

차원 축소 알고리즘을 이용한 한국어 자음간 거리 매핑 연구

우송대학교언어치료·청각재활학부¹ · 서울대학교의과대학 이비인후과학교실, 서울대학교의과대학 감각기관연구소²
포항공대 컴퓨터공학과, 정보전자융합공학부, 창의IT융합공학과³ · 일리노이대 컴퓨터공학과⁴ · 포항공대 컴퓨터공학과⁵
장 선 아¹ · 오 승 하² · 최 승 진³ · 김 민 제⁴ · 박 선 호⁵

ABSTRACT

A Study of mapping distance among Korean consonants using dimension reduction algorithm

Son-A Chang¹, Seung-Ha Oh², Seung-Jin Choi³, Min-Jae Kim⁴ and Sun-Ho Park⁵

¹School of Speech Therapy and Aural Rehabilitation, Woosong University, Daejeon, Korea

²Department of Otorhinolaryngology, Seoul National University College of Medicine, Sensory Organ Research Institute, Seoul National University Medical Research Center, Seoul, Korea

³Department of Computer Science and Engineering, Division of IT Convergence Engineering (WCU Program), Department of Creative Excellence Engineering (by courtesy), Pohang University of Science and Technology, Pohang, Korea

⁴Department of Computer Science, University of Illinois at Urbana-Champaign, Illinois, USA

⁵Department of Computer Science and Engineering, Pohang University of Science and Technology, Pohang, Korea

There have been increasing needs to understand speech signals to reveal auditory processing mechanism. We investigated an optimal dimension reduction algorithm and features for reflecting distance between Korean consonants. First, we reduced features of consonant articulatory phonetics from 15 to 5-6 using three dimension reduction algorithms. This revealed 90% of the distance distribution of Korean consonants. Next, we applied the Isomap method to show the most similar acoustic features of speech banana. Results showed a useful mapping of consonant articulatory distances. For example, we found that consonant /ㅂ(p)/ and /ㅌ(t)/ share 13 out of 15 features of articulatory phonetics, with only the features of bilabial and coronal being different, suggesting that they are the most difficult to discriminate from each other and reflected in the shortest distance between these two consonants in the Isomap method. However, the Isomap method was not effective in distinguishing the acoustic features of fortis from aspirated. This study shows that the use of reduction algorithms and the Isomap method is effective in consonant mapping of articulatory features. These findings suggest that they are useful tools in phonetics research when combined with acoustic phonetics data and real voice data.

KEY WORDS : Sound processing, Dimension reduction algorithm, Speech features, Articulatory phonetics, Acoustic phonetics, Consonant distribution

INTRODUCTION

인공 뇌 구현의 일환으로 음성 처리에 관한 두뇌의 확률적 모델 연구의 필요성이 제기되어 왔다. 이에 따라 기존

음성 처리 방법의 기술적 한계를 극복하고 새로운 음성 처리 알고리즘을 개발하는 데에 뇌의 청각 영역에 기반을 둔 음성 처리 연구의 필요성이 대두되게 되었다. 이러한 배경 하에 본 연구에서는 인간 청각 영역의 모델링에 기초한 음원 분리 및 복원 기술 개발을 위한 기초 연구로 우리말 자음이 어떻게 분포되어 있는지 살펴보고자 하였다. 즉 우리말 자음의 변별 자질들을 기계학습의 방법론인 차원 축소 알고리즘들에 적용하여 우리말 자음 음소들이 어떠한 분포 형태와 자음간 거리를 가지고 있는지 분석하였다.

인간 뇌신경 정보처리 과정은 감각수용 차원에서 신경

논문접수일: 2012년 05월 20일

논문수정일: 2012년 06월 14일

게재확정일: 2012년 06월 15일

교신저자: 오승하, 110-744 서울 종로구 연건동 28, 서울대학교 의과대학 이비인후과학교실

전화: (02) 2072-2442, 전송: (02) 2072-2537

E-mail : shaoh@snu.ac.kr

학적 병목 현상이라는 정보의 선별 및 차단 메커니즘을 가지고 있으며, 청각적인 투입 정보 역시 청각 경로 상에서 신경학적 병목을 통과하게 된다(Clark, 2003). 이를 통해 감각에서 수용되는 청각 정보가 그대로 뇌로 전달되는 것이 아니라 신경학적인 전달 과정에서 분리, 선별 또는 통합된다는 것을 알 수 있다. 인간 청각계의 정보 통합과 선별 과정을 기계적으로 구현하는 방법의 하나로, 본 연구에서는 차원 축소 알고리즘(dimension reduction algorithm)을 적용하여 보고자 한다. 차원 축소 알고리즘이란 데이터의 각 특징에 하나의 차원을 할당한 후 특정 차원 축소 알고리즘에 적용하여 차원들의 개수를 줄이게 된다. 이렇게 차원 축소 알고리즘은 관찰된 많은 수의 차원을 대표하는 최소의 자질을 추출하기 위해 사용된다(Cox & Cox, 2001).

차원 축소는 크게 선형적인 방식과 비선형적인 방식으로 나눌 수 있다. 선형적인 방식에 비해 비선형적인 방식이 가지는 특징은, 비선형적으로 분포한 데이터의 매니폴드(manifold)를 고려하여 차원 축소를 할 수 있다는 점이다. 이는 특히 고차원 데이터와 같이 샘플이 복잡하게 분포하는 경우에 데이터의 분포를 잘 반영하는 차원 축소 결과를 얻어낼 수 있고, 이러한 알고리즘을 통칭해서 매니폴드 학습(manifold learning)이라고 한다(Cox & Cox, 2001; Tenenbaum et al., 2000).

Roweis & Saul(2000)은 이와 같은 개념의 이해를 돕기 위해 비선형 차원 축소 문제에서 많이 쓰이는 Swiss roll data를 통하여 도식적으로 설명하였다. 다시 말하면, 3차원 상의 Swiss roll에서 생성된 샘플들의 분포를 2차원 상의 매니폴드로 투사(projection)하여 Swiss roll 데이터의 매니폴드 구조(manifold structure)를 유지한 채 2차원으로 차원 축소가 수학적으로 가능함을 보여주었다. 즉 매니폴드 학습이란 주어진 데이터의 매니폴드 구조를 파악하고 이를 최대한 유지하면서 보다 낮은 차원으로 매핑(mapping)하는 일련의 알고리즘을 통칭한다. 이러한 차원 축소 알고리즘은 말소리와 같이 다양한 요소로 이루어져 있으며 다면적인 현상의 핵심적인 자질이 무엇인지 이해하고 말소리 음소간의 관계구조를 파악하기 위해 적용될 수 있을 것이다.

본 연구에서는 선형 차원 축소 알고리즘으로 주성분분석(Principal Component Analysis, PCA)을 사용하였고, 비선형 매니폴드 학습 알고리즘으로 Isomap, Laplacian Eigenmap를 사용하였다(Tenenbaum et al., 2000; Belkin & Niyogi, 2003). PCA 방식은 고차원의 공간에서 측정된 변수들을 선형적으로 축소하여 가장 잘 반영하는

가장 낮은 차원의 데이터 샘플들을 추출하는 방식을 말한다. PCA 방식은 입체 공간에 복잡하게 존재하는 데이터들 간의 거리를 선형적으로 축소하여 실제 데이터 포인트 간의 거리를 왜곡할 수 있다. 따라서 Roweis & Saul(2000)의 연구처럼 입체 공간 상에 분포하는 데이터 간의 실제 거리를 잘 반영하기 위해서는 지역적인 구조(local structure)를 반영하는 분석 방법이 필요하다. 이러한 분석 방법들의 종류로 Isomap, LLE (locally linear embedding), Eigenmaps 등이 있다. 이러한 알고리즘들 중 Isomap과 Laplacian Eigenmaps은 말소리와 같이 고차원의 데이터를 중요한 몇 가지의 차원으로 축소하여 매핑할 지라도 지역적인 거리를 반영할 수 있는 알고리즘으로 실제 데이터 포인트 간의 거리를 더 잘 반영한다고 할 수 있다(Tenenbaum et al., 2000). 아직까지 말소리 데이터를 적용하여 선형 차원 축소 방식과 비선형 차원 축소 방식이 얼마나 다른 거리로 말소리를 분포시키는지 비교한 연구는 많지 않다. 게다가 한국어 말소리에 이러한 분석을 적용하는 것은, 음성 합성과 분리 연구나 인공와우 어음처리 연구 등에 유용한 정보를 제공할 수 있을 것이다.

Miller & Nicely(1955)는 영어 자음 음소의 지각 실험을 통해 유성성(voicing), 비성(nasality), 파찰성(affrication), 지속성(duration), 조음 위치(place of articulation)의 다섯 가지 자질은 지각에 독립적으로 영향을 미치는 자질들이라고 설명하고 있다. Cole & Fanty(1990)는 ISOLET Spoken Letters Data를 이용하여 EAR (English Alphabet Recognizer) 모듈을 바탕으로 알파벳 음소들의 조음적 자질 측정을 첨가한 후 Name Retrieval Modules를 이용한 기계 학습(computational learning) 모델로 구두 철자 인식(spoken letter recognition) 실험을 하였다. 이들은 B와 V같이 변별이 어려운 음소들을 변별하기 위한 구체적인 조음적(articulatory) 자질들을 첨가했을 때 (예; 유성 자질, 전-유성 자질), 주파수 분포(spectral) 자질만으로 수행했을 때보다 10% 향상된 결과를 보였다. 또한 구어를 분명하게 분절(segmentation)했을 때 음소 지각에 중요한 자질을 제공하였다. 이렇게 조음 음성학적 자질은 말소리의 산출과 지각에 대한 이해에 중요한 단서를 제공할 수 있다.

조음 음성학(articulatory phonetics)적 자질이란, 음소 또는 음운이 가지고 있는 다른 음소 또는 음운과 구별되는 고유한 특징이나 속성을 말한다(신지영 & 차재은, 2003). 조음 음성학은 공기가 성도(vocal tract)를 통과하는 동안 조음 기관이 어떻게 상호 작용하여 말소리를 만들어 내는가를 설명한다. 조음 음성학은 말소리의 음운적

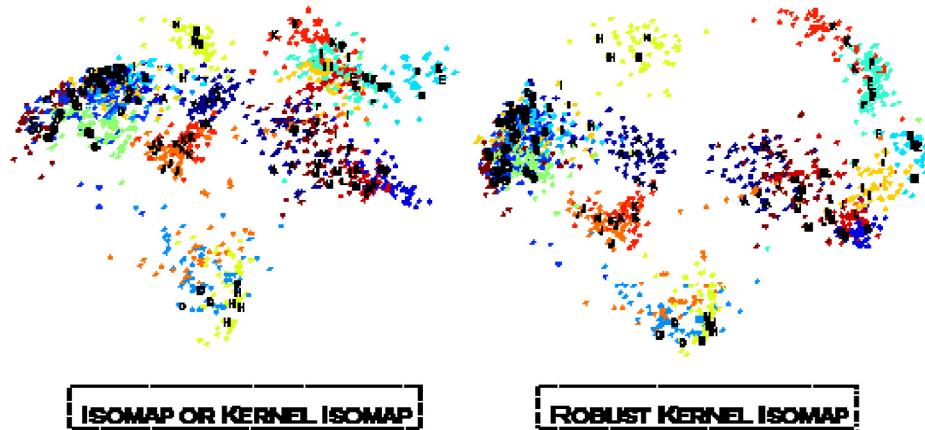


Figure 1. Kernel Isomap과 Robust Kernel Isomap을 이용한 영어구두철자의 매니폴드 학습 결과: 매니폴드 학습 방식을 사용하여 영어알파벳 간의 거리 매핑을 시도하였을 때 차원 축소 알고리즘의 종류에 따라 약간씩 다른 결과가 나오는 것을 보여주는 데이터

행동을 설명하고 예측하고자 하며, 말소리의 특징을 분석하고 이해하는 데에 많은 정보를 제공한다.

말소리의 변별 자질은 Jacobson 이후 Chomsky & Halle가 제시한 SPE (Sound Pattern of English) 체계를 기본으로 사용되어 왔으며, 이 SPE 체계의 변별 자질은 우리말 소리 자질을 설명하는 데에 잉여적인 자질들을 포함하고 있다. 또한 한국어에서 자음을 구분하는 변별 자질의 집합과 관계에 대해서는 국내 학자들 사이에 완전한 합의가 이루어져 있지 않다(배문정 & 김정오, 2002). 그럼에도 불구하고 이러한 자질들이 음소들의 지각에 영향을 미칠 수 있다면 어떠한 자질들이 우리말 자음들의 특징을 가장 잘 반영하고 있는지를 알아보는 것은 의미 있는 일일 것이다.

조음 음성학적 자질들 외에 음향 음성학(acoustic phonetics)적 자질들 또한 음소의 분석에 중요한 틀을 제공하여 준다. 음향 음성학은 말소리의 음향적 특성, 즉 물리적인 속성에 따른 자질들을 분석 대상으로 한다. 화자의 조음 기관의 위치와 모양의 변화에 따라 서로 다른 음파가 생성되고 공기를 타고 전달되는 음파의 압력 변화가 청자에게 소리로 지각된다. 이때 전달되는 소리는 시간-강도-주파수 간의 특징적인 상호작용의 모습으로 현상화될 수 있다. 이러한 특징을 시각화하여 말소리의 분석에 사용되는 것이 음향 음성학적 자질들이다. 음향 음성학적 자질들은 조음 음성학적인 방법만으로는 현상화되기 어려운 말소리의 특징들을 나타내 줄 수 있을 것이다.

또한 말소리의 자질들뿐만 아니라 분석의 도구도 음소들의 특징을 현상화하는 데에 영향을 미칠 수 있다. 본 연구의 예비 연구에서 위의 연구와 같은 ISOLET database의 영어 구두 철자를 Isomap 방식의 일종인 Kernel Isomap을 이용하여 거리 매핑(distance mapping)을 하였다(Choi & Choi, 2005; Choi & Choi, 2007). Fig. 1을 보면 알고리즘에 따라 음소간 거리가 조금씩 달라지고 있는 것을 관

찰할 수 있다.

따라서, 서로 다른 차원 축소 알고리즘에 우리말 분석에 사용되어 오고 있는 조음 음성학적 자질들과 음향 음성학적 자질들을 적용해 보는 것은 우리말 자음에 대한 지각적 이해에 도움이 될 것이다. 본 연구는 우리말 자음들이 3가지 차원 축소 알고리즘을 통해 어느 정도의 거리를 가지고 분포하는지, 어떤 차원 축소 알고리즘이 우리말 자음의 분포를 가장 잘 반영하는지, 그리고 우리말 자음의 분포가 조음 음성학적 자질과 음향 음성학적 자질에 따라 어떻게 분포되는지 알아보려고 하였다.

이를 위하여 두 가지 실험을 하였으며, 실험 1에서는 조음음성학적 자질만을 이용하여 각 변별 자질들을 하나의 차원으로 놓고 3가지 차원 축소 알고리즘을 사용하여 매핑하였을 때 각 음소들 간의 거리가 어떻게 분포되는지 알아보고 영어 speech banana에 우리말 자음을 대응한 결과와 거리 분포가 어떠한 차이가 있는지 알아보았다.

실험 2에서는 speech banana와 같은 지각 데이터와 가장 유사한 형태를 보인 Isomap 차원 축소 방식에 음향 음성학적 자질만을 적용시켰을 때, 조음음성학적 자질을 이용했을 때의 결과와 어떠한 차이가 있는지 비교하여 보았다.

MATERIALS AND METHODS

1. Experiment 1

Procedures

본 연구에서는 우리 말소리 중 자음의 조음 음성학적 자질들을 분류했으며, 김정오(2002), 차재은(2003), 이호영(1996), 엄정희(1994)가 분류한 체계를 참조하여 재 분류하였다(Table 1). 19개의 자음 음소들에 대해 총 15개의 변별 자질들이 분류, 적용되었다. 본 연구에서는 변별 자질

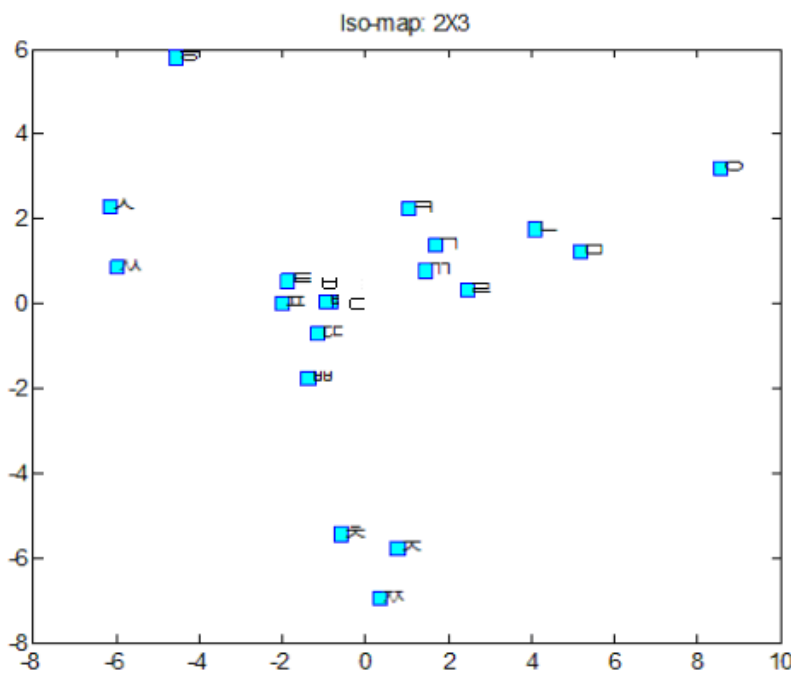
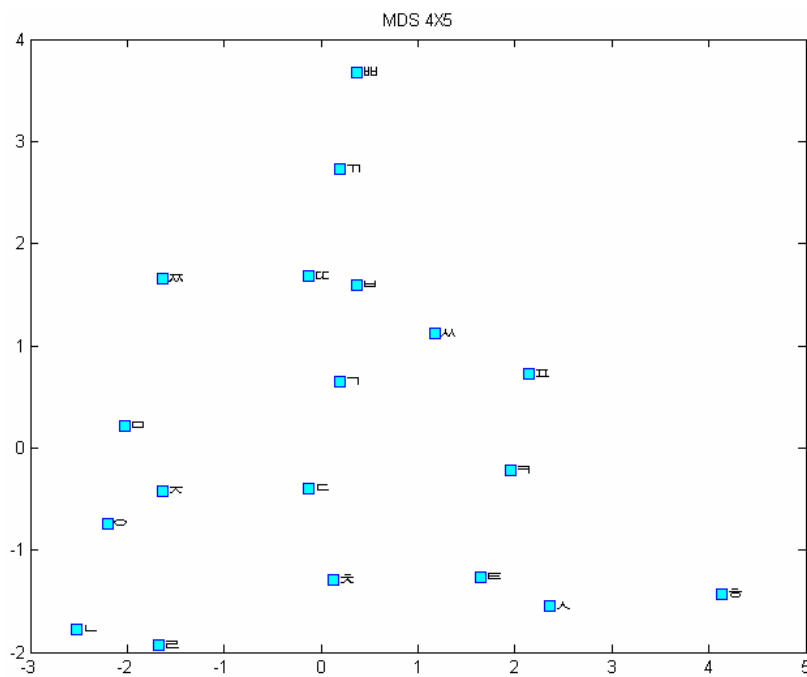


Table 2. 세가지 매핑 방식에 따른 우리말 자음 음소간 거리와 자질 분석표

Mapping 방식		음소 자질간 거리		
		최단 거리	최장 거리	
PCA	ㅂ	양순성, 전방성	ㄴ	공명성, 비성, 전방성, 설정성
	ㅈ	긴장성, 전방성, 설정성	ㅎ	협착기식, 기식성, 지속성, 저설성, 후설성
Isomap	ㅂ	양순성, 전방성	ㅅ	기식성, 지속성, 전방성, 설정성
	ㅈ	전방성, 설정성	ㅇ	공명성, 비성, 고설성, 후설성
Laplacian	ㅇ	공명성, 비성, 고설성, 후설성	ㅍ	기식성, 양순성, 전방성
Eigenmap	ㅎ	협착기식, 기식성, 지속성, 저설성, 후설성	ㄴ	공명성, 비성, 전방성, 설정성

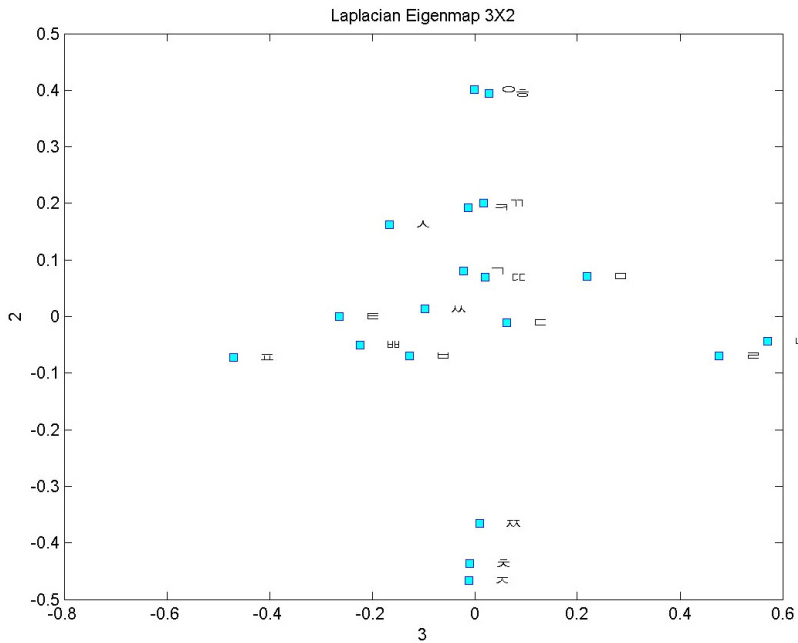


Figure 5. Laplacian Eigenmap 방식을 이용한 우리말 자음 음소 매핑: 조음 음성학적 자질 분석

Results

PCA 방식, Isomap, Laplacian Eigenmap 방식을 이용하여 우리말 자음 음소들의 변별 자질들을 차원 축소한 결과, 총 15개의 변별 자질들이 PCA의 경우 6개의 축으로, Isomap과 Laplacian Eigenmap의 경우 5개의 축으로 차원 축소되었다. 즉, 5-6개의 축으로 차원 축소하였을 때 기존의 데이터가 가지는 정보를 90% 이상 설명할 수 있었다. 이렇게 축소된 차원으로 기존의 데이터를 공간에 배열하는 PCA, Isomap, Laplacian Eigenmap 을 이용하여 우리말 자음 음소를 공간에 배열한 결과는 다음과 같다 (Fig. 3-5). 이때 4 X 5는 4번째와 5번째 vector space 라는 뜻이고 4가 가로축, 5가 세로축을 뜻한다.

위의 그림들은 음소들은 보유하고 있는 자질들의 차원을 축소하여 얻어진 음소들간의 거리를 나타낸다. 세 개의 그림은 차원축소 알고리즘에 따라 조금씩 다른 음소들간 거리를 나타내고 있었다. PCA 방식에서는 /ㄱ/과 /ㄷ/, ISO-MAP 방식에서는 /ㄱ/과 /ㄷ/, Laplacian Eigenmap에서는 /ㅇ/과 /ㅎ/이 가장 가까운 거리를 보였다. 또한 가장 먼 거리는 PCA 방식에서 /ㄴ/과 /ㅎ/, Isomap 방식에서는 /ㄴ/과 /ㅇ/, Laplacian Eigenmap에서는 /ㅍ/과 /ㄴ/이 가장 먼 거리를 보였다.

가장 단거리를 보인 음소들과 가장 장거리를 보인 음소들이 가진 자질들을 정리한 것이 Table 2이다. 이 표에 의하면 PCA와 Isomap이 유사한 결과를 보이고 Laplacian Eigenmap이 다소 다른 결과를 보인 것을 알 수 있다. PCA와 Isomap에서 음소간 거리가 짧은 경우 전방성 자

질을 공유하고 있었으며 공명성과 기식 및 긴장성은 대조되는 자질 구성을 보이고 있었다. 이에 비해 Laplacian Eigenmap에서는 공명성/비성성이 기식성/지속성과 유사 자질인 것으로 나타나면서 기식성과 공명성이 가장 장거리 음소들이 공유하는 자질로도 나타나는 등 비일관된 자질 구성을 보여주었다.

다음으로 Fig. 2의 Speech banana에서 자음만 추출한 후 분포되어 있는 음소들을 전체적으로 중심을 축으로 180도 회전하여 보았다(Fig. 6). 회전의 이유는 차원 축소를 통해 얻은 음소의 지도와 비교를 가능하게 하기 위함이다.

이렇게 만들어진 영어 음소 자음 분포를 Isomap 지도의 축에 대입하면서 우리말 음소에 대응이 가능한 음소들을 표시하여 보았다 (Fig. 7). 우리말 음소에 대한 Speech banana는 아직 만들어지지 않았기 때문에, 자질 분석을 통해 영어의 음소와 가장 유사한 자질 구성을 보이는 우리말 음소들을 대입하여 분석하기로 하였다. 왼쪽부터 s는 [s]으로, k는 [k], g는 [g], h는 [h], ch(^h)는 [tʃ], p는 [p]으로 대입했으며, r은 [r], d는 [d], b는 [b], n은 [n], m은 [m], ng(ŋ)는 [ŋ], z는 [z]으로 대입하였다. 이 외에 우리말 자음에 대입할 수 없는 영어 음소는 영문만 표기하였다. Fig. 7의 가로와 세로축은 각각 주파수와 강도를 나타내며 주파수의 경우 화살표 방향으로 높아지고 강도의 경우 화살표 방향으로 커진다.

영어 Speech banana와 Isomap learning을 통한 자음 지도를 비교하여 보면 (Fig. 8), 각각의 지도에서 큰 원과

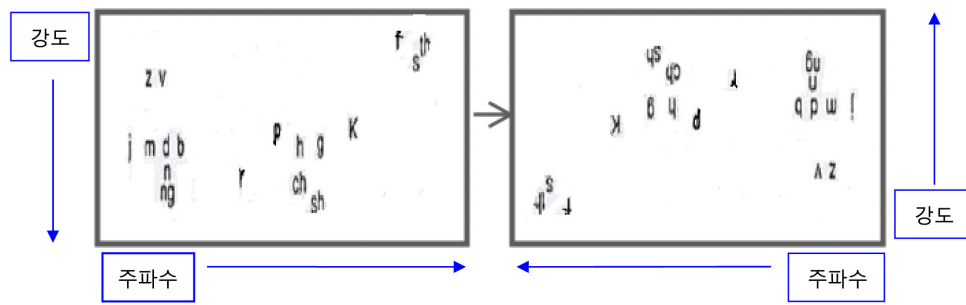


Figure 6. 영어 speech banana의 자음 분포

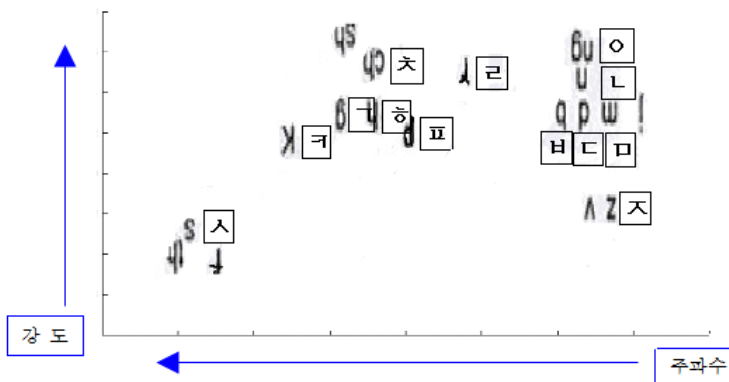


Figure 7. 영어 자음 음소의 주파수-강도 축 분포를 Isomap 차원축에 대응한 후 우리말 자음 음소를 대응한 결과

작은 원으로 표시된 것끼리 서로 비슷한 위치에 분포하고 있는 것을 볼 수 있다. 자세히 살펴보면, [ɹ]이 가장 왼쪽에 위치하고 중간에 [ɰ, ɱ, ɸ] (무성음)이 위치하며 오른쪽 면에 [ㄹ, ㄴ, ㄷ, ㄴ] (유성음)이 위치하는 것이 유사한 것을 볼 수 있다. 그러나 우리말 음소 [ㄱ, ㅋ, ㆁ, ㄷ, ㅌ, ㅈ]의 위치는 영어 Speech banana에 대입하였을 때 Isomap에서의 위치와 매우 다른 것을 볼 수 있다. 즉, 영어의 [b, d]가 비음/유음과 함께 위치하고 있으며, 왼쪽에 치우쳐 있는 ㄷ(ㄷ^h)에 대응되는 [ch]는 중간에 위치하고 있고, [ɹ]이 왼쪽에 위치한 우리말에 비해 [ㄱ]이 오른쪽에 위치하고 있는 것이 다른 점이었다.

또한 주파수와 강도 축을 Isomap 결과에 대입하여 본다면, 영어의 speech banana에서 [s]는 작은 강도, 높은 주파수에 위치하는데 Isomap의 우리말 [ㅅ]은 이에 비해 큰 강도의 위치에 분포하였으며, 영어의 유성 자음과 비음들 [b, d, m, n, ɱ]은 보다 큰 강도, 낮은 주파수에 위치하는데 Isomap의 비음들 [ㄹ, ㄴ, ㄴ]은 Speech banana의 위치에 비하면 보다 작은 강도의 위치에 분포하는 것을 볼 수 있었다.

Discussion

본 연구에서 살펴본 세가지 매니폴드 학습 방법론에 의해 새롭게 추출된 자질들은 단순히 기존의 15개 자질들

중 중요한 몇 가지일 수도 있고, 또는 15개 자질 중 중요한 몇 가지가 혼합되어 하나의 새로운 자질들을 나타내고 있을 수도 있다. 예를 들어, Isomap 결과에서 가장 가까운 거리를 보인 /ㄷ/과 /ㄷ/의 방향성은 실제로 원래의 15개 자질 중 양순성과 설정성 외에는 모두 같은 자질을 가짐으로써 이 두가지 자질은 자질 간 변별성이 가장 적다고 해석될 수 있을 것이다. 또한 가장 먼 거리를 보인 /ㅅ/과 /ㅇ/은 협착기식, 지속성, 전방성, 설정성을 갖는 /ㅅ/과 공명성, 비성, 고설성, 후설성을 갖는 /ㅇ/이 우리말 자음에서 가장 자질 공유가 적다. 이것이 매니폴드 학습 방법론으로 새롭게 추출된 자질들을 근거로 한 음소간 거리에 반영된다는 것은 기계 학습 방법이 음소들의 특성을 이해하고 수학적 방법으로 음소들의 특성을 계산하여 배열하는 것이 음소 변별에 도움을 줄 수 있다는 것을 의미한다. 이는 인공와우 어음처리 방식의 개발과 같은 음성 공학과 접목될 수 있는 가능성을 또한 뜻할 것이다.

15개의 원래 자질 중, 양순성과 설정성은 가장 가까운 거리를 갖는 음소들에서 주로 공유하는 자질들은 조음의 위치(place)와 관련된 것으로 우리말 변별에 큰 영향을 끼치지 못하는 것으로 볼 수 있다. 이는 조음 위치가 우리말 자음 음소에서 가장 변별하기 어려운 자질임을 나타내는 것이기도 하다. 반면에 공명성 대 기식성, 설정성 대 저설 및 후설성, 설정성 대 고설 및 후설성 등 조음의 방식

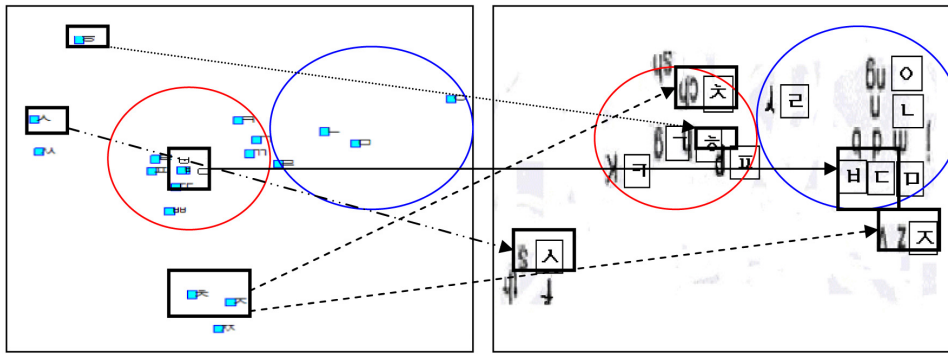


Figure 8. Isomap 지도와 speech banana 지도의 비교

Table 3. 우리말 자음들의 음향 음성학적 변별 자질

	ㄱ	ㄲ	ㅋ	ㄴ	ㄷ	ㄸ	ㅌ	ㄹ	ㅍ	ㅂ	ㅃ	ㅅ	ㅆ	ㅇ	ㅈ	ㅊ	ㅊ	ㅊ	
위치(place)	3	3	3	1	1	1	1	1	0	0	0	0	1	1	3	2	2	2	4
비성(nasality)	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
유성성(voicing)	1	0	0	2	1	0	0	2	2	1	0	0	0	0	2	1	0	0	0
협착기식(frication)	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1
지속성(duration)	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1
envelope	1	0	0	3	1	0	0	3	3	1	0	0	2	2	3	1	0	0	2

(manner)와 관련된 차이와 혀의 위치와 관련된 자질이 가장 중요한 변별 자질 특성이라고 볼 수 있을 것이다. 이는 또한 조음 방식이 우리말 자음의 변별을 돕는 자질이라고 할 수 있다. 특히 이 변별 자질들 중 공명 및 비성과 기식, 협착기식이 특히 많은 영향을 미치고 있었다. 그러나 본 분석의 변별 자질들은 개별적으로 거리 분포에 기여한 것이 아니라, 차원이 축소되며 새로운 자질 변수로 변환될 수 있기 때문에 이러한 해석에는 제한이 있을 것이다. 그럼에도, 분석의 근거가 된 자질들의 특성이 음소간 거리 분포에 영향을 미친 것이므로, 변별을 돕는 자질을 특성을 극대화하고 조음 위치와 같은 자질을 어떻게 음성적인 차이로 기술해 내는가 등에 대한 연구는 앞으로의 음성 처리 방식에 도움이 되는 정보를 줄 수 있을 것이다.

Speech banana에서 볼 수 있는 지각적인 음소 분포와 차원 축소 알고리즘을 통해 추출한 음소간 거리가 상당히 부분 유사한 것을 알 수 있었다. 저주파수에 위치하는 유음/비음 [ㄹ, ㅁ, ㄴ, ㅇ], 중주파수에 위치하는 파열음 [ㅋ, ㆁ, ㅍ], 고주파수에 위치하는 마찰음 [ㅅ]이 영어 Speech banana의 분포와 상당히 비슷했다. 이는 차원 축소 알고리즘에서 산출한 음소간 거리가 지각적인 차이를 반영할 수 있다는 것을 보여주는 것이다. 그러나 [ㅎ, ㅈ, ㅊ, ㅊ, ㅊ]

ㅂ]은 영어 음소에 대입했을 때 다른 분포를 보이고 있었다. 이는 우리말에서는 무성음인 [ㅈ, ㅊ, ㅊ]이 영어 음소 [z, d, b]와 유사하므로 대입을 하였으나, [z, d, b]는 유성음이므로 따라서 더 낮은 주파수에서 지각될 수 있음을 시사하는 것이라고 볼 수 있다. 또한 영어 음소의 지도를 우리말에 대입한 것으로는 영어에서는 변이음으로 존재하는 경우 [ㄱ, ㄴ, ㅅ, ㅆ, ㅆ]에 대한 비교는 제공하지 못하였다. [ㅎ]의 경우, 영어 음소와 우리말 음소가 본질적으로 같으나 조음 음성학적 자질 분류가 달랐기 때문에 지도에서 다르게 위치한 것으로 보인다.

또한 우리말 특유의 자질인 긴장성을 갖는 경우 [ㄱ, ㄴ, ㅅ, ㅆ, ㅆ]이 영어 Speech banana에서는 나타나고 있지 않고, 영어에만 있는 [th(θ), f, v, j(ɟ)]와 같은 음소들은 우리말 지도에서 관찰할 수 없었다 (본 연구에서는 [ㅈ]을 [z]로 보았으나 [j(ɟ)]와 더 유사하다고 볼 수도 있을 것이다).

몇몇 음소들의 분포와 절대적 위치는 다른 양상을 보였는데, 이는 주파수와 강도 축으로 산출된 Speech banana와 vector space에 분포한 Isomap지도는 산출 방식이 다른 데에 기인할 것이다. 그러나 음소간의 상대적인 거리와 분포에 두 가지 매핑 방법이 유사점을 보이고 있다는 것은

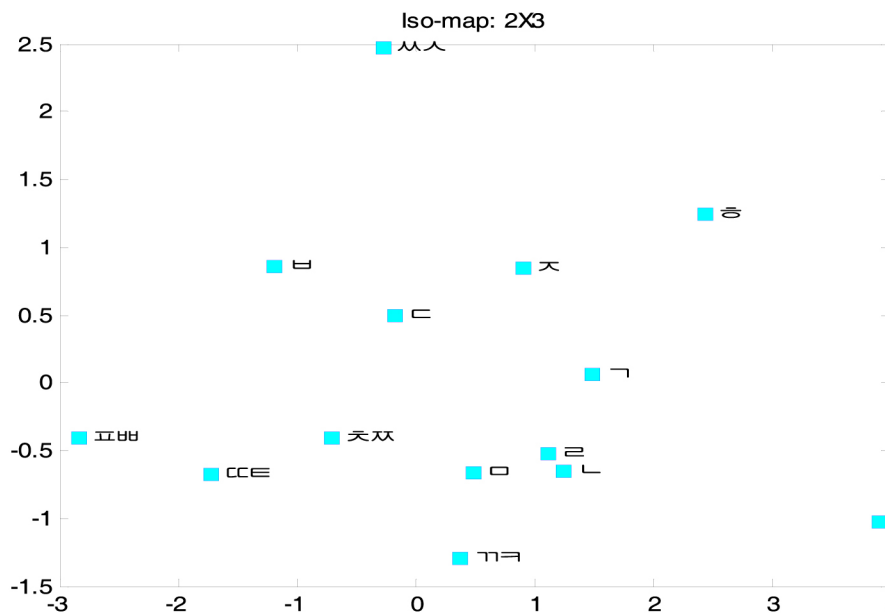


Figure 9. Isomap 방식을 이용한 우리말 자음 음소 매핑: 음향 음성학적 자질 분석

시사하는 바가 크다고 할 수 있다. 즉 지각적인 특성에 근거한 음소간의 차이가 조음음성학적인 자질로 분석한 차이와 비견할 수 있으며, 이는 주파수와 강도만이 아닌 조음음성학적 자질을 분석함으로써 음소의 성질에 대한 이해를 증진시키고 음소들간의 변별에 대한 새로운 관점을 제공한다는 의미가 있을 것이다.

2. Experiment 2

Procedures

조음 음성학적 자질 분류를 이용한 실험 1에서 세가지 차원 축소 방식 중 speech banana와 가장 유사한 형태의 자음 분포를 보이는 Isomap 방식을 우리말 자음의 음향 음성학적 자질에 적용시켜 보았다. 음향 음성학적 자질 분류는 Dorman(1993), Preece & Tyler(1992), Tyler & Moore(1992) 등의 기준을 참조하여 재분류하였다. 우리말 자음의 음향 음성학적 자질 분포는 Table 3과 같다.

Results

Isomap을 우리말 자음의 음성학적 자질 분류에 적용한 결과 총 6개의 자질들이 3개의 차원으로 축소되었으며 우리말 음소간의 거리를 가장 잘 나타낸 지도는 조음 음성학적 데이터와 같은 2x3이었다(Fig. 9).

그림에서 볼 수 있듯이, 음성학적 자질만으로 분류했을 때에는 경음과 격음(ㄱ/ㅋ, ㄷ/ㅌ, ㅌ/ㅍ, ㅈ/ㅊ)간의 구별과 마찰음(ㄴ/ㄹ)간의 구별이 없음을 볼 수 있다. 이는 영어 음소 분류에 주로 사용된 음향 음성학적 자질 분류만으로는 우리말 자음을 바르게 분석할 수 없으므로, 우리말 음

소에 적합한 음향 음성학적 변별 자질을 확립하는 것이 필요함을 나타낸다고 할 수 있다.

경음과 격음 이외의 평음들의 분포는 조음 음성학적인 분석 지도와 유사한 분포를 보이고 있었다. 그러나 음향 음성학적 분석에 의한 지도는 자음들간의 거리가 조음 음성학적인 분석에 비해 더 멀리 떨어져 있음을 관찰할 수 있었다.

Discussion

영어 음소를 근거로 한 음성학적 자질 분류는 우리말 음소 간의 거리를 모두 반영하지 못하였으며 음소들간의 상대적인 거리 또한 조음음성학적 자질로 분석했을 때나 Speech banana에 나타난 것과 상당한 차이가 있었다. 그러나 조음음성학적 자질(15개)에 비해 적은 숫자의 자질(6개)이 분석에 사용되었고 우리말 음소 변별에 중요한 VOT (voice onset time)나 모음 전이(vowel transition)과 같은 자질들이 분석에 포함되지 않은 것이 음소 분포가 다른 이유가 될 수 있을 것이다. Kim(2008)은 한국어 음소 지각에 파열성을 가진 음소들의 변별이 얼마나 중요한지 강조하며, 연음, 경음, 격음의 변별은 VOT뿐만 아니라 후행 모음의 기본 주파수와 음질(제 1 배음과 제 2 배음의 진폭 차)이 중요한 역할을 한다고 하였다. 본 분석에서 연음은 변별적 위치를 가지고 있는데 비해 경음과 격음이 분리되지 못한 것은 자음의 음성학적 자질만을 분석에 투입하였고 후행 모음과 관련된 자질은 없었던 것이 원인일 것이다.

음성학적 분석은 말소리의 음향적 특성을 기계적으로 측정하여 수치를 산출해 내는 것과 관련된다. 특히 경음과

격음의 변별에 영향을 주는 것은 후행 모음의 기본 주파수 이므로(Kim, 2008), 음향 특성의 수치를 차원 축소 알고리즘에 적용하여 조음음성학적 자질이나 음성학적 자질로 분석했을 때와 차이를 보는 연구가 후속되어야 할 것이다.

DISCUSSION AND CONCLUSIONS

본 연구의 결과는 우리말 자음 음소들이 가지는 변별적 자질이 음소마다 차이가 있으며 영어 음소와도 차이가 있음을 보여준다고 할 수 있다. Cole & Fanty(1990)는 음향적인 말소리 정보에 조음 자질적 정보가 부가되었을 때 알파벳 구어 철자의 재인 정확도가 증가되었다고 하였다. 이들의 결과와 같이, 우리말 자음 음소에서도 영어 음소에 서처럼 음향 음성학적인 변별 자질만으로는 음소 지도를 적절히 만들 수 없었다. 조음 음성학적인 자질 분류가 우리말 음소의 분포를 어느 정도 반영할 수 있었으며, 음향 음성학 자질만으로는 변별이 되지 않는 음소들간의 변별적 단서를 제공할 수 있었다. 이는 영어 음소 분류에 사용된 음향 음성학적인 자질들은 우리말 음소를 분석하는 데에 미흡하기 때문이라고도 할 수 있을 것이다.

특히 음향 음성학적 분석만으로는 우리말에 특징적인 음소들인 경음과 격음이 구분되지 못하였다. 경음/격음 쌍(예, ㄱ/ㅋ)은 영어 음소에는 없는 최소대립 쌍(minimally distinct pair)으로, 조음 음성학적 지도에서도 음소간 거리가 짧았으며 이는 음소들간의 거리가 변별의 어려움 정도를 나타내주는 증거가 될 수 있을 것이다. 또한 각 자음들 간의 거리가 조음 음성학적인 지도 분석에서 보였던 것보다 더 멀리 떨어져 있음으로써 우리말 자음 음소들간의 거리를 파악하는 데는 기존의 음향 음성학적인 자질들이 불충분함을 다시 확인할 수 있었다.

본 연구를 통해 Miller & Nicely(1955)가 대표적이고 독립적인 영어 자음 음소 변별 자질로 제시한 다섯 가지 자질 중에 유성성은 우리말 자음의 분석에 있어서 변별적 역할에 큰 도움을 주지 못하며 우리말 자음에 특징적인 경음/격음의 구분이 필수적임을 확인할 수 있었다. 또한 조음 위치와 관련된 자질들이 변별에 큰 영향을 못 미치며 조음 방법과 혀의 위치와 관련된 자질들이 큰 변별적 역할을 하고 있음을 알 수 있었다. 음향 음성학적인 자질들은 우리말 소리의 특징을 잘 반영하는 자질들을 추출하여 적용하는 것이 필요하였다. 영어 음소에 적용되는 말소리의 특징은 특히 우리말의 경음과 격음을 분리하지 못하였다. 따라서 음향 음성학적 지도를 산출할 때 자음 자체의 자질들만이 아닌 후행 모음에 대한 정보가 부가되는 것이 한국

어 자음 음소의 거리 분포를 파악하는 데에 도움이 될 것이다.

조음 음성학과 음향 음성학적인 우리말 소리에 대한 정보의 음성 공학적 처리에 도움을 줄 수 있을 것이며, 우리말 음소의 변별적 차이를 수학적 공식으로 추출하여 낼 수 있다면, 인공와우와 같은 음성 처리 알고리즘에 도움을 줄 수도 있을 것이다. 수많은 어음 정보를 효율적으로 줄이고 추출하여 재구성, 전달해야 하는 인공와우 장치에 차원 축소 알고리즘은 최적의 자질을 추출하는 데에 도움을 줄 수 있을 것이며, 조음 음성학적인 자질에 대한 분석을 추가하는 것이, 음향학적인 음성 정보의 분석만으로 이루어진 현재의 인공와우 알고리즘을 더 발전시킬 수 있는 계기가 될 수도 있을 것이다.

본 연구 결과, 우리말 음소의 분석에 중요한 음향 음성학적인 자질들을 추출해내어 차원 축소 알고리즘을 이용한 매니폴드 학습을 사용하면 사용한 차원, 변별 자질들이 우리말 음소의 분석에 적합한지 아닌지를 어느 정도 판단할 수 있음을 알 수 있었다. 우리말 소리 자질에 대한 정확한 지식을 갖게 되면 인공와우와 같은 어음 처리 방식을 사용한 기계의 기능을 향상시켜 인공와우의 사용자들로 하여금 보다 향상된 어음 지각 능력과 말-언어 발달을 가져올 수 있을지 모른다. 본 연구는 인공와우를 우리말에 맞게 개발해야 할 필요성을 제기함과 동시에, 우리말 음소의 특징 및 음원으로서의 자질에 대한 정보를 제공하였다. 앞으로 본 연구의 결과를 영어의 조음 음성학적인 자질 분석에 따른 차원 축소 알고리즘을 음성 자료로 시행한 Choi & Choi(2005, 2007)의 연구와 비교 분석하고 한국어 음성 자료를 구축하여 음성 자료를 이용한 자질 축소 알고리즘을 적용하여 보는 연구가 필요하겠다.

ACKNOWLEDGEMENT

본 연구는 산업자원부 핵심기술개발사업 Brain Neuroinformatics Program의 사업비(10023917-2007-4) 일부와 교육과학기술부 뇌과학원천기술개발사업 세부 과제 사업비 지원(20110030102)에 의해 수행되었습니다.

REFERENCES

- 배문정 & 김정오 (2002). 한국어 자음의 지각 구조. *한국심리학회지: 실험 및 인지*, 12, 375-408
- 신지영 & 차재은 (2003). 우리말 소리의 체계. 서울: 한국문화사.
- 엄정희 (1994). 정상 말소리의 발달(II) : 3, 4, 5세 아동. 아동의 조

- 음장에 치료. 한국언어병리학회 편. 서울: 군자출판사.
- 이호영 (1996). 국어 음성학. 서울: 태학사.
- Belkin, M. & Niyogi, P. (2003). Laplacian Eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15, 1373-1396.
- Choi, H. & Choi, S. (2005). Kernel Isomap on noisy manifold. in Proc. IEEE Int'l Conf. Development and Learning (ICDL), 208-213, Osaka, Japan, July 19-21.
- Choi, H. & Choi, S. (2007). Robust kernel Isomap. *Pattern Recognition*, 40(3), 853-862.
- Cole, R. & Fanty, M. (1990). "Spoken letter recognition", HLT '90. Proceedings of the workshop on Speech and Natural Language, 385-390.
- Clark, G. M. (2003). *Cochlear implants: fundamentals and applications*. New York, NY: Springer.
- Cox, T. F. & Cox, M. A. (2001). *Multidimensional Scaling* (2nd Ed). London: Chapman and Hall.
- Dorman, M. (1993). Speech perception by adults. In R. Tyler (1st Ed). *Cochlear Implants, Audiological Foundations* (pp.317-329). San Diego, CA: Singular Publishing Group.
- Kim, M. J. (2008). University of Washington, Doctoral dissertation, *Phonetic and Phonological Development of Word-initial Korean Obstruents in Young Korean Children*, Seattle.
- Miller, G. A. & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27(2), 338-352.
- Preece, J. & Tyler, R. (1992). Consonant confusions by users of three cochlear implant devices. *Seminars in Hearing*, 13(3), 226-238.
- Roweis, S. T. & Saul, L. K. (2000). Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science*, 290, 2323-2326.
- Tenenbaum, J. B., de Silva, V., & Langford, J. C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science*, 290, 2319-2323.
- Tyler, R. & Moore, B. (1992). Consonant recognition by some of the better cochlear implant patients. *Journal of the Acoustical Society of America*, 92, 3068-3077.
- A Net Presence & Pixel Kitchen (1996-2001). Where do we go from Hear? info@gohear.org