論文 89-26-9-1

한글 음절의 초성, 중성, 종성 단위의 발생확률, 엔트로피 및 평균상호정보량

(Entropy and Average Mutual Information for a 'Choseong', a 'Jungseong' and a 'Jongseong' of a Korean Syllable)

李 在 弘*, 吳 相 縣**

(Jae Hong Lee and Sang Hyun Oh)

要 約

한글의 음절은 그 발생의 확률적 성질에 따라 확률변수로 간주된다. 한글의 음절은 확률변수들로 간주되는 초성, 중성, 종성으로 나누어진다. 이 세 확률변수에 대하여 모든 가능한 결합 확률과 조건부 확률이 한글의 음절별 누적 빈도수 통계로부터 계산된다. 세 확률변수에 대하여 모든 가능한 결합 엔트로피와 조건부 엔트로피가 결합확률과 조건부 확률로부터 계산된다. 또한 세 확률변수에 대하여 모든 가능한 평균상호정보량이 계산된다. 이들 엔트로피아 평균상호정보량으로부터 한글 음절의 성질을 정보이론적으로 분석한다.

Abstract

A Korean syllable is regarded as a random variable according to its probabilistic property in occurrence. A Korean syllable is divided into a 'choseong,' a 'jungseong,' and a 'jongseong' which are regarded as random variables. From the cumulative frequency of a Korean syllable all possible joint probabilities and conditional probabilities are computed for the three random variables. From the joint probabilities and the conditional probabilities all possible joint entropies and conditional entropies are computed for the three random varibles. Also all possible average mutual informations are calculated for the three random variables. Average mutual information between two random variables hss its biggest value between choseong and jungseong. Average mutual information between a random variable and other two random variables has its biggest value between jungseong and choseong-jongseong.

*正會員, **準會員, 서울大學校 電子工學科
(Dept. of Elec. Eng., Seoul Nat'l Univ.)
**準會員, (株) 金星社.
(GoldStar Co., Ltd.)

接受日字 : 1989年 5月 29日
(※ 이 논문은 1988년도 문교부 지원 한국 학술진흥재단의 자유공모과제 학술연구조성비에 의하여 연구되었음.)

I. 서 론

언어는 인간의 의사전달의 가장 중요한 수단이다. 언어가 문자화 된것이 글이다. 글은 기호들이 연속적으로 나열된 기호열로 볼 수 있다. 기호열은 정보를 내포하는데 기호열이 내포하는 정보량을 계량하려는 시도가 있어왔다. 즉, Shannon이 정보이론에 기초하여[1], 영어 기호열의 정보량 척도로서 엔트로피(entropy)를 제안한 이래 몇몇 언어에 대하여 그 엔트로피

가 계산되었다.[2, 3] 한글에 있어서는 Shannon의 계산방법을 이용한 자소단위의 엔트로피에 관한 연구가 몇 차례 있었다.[4~6] 이들 연구에서는 한글의 24자모를 영어의 26자의 알파벳에 대응되는 것으로 보고 엔트로피를 구하였다.

한글은 표음문자로서 모아쓰기를 한다. 모아쓰기를 함으로써 초성, 중성, 종성이 모여서 한 개의 도형적 合字즉, '음절'(以下 '음절'로 표기)을 구성하는데 이러한 도형상 특성은 다른 언어에는 찾아볼 수 없는 것이다. 이러 특성을 고려할때, 초성, 중성 또는 종성을 단위로 한 발생의 불확실성(uncertainty) 즉, 엔트로피와 초성이 중성에 대하여 제공하는 정보량, 초성과 중성이 종성에 대하여 제공하는 정보량등의 계산분석 결과는 한국어 음성인식및 합성, 자연언어처리, cryptography, 언어학, 음성학 등에 유용할 것이다. 그러나 이러한 점을 고려한 논문은 몇 년전까지는 전무하였으며 최근 부분적인 연구 결과가 발표된 바 있다.[7]

한글의 엔트로피 계산에는 한글의 낱말 또는 음절 단위의 빈도 분포가 필요하다. 1945년 이래 한글의 낱말 또는 음절 단위의 빈도 분포의 체계적 조사는 빈약하였다.[8, 9] 조사의 내용에 있어서도 다른 언어에 대한 보다 체계적인 분포 조사와 통계적 분석에 비하여 빈약하였다.[10] 한편, 한글사전의 표제어에 나타난 음소의 빈도분포를 구한 연구가 있었으나,[11] 사전에 나타난 각 표제어의 사용빈도가 서로 다르기 때문에 사전 표제어에 나타난 음소의 빈도분포와 실제 사용되는 음소의 빈도분포는 차이가 있을 것이다.

이 논문에서는 먼저 한글의 음절별 발생빈도수 분포로 부터 구한 음절별 발생 확률로 부터 초성, 중성, 종성 단위의 조건부 발생확률(conditional probabilitiy)을 구하고 그로부터 초성, 중성, 종성 단위의 조건부 엔트로피를 구한다. 또한 초성, 중성, 종성 단위간의 평균상호정보량(average mutual information)을 구하고 비교한다. 이 결과로부터 초성, 중성 또는 종성이 무슨 자소인가를 아는 것이 같은 음절 내의 다른 단위(초성, 중성, 종성)가 무슨 자소인가에 대하여 얼마만한 정보를 제공하는지를 분석한다.

Ⅱ. 한글의 엔트로피와 정보량

한글의 도형적 合字(이하 '음절')를 구성하는 초성, 중성, 종성은 각각 가능한 자소집합으로부터 어떤 확률로 발생하는 확률변수(random variable)로 간주할 수 있다. 초성, 중성, 종성을 표시하는 확률변수를 각각 X, Y, Z라 정의하고 그 가능한 자소의 집합 즉, 표본 공간(sample space)을 각각 A_X, A_Y, A_Z 초성 x의 발생확률, 중성 y의 발생확률, 종성 z의 발생확률을 각각 p(x), p(y), p(z)라 하자. 그러면 A_X, A_Y, A_Z는 다음과 같이 주어진다.

A_X={ㄱ, ㄲ, ㄴ, ㄷ, ㄸ, ㄹ, ㅁ, ㅂ, ㅃ, ㅅ, ㅆ, ㅇ, ㅈ, ㅉ, ㅊ, ㅋ, ㅌ, ㅍ, ㅎ}

A_Y={ㅏ, ㅐ, ㅑ, ㅒ, ㅓ, ㅔ, ㅕ, ㅖ, ㅗ, ㅘ, ㅙ, ㅚ, ㅛ, ㅜ, ㅝ, ㅞ, ㅟ, ㅠ, ㅡ, ㅢ, ㅣ}

A_Z={공백소, ㄱ, ㄲ, ㄳ, ㄴ, ㄵ, ㄶ, ㄷ, ㄹ, ㄺ, ㄻ, ㄼ, ㄽ, ㄾ, ㄿ, ㅀ, ㅁ, ㅂ, ㅄ, ㅅ, ㅆ, ㅇ, ㅈ, ㅊ, ㅋ, ㅌ, ㅍ, ㅎ}

공백소(blank)는 자소가 결여된 상태 즉, 확률변수의 outcome이 null인 상태를 표시하며 종성에만 발생할 수 있고 초성, 중성에는 발생하지 않는다. 초성과 종성에 공통으로 사용되는 자소의 집합은 다음과 같이 주어진다.

$A_X \cap A_Z$={ㄱ, ㄲ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅆ, ㅇ, ㅈ, ㅊ, ㅋ, ㅌ, ㅍ, ㅎ}

한 음절 내의 초성, 중성, 종성간의 발생의 상관관계를 분석하기 위하여 먼저 모든 가능한 엔트로피, 결합 엔트로피, 조건부 엔트로피를 구한다. 엔트로피의 유형에는 초성의 엔트로피 H(X), 초성, 중성(X, Y)의 결합 엔트로피 H(X, Y), 초성, 중성, 종성(X, Y, Z) 즉, 음절의 엔트로피 H(X, Y, Z), 중성 Y가 주어질 때 초성의 조건부 엔트로피 H(X|Y), 종성 Z가 주어질 때 초성, 중성, (X, Y)의 조건부 엔트로피 H(X, Y|Z), 중성, 종성(Y, Z)가 주어질 때 초성의 조건부 엔트로피 H(X|Y, Z)가 있으며 각각 다음의 식으로 구해진다.[11, 12]

$$H(X) = -\sum_{x} p(x) \log p(x) \quad (1)$$

여기서 p(x)는 초성 x의 발생확률이다.

$$H(X, Y) = -\sum_{x,y} p(x, y) \log p(x, y) \quad (2)$$

여기서 p(x, y)는 초성, 중성(x, y)의 결합확률이다.

$$H(X, Y, Z) = -\sum_{x,y,z} p(x, y, z) \log p(x, y, z) \quad (3)$$

여기서 p(x, y, z)는 초성, 중성, 종성 즉, 음절(x, y, z)의 결합확률이다.

$$H(X|Y) = -\sum_{x,y} p(x, y) \log (x|y) \quad (4)$$

여기서 p(x|y)는 중성이 y때 초성이 x일 조건부 확률이다.

$$H(X, Y|Z) = -\sum_{x,y,z} p(x, y, z) \log p(x, y|z) \quad (5)$$

여기서 p(x, y|z)는 종성이 z일때 초성이 x이고 중성이 y일 조건부 확률이다.

$$H(X|Y,Z) = -\sum_{X,Y,Z} p(x,y,z) \log p(x|y,z) \quad (6)$$

여기서 $p(x|y,z)$는 중성이 y이고 종성이 z일때 초성이 x일 조건부 확률이다.

이러한 유형의 엔트로피들 간에 다음의 관계가 성립한다.[12,13]

$$H(X,Y) = H(X) + H(Y|X) \quad (7)$$

$$H(X,Y,Z) = H(X,Y) + H(Z|X,Y) \quad (8)$$

$$H(X,Y,Z) = H(X) + H(Y,Z|X) \quad (9)$$

이 엔트로피로부터 평균상호정보량이 다음과 같이 구해진다.

$$I(X;Y) = H(X) - H(X|Y) \quad (10)$$

$$I(X,Y;Z) = H(X,Y) - H(X,Y|Z) \quad (11)$$

$$I(X;Y,Z) = H(X) - H(X|Y,Z) \quad (12)$$

평균상호정보량간에 다음의 관계가 성립한다.[12,13]

$$I(X;Y) = I(Y;X) \quad (13)$$

$$I(X;Y,Z) = I(Y,Z;X) \quad (14)$$

Ⅲ. 계산 및 결과

초성, 중성, 종성의 발생확률, 결합확률 및 조건부 확률을 1985년 유 재원이 조사한 1,547개 한글 음절의 빈도수[9]로부터 계산하고, 결합확률과 조건부 확률로부터 초성, 중성, 종성의 결합 엔트로피와 조건부 엔트로피 및 평균 상호 정보량을 계산한다.

1. 초성, 중성, 종성의 발생 확률, 결합 확률과 조건부 확률

음절(x, y, z)의 발생확률 $p(x,y,z)$는 '전체 음절별 누적 빈도수'로부터 계산된다. $p(x,y,z)$의 계산값은 음절(x, y, z)에 따른 데이타 앙이 방대하여 여기에 싣지 않는다.

음절별 발생빈도수 분포 통계로부터 음절(x, y, z)의 발생확률 $p(x,y,z)$가 계산되고, $p(x,y,z)$로 부터 초성의 발생확률 $p(x)$가 다음 식으로 주어진다.

$$p(x) = \sum_{Y,Z} p(x,y,z) \quad (15)$$

중성의 발생확률 $p(y)$와 종성의 발생확률 $p(z)$ 도 같은 방법으로 구해진다. $p(x)$, $p(y)$와 $p(z)$의 계산값을 표 1 에 보인다.

표　1. 초성, 중성, 종성의 발생확률
Table　1. Probabilities of a choseong, a jungseong, and a jongseong.

x	ㄱ	ㄲ	ㄴ	ㄷ	ㄸ	ㄹ	ㅁ	ㅂ	ㅃ	ㅅ	ㅆ	ㅇ	ㅈ	ㅉ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
p(x)	.12519	.03043	.04511	.07900	.02138	.09746	.06394	.07112	.00986	.07769	.01289	.11429	.09593	.01419	.03107	.00980	.01779	.00986	.06365

(a) 초성의 발생확률 p(x)

y	ㅏ	ㅐ	ㅑ	ㅒ	ㅓ	ㅔ	ㅕ	ㅖ	ㅗ	ㅘ	ㅙ	ㅚ	ㅛ	ㅜ	ㅝ	ㅞ	ㅟ	ㅠ	ㅡ	ㅢ	ㅣ
p(y)	.24227	.05231	.00910	.00009	.13293	.01872	.02091	.00156	.09466	.00609	.00170	.00762	.00266	.12027	.00209	.00063	.01110	.00131	.08739	.00192	.18468

(b) 중성의 발생확률 p(y)

z	공백소	ㄱ	ㄲ	ㄳ	ㄴ	ㄵ	ㄶ	ㄷ	ㄹ	ㄺ	ㄻ	ㄼ	ㄽ	ㄾ	ㄿ	ㅀ	ㅁ	ㅂ	ㅄ	ㅅ	ㅆ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
p(z)	.50180	.00149	.00125	.00027	.06531	.00051	.00056	.00400	.12298	.00226	.00032	.00080	.00005	.00012	.00002	.00044	.05238	.02106	.00192	.03148	.00035	.08930	.00473	.00348	.00019	.00707	.00363	.00221

(c) 종성의 발생확률 p(z)

음절(x, y, z)의 발생확률 p(x, y, z)로부터 초성 x와 중성 y의 결합확률 p(x, y)가 다음 식으로 주어진다.

$$p(x, y) = \sum_{z} p(x, y, z) \tag{16}$$

초성 x와 종성 z의 결합확률 p(x, z)와 중성 y와 종성 z의 결합확률 p(y, z)도 같은 방법으로 구해진다. p(x, y), p(x, z)와 p(y, z)의 계산값은 여기에 싣지 않는다.

음절(x, y, z)의 발생확률 p(x, y, z)로부터 중성이 y일때 초성이 x일 조건부 확률 p(x|y)가 다음 식으로 주어진다.

$$p(x|y) = \frac{p(x, y)}{p(y)} = \frac{\sum_{z} p(x, y, z)}{\sum_{x,z} p(x, y, z)} \tag{17}$$

종성이 z일 때 초성이 x일 조건부 확률 p(x|z), 초성이 x일 때 중성이 y일 조건부 확률 p(y|x), 종성이 z일 때 중성이 y일 조건부 확률 p(y|z), 초성이 x일 때 종성이 z일 조건부 확률 p(z|x)와 중성이 y일 때 종성이 z일 조건부 확률 p(z|y)도 같은 방법으로 구해진다. p(x|y), p(x|z), p(y|x), p(y|z), p(z|x)와 p(z|y)의 계산값을 각각 표 2 에 보인다.

음절(x, y, z)의 발생확률 p(x, y, z)로부터 종성이 z일 때 초성이 x이고 중성이 y일 조건부 확률 p(x, y|z)가 다음 식으로 주어진다.

$$p(x, y|z) = \frac{p(x, y, z)}{p(z)} = \frac{p(x, y, z)}{\sum_{x,y} p(x, y, z)} \tag{18}$$

중성이 y일 때 초성이 x이고 종성이 z일 조건부 확률 p(x, z|y)와 초성이 x일 때 중성이 y이고 종성이 z일 조건부 확률 p(y, z|x)도 같은 방법으로 구해진다. p(x, y|z), p(x, z|y)와 p(y, z|x)의 계산값은 초성 x, 중성 y, 종성 z에 따른 데이타 양이 방대하여 여기에 싣지 않는다.

음절(x, y, z)의 발생확률 p(x, y, z)로부터 중성이 y이고 종성이 z일 때 초성이 x일 조건부 확률 p(x|y, z)가 다음식으로 주어진다.

$$p(x|y, z) = \frac{p(x, y, z)}{p(y, z)} = \frac{p(x, y, z)}{\sum_{x} p(x, y, z)} \tag{19}$$

초성이 x이고 종성이 z일 때 중성이 y일 조건부 확률 p(y|x, z)와 초성이 x이고 중성이 y일 때 종성이 z일 조건부 확률 p(z|x, y)도 같은 방법으로 구해진다. p(x|y, z), p(y|x, z)와 p(z|x, y)의 계산값은 초성 x,

표 2. 초성, 중성, 종성의 조건부 확률 p(x|y)型
Table 2. Conditional probabilities for a choseong, a jungseong, and a jongseong p(x|y) type.

y \ x	ㄱ	ㄲ	ㄴ	ㄷ	ㄸ	ㄹ	ㅁ	ㅂ	ㅃ	ㅅ	ㅆ	ㅇ	ㅈ	ㅉ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
ㅏ	.09321	.02278	.06781	.06854	.01959	.05080	.06998	.09230	.00945	.06861	.01683	.06249	.09941	.01700	.02505	.00714	.01707	.00945	.16977
ㅐ	.10760	.03030	.08054	.15006	.03354	.06450	.08070	.09399	.02252	.13725	.00940	.02090	.03581	.01215	.03565	.00535	.01847	.02252	.04132
ㅑ	.04007	.02516	.08947	-	-	.04660	.23299	.02237	.01491	-	-	.48183	-	-	-	.01678	-	.01491	.02703
ㅒ	.36364	-	-	-	-	-	-	-	-	-	-	.63636	-	-	-	-	-	-	-
ㅓ	.19622	.02583	.02615	.06964	.02857	.09062	.05408	.06919	.01282	.07276	.01333	.11504	.11090	.01588	.03539	.00829	.02315	.01282	.02066
ㅔ	.08741	.02491	.03623	.09194	.04710	.21286	.07201	.03578	-	.12726	.00091	.07246	.10145	.00543	.02627	.00408	.01721	-	.03668
ㅕ	.18119	.00567	.04824	-	-	.11958	.08837	.14552	.04621	.00122	-	.24970	.00203	-	.01216	.01621	-	.04621	.02513
ㅖ	.41848	-	.00543	-	-	.20109	-	-	-	-	-	.32609	-	.00543	-	.01087	-	-	.01087
ㅗ	.12331	.06268	.05409	.13674	.02087	.04639	.06000	.07871	.00546	.10674	.00734	.08185	.06197	.01854	.02669	.02266	.03949	.00546	.03143
ㅘ	.25905	.04596	-	.00418	.00557	.00279	-	.00975	-	.00696	.00418	.28412	.02646	.01114	-	.01671	-	-	.32312
ㅙ	.22500	.12000	-	.07000	.01000	-	-	-	-	.02500	.04500	.32500	.02500	-	-	.02500	.01000	-	.12000
ㅚ	.05784	.04894	.01335	.20912	.00779	.01446	.00556	.01335	-	.23804	.00556	.21691	.04116	.00556	.00334	.00222	.02892	-	.08788
ㅛ	.03185	-	.00955	-	-	.04777	.00637	-	.07006	.00955	-	.65605	-	-	-	-	-	.07006	-
ㅜ	.09240	.03996	.03538	.07612	.02347	.04257	.15104	.12363	.01318	.07464	.00959	.08549	.10058	.01734	.03073	.00430	.01818	.01318	.01219
ㅝ	.05285	.07317	.00407	.00813	-	-	.01626	.00407	-	.00407	-	.74797	-	-	-	.04472	-	-	.04472
ㅞ	.05405	.29730	.01351	.06757	-	-	-	-	-	-	.05405	.37838	.05405	-	-	.01351	.02703	-	.04054
ㅟ	.23511	.00992	.02366	.24656	.04885	.00076	.00076	-	-	.03511	-	.18626	.08779	-	.01069	.02824	.01832	-	.06794
ㅠ	.01290	-	.05806	-	-	.07097	-	-	.07742	-	-	.67097	-	-	-	-	-	.07742	.10968
ㅡ	.20652	.05898	.05199	.17664	.04792	.15191	.00204	.00058	.00146	.08798	.01630	.10069	.00621	.00126	.00407	.02406	.02716	.00146	.03250
ㅢ	-	-	.12335	-	.04846	-	-	-	-	-	.05727	.29075	-	-	-	-	.00881	-	.47137
ㅣ	.09924	.01019	.02084	.00812	.00252	.21316	.03323	.04434	.00578	.06885	.01611	.18094	.18012	.01708	.05935	.00413	.00280	.00578	.02217

(a) 조건부 확률 p(x|y)

x \ z	ㄱ	ㄲ	ㄴ	ㄷ	ㄸ	ㄹ	ㅁ	ㅂ	ㅃ	ㅅ	ㅆ	ㅇ	ㅈ	ㅉ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
공백소	.14457	.02262	.04750	.06409	.01500	.12320	.05418	.06054	.00914	.06835	.00953	.13102	.07632	.00792	.03222	.00706	.01326	.00914	.10018
ㄱ	.05056	.02923	.01602	.10403	.06928	.09747	.09092	.07885	.01259	.07563	.04057	.04816	.14969	.06907	.02309	.01061	.01571	.01259	.00759
ㄲ	.04762	.20408	.13605	.17007	-	-	.06122	.20408	-	.14966	-	.01361	-	-	-	-	-	-	-
ㄳ	-	-	.06250	-	-	-	.12500	.03125	-	.78125	-	-	-	-	-	-	-	-	-
ㄴ	.13108	.05036	.07644	.05308	.01116	.04166	.07060	.09241	.01415	.12421	.00753	.11188	.06827	.00519	.02051	.01869	.00467	.01415	.04789
ㄵ	-	-	-	-	-	-	-	-	-	-	-	1.00000	-	-	-	-	-	-	-
ㄶ	-	.24242	-	-	-	-	.06061	-	-	-	-	.12121	.28788	.01515	.25758	-	-	-	.01515
ㄷ	.20763	-	.01059	.16737	.06144	-	.08686	.33475	.02331	.03814	.00636	.02119	.02331	-	.00424	.00212	.00424	.02331	.00847
ㄹ	.10904	.02757	.04887	.11228	.02026	.00241	.10505	.08885	.01137	.10112	.01489	.09602	.11973	.01330	.02557	.01151	.02585	.01137	.02674
ㄺ	.15730	-	.08240	.11985	-	-	.05993	.10112	-	.02622	-	.21348	-	-	.02247	-	.01498	-	.20225
ㄻ	.13158	.02632	-	.07895	-	-	-	.02632	-	.34211	-	.15789	.21053	-	-	-	-	-	-
ㄼ	-	-	.59574	.10638	.04255	-	-	.10638	-	.01064	-	.03191	-	.10638	-	-	-	-	-
ㄽ	.50000	-	-	.33333	-	-	-	-	-	-	-	.16667	-	-	-	-	-	-	-
ㄾ	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.00000
ㄿ	-	-	-	-	-	-	-	-	-	-	-	1.00000	-	-	-	-	-	-	-
ㅀ	.17308	.25000	-	.03846	.23077	-	-	-	-	.05769	-	.25000	-	-	-	-	-	-	-
ㅁ	.14274	.05130	.05082	.05098	.01974	.16686	.02395	.04046	.00421	.09775	.01262	.15617	.06020	.01003	.05583	.01376	.01214	.00421	.02039
ㅂ	.13309	.05388	.03378	.03619	.00080	.14073	.01166	.08926	.00442	.03096	.01246	.07318	.27986	.01086	.01086	-	.04302	.00442	.03337
ㅃ	.17621	-	-	-	-	-	-	-	-	-	-	.82379	-	-	-	-	-	-	-
ㅅ	.20355	.02342	.01966	.09666	.01750	.07889	.07862	.08185	.01131	.05843	.00377	.12062	.08185	.00135	.04173	.00781	.01104	.01131	.03716
ㅆ	.09756	.04878	-	-	.02439	.04878	-	-	-	.02439	.04878	.65854	-	.04878	-	-	-	-	-
ㅇ	.05857	.02088	.01937	.14221	.03313	.10860	.04955	.06816	.01291	.08458	.01481	.06750	.14486	.01718	.04215	.01994	.04936	.01291	.02706
ㅈ	.15771	.02509	.12186	.02330	-	.01075	.38530	.05735	-	-	-	.01613	.17025	.01254	.01971	-	-	-	-
ㅊ	.00244	.59756	.06098	.04878	-	-	.01951	.13171	-	.03659	-	.07317	.00976	.01951	-	-	-	-	-
ㅋ	-	-	.60870	-	-	-	-	.04348	-	-	-	.34783	-	-	-	-	-	-	-
ㅌ	.24820	.12350	.04676	.00120	-	.00120	.11031	.28297	-	.05635	-	.01319	.00480	-	-	-	-	-	.05276
ㅍ	.06075	-	.07243	.04439	-	.03738	-	.00467	-	.06308	-	.59813	.11215	-	-	-	.00701	-	-
ㅎ	.06130	-	.38697	.07663	.01533	.26054	.05364	.00383	.00383	-	.03448	.04598	.03831	.01916	-	-	-	.00383	-

(b) 조건부 확률 p(x|z)

x \ y	ㅏ	ㅐ	ㅑ	ㅒ	ㅓ	ㅔ	ㅕ	ㅖ	ㅗ	ㅘ	ㅙ	ㅚ	ㅛ	ㅜ	ㅝ	ㅞ	ㅟ	ㅠ	ㅡ	ㅢ	ㅣ
ㄱ	.18039	.04496	.00291	.00027	.20836	.01307	.03027	.00521	.09324	.01259	.00305	.00352	.00068	.08877	.00088	.00027	.02086	.00014	.14416	-	.14640
ㄲ	.18134	.05209	.00752	-	.11281	.01532	.00390	-	.19499	.00919	.00669	.01226	-	.15794	.00501	.00613	.00362	-	.16936	-	.06184
ㄴ	.36415	.09339	.01804	-	.07704	.01503	.02236	.00019	.11349	-	-	.00225	.00056	.09433	.00019	.00019	.00582	.00169	.10071	.00526	.08531
ㄷ	.21019	.09936	-	-	.11717	.02178	-	-	.16384	.00032	.00150	.02017	-	.11588	.00021	.00054	.03466	-	.19539	-	.01899
ㄸ	.22205	.08208	-	-	.17764	.04124	-	-	.09239	.00159	.00079	.00278	-	.13204	-	-	.02538	-	.19588	.00436	.02181
ㄹ	.12629	.03462	.00435	-	.12360	.04088	.02566	.00322	.04506	.00017	-	.00113	.00130	.05254	-	-	.00009	.00096	.13621	-	.40393
ㅁ	.26515	.06602	.03314	-	.11242	.02108	.02890	-	.08882	-	-	.00066	.00027	.28410	.00053	-	.00013	-	.00278	-	.09598
ㅂ	.31442	.06913	.00286	-	.12932	.00942	.04279	-	.10477	.00083	-	.00143	-	.20906	.00012	-	-	-	.00072	-	.11514
ㅃ	.23216	.11952	.01376	-	.17283	-	.09802	-	.05245	-	-	-	.01892	.16079	-	-	-	.01032	.01290	-	.10834
ㅅ	.21397	.09242	-	-	.12450	.03066	.00033	-	.13006	.00055	.00055	.02335	.00033	.11555	.00011	-	.00502	-	.09896	-	.16367
ㅆ	.31624	.03813	-	-	.13741	.00131	-	-	.05391	.00197	.00592	.00329	-	.08941	-	.00263	-	-	.11045	.00855	.23077
ㅇ	.13247	.00957	.03835	.00052	.13381	.01187	.04569	.00445	.06779	.01513	.00482	.01446	.01528	.08997	.01365	.00208	.01810	.00771	.07699	.00490	.29239
ㅈ	.25104	.01953	-	-	.15366	.01979	.00044	-	.06115	.00168	.00044	.00327	-	.12609	-	.00035	.01016	-	.00566	-	.34674
ㅉ	.29032	.04480	-	-	.14875	.00717	-	.00060	.12366	.00478	-	.00299	-	.14695	-	-	-	-	.00777	-	.22222
ㅊ	.19536	.06003	-	-	.15143	.01583	.00819	-	.08131	-	-	.00082	-	.11896	-	-	.00382	-	.01146	-	.35280
ㅋ	.17647	.02855	.01557	-	.11246	.00779	.03460	.00173	.21886	.01038	.00433	.00173	-	.05277	.00952	.00087	.03201	-	.21453	-	.07785
ㅌ	.23249	.05431	-	-	.17294	.01810	-	-	.21010	-	.00095	.01239	-	.12292	-	.00095	.01143	-	.13340	.00095	.02906
ㅍ	.23216	.11952	.01376	-	.17283	-	.09802	-	.05245	-	-	-	.01892	.16079	-	-	-	.01032	.01290	-	.10834
ㅎ	.64616	.03396	.00386	-	.04315	.01079	.00826	.00027	.04674	.03090	.00320	.01052	-	.02304	.00146	.00040	.01185	.00226	.04461	.01425	.06432

(c) 조건부 확률 p(y|x)

y / z	ㅏ	ㅐ	ㅑ	ㅒ	ㅓ	ㅔ	ㅕ	ㅖ	ㅗ	ㅘ	ㅙ	ㅚ	ㅛ	ㅜ	ㅝ	ㅞ	ㅟ	ㅠ	ㅡ	ㅢ	ㅣ
공백소	.21846	.07842	.00453	.00019	.10007	.03189	.01362	.00284	.07727	.00478	.00204	.01329	.00363	.10083	.00147	.00059	.01781	.00113	.05806	.00299	.26628
ㄱ	.27647	.02757	.03714	-	.26027	.00187	.01716	-	.11079	.00551	.00166	.00052	.00187	.17143	.00083	.00187	.00031	.00302	.04588	-	.04182
ㄲ	.44898	-	-	-	.22449	-	.06122	-	.20408	-	-	-	-	.06122	-	-	-	-	-	-	-
ㄳ	.71875	-	-	-	.15625	-	-	-	.12500	-	-	-	-	-	-	-	-	-	-	-	-
ㄴ	.27476	.00870	.00260	-	.12067	.00493	.05646	.00013	.08332	.01272	.00065	.00623	-	.13757	.00805	.00065	.00156	.00065	.17599	.00519	.09929
ㄵ	.88333	-	-	-	.11667	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ㄶ	.74242	-	-	-	-	-	-	-	.01515	-	-	-	-	-	-	-	-	-	.24242	-	-
ㄷ	.41949	-	-	-	.15678	-	.01907	-	.12500	-	-	-	-	.12288	-	-	-	-	.13347	-	.02331
ㄹ	.25620	.00579	.00352	-	.14302	.00041	.01344	-	.09526	.00696	.00007	.00021	.00014	.19300	.00489	-	.00055	.00062	.13648	-	.13944
ㄺ	.31835	-	-	-	.12734	-	-	-	.02996	-	-	-	-	.13858	-	-	-	-	.35581	-	.02996
ㄻ	.44737	-	-	-	.15789	-	-	-	.15789	-	-	-	-	.13158	-	-	-	-	.05263	-	.05263
ㄼ	.22340	-	.01064	-	.74468	-	.02128	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ㄽ	-	-	-	-	-	-	-	-	1.00000	-	-	-	-	-	-	-	-	-	-	-	-
ㄾ	.35714	-	-	-	-	-	-	-	.14286	-	-	-	-	.50000	-	-	-	-	-	-	-
ㄿ	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.00000	-	-
ㅀ	.15385	-	-	-	-	-	-	-	.25000	-	-	-	-	.40385	-	-	-	-	.09615	-	.09615
ㅁ	.24794	.02104	.00421	-	.09112	.00922	.01651	-	.07671	.00016	-	.00227	.00032	.06820	-	-	.00324	-	.26849	.00049	.17009
ㅂ	.28548	.01045	.00040	-	.23884	.00080	.04624	.00040	.10334	-	-	.00080	-	.05388	-	-	.00402	-	.05147	-	.20386
ㅄ	.17621	-	-	-	.81498	-	.00881	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ㅅ	.09316	.06320	.00061	-	.14243	.02908	.03256	.00539	.08724	.00377	.00108	.00889	.00135	.13086	.00027	-	.05223	-	.15670	.00054	.17044
ㅆ	.12195	.14634	-	-	.07317	.04878	.07317	.14634	-	.04878	-	-	-	-	-	-	-	-	-	-	.34146
ㅇ	.30444	.05952	.03161	-	.17667	.00845	.03560	-	.17353	.01576	.00503	.00066	.00684	.11667	.00152	.00152	.00085	.00256	.03759	.00047	.02069
ㅈ	.57527	.02330	-	-	.12186	-	-	-	.06272	-	-	-	-	.04839	-	-	-	-	.07168	-	.09677
ㅊ	.09512	-	-	-	.00488	-	.01220	-	.65854	-	-	-	-	.04634	-	-	-	.04390	.00488	-	.13415
ㅋ	-	-	-	-	.30435	-	.60870	-	-	-	-	-	-	-	.04348	-	-	-	.04348	-	-
ㅌ	.36331	.00240	.00959	-	.10552	-	.06993	-	.08393	-	-	-	-	.14868	-	-	-	-	.10192	-	.09472
ㅍ	.22430	-	-	-	.12850	-	.06776	-	.07243	-	-	-	-	.02336	-	-	-	-	.03738	-	.44626
ㅎ	.34483	-	.01916	-	.25670	-	.01916	-	.32950	-	-	-	-	-	-	-	-	-	.01149	-	.01916

(d) 조건부 확률 p(y|z)

x / z	공백소	ㄱ	ㄲ	ㄳ	ㄴ	ㄵ	ㄶ	ㄷ	ㄹ	ㄺ	ㄻ	ㄼ	ㄽ	ㄾ	ㄿ	ㅀ	ㅁ	ㅂ	ㅄ	ㅅ	ㅆ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
ㄱ	.57950	.03291	.00047	-	.06839	-	-	.00664	.10712	.00284	.00034	-	.00020	-	-	-	.00061	.05972	.02241	.00271	.05119	.00027	.04178	.00596	.00007	-	.01402	.00176	.00108
ㄲ	.37298	.07827	.00836	-	.10806	-	.00446	-	.11142	-	.00028	-	-	-	-	-	.00362	.06830	.03733	-	.02423	.00056	.06128	.00390	.06825	-	.02880	-	-
ㄳ	.52837	.02894	.00376	.00038	.11067	-	-	.00094	.13322	.00413	-	.01052	-	-	-	-	-	.05900	.01578	-	.01372	-	.03833	.01278	.00470	.00263	.00733	.00582	.01898
ㄴ	.40708	.10730	.00268	-	.04388	-	-	.00848	.17479	.00343	.00032	.00107	.00021	-	-	-	.00021	.03380	.00966	-	.03852	-	.16073	.00139	.00215	-	.00011	.00204	.00215
ㄵ	.35210	.26408	-	-	.03410	-	-	.01150	.11657	-	-	.00159	-	-	-	-	.00476	.04837	.00079	-	.02577	.00040	.13838	-	-	-	-	-	.00159
ㄶ	.83434	.08150	-	-	.02782	-	-	-	.00304	-	-	-	-	-	-	-	-	.08968	.03044	-	.02548	.00017	.09950	.00052	-	-	.00009	.00139	.00591
ㄷ	.42516	.11587	.00119	.00053	.07212	-	.00053	.00544	.20204	.00212	-	-	-	-	-	-	-	.01962	.00384	-	.03871	-	.06920	.02850	.00106	-	.01220	-	.00186
ㄹ	.42718	.09035	.00358	.00012	.08486	-	-	.01883	.15364	.00322	.00012	.00119	-	-	-	-	-	.02980	.02648	-	.03623	-	.08558	.00381	.00844	.00012	.02813	.00024	.00012
ㄺ	.46518	.10404	-	-	.09372	-	-	.00848	.14187	-	-	-	-	-	-	-	-	.02236	.00946	-	.03611	-	.11694	-	-	-	-	-	.00086
ㄻ	.44146	.07832	.00240	.00273	.10442	-	-	.00196	.16007	.00078	.00142	.00011	-	-	-	-	.00033	.06590	.00840	-	.02368	.00011	.09722	-	.00164	-	.00513	.00295	-
ㄼ	.37081	.25641	-	-	.03813	-	-	.00197	.14201	-	-	-	-	-	-	-	-	.05128	.02038	-	.00920	.00131	.10256	-	-	-	-	-	.00592
ㄽ	.57529	.03434	.00015	-	.06394	.00445	.00059	.00074	.10332	.00423	.00045	.00022	.00007	-	.00015	.00098	.07158	.01350	.01387	.03323	.00200	.05274	.00087	.00223	.00059	.00082	.01899	.00089
ㄾ	.39922	.12715	-	-	.04648	-	.00168	.00097	.15349	-	.00071	-	-	-	-	-	-	.03287	.06150	-	.02686	-	.13484	.00839	.00035	-	.00035	.00424	.00068
ㄿ	.28017	.39685	-	-	.02389	-	.00060	-	.11529	-	-	.00597	-	-	-	-	-	.03704	.31613	-	.00299	.00119	.10812	.00418	.00478	-	-	-	.00299
ㅀ	.52033	.06057	-	-	.04311	-	.00464	.00055	.10123	.00184	-	-	-	-	-	-	-	.09413	.30737	-	.04229	-	.12115	.00300	-	-	-	-	-
ㅁ	.36159	.08824	-	-	.12457	-	-	.00087	.14448	-	-	-	-	-	-	-	-	.07353	-	-	.02509	-	.18168	-	-	-	-	-	-
ㅂ	.37399	.07194	-	-	.01715	-	-	.00095	.17866	.00191	-	-	-	-	-	-	-	.03573	.05098	-	.01953	-	.24774	-	-	-	-	.00143	-
ㅄ	.46518	.10404	-	-	.09372	-	-	.00946	.14187	-	-	-	-	-	-	-	-	.02236	.00946	-	.03611	-	.11694	-	-	-	-	-	.00086
ㅅ	.78972	.00872	-	-	.04914	-	.00013	.00053	.05167	.00719	-	-	-	.00186	-	-	-	.01678	.01105	-	.01838	-	.03795	-	-	-	.00586	-	-

(e) 조건부 확률 p(z|x)

x / z	공백소	ㄱ	ㄲ	ㄳ	ㄴ	ㄵ	ㄶ	ㄷ	ㄹ	ㄺ	ㄻ	ㄼ	ㄽ	ㄾ	ㄿ	ㅀ	ㅁ	ㅂ	ㅄ	ㅅ	ㅆ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
ㅂ	.45248	.09097	.00231	.00080	.07407	.00185	.00171	.00693	.13066	.00297	.00059	.00073	-	.00017	-	.00028	.05360	.02484	.00140	.01211	.00017	.11221	.01123	.00136	-	.01060	.00336	.00315
ㅃ	.75223	.04294	-	-	.01086	-	-	-	.01361	-	-	-	-	-	-	-	-	.02107	.00421	-	.05007	.00097	.10160	.00211	-	-	.00032	-	-
ㅅ	.24977	.33271	-	-	.01864	-	-	-	.04753	-	-	.00093	-	-	-	-	-	.02423	.00093	-	.00280	-	.31034	-	-	-	.00746	-	.00466
ㅆ	1.00000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ㅇ	.37778	.15956	.00210	.00032	.05924	.00045	-	.00472	.13233	.00217	.00038	.00448	-	-	-	-	-	.03590	.03788	.01180	.03374	.00019	.11868	.00434	.00013	.00045	.00561	.00351	.00427
ㅈ	.85507	.00815	-	-	.01721	-	-	-	.00272	-	-	-	-	-	-	-	-	.02582	.00091	-	.04691	.00091	.04031	-	-	-	-	-	-
ㅊ	.32671	.06868	.00365	-	.17633	-	-	.00365	.07904	-	-	.00091	-	-	-	-	-	.04135	.04662	.00091	.04905	.00122	.15201	-	.00203	.00567	.03040	.01176	.00203
ㅋ	.84783	-	-	-	.00543	-	-	-	-	-	-	-	-	-	-	-	-	-	.00543	-	.10870	.03261	-	-	-	-	-	-	-
ㅌ	.40960	.09537	.00269	.00038	.05749	-	.00009	.00528	.12376	.00072	.00054	-	.00054	.00018	-	.00116	.04245	.02301	-	.02901	-	.16370	.00313	.02418	-	.00627	.00278	.00770
ㅍ	.39415	.07382	-	-	.13649	-	-	-	.14067	-	-	-	-	-	-	-	-	.00139	-	-	.01950	.00279	.23120	-	-	-	-	-	-
ㅎ	.60500	.08000	-	-	.02500	-	-	-	.00500	-	-	-	-	-	-	-	-	-	-	-	.02000	-	.26500	-	-	-	-	-	-
ㅏ	.87542	.00558	-	-	.05339	-	-	-	.00334	-	-	-	-	-	-	-	-	.01557	.00222	-	.03671	-	.00779	-	-	-	-	-	-
ㅐ	.68471	.05732	-	-	-	-	-	-	.00637	-	-	-	-	-	-	-	-	.00637	-	-	.01592	-	.22930	-	-	-	-	-	-
ㅑ	.42071	.11615	.00063	-	.07471	-	-	.00409	.19735	.00261	.00035	-	.00049	-	-	.00148	.03841	.00944	-	.03425	-	.08662	.00190	.00134	-	.00874	.00070	-
ㅒ	.35366	.03252	-	-	.25203	-	-	-	.25862	-	-	-	-	-	-	-	-	-	-	-	.00407	-	.06504	-	-	.00407	-	-	-
ㅓ	.47297	.24324	-	-	.06757	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.21622	-	-	-	-	-	-
ㅔ	.80458	.00229	-	-	.00916	-	-	-	.00611	-	-	-	-	-	-	-	-	.01527	.00763	-	.14809	-	.00687	-	-	-	-	-	-
ㅕ	.43226	.18710	-	-	.03226	-	-	-	.05806	-	-	-	-	-	-	-	-	-	-	-	-	-	.17419	-	.11613	-	-	-	-
ㅖ	.33340	.04278	-	-	.13154	-	.00155	.00611	.19207	.00922	.00019	-	-	-	-	.00019	.00049	.16093	.01242	-	.05646	-	.03841	.00388	.00019	.00010	.00825	.00155	.00029
ㅗ	.77974	-	-	-	.17621	-	-	-	-	-	-	-	-	-	-	-	-	.01322	-	-	.00881	-	.02203	-	-	-	-	-	-
ㅘ	.72354	.01845	-	-	.03511	-	-	.00056	.09286	.00037	.00009	-	-	-	-	-	.00023	.04824	.02327	-	.02906	.00064	.01001	.00248	.00252	-	.00383	.00877	.00023

(f) 조건부 확률 p(z|y)

중성 y, 종성 z에 따른 데이타 양이 방대하여 여기에 싣지 않는다.

2. 초성, 중성, 종성의 엔트로피와 결합 엔트로피

초성 X의 엔트로피 H(X)가 식(1)과 식(15)으로 부터 구해지며 중성 Y의 엔트로피 H(Y)와 종성 Z의 엔트로피 H(Z)도 같은 방법으로 구해진다. H(X), H(Y)와 H(Z)의 계산값을 표 3 에 보인다. 표 3 에서 초성의 엔트로피가 가장 크고, 종성의 엔트로피가 가장 작음을 알 수 있다.

표 3. 초성, 중성, 종성의 엔트로피
Table 3. Entopy for a choseong, a jungseong, and a jongseong.

엔 트 로 피	bits
H (X)	3.897
H (Y)	3.118
H (Z)	2.518

초성-중성 (X, Y)의 결합 엔트로피 H(X, Y)가 식(2)와 식(16)로부터 구해지며 초성-종성(X, Z)의 결합 엔트로피 H(X, Z)와 중성, 종성 (Y, Z)의 결합 엔트로피 H(Y, Z)도 같은 방법으로 구해진다. H(X, Y), H(X, Z)와 H(Y, Z)의 계산값을 표 4 에 보인다. 표 4 에서 초성-중성의 엔트로피가 가장 크고, 중성-종성의 엔트로피가 가장 작음을 알 수 있다.

표 4. 초성, 중성, 종성의 결합 엔트로피 H(X, Y)型
Table 4. Joint entopy for a choseong, a jungseong, and a jongseong. H(X, Y) type.

엔 트 로 피	bits
H (X, Y)	6.715
H (X, Z)	6.229
H (Y, Z)	5.459

음절(X, Y, Z)의 엔트로피 H(X, Y, Z)가 식(3)에 의해 구해지며 그 값은 8.679이다.

3. 초성, 중성, 종성의 조건부 엔트로피

중성 Y가 주어질 때 초성 X의 조건부 엔트로피 H(X|Y)가 식(4)와 (17)로부터 구해진다. 종성 Z가 주어질 때 초성 X의 조건부 엔트로피 H(X|Z), 초성 X가 주어질 때 중성 Y의 조건부 엔트로피 H(Y|X), 종성 Z가 주어질 때 중성 Y의 조건부 엔트로피 H(Y|Z), 초성 X가 주어질 때 종성 Z의 조건부 엔트로피 H(Z|X)와 중성 Y가 주어질 때 종성 Z의 조건부 엔트로피 H(Z|Y)도 같은 방법으로 구해진다. H(X|Y), H(X|Z), H(Y|X), H(Y|Z), H(Z|X)와 H(Z|Y)의 계산값을 표 5에 보인다. 표 5에서 종성이 주어질 때 초성의 조건부 엔트로피가 가장 크고 중성이 주어질 때 종성의 조건부 엔트로피가 가장 작음을 알 수 있다.

표 5. 초성, 중성, 종성의 조건부 엔트로피 H(X|Y)型
Table 5. Conditional entopy for a choseong, a jungseong, and a jongseong H(X|Y) type.

엔 트 로 피	bits
H (X|Y)	3.598
H (X|Z)	3.712
H (Y|X)	2.836
H (Y|Z)	2.947
H (Z|X)	2.350
H (Z|Y)	2.341

종성 Z가 주어질 때 초성-중성(X, Y)의 조건부 엔트로피 H(X, Y|Z)가 식(5)와 (18)로부터 구해진다. 중성 Y가 주어질 때 초성-종성 (X, Z)의 조건부 엔트로피 H(X, Z|Y)와 초성 X가 주어질 때 중성-종성 (Y, Z)의 조건부 엔트로피 H(Y, Z|X)도 같은 방법으로 구해진다. H(X, Y|Z), H(X, Z|Y)와 H(Y, Z|X)의 계산값을 표 6에 보인다. 표 6에서 종성이 주어질 때 초성-중성의 조건부 엔트로피가 가장 크고 초성이 주어질 때 중성-종성의 조건부 엔트로피가 가장 작음을 알 수 있다.

표 6. 초성, 중성, 종성의 조건부 엔트로피 H(X, Y|Z)型
Tabele 6. Conditional entopy for a choseong, a jungseong, and a jongseong H(X, Y|Z) type.

엔 트 로 피	bits
H (X, Y|Z)	6.161
H (X, Z|Y)	5.561
H (Y, Z|X)	4.799

중성-종성 (Y, Z)가 주어질 때 초성 X의 조건부 엔트로피 H(X|Y, Z)가 식(6)과 (19)로부터 구해지며 초성-종성(X, Z)가 주어질 때 중성 Y의 조건부 엔트로피 H(Y|X, Z)와 초성-중성 (X, Y)가 주어질 때 종성 Z의 조건부 엔트로피 H(Z|X, Y)도 같은 방법으로 구해진다. H(X|Y, Z), H(Y|X, Z)와 H(Z|X, Y)의 계산값을 표 7 에 보인다. 표 7 에서 중성과 종성이 주어질 때 초성의 엔트로피가 가장 크고, 초성과 중성이 주어질 때 종성의 조건부 엔트로피가 가장 작음을 알 수 있다.

표 7. 초성, 중성, 종성의 조건부 엔트로피 H(X|Y, Z) 型.

Table 7. Conditional entopy for a choseong, a jungseong, and a jongseong H(X|Y, Z) type.

엔 트 로 피	bits
H (X|Y, Z)	3.220
H (Y|X, Z)	2.449
H (Z|X, Y)	1.967

4. 초성, 중성, 종성간의 평균상호정보량

초성 X와 중성 Y간의 평균상호정보량 I(X;Y) 이 식(10)으로부터 구해지며 초성 X와 종성 Z간의 평균상호정보량 I(X;Z)와 중성 Y와 종성 Z간의 평균상호정보량 I(Y;Z)도 같은 방법으로 구해진다. I(X;Y) I(X;Z)와 I(Y;Z)의 계산값을 표 8 에 보인다. 표 8에서 초성과 중성간의 평균상호정보량이 가장 크고, 초성과 종성 간의 평균상호정보량이 가장 작음을 알 수 있다.

이것은 한 음절 내의 초성이 무슨 자소인가를 아는 것이 같은 음절 내의 중성이 무슨 자소인가에 대하여 비교적 큰 정보를 제공하는 반면, 초성이 무슨 자소인가를 아는 것은 같은 음절 내의 종성이 무슨 자소인가에 대하여 비교적 작은 정보를 제공함을 뜻한다.

표 8. 초성, 중성, 종성간의 평균상호정보량 I(X;Y) 型.

Table 8. Average mutual information for a choseong, a jungseong, and a jongseong I(X;Y) type.

평균상호정보량	bits
I (X;Y)	0.282
I (X;Z)	0.168
I (Y;Z)	0.177

초성 X와 중성-종성 (Y, Z)간의 평균상호정보량 I(X;Y, Z)가 식(11)로부터 구해지며 중성 Y와 초성-종성 (X, Z)간의 상호평균정보량 I(Y;X, Z)와 종성 Z와 초성-중성 (X, Y)간의 평균상호정보량 I(Z;Y, Y)도 같은 방법으로 구해진다. I(X;Y, Z), I(Y;X, Z)와 I(Z;X, Y)의 계산값을 표 9 에 보인다. 표 9 에서 종성과 초성-종성 간의 평균상호정보량이 가장 크고, 종성과 초성-중성간의 평균상호정보량이 가장 작음을 알 수 있다.

이것은 한 음절 내의 중성이 무슨 자소인가를 아는 것이 같은 음절 내의 초성-종성이 무슨 자소들인가에 대하여 비교적 큰 정보를 제공하는 반면, 종성이 무슨 자소인가를 아는 것은 같은 음절 내의 초성-중성이 무슨 자소들인가에 대하여 비교적 작은 정보를 제공함을 뜻한다.

표 9. 초성, 중성, 종성간의 평균상호정보량 I(X;Y, Z) 型.

Table 9. Average mutual information for a choseong, a jungseong, and a jongseong I(X;Y, Z) type.

평균상호정보량	bits
I (X;Y, Z)	0.659
I (Y;X, Z)	0.668
I (Z;X, Y)	0.554

Ⅳ. 결 론

한글의 음절을 초성, 중성, 종성 단위로 나누고 각각 확률변수로 간주하고 음절의 발생확률을 음절별 발생 빈도수 분포로부터 계산하였다. 음절의 발생확률로 부터 초성, 중성, 종성의 발생확률, 결합확률과 조건부 확률을 계산하였다.

초성, 중성, 종성의 발생확률, 결합확률과 조건부 확률로부터 초성, 중성, 종성의 엔트로피를 계산하고 초성, 중성, 종성의 결합에 대한 결합 엔트로피를 계산하고 초성, 중성, 종성 간의 조건부 엔트로피를 계산하였다. 각 유형의 조건부 엔트로피에 있어서 종성이 주어질 때 초성의 조건부 엔트로피, 종성이 주어질 때 초성-중성의 조건부 엔트로피, 중성

-종성이 주어질 때 초성의 엔트로피가 각각 가장 크다.

또한 초성, 중성, 종성간의 평균상호정보량을 계산하였다. 두 확률변수 간의 평균상호정보량은 초성과 중성 간이 가장 크고 초성과 종성 간이 가장 작으며, 한 확률변수와 다른 두 확률변수간의 평균상호정보량은 중성과 초성-종성 간이 가장 크고 종성과 초성-중성 간이 가장 작다.

이 논문에 실린 초성, 중성, 종성의 발생확률, 조건부 확률, 엔트로피, 평균상호정보량의 계산값이 한국어 음성인식 및 합성, 자연언어처리, cryptography, 언어학, 음성학 등에 이용될 것을 기대한다.

參 考 文 献

[1] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379-413, 623-656, July and Oct. 1948.

[2] C. Shannon, "Prediction and entropy of printed English," *Bell Syst. Tech. J.*, vol. 30, pp. 50-64, Jan. 1951.

[3] R. Manfrino, "Printed Portuguese entropy - statistical calculation," *IEEE Trans. Inform. Theory*, vol. IT-16, p. 122, Jan. 1970.

[4] 이주근, 최홍문, "한국어 음절의 entropy에 관한 연구," 전자공학회지, 제 11 권, 제 3 호, pp. 15-21, 1974년 6 월.

[5] 안수길, 안지환, "공백소를 포함한 한글 자소 발생 확률과 엔트로피", 전자공학회지, 제 17권, 제 2 호, pp. 23~28, 1980년 4 월.

[6] 남궁건, 한글 낱말의 발생빈도 분포의 엔트로피에 관한 연구. 석사학위 논문, 서울대학교, 1979.

[7] 이재홍, 오상현, "한글의 초성, 중성, 종성 단위의 조건적 발생확률과 엔트로피," 한국통신학회 추계학술발표회 논문집, pp. 53~56, 1987년 11월.

[8] 문교부, 우리말 말수 사용의 찾기 조사. 문교부, 1956.

[9] 유재원, 우리말 역순사전. 정음사, 1985.

[10] J.B. Carroll, P. Davies, and B. Richman, Word Frequency Book, Houghton Mifflin, 1971.

[11] 이용주, 김경태, 조철우, 이성구, "대용량 발음사전 표제어에 나타난 음소의 통계적 성질," 대한전자공학회 교환 및 통신연구회 합동학술발표회 논문집, pp. 117~121, 1987년 11월.

[12] R. McEliece, The Theory of Information and Coding, Addison-Wesley, 1977.

[13] R. Gallager, Information Theory and Reliable Communication, Wiley, 1968.

著 者 紹 介

李 在 弘 (正會員)　第25卷 第 3 號 參照.
현재 서울대학교 전자공학과 조교수.

吳 相 縣 (準會員)　第25卷 第 3 號 參照.
서울대학교 전자공학과 석사학위 취득.
현재 (주)금성사 근무