 이어지게 해야 한다.

이런 방향의 개선은 여러 방법으로 시도해 볼 필요가 있다. 우리가 극단적인 사회주의만 아니면 사회 여건과 경제구조의 향상을 위해 자본주의를 보완할 수 있는 많은 아이디어를 살려 볼 수 있기 때문이다.

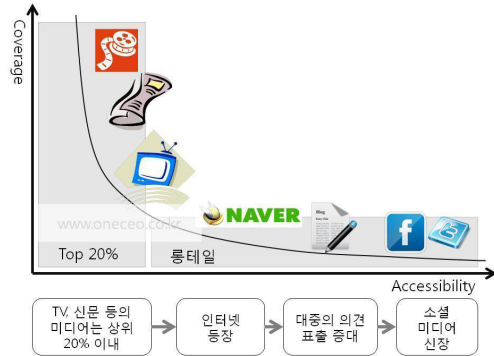
통계학에서, 파레토 분포(Pareto distribution)는 사회과학에서 널리 볼 수 있는 확률분포로 사회에서 부의 분포를 나타내기 위해 사용하였다. 사회에서는 부의 불공평한 분포로 인해 대부분의 부가 소수에 의해 소유되는데 파레토 분포는 이를 효과적으로 나타낸다.

그러나 위의 아마존의 롱테일 효과 등에서 보았듯이, 롱테일에 속하는 다수자의 위상이 점차 부각되며 그 중요성도 인식되고 있다. 이는 마치 경제 사회의 구조가 원천적으로는 파레토식 80 대 20 분포로 구성되어져 있지만, 즉 자본주의적 분포로 되어 있지만, 사회구조 개선을 위해 사회주의적 경제 정책으로 부의 분배를 시도해 왔다는 현실이 깔려 있어서, 그 맥이 롱테일 효과로도 이어졌다 할 수 있겠다.

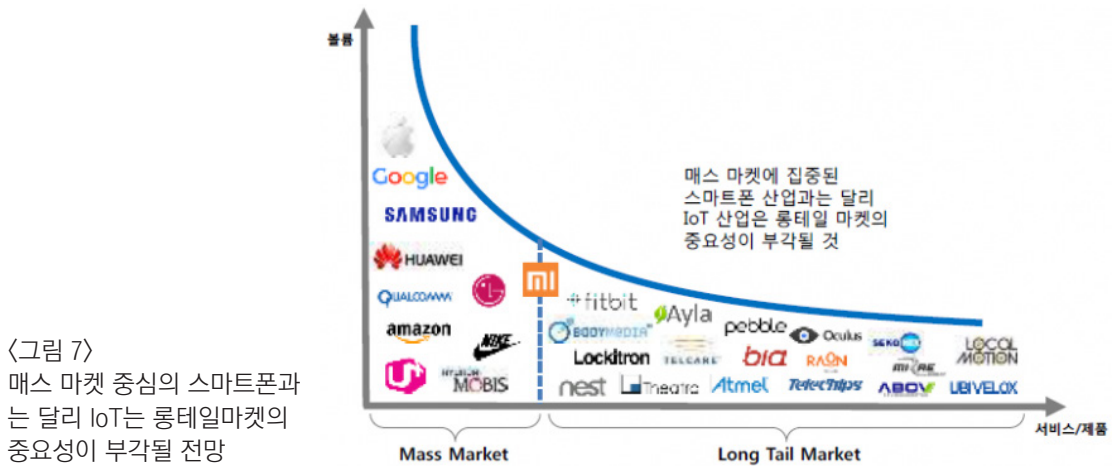


자료: 조봉수, <디지털 컨슈머 & 마케팅 전략>, 에이콘, 2014

<그림 5> 파레토 개선이 가능한 상태의 예



<그림 6>



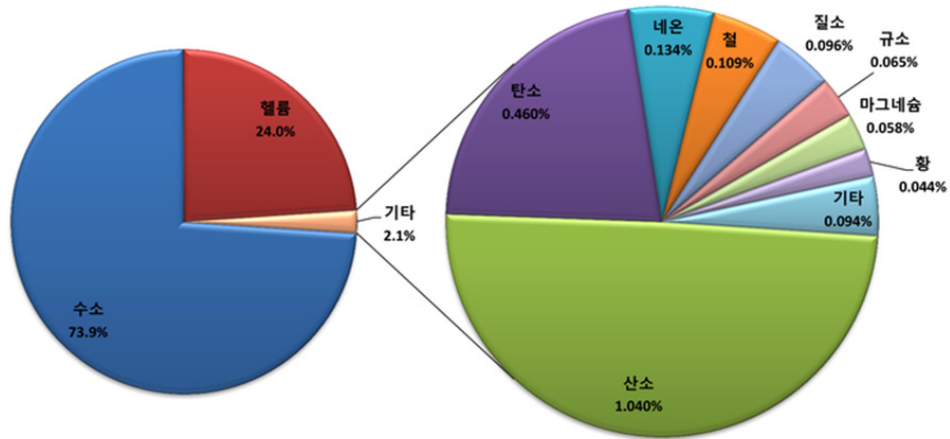
<그림 7>
매스 마켓 중심의 스마트폰과는 달리 IoT는 롱테일마켓의 중요성이 부각될 전망

그 중간과정에서 20세기 말 자본주의에 공산주의가 완패하는 역사가 완전히 드러났었다. 아마도 그 이유는 부의 분포에 대한 개선 노력에도 불구하고 원래 부는 80 대 20으로 분포되어 있는 상태가 본연적인 구조라는 데 있을 것이다. 아무리 분포 개선 평균화의 뜻이 좋더라도 명분과 이념이 현실을 떠나 본연의 자연스러운 상태를 깨뜨리는 방향으로만 가서는 큰 탈이 나는 법이다. 물론 그 외에 공산주의는 당 간부층의 독점구조에서 오는 자체 모순적인 불평등, 부패, 비효율성 등의 많은 문제점들을 가지고 있었기 때문에 자본주의의 본연적 힘을 감쇄(減殺)하지 못했던 것이다.

3. 타 분야의 L자형 예들

- 1) 우주의 구성 요소: 은하계 초기의 원소 조성은 수소와 헬륨이 압도적이었고, 2.1%의 기

타 요소 속에 산소와 탄소가 다시 압도적이었는데 바로 이 두 원소가 태양의 주성분이 되었던 것이다. 다음은 은하계의 원소 조성이다.



한편 지구 구성 원소의 질량비(%)는 다음과 같다.

지각 - 산소 46.6, 규소 27.7; 지구 전체 - 철 35, 산소 30

지구의 지각은 산소와 규소, 지구 핵은 압도적 철과 약간의 니켈, 지구 전체로는 철과 산소가 역시 편재(偏在) 현상을 보인다.

구분	지각(%)	핵(%)	지구 전체(%)
산소(O)	46.6	-	30
규소(Si)	27.7	-	15
알루미늄(Al)	8.1	-	1
철(Fe)	5.0	88.6	35
칼슘(Ca)	3.6	-	1
나트륨(Na)	2.8	-	-
칼륨(K)	2.6	-	-
마그네슘(Mg)	2.1	-	13
니켈(Ni)	-	8.5	2
황(S)	-	-	2
기타	1.5	2.9	1

2) 색채어의 출현 빈도: 흰색 검정색 붉은색이 상위에 있는 언어들이 많다.

English	German	French (freq.)	Italian(freq. rank)	Portuguese(freq. rank)
white 1	weiss 4	blanc(he) 2 23674 2	bianco 17,973: 2,435 4	branco 22,066: 2,504 1
black 3	schwarz 2	noir 1 25476 1	nero 26,786: 1,650 2	preto 9,475: 5,230 6
red 2	rot 1	rouge 3 13222 3	rosso 29,117: 1,508 1	vermelho 13,517: 3,880 3
green 5	grün 3	Vert 4 8172 5	verde 25,444: 1,739 3	verde 18,970: 2,925 2
blue 4	blau 5	bleu 5 10915 4	blu	azul 13,261: 3,929 4
yellow 6	gelb 7	jaune 7 5369 6	giallo	amarelo 13,220: 3,940 5
gray 7	grau 6	gris 6 6438 5	grigio	cinza 3,390: 11,323 7

3) 국가별 인구 순위: 중국, 인도가 압도적으로 우위, 미국, 인도네시아, 파키스탄이 다음 층위에.

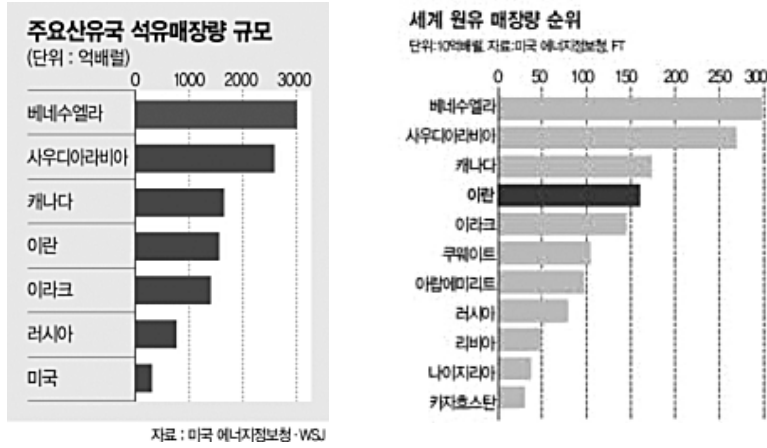
(단위: 천명, %)

순위	1950년			2009년			2025년			2050년		
	국가	인구	비율	국가	인구	비율	국가	인구	비율	국가	인구	비율
1	중국	544,951	21.5	중국	1,345,751	19.7	중국	1,453,140	18.1	인도	1,613,800	17.6
2	인도	371,857	36.2	인도	1,198,003	37.2	인도	1,431,272	36.0	중국	1,417,045	33.1
3	미국	157,813	42.5	미국	314,659	41.9	미국	358,735	40.5	미국	403,932	37.5
4	러시아	102,702	46.5	인도네시아	229,965	45.2	인도네시아	263,287	43.8	파키스탄	335,195	41.2
5	일본	82,824	49.8	브라질	193,734	48.1	파키스탄	246,286	46.8	나이지리아	289,083	44.4

4) GDP: 미국(205,130억 불), 중국(134,573억 불)이 압도적으로 우위, 일본(50,706억 불), 독일(40,291억 불)이 다음 층위에, 나머지는 긴 꼬리 모양.

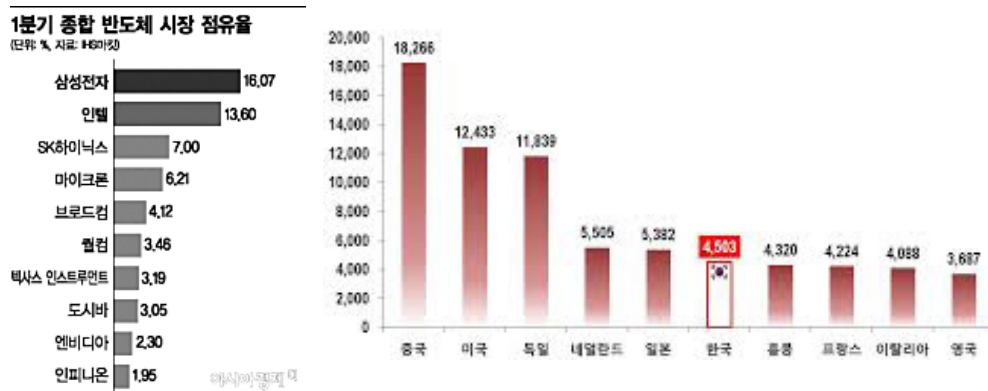
순위	국가	GDP(백만\$)
1	미국	20,513,000
-	유럽연합	18,769,286
2	중국	13,457,267
3	일본	5,070,626
4	독일	4,029,140
5	영국	2,808,899
11	대한민국	1,655,608

5) 석유 매장량: 베네수엘라, 사우디아라비아, 캐나다, 이란, 이라크 등의 순.



OPEC 석유 수출 및 생산(2014년): 석유 매장량이 약간 다른 통계에 따라 달라진다. 특히 근래 미국의 셰일오일 개발로 순위가 더 바뀌고 있다.

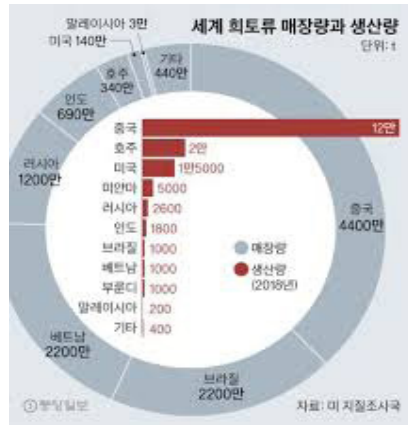
6) 무역, 산업 분야



세계 수출입 순위 톱 10(한국 6위, 2018)

7) 기타 세일 오일 · 가스, 희토류 등 지하자원 자료들

세계의 지하자원에 대한 통계가 위키백과에 여러 가지로 나와 있는데 그 상세 내용은 조금씩 다르다. 희토류는 생산시 공해를 마다않는 중국이 앞서나 다른 나라에도 매장량은 많다.



희토류

순위	국가명	매장량	점유율
1	중국	3610000	17.4%
2	미국	2440000	11.8%
3	아르헨티나	2270000	11.0%
4	멕시코	1930000	9.3%
5	캐나다	1620000	7.8%
6	멕시코	1540000	7.4%
7	남아공	1370000	6.6%
8	유럽연합	1330000	6.4%
9	오스트레일리아	1212870	5.9%
10	사우스 아프리카	1100000	5.3%
11	기타	2300000	11.1%
합계		20722870	100%

☆ 미국에 매장된 세일가스는 전 세계 60년 사용량

2018년 수출입 동향

8) 성씨의 구성 빈도: 김, 이 박 최씨의 합계: 49.3%, 김: 21.5%, 이: 14.7%, 박: 8.4%, 최: 4.7%

순위 (본관별 성씨) 인구수 시조 대표적인 인물 (현조) 현대 유명 인사

1. 김해김씨(金海金氏)-4,356,573 김수로 김유신 김홍도 김대건 김대중 김종필 김활란 김준연 김현철 김성곤 김용철 김기창 9.0%
2. 밀양박씨(密陽朴氏)-3,230,367 혁거세 박언침 박강생 박충원 박충훈 박중화 박순천 박두병 박경원 박대선 박흥식 박동진 6.2%
3. 전주이씨(全州李氏)-2,788,578 이 한 이성계 이원익 이하응 이승만 이범석 이재형 이기봉 이은상 이한기 이회창 이길상 5.3%
4. 경주김씨(慶州金氏)-2,132,423 김알지 김부식 김정희 김홍집 김정렬 김기승 김천애 김학렬 김동익 김석원 김성환 김중서
5. 경주이씨(慶州李氏)-1,776,938 이알평 이제현 이항복 이상설 이시영 이병철 이종찬 이정우 이민우 이건희 이규호 이상룡
6. 경주최씨(慶州崔氏)-994,994 최치원 최승로 최익현 최제우 최현배 최문화 최복현 최기철 최동호 최중환 최태영 최영철

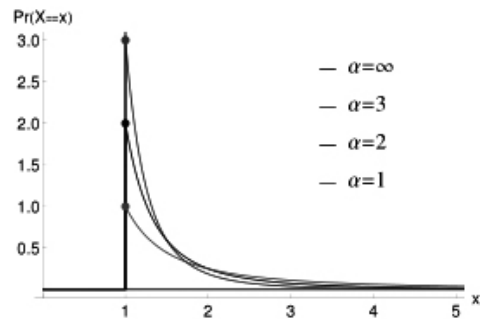
4. 결론

앞서 1장 끝 소결론에서 언어도 자본주의적 행태가 보인다고 했었다. 이 말은 L자형 곡선에서 앞머리에 편중된 20%가량의 소수가 부를 독과점한다든가 하는 자본주의 자유경제의

모습과 흡사하다는 뜻이다. 본성적으로 인간사에 관여된 만물에서 이런 편중 현상이 보편적으로 보인다. 이 현상을 원래 불공평하게 잘못된 상태라고 보면 마치 본성적으로 악하다는 ‘성악설’로 기울 수도 있겠다. 그래서 2장 끝에서 언급한 바, 파레토 비효율로 파레토 개선 (Pareto improvement)이 가능한 상태를 꿈꾸게 될 수 있겠다. 이상을 확률 밀도 함수와 누적 분포 함수로 아래와 같이 표시할 수 있는데, 이 둘은 같은 내용을 달리 표현한 것일 뿐이다. 이에 병행해 언어 자료가 보이는 결과를 함께 표시해 보겠다.

1) 확률 밀도 함수

옆의 그림과 같이 확률밀도를 단순히 나열해서 그 정도의 차이를 보이는 L자형 곡선 도표들이 앞으로 많이 소개될 것이다. 그런데 이 확률을 누적시켜 그 합으로 분포를 파악해 볼 수도 있다. 다음 페이지에 보이는 <그림 8>은 그러한 일종의 역상을 나타내는 것이며, 표현상 L자형을 뒤집은 것일 뿐 내용상 아무 차이가 없다.



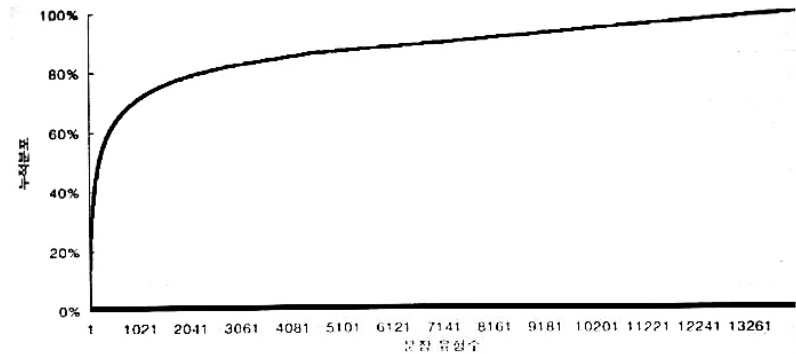
2) 누적 분포 함수

필자는 이 글에서 개선의 방법까지는 일일이 제시하지 않겠다. 다만 그런 방향으로 가고 싶은 사람들이 있다면, 꼭 명심해야 할 사항이 있다는 말을 해 두려는 것이다. 현재 정권의 정책 수립자들이나 그 동조자들이 꽤 있으니, 그 세력들에 의해 전통적 독과점 구조를 개선하려는 인위적 작위는 끊임없이 있어 왔고 앞으로도 계속될 것이다. 가령 ‘소득주도성장’ 정책 하에 최저임금 2년간 30% 인상, 주 52시간 근무제, 비정규직의 정규직화, 부유층과 대기업에 대한 중과세, 재건축 초과이익환수제 및 분양가 상한제, 건강보험 보장 확대, 심지어 우수한 자사고까지 일반고화와 대학 서열화 완화 정책이 적정치와 속도를 과도하게 넘기며 난무하고 있다.

물론 이 단계에서 어느 입장에 찬성한다는 뜻을 펴려는 취지도 아니다. 다만 우리 주변에 일어나는 세상 제만사의 본성은 잘 파악하고 일을 해 나가야 한다는 생각이다. 역사적으로도 이 본성은 결국 공산주의 통제계획경제가 자본주의 자유시장경제에 완패하는 결과를 초래했다. 과연 좌파의 사회주의적 경향에 경도하여 거의 일방적으로 평등 분배의 포퓰리즘 정책들만 집중적으로 추진하다가 다시 역사적 과오를 범할 수 있으니 부디 L자형 지표들이 출발

cf. 홍정하 김주영 강범모
(2007)

〈그림 8〉
문장표현 주요 성분구성 유
형의 누적빈도 분포



점에 있다는 사실을 직시하자는 말이다. 그 사실을 잊고 무리하게 __형으로만 가려 한다면 다 같이 빈한해질 뿐이다.

이념에 투철하려는 고집보다는 사실을 잘 살펴보고 부작용이 많으면 현실에 맞게 고쳐 나가는 자가 현명한 것이다. 경제정책에 정치를 과잉 개입시키면서 시장을 이기는 정부는 없는 법이고 사회주의가 완전 성공한 역사는 없다. 자영업 기반 붕괴로 일자리가 없어지고 분배도 악화되고 소비, 투자 축소로 출산율 0, 성장률 1%, 제2의 IMF 위기감을 맞고 있는 현정권 시대에 우리는 적어도 현명해지려고 애쓰는 지도자가 이 글처럼 잘 성찰해서 우리 운명에 조심스레 관여하기를 바라고 싶은 것이다.

참고문헌

1. 문영호(2001) *조선어 어휘 통계학*, 박이정.
2. Lee, Sang Oak (1989b) A Glottometrical Study of Korean Lexicon, *Harvard Studies in Korean Linguistics III*, Dept. of Linguistics, Harvard Univ., Hanshin Pub. 159-166.
3. 이상억(1990b) 현대 국어 음변화 규칙의 기능부담량, *어학연구* 26.3, 441-467.
4. 이상억(1993) 통사적 현상의 기능부담량에 대하여: 현대 국어에 대한 계량적 조사, *서울대학교원 국어연구회 편, 안병희 선생 회갑기념논총: 국어사 자료와 국어학의 연구*, 문학과 지성사, 475-487.
5. 이상억(2005) "Preferred Colors in Various Cultures based on Lexical Frequency: Comparison among Korean, English, French, Russian, Romanian, Chinese, Japanese, and Mongolian," *The 30th Anniversary Publication of International Circle of Korean Linguistics*. Seoul: Pagijong Press. 185-210.
6. 홍정하, 김주영, 강범모(2007) 세종 구문분석 말뭉치의 구축과 통사 범주 및 기능의 통계적 분포, 정부 교육인적자원부 및 한국학술진흥재단의 지원 연구.
7. Meyerstein, R. S. (1970) Functional Load, *Janua Linguarum* Nr. 99, Mouton.