

# 일상에서 찾아볼 수 있는 개념 ‘엔트로피-’

(表現의 정수는 의미밀도 향상과 문맥의 모순 제거)

공과대학 전기컴퓨터공학부 명예교수 안 수 길

지난 번 글을 통해 필자는 30년 전 한국의 소설과 신문 등 “우리의 어문생활에는 상당히 많은 ‘술어/術語’와 중견어/中堅語들이 들어 있다”는 이야기를 했다. 오늘날 ‘짐작이 된다’나, ‘일부로’, ‘사귀다’ 등의 표현이 있는 데도 ‘추정/推定’, ‘고의적’, 그리고 ‘교제하다’와 같은 또 다른 표현도 많이 쓰이는 것이 그 구체적인 예가 될 것이다.

이와같이 우리 언중/言衆은 術語급 단어들도 쉽게 배워 큰 어휘를 구사하는 성향이 있는데 이러한 대어휘/大語彙 구사 성향이 英語의 경우보다는 심하기 때문에 한국말 Zipf의 곡선이 높게 되고 따라서 평균정보 전달/표현효율(엔트로피- : Entropy)이 더 높은, 다시말해 <어문/語文 면에서 더> 문화적인 족속이었다는 자랑(“기초어휘에서 맴도는 어문생활 - Zipf의 법칙” 서울대학교 명예교수회보’ 2007 - 제3호 참조)도 하고, 더 나아가서 ‘우리가 우수하다’는 주장을 받아드리기에 마음이 내키지 않는 사람들을 위해서는 “우리가 우수했다는 것보다는 서양/西洋사람들이 생각키보다 ‘문자(elegant wording)를 안 쓰고’ 더 흔한 단어/單語를 쓰며, 전치사/前置詞 등을 동원해서 숙어/熟語나 재귀동사/再歸動詞 등 형태로 쉬운 단어의 결합을 씬으로 해서, 아직 없거나 있어도 보급되지 못한 術語들을 대신하곤 하기 때문에 기초단어에서 맴도는 경향이 강하다고 생각하면 된다”고 설명하기도 했다.

Zipf의 곡선 기울기와 언어효율의 개념과약을 위해서 반드시 알고 있어야 하는 것이 ‘Entropy의 개념’이다. 그래서 이번에는 ‘Entropy’라는 단어의 설명을 시도하는 한편, 글을 씬에 있어서 같은 내용(의미)을 전달, 또는 표현하면서도 글의 총량을 줄이고, 單語들 間 모순을 제거, 문장의 一義化를 이룩함으로써 문맥判斷 때 소모하게 되는 知的 人力소비와, 이에 따른 피로를 줄이는 문제를 喚起시키기로 한다.

‘Entropy’는 열역학에서 나온 개념으로 이를 일상생활에 응용하자면 “무질서의 尺度”라고 표현하면 우선 이해하기가 쉬운 것이다. ‘Entropy’란 “무질서하면 수치가 올라가고, 질서를 세우면 오히려 數値가 줄어드는 量”이다.

## ‘Entropy’는 무질서의 척도

따라서 秩序 존중을 우선으로 생각하는 사람들을 위해 ‘Entropy’와는 반대로 움직이는 <질서가 향상되면 수치가 커지는> 새 지수/指數 “네겐트로피-”를 만들어 좋은 일에는 따라 올라가게 (Entropy의 反數)하자는 제안을 한 사람(Leon Brillouin)도 있다.

그런데 그렇게 해서 ‘네겐트로피’는 ‘善’, 그리고 ‘Entropy’는 ‘惡’이고 따라서 “우리는 네겐트로피 - 의 증진을 위해서 노력한다”로 끝나면 좋겠지만 현실은 그렇게 간단치가 않아서 네겐트로피 - 數値가 올라가는 일보다는 Entropy가 높은 것을 더 평가하는 일이 많아지고 있다. 예를 들어 戰爭을 할 때 뻔히 보이는 것보다는 감춰지는 것이 중요하고, 이 경우 “뻔한 것”은 Entropy 감소로 나타난다. 敵이 아방의 전략을 알아차린 눈치면 허위첩보를 흘려 적에게 ‘읽어냈다’고 생각한 것에) 확신을 못 갖게 하면 적방의 Entropy는 증가한다.

‘네겐트로피’는 ‘질서’를 뜻하니까 규칙이 있고, 그 규칙에 따라 여러 가지 추리가 가능하게 된다. 이러한 추리가능성(즉, 高 네겐트로피-)은 敵에게 ‘읽힌다’는 뜻이고 따라서 아방의 敗北를 뜻하기도 해서 Entropy를 높게 유지하는 것이 필요하게 되는 것이다.

일상생활에서도 사람들은 ‘High entropy’의 상황에 더 관심을 갖는 경향이 있다. 사람들을 추리불능한(높은 entropy) 상태로 두어두면 궁금해서 더 쉽게 돈을 내는 경향이 있다는 것도 알려지고 있다.

‘네겐트로피-가 높다’, 다시말해 ‘질서 정연하다’는 것은 위에서 말한 바와 같이 그 질서 또는 規則를 통해 정확한 추리가 가능하기 때문에 네겐트로피-가 크면 그 질서 속에서 안심은 하지만 ‘네겐트로피- 100%의 질서 사회’는 너무나 뻔해서 흥미면에 있어서는 떨어지는 것이다.

## 사람 마음의 양면성/兩面性

사람들이 생활에 꼭 필요한 사물에 대해서 代價를 치루는 것은 당연하지만 지금은 호기심을 위해서 더 많은 돈을 쓰는 경향도 있다. 예를 들어 <내용이 뻔한: 高 네겐트로피-> 修身교과서를 어른들을 위해서 출판을 했다가는 안 팔려서 재정적으로 망하게 되지만 ‘稀貴한 일, 있을 수 없는 일’에 관해서 소설이나 드라마를 쓰면 잘 팔려 돈을 모으게 되는 것이다.

통신/通信, 따라서 방송에 관해서도 ‘뻔히 알만한 일’을 전달하면 별로 주의를 끌지 못하지만 非정상적인 일, 또는 완전히 상상을 초월한 稀貴한 일을 알려주면 거기에 들게 되는 통신료에 관해서는 군소리가 없다.

이 경우 빠른 것을 뜻하는 ‘네겐트로피’는 가치가 없고 그 반대인 ‘Entropy’, 즉 도입부에서 ‘무질서의 척도’라고 했던 指數에게만 가치가 부여되는 것이다. 즉 통신의 본질은 ‘짐작이 안 되는 부분을 전달하는 것’이고, 곧 추리가 되는 질서(규칙)성분은 信號의 낭비이며 通信 때 쓰게 되는 펄스(pulse) 數만 키우는 꼴이 된다(情報理論에서 펄스단위는 binit이고, 흔히 알려진 “bit”는 ‘實(알맹이) 情報의 단위’로 이 實 알맹이 정보(net information)는 낭비부분이 없을 때, 즉 實정보가 커짐에 따라 bit 수가 커져서 binit 수치에 근접하게 되지만 binit 수를 넘지는 못한다. 다시 말해 浪費가 전혀 없을 때 bit 수는 가득 차게 된다.)

진정한 정보/情報, 즉 ‘추리不能부분(high entropy)’이 아니고 반대로 추리가 가능한 성분을 우리는 ‘Redundancy’라 부르고 40년 전에 대한電子공학회에서는 이 ‘Redundancy’를 “용장/冗長部分”으로 用語제정을 해서 발표했지만 당시 기자들 눈에 띄지않고 지나다가 누군가가 쓰기 시작한 ‘군더더기’라는 말이 더 보급이 된 것이다.

### 字素/音素의 數(Repertory)가 많으면 정보효율이 높다

한 언어에 있어서 그 바탕이 되는 요소(소리의 경우는 서로서로 구분을 할 수 있게 해주는 바 音韻)의 개수 <영어는 40~60韻>, 글자의 경우는 알파벳의 글자(西歐언어의 경우는 26)의 개수는 그 언어의 표현력을 위해서 중요하다. 거칠게 말해서 이 要素數가 클수록 그 언어의 표현력이 커지고, 또 그 평균표현력이 상기한 바 ‘entropy’이다. 즉 이 ‘26자(英)’라는 기초 판독(reading, 또는 hearing)要素의 가지數가 크면 그 中에서 어느것이 發生할지 몰라서 불확실성이 커지기 때문에 그 언어의 정보표현능력(entropy)이 커지는 것이다(다음에 나올 단어가 짐작이 돼버리는 초등교생 표현보다는 興味가 유지된다).

‘바탕이 되는 요소’가 많을수록(문장의 경우는 大어휘일수록) 유리하고 이와 반대의 극단이라고 말할 수 있는 ‘모-르스 기호: Morse code’의 경우를 생각한다면 그 옛날 이를 활용했던 電信시대에도 전보를 통해 完全意思소통은 했지만 원래 모-르스 기호에는 점(dot)과 선(dash)이란 두 elements (要素)밖에 없기 때문에 이 element들 각각의 發生확률이 커서 그 하나하나가 가져온 평균정보는 무척 적었던 것이다.

다시 말해 모-르스 기호(2부호)의 경우는 요소가 다양하지 못해서(발생확률이 너무 크고, 따라서) 實 情報量이 적어서 더 긴 시간동안을 交신해야 했던 것으로 해석할 수가 있다.

반대로 4부호 시스템이나 16부호시스템을 쓰면 한 부호 당 평균정보가 커져서 짧은 시간에 같은 實정보량을 보낼 수가 있다. 이는 또한 近世에 와서 學術이 다양하게 발달한 서구사회에서 라틴語 단어나 希臘文字를 도입 병용하면서 단어수와 기호수를 키울 수밖에 없었던(숨어있는)

이유이기도 하다.

발생할 수 있는 부호 종류가 많은 경우 각각의 發生確率이 적어서 그 하나하나의 기호가 가져오는 정보는 커지고 따라서 짧은 시간에 의사소통이 되는 것이다.

그러한 뜻에 있어서 韓國音韻(91韻: 東國正韻)과 한글(자소는 24이지만 판독에 쓰는 ‘조합된 글자’ 數는 극히 크다)과 역시 글자 수가 많은 漢字는 정보량에 있어서 극히 유리하다고 하겠다. 문제는 후손인 우리의 ‘韻’ 判定능력이 급격히 떨어져가고 있다는 사실이고, 거기에 더 해서 이 사회에 그러한 ‘국어의 퇴화’를 중요하지 않는 것으로 생각하게 하는 흐름이 있다는 점이다.

대화를 할 때는 표정도 보고 성조/聲調 등 side information이 있어서 말을 무리할 정도 짧힐 수가 있지만, 문서로만 다루게 될 경우는 완전한 이해를 위해서는 文章만으로도 오해의 여지가 없게, 다시 말해 ‘말이 되게(論理性)’ 해줘야 하기 때문에 文語體가 생겼던 것이다. 그런데 작금과 같이 口語體만이 인정되는 분위기에서는, 표정과 말투가 나타나는 TV, 그리고 최소한 성조/聲調와 발음의 장단/長短이 도와줄 수 있는 라디오에 비해서 신문과 잡지 등은 크게 불리하게 되기 때문에 관계자, 그리고 많은 경우에 그 활동이 時空을 초월해야 하는 學者, 그리고 전문가들은 文語體를 꼭 지켜내야 할 것이다.

그래도 일반론적으로 말해서 音素종류가 한정된 ‘소리 요소(音文의 경우)’는 정보가 적고, 그 조합이 무궁무진할 수 있는 ‘자형요소(義文의 경우)’는 가져오는 정보가 크다는 것은 확실한 사실이어서 ‘Herald: 왕의 布告官’이 소리로 말하는 유럽에 비해 거리마다 ‘榜’을 붙이는 극동의 정보전달 방식이 그러한 뜻에 있어서 더 우수한 것이었다.

또 한편에 있어서 시대의 발달에 따라 인간의 활동내용이 커지고 複雜하게 되어가고 있고, 그래서 의사소통이 上述한 바와 같이 음운판별력/音韻判別力이 떨어져 가는 속에서 ‘소리言語(音文)’만 갖고는 ‘정확한 의사소통’은 불가능하게 되어가고 있기 때문에, 머리속에 漢字자형을 알고 있어서 同音異義의 경우와 같은 소리정보의 混同을 풀어낼 수 있는 극동의 語文方式(義文)이 더 유리하다고 말 할 수가 있을 것이다.

## ‘게임’ 결과를 모르기 때문에 경기장에 간다.

돈을 받게 되는 ‘High Entropy’의 사례로는 스포츠 메취를 이야기할 수 있다. 즉 승부/勝負가 뻔한 경기에는 경기장 入場客은커녕 시청률까지도 떨어지게 되는 데 이 때 문제가 되는 것은 경기 결과가 ‘누구나가 다 추리할 수 있을 정도’ 뻔하다는 것으로, 다시 말해 ‘entropy가 낮다’는 것이기 때문에, 이 현상자체가 “Entropy=돈”이라는 경향의 하나의 나타남(發現)이라고 말할 수가 있을 것이다.

이제 우리의 이야기를 元 出發點인 熱力學의 경우로 돌려오자면 蒸氣機關의 ‘안’과 ‘밖’의 경우와 같이 한편은 온도가 높고 한편은 온도가 낮은 경우 ‘규칙성이 있는 썸’이니 entropy가 낮다. 즉 양자를 쉽게 구별할 수가 있다. 그런데 우리가 熱機關을 통해 그 에너지를 빼 쓰게 되면 兩者의 온도차가 줄어들게 되어 entropy는 증가한다.

이 경우 ‘한 편은 뜨겁고, 다른 한 편은 안 뜨겁다’와 같이 ‘뚜렷했던 것이 구분이 없어지는’ 상태, 또는 내연기관/內燃機關 내의 분자(속도) 분포와 내연기관 외부의 분자(속도) 분포 간 차이가 없는 상태에서의 entropy는 가장 크게 되는데, 이와 같이 熱境界의 시각으로 봐서 쓸모가 없게 된 상태를 ‘熱的死亡’이라고 말한다. ‘사망’을 이야기하는 것은 그 이상 이용할 것이 없기 때문이다.

### “Entropy는 커지기만 한다”는 데 ...

닫힌 체계(외부에서 에너지 공급이 없는 조건)에서 “Entropy는 증가하기만 한다”라는 ‘entropy의 일방적 증가’를 지적하는 것이 熱力學 제 2법칙이다. 즉 제2법칙은 “閉鎖系에서 entropy는 올라가기만 한다”로 표현되는 것이다.

이와 같이 하나의 閉鎖系는 熱力學 제2법칙에서 벗어날 수가 없지만 외부에서 에너지가 투입될 경우 반대로 entropy가 내려가게 할 수 있기 때문에(그렇다고 해서 에너지만 투입하면 浪費를 해도 net entropy가 내려가는 것은 아닌 사정은 위에서 설명한 바 binit가 전부 bit증가로 이어지는 것은 아닌 사정과 같다) 秩序 상실이 運命지워져 있는 인류가 다시 질서를 증가시키는 가능성이 아주 없는 것은 아니다.

한국 정치상황은 조선조 창건/創建 때의 유교/儒敎 도입과 같이 강력한 新制度를 시작하는 경우, 그리고 근자에서는 大選과 政權交替 때에만 네겐트로피가 커지곤 했다. 유교와 같은 하나의 질서가 허물어지는 것은 eEntropy의 증가로 볼 수 있다.

국제적인 흐름에 따라 한국에서도 10년간의 노력을 쏟아 에스컬레이터 승용 시 우측으로 붙어서게 만들어 놓은 것을 서울의 모 지하철공사가 “두 줄로 다녀야(자기 회사 영업권 내에서) 安畵事故 인원 수가 줄어든다”고 다시 옛날의 혼란으로 유도하고 있는 것은 모처럼 내려놓은 entropy를 다시 올리는 것이 된다. 이는 자기 회사 영업권내 판단 기준을 위해서는 안전할지 모르지만 급한 사람들에게는 전체효율을 크게 떨어뜨리는 결과가 된다. “년간 418명이라는 안전사고 증가”를 거론하면서 게다가 개똥을 하겠다고 공공예산을 들여 여러 宣傳포스터를 만들어 사람들을 세뇌하고 있는데 그보다는 지금까지 10년의 秩序確立 노력과 국제사회전체가 경험을 쌓아서 얻어낸 ‘한편으로 다가서기’의 지혜를 일단 존중하고 (이미 적자인) 예산을

아껴쓰는 검손을 배우는 것이 선행되어야 할 것이다.

Entropy는 방치해봐도 저절로 올라가는 것이어서 쏠세계가 ‘한편으로 다가서기’의 질서를 유지하느라 노력하고 있는 속에서, 서울지하철이 여러 포스터-를 제작/게시해서 사회전반에 감히 작용해 가면서 ‘Entropy를 힘들여 올라가게 유도하는 것’은 無知의 극단이라고 해야 할 것이다.