

한국어 음성 자료의 수집과 정리

이호영

서울대학교 언어학과 교수

1. 머리말

아마도 많은 사람들이 고등학교나 대학교에서 훈민정음으로 기록된 15세기 문헌을 배우면서 당시 우리말의 발음이 어떠했을까 궁금하게 생각했을 것이다. 다행스럽게도 훈민정음이 표음문자이기 때문에 우리는 15세기 우리말의 발음을 상당히 잘 파악하고 있다. 그러나 국어학자들은 모음의 음가, 겹자음과 겹모음의 음가, 성조, 억양 등 15세기 우리말의 발음을 좀더 정확하게 알기를 원한다. 뿐만 아니라 15세기 주요 지역 방언의 모습이 어떠했는지 알기를 원한다. 15세기에는 녹음기가 없었기 때문에 불가능했던 일이지만 지금 15세기 우리말을 녹음해 놓은 음성 자료가 발굴된다면 이 자료는 당장 국보로 지정될 것이다.

녹음기가 발명되고 나서부터 인간은 말소리를 영구히 기록할 수 있게 되었다. 지금 우리는 20세기 전반에 녹음된 음성 자료는 조금밖에 가지고 있지 않지만 1950년대 이후에 보존된 라디오와 텔레비전 방송 자료들은 많이 가지고 있다. 그리고 지난 1980년대 한국정신문화원에서 수집한 방언 조사 자료들과 1997년에 국립국어연구원에서 실시한 서울토박이말 연구에서 수집된

음성 자료들을 가지고 있다.¹⁾ 이 자료들을 잘 정리해 두면 후세의 국어학자들은 20세기 우리말의 발음에 관해 상당히 자세하게 알게 될 것이다.

그러나 지금 우리가 구할 수 있는 대부분의 음성 자료들은 음질이 떨어지기 때문에 정밀한 음향 분석이 거의 불가능하며, 이 자료들만으로는 표준말의 기준이 되는 서울말의 세대별, 직업별, 성별 발음의 특성을 체계적으로 파악할 수 없고, 서울 이외의 다른 지역 방언들과 해외 동포들의 발음을 체계적으로 파악할 수 없다. 따라서 우리는 지금 구할 수 있는 음성 자료들을 모아 체계적으로 정리하는 작업을 수행하는 동시에 서울말의 세대별, 직업별, 성별 발음의 특성을 체계적으로 파악할 수 있는 음성 자료와 서울 이외의 다른 지역 방언들과 국외 동포들의 발음을 체계적으로 파악할 수 있는 음성 자료를 수집하고 정리하는 작업을 시작해야 할 것이다. 따라서 이 논문에서 필자는 어떤 음성 자료들을 수집해야 하고, 이 음성 자료들을 어떻게 수집하고 어떻게 정리할 것인가 논의하고자 한다.

2. 어떤 음성 자료를 수집할 것인가?

우리가 지금 이 시점에 사명감을 가지고 수집해야 할 음성 자료는 크게 두 종류로 나눌 수 있다. 하나는 음성 합성기와 음성 인식기의 개발에 필요한 음성 자료이고, 다른 하나는 문화재로서 보존하고 학술 연구에 이용할 수 있는 음성 자료이다.

음성 합성 및 인식 기술은 지난 1월에 열린 스위스 다보스 세계경제포럼에서 미래를 여는 7가지 기술 중의 하나로 선정되고, 올해 미국 MIT대학에서 발간된 '테크놀로지 리뷰(Technology Review)'에서 10년 내 세상을 바꿀 10가지 기술 중의 하나로 될 만큼 중요한 기술로서 국가 경쟁력 강화를 위해 정부 차원에서 집중 육성해야 하는 기술이다. 다행스럽게도 올 들어 정

1) 한국정신문화원(1987~1995), 국립국어연구원(1997, 1998) 참조.

보통신부와 산업자원부에서 경쟁적으로 음성 합성 및 인식 기술 개발에 많은 예산을 지원하기로 결정했는데, 이 국책 사업의 일부로서 음성 합성기 및 인식기 개발을 위한 음성 데이터베이스 구축 사업이 올 하반기부터 진행될 예정이다.

따라서 이 논문에서는 문화재로서 보존하고 학술 연구에 이용할 수 있는 음성 자료의 수집에 관해서만 논의하고자 한다.

2.1. 서울말 음성 자료

서울은 우리나라 정치·경제·사회·문화의 중심지이고, 표준말은 서울말에 기반을 두고 있다. 따라서 서울말은 다른 방언보다도 우선적으로 연구되어야 마땅하다. 그러나 서울말에 관한 음성학적 연구는 의외로 부진하다. 예를 들어 지난 100년간 서울말의 발음이 제법 많이 변화했음에도 불구하고 세대에 따른 발음 변화에 관한 체계적인 연구가 이루어지지 않고 있으며, 서울말에 대한 체계적인 음향음성학적 연구도 아직은 미흡한 상태이다. 따라서 세대간의 발음 변화를 잘 파악할 수 있고, 각 세대의 발음을 체계적으로 연구할 수 있는 음성 자료를 수집해야 한다. 세대간의 발음 변화를 파악하기 위해서는 제보자를 20대, 30대, 40대, 50대, 60대 이상으로 나누어 음성 자료를 수집해야 할 것이다.

2.2. 방언의 음성 자료

서울 이외의 다른 지역의 방언 자료를 수집하는 것도 매우 중요한 일이다. 각 지역 방언의 음성 자료는 후세에 중요한 문화재로서 인식될 것이기 때문이다. 방언을 어느 정도까지 세밀하게 쪼개서 음성 자료를 수집할 것이냐는 얼마나 많은 예산을 확보할 수 있느냐에 달려 있다. 시·군·면 단위까지 세밀하게 쪼개서 수집하면 좋겠지만 여건이 허락하지 않으면 군 단위까지 나눠

서 수집하면 될 것이다. 세대간의 발음 변화를 파악하기 위해서는 제보자를 적어도 20대, 40대, 60대로 나누어 음성 자료를 수집해야 할 것이다.

2.3. 국외 동포의 음성 자료

현재 약 570만 국외 동포들이 140여 개국에 살고 있는데, 이들의 언어에 대해서는 산발적인 연구가 있었을 뿐 체계적인 연구는 한 번도 이루어진 적이 없다. 특히 우리말의 고향을 가장 많이 간직하고 있는 카자흐스탄 동포들의 말은 다른 지역 동포들의 말보다도 우선적으로 수집해서 정리해 두어야 한다. 대부분의 나이 어린 국외 동포들이 우리말을 제대로 구사하지 못하고 있는 현실을 고려하면 서둘러서 국외 동포들의 말을 수집해야 할 것이다. 국외 동포들의 말도 출신 지역에 따라, 거주 지역에 따라 차이가 날 수 있지만 이러한 차이가 확연하게 드러나지 않으면 무시해도 좋을 것이다. 그리고 세대간의 발음 변화를 파악하기 위해서는 제보자를 최소한 30대와 60대로 나누어 음성 자료를 수집해야 할 것이다.

3. 음성 자료의 수집 방법

3.1. 수집할 단어와 문장의 선정

음성 자료는 방언이나 세대에 따른 어휘의 차이를 조사하기 위해서가 아니라 발음의 차이를 조사하기 위해 수집하는 것이므로 모든 음소들이 여러 환경에서 골고루 나오는 어휘들과 모든 음소와 문장 유형이 고루 안배된 문장들을 먼저 선정해야 한다. 전자를 음성 공학에서는 ‘음성적으로 균형 잡힌 단어들(phonetically balanced words)’이라 하고 후자를 ‘음성적으로 균형 잡힌 문장들(phonetically balanced sentences)’이라 한다. 음성 공학자들은 음성적으로 균형 잡힌 단어와 문장들을 만들어서 음성 합성기와 음성 인식기

의 개발에 이용하고 있는데,²⁾ 이미 개발된 단어와 문장 목록을 검토하고 부족한 점은 개선해서 사용하면 될 것이다. 방언의 음성 자료를 수집할 때는 조사 자료로 선정된 단어와 문장들을 조사하는 지역의 방언형으로 바꾸는 작업을 먼저 해야 한다.

단어나 문장의 낭독은 자연스런 발화와 음성학적으로 차이가 나므로 자연스런 발화 자료도 함께 수집해야 한다. 자연스런 발화를 유도하기 위해 “학교 다닐 때 가장 기억에 남는 일이 있으면 이야기 좀 해 주세요.,” “자녀분들은 뭘 하는지 말씀해 주세요.” 등과 같이 제보자가 즐거운 마음으로 대답할 수 있는 질문 목록을 만들어야 한다.

3.2. 제보자의 선정

제보자는 가능하면 3대 이상 한 지역에 거주하고 외부 지역에서 장기간 생활한 경험이 없는 사람이어야 한다. 같은 지역에 사는 사람들도 교육 수준과 직업에 따라 조금씩 다른 말을 구사하므로 가능하면 비슷한 교육 수준과 직업을 갖고 있는 제보자를 구해야 한다. 방언 연구자들은 남성 제보자를 선호하는데, 음성 자료를 체계적으로 수집하기 위해서는 남성 제보자와 여성 제보자를 같은 수로 선정하는 것이 바람직하다.

3.3. 녹음

지금까지 방언 연구자들은 소형 아날로그 녹음기와 비전문가용 마이크를 이용해 음성 자료를 녹음한 다음 음향 분석기를 이용하지 않고 자신들의 청취 판단으로 음성 자료를 분석해 왔다. 영국의 위대한 음성학자인 대니얼 존스(Daniel Jones) 교수가 주장했듯이 인간의 귀가 가장 훌륭한 음성 분석 도

2) 음성 공학적인 관점에서 구축된 음성 데이터베이스에 대해서는 이용주 외(1995), 이호영(1995), 지민제(1995), 이용주(1996) 등 참조.

구입에 틀림없다. 그러나 청취 판단은 오류의 가능성이 있으며 객관적으로 검증되기 어렵다는 심각한 문제점을 가지고 있다. 따라서 방언 연구, 특히 음성 연구의 과학화를 위해서는 반드시 음향 분석기를 이용해서 음향 분석을 해야 한다. 음향 분석기를 이용하면 귀로는 판단하기 어려운 중요한 말소리의 음향적 특징을 파악할 수도 있다. 그런데 소형 아날로그 녹음기와 비전문가용 마이크로 수집한 음성 자료는 녹음 품질이 많이 떨어지기 때문에 음향 분석기를 이용한 음향 분석에 사용하기에 적당하지 못하다. 물론 녹음 품질이 떨어질수록 음성 자료로서의 가치도 떨어지게 된다.

녹음 장비의 성능은 보통 신호 대 잡음비(signal-noise ratio)로 평가한다. 녹음 장비의 신호 대 잡음비가 높을수록 주변 소음이 적게 들어가서 고품질의 녹음이 가능해진다. 신호 대 잡음비는 데시벨(dB)을 단위로 해서 표시하는데, 미국의 국립음성언어센터(National Center for Voice and Speech)에서는 전문적인 음향 분석을 위해서 85데시벨 이상의 신호 대 잡음비를 갖고 있는 녹음 장비를 사용할 것을 권고하고 있다(Titze 1994). 그런데 좋은 품질의 아날로그 녹음기의 신호 대 잡음비도 68데시벨 정도밖에 되지 않기 때문에 아무리 좋은 녹음실에서 전문가용 마이크를 이용해 녹음하더라도 만족스런 음성 자료를 얻을 수 없다.

85데시벨 이상의 신호 대 잡음비를 갖고 있는 녹음 장비로는 DAT와 KAY사의 CSL 등이 있다. CSL은 컴퓨터에 연결해서 사용하는 음성 녹음 및 분석을 위한 음향 분석기인데, 이 기계는 부피가 다소 클 뿐만 아니라 컴퓨터와 함께 사용해야 하므로 이동성이 현저히 떨어진다. 따라서 방언 음성 자료의 수집을 위해서는 이동성이 탁월한 DAT를 사용하는 것이 바람직하다.

마이크도 녹음의 품질에 큰 영향을 미친다. 고품질의 녹음을 하려면 Shure SM-48과 같은 XLR 타입의 고성능 마이크를 사용해야 한다. 지금까지 방언 조사를 할 때는 제보자가 마이크를 보고 긴장하는 것을 방지하기 위해 마이크를 숨기고 녹음했다. 그러나 마이크가 제보자의 입으로부터 멀리 떨어져 있으면 주변 소음이 많이 들어가기 때문에 DAT와 고성능 마이크를 사용하더

라도 좋은 품질의 녹음을 얻을 수 없다. 따라서 녹음할 때에는 마이크를 입에서 10센티미터 이내로 가까이 대고 녹음해야 한다. 장시간 녹음을 할 경우에는 마이크를 손에 들고 있을 수 없으므로 마이크 받침대에 고정시키거나 AKG Acoustics사의 C-410과 같이 머리에 착용하는 고성능 마이크를 사용하는 것이 바람직하다.

녹음할 때 녹음 환경도 중요한데, 가능하면 제보자를 방음 처리가 잘 된 녹음실로 초청해서 녹음하는 것이 바람직하고, 이것이 여의치 않을 때는 주변 소음이 없는 방에서 녹음해야 한다.

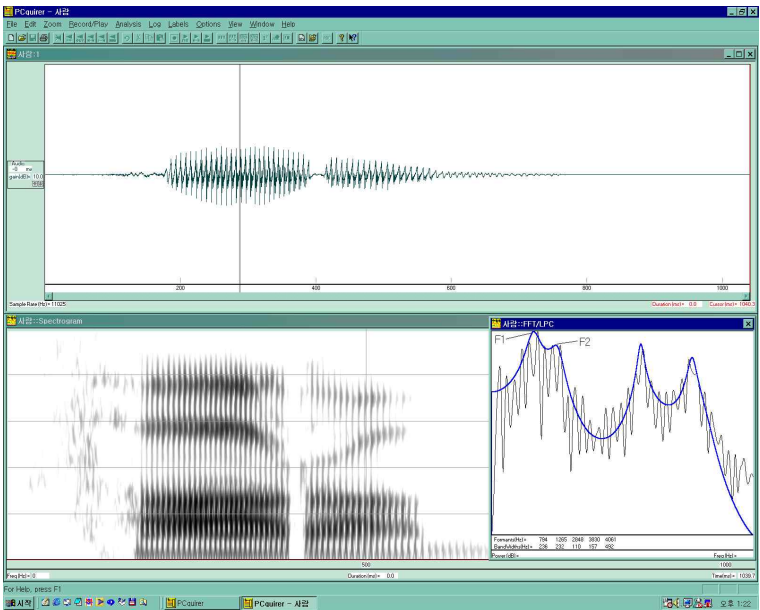
4. 음성 자료의 정리

4.1. 음성 파일 만들기

지금까지의 방언 조사에서는 아날로그 녹음기를 이용하여 음성 자료를 반복 청취하면서 분석하는 방법이 사용되었다. 그러나 이와 같은 연구 방법은 몇 가지 심각한 문제점을 가지고 있다. 첫째, 원하는 부분을 검색하려면 빨리 감기, 되감기, 재생 등의 기계 조작을 반복해야 하기 때문에 시간과 노력이 많이 소요된다. 둘째, 음성 자료 분석을 위해 기계 조작을 반복하다 보면 녹음기가 쉽게 고장나서 경제적인 손실이 발생한다. 셋째, 귀로 분간하기 어려운 발음 내지 음향 자질을 분석할 수 없으며 음성 분석의 과학화를 이룰 수 없다. 이와 같은 문제점들은 컴퓨터 음향 분석기를 이용해 음성 자료를 컴퓨터 파일로 변환한 다음 분석하면 쉽게 해결된다.

음성 자료의 정리와 분석을 위해서 사용할 수 있는 컴퓨터 음향 분석기로는 KAY사의 CSL와 같이 독립된 신호 처리 하드웨어를 가지고 있는 음향 분석기와 KAY사의 Multi-Speech나 Scicon사의 PCquirer/PitchWorks와 같이 사운드 카드를 이용하는 음향 분석기가 있다.³⁾ 사운드 카드는 녹음을 위해 개발된 제품이 아니고 음악 재생을 위해 개발된 제품이므로 신호 대 잡

음비가 70데시벨에 불과해⁴⁾ 고품질의 음성 자료를 만들기 위해서는 CSL을 사용해야 한다. 그러나 CSL은 1200만원이 넘는 고가의 장비이므로 한 연구기관에서 하나만 구입해서 음성 자료를 녹음하고 이를 음성 파일로 변환하는데 사용하고 실제 음성 분석은 Multi-Speech나 PCquirer/PitchWorks와 같은 저가의 고성능 음향 분석 소프트웨어를 사용하는 것이 바람직하다.⁵⁾



〈그림 1〉 ‘사람’의 파형, 스펙트로그램, FFT/LPC 스펙트럼

- 3) Multi-Speech와 PCquirer/PitchWorks는 CSL로 만든 파일을 읽을 수 있다.
- 4) 그러나 신호 대 잡음비가 70데시벨만 되어도 거의 대부분의 음향 분석을 만족스럽게 수행할 수 있다.
- 5) 최근에는 상당히 좋은 음향 분석 소프트웨어들이 무료로 배포되고 있다. 이 소프트웨어들은 Multi-Speech나 PCquirer와 같은 프로그램을 구입할 예산이 없는 사람들에게는 좋은 분석 도구가 된다. 무료 소프트웨어를 사용하더라도 Sound Blaster DE 정도의 고급 사운드 카드와 Shure SM-48과 같은 마이크는 반드시 갖춰야 한다.

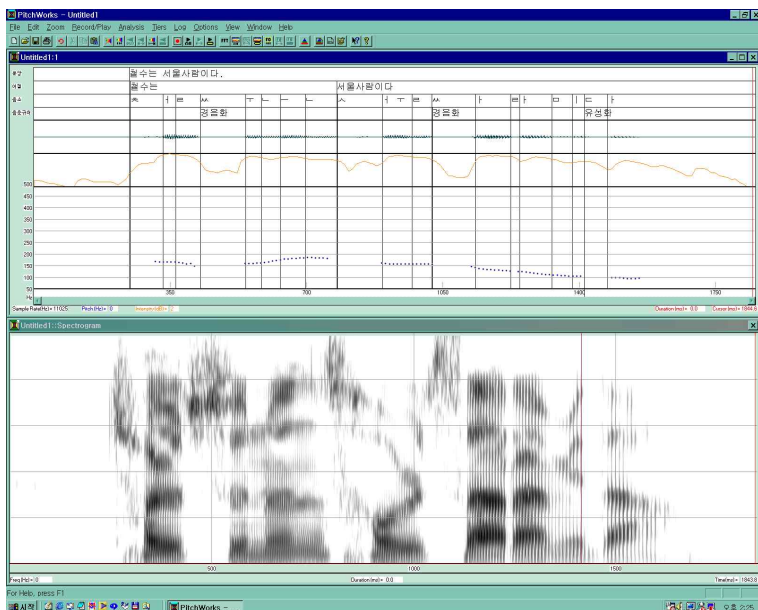
〈그림 1〉은 컴퓨터 음향 분석 소프트웨어인 Scicon사의 PCquirer란 소프트웨어를 이용해서 ‘사람’이란 단어를 분석한 것이다. 〈그림 1〉의 상단에는 ‘사람’의 파형이 나와 있고 하단에는 ‘사람’의 스펙트로그램이 나와 있으며, 하단 오른쪽에는 첫 모음 중간 부분의 FFT/LPC(Fast Fourier Transform/Linear Prediction Coefficient) 스펙트럼이 나와 있다. 거의 대부분의 음향 정보는 스펙트로그램에서 얻을 수 있으며, 모음의 음가를 결정하는 포먼트 주파수 정보는 FFT/LPC 스펙트럼에서 얻을 수 있다.

〈그림 1〉과 같이 음성 자료를 음성 파일로 저장해 놓으면 정밀한 음향 분석이 가능해지며, 원하는 파일을 쉽고 빠르게 검색할 수 있고, 마우스의 클릭만으로 원하는 파일을 무한대로 반복 청취할 수 있게 된다. 뿐만 아니라 음성 자료를 디지털 신호로 저장해 놓으므로 시간이 지나도 음질의 변화가 생기지 않는 장점이 있다.

4.2. 분절 및 전사

음성 자료를 녹음하고 음성 파일로 변환한 다음에는 음성을 문장, 어절, 음절, 음소 단위로 분절하는 작업이 필요하다. 〈그림 2〉는 Scicon사의 PitchWorks를 이용해서 “철수는 서울 사람이다.”라는 문장의 음파를 문장, 어절, 음소 단위로 분절하고 한글을 이용해 전사한 예를 보여준다. 음소열 아래에는 ‘경음화’, ‘유성화’와 같이 관찰된 음운 규칙을 기록했다.⁶⁾ 물론 연구자의 의도에 따라 각 열에 어떤 정보를 기록할 것인지 정할 수 있다. 예를 들어 1열에는 문장을 맞춤법에 따라 적고, 2열에는 억양 정보, 3열에는 세기(intensity)에 관한 정보, 4열에는 음소의 길이에 관한 정보를 적을 수 있다. 또한 10열까지 서로 다른 정보를 적을 수 있다. 이와 같은 방법으로 음성 자료를 분석해 놓으면 다른 연구자가 수정할 수도 있으므로 음성 자료가 점점 더 정확하게 분석될 수 있다.

6) 국어 음성의 전사와 음운 규칙에 관해서는 이호영(1996) 참조. 국어 음성의 음향적 특징에 관해서는 신지영(2000) 참조.



〈그림 2〉 “철수는 서울 사람이다.”의 분절과 전사

음성 자료를 분절하고 전사하는 작업에는 많은 인력과 시간이 소요되므로 예산 범위 내에서 작업 범위를 정해야 하며, 음성 자료 분석 작업에 들어가기 전에 연구자들이 어떻게 분절하고 어떤 정보를 기록할 것인지 결정해야 한다.

5. 맺음말

지금 서울대학교 규장각에는 많은 고문서들이 체계적으로 보관되어 있다. 조선시대에 규장각을 설립해 중요한 문서들을 국가적인 차원에서 보관한 것은 이 문서들의 중요성을 인식했기 때문이다. 조선시대에는 녹음기가 없었기 때문에 음성 자료는 하나도 남아있지 않다. 아마도 녹음된 음성 자료가 있다면 국보로 지정될 것이다.

조선시대에 조정에서 왕조실록을 편찬했던 것과 마찬가지로 우리는 음성 자료를 체계적으로 녹음하고 정리하는 작업을 주기적으로 해서 후손에게 물려주어야 한다. 이 자료는 『조선왕조실록』 못지않은 중요한 문화재가 될 것이다. 이 작업에는 많은 인력과 예산이 소요되므로 한 개인이 할 수 있는 일이 아니다. 정부의 지원 하에 국립국어연구원에서 추진해야 하는 일이다.

외국의 선진국들은 음성 합성기와 음성 인식기의 개발을 위해 방대한 양의 음성 데이터베이스를 구축하고는 있지만 음성언어를 하나의 문화재로 인식하고 이를 체계적으로 녹음하고 정리하는 나라는 아직 없는 것 같다. 그러나 우리는 세계 제일의 음소 문자를 만든 민족이고 『조선왕조실록』과 같은 방대한 역사 기록을 보존하고 있는 문화 민족이다. 따라서 우리의 이 빛나는 전통을 계승한다는 차원에서 음성 자료를 체계적으로 수집하고 정리하는 작업을 추진해야 할 것이다.

참고 문헌

- 국립국어연구원(1997). 『서울토박이말 자료집』 I.
- 국립국어연구원(1998). 『서울토박이말 자료집』 II.
- 신지영(2000). 『말소리의 이해』. 서울 : 한국문화사.
- 이용주 외(1995). 「해외 음성 DB 구축 동향」, 『제12회 음성통신 및 신호처리 워크샵 논문집』, 한국음향학회.
- 이용주(1996). 「음성 데이터베이스의 현황 및 과제」, 『1996년 제13회 음성통신 및 신호처리 워크샵 논문집』. 한국음향학회.
- 이호영(1995). 「대화체 음성 및 운율 DB」, 『제12회 음성통신 및 신호처리 워크샵 논문집』, 한국음향학회.
- 이호영(1996). 『국어음성학』. 서울 : 태학사.
- 지민제(1995). 「한국어 음성의 레이블링 기준에 관하여」, 『제12회 음성통신 및 신호처리 워크샵 논문집』, 한국음향학회.

- 한국정신문화연구원(1987ㄱ). 『한국방언자료집』 III(충청북도편).
- 한국정신문화연구원(1987ㄴ). 『한국방언자료집』 V(전라북도편).
- 한국정신문화연구원(1989). 『한국방언자료집』 VII(경상북도편).
- 한국정신문화연구원(1990ㄱ). 『한국방언자료집』 IV(충청남도편).
- 한국정신문화연구원(1990ㄴ). 『한국방언자료집』 II(강원도편).
- 한국정신문화연구원(1991). 『한국방언자료집』 VI(전라남도편).
- 한국정신문화연구원(1993). 『한국방언자료집』 VIII(경상남도편).
- 한국정신문화연구원(1995ㄱ). 『한국방언자료집』 IX(제주도편).
- 한국정신문화연구원(1995ㄴ). 『한국방언자료집』 I(경기도편).
- Titze, Ingo R.(1994). *Workshop on Acoustic Voice Analysis: Summary Statement*, National Center for Voice and Speech.