

LPC를 이용한 평안방언의 음향지표에 관한 연구

송철규*, 이명호*, 김영배**

= Abstract =

A Study for Acoustic Cues of Pyoung-An Do Dialect Using LPC

Chul Gyu Song*, Myoung Ho Lee*, Young Bae Kim**

This paper deal with the acoustic cues of Pyoung-An Do dialect using linear prediction. Also, this paper describes a statistical comparison between standard tone speech data and Pyoung-An Do dialects. The analysis done mainly focused on the distribution of formants and pitch periods accord to accents variation. For the purpose of objective comparison, the experiments are performed by extract formant LPC spectrum and pithch periods from average magnitude difference function waveforms.

Summing up the results, In disyllable words(VCV pattern), prepositioned vowels have longer phonation time than postpositioned vowels and the intrinsic phonation time is more longer in the low vowels than in the high ones. The africative consonants show the mixed characteristics of the plosive and fricative consonants.

The remarkable acoustic cues are the low frequency noise-like waves just before the 1st formants in the plosive consonants, the high frequency noise-like waves in the fricative consonants, and phonation time is not affected by the kinds of prepositioned or postpositioned vowels.

1. 서 론

현재 통용되고 있는 한국어의 모음은 9개의 단모음과 12개의 중모음 등 21개이며, 자음은 9개의 파열음과 3개의 파찰음, 3개의 마찰음, 3개의 비음 및 1개의 유음등 19개로 되어 있다[1].

자음에 대한 음향적 연구는 자음의 복잡한 성질과 불규칙적인 시간적 관계등으로 인한 연구방법의 곤란성 때문에 연구업적이 비교적 적다.

특히 방언이 청음인지에 관한 결정적인 역할을 하는 음향지표(acoustic cues)[2]에 관한 연구는 전무한 실정이므로 방언의 어미와 조사가 연결된 어음재료에 대한 음향지표를 관찰하는 것은 방언 연구의 중요한 기초자료가 된다.

최근 한국어 단어인식에 대한 연구가 활발히 진행되고 있으나, 조사 및 어미변화, 조음현상 때문에 각 단어들의 어미경계가 뚜렷하지 않아 음성 연구에 있어서 패턴의 표준화가 어려운 점이 있다. 그러므로 연속음을 분석하기 위해서는 음소나 음

<접수 : 1992년 9월 1일>

* 연세대학교 전기공학과

** 동국대학교 국어국문학과

* Yonsei Univ. Dept. of Electrical Engineering

** Dong Kug Univ. Dept. of Korean Language and Literature

절단위로 인식처리하는 것이 바람직한 방법으로 제안되고 있다(3). 또한 음운론적인 연구는 음향학적인 1차적인 관심사가 되어오지 못하였으나 근래 음성을 합성·인식·이해하는 여러 작업에서 음운 규칙과 음운부의 처리과정이 설정되어있는 현실을 볼 때 방언이 갖는 음운학적인 특성의 고찰은 흥미있는 연구라 할 수 있다. 따라서 본 연구에서는 평안 남북도 10개의 지점의 어음재료와 서울 표준어에 대한 비교 고찰을 시도하였다.

2. 평안남·북도 방언 해석 방법

평안방언의 특징을 추출하기 위해 먼저 녹음된 음성을 16비트의 해상도를 갖는 A/D변환기를 이용하여 10KHz로 샘플링한다. A/D변환된 음성샘플은 해밍윈도우(Hamming window)함수를 곱한 후 LPC 방법(4)으로 음성해석을 하였다. 이로부터 성도를 특징짓는 PARCOR(partial correlation)계수를 Le-Roux(5)방법으로 구하고 예측계수 $\{a_i\}$ 를 추출한다. 예측계수로 부터 예측기를 형성하고 입력신호는 inverse filtering하여 예측오차를 발생시켜 피치(pitch)검출에 이용된다.

2.1 선형예측방법

윈도우를 통과한 음성 샘플중에서 시간 $t=nT$ 에서의 $S(nT)$ 의 예측신호 $\hat{S}(nT)$ 는 과거 p 개의 샘플로 나타낼 수 있고, 오차는 식(2.2)와 같다.

$$\hat{S}(n) = \sum_{k=1}^p a_k S(n-k) \quad (2.1)$$

$$e(n) = S(n) - \sum_{k=1}^p a_k S(n-k) \quad (2.2)$$

다음의 그림 2.1은 선형예측법의 구성을 보여주고 있다.

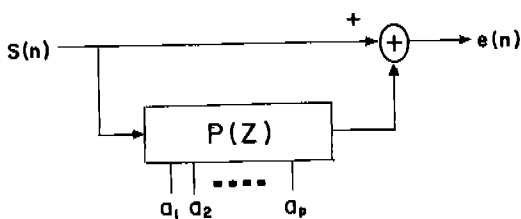


그림 2.1 선형예측법의 블럭선도

Fig. 2.1 Block diagram of linear prediction

2.2 피치 검출

LPC inverse filtering된 잔차(residual error)신호로부터 AMDF(average magnitude difference function)를 구해서 그 파형에서 주기를 검출하는 방법을 사용하고(6-8), 입력음성을 가지고 피치를 검출할 경우 성도의 포먼트 영향 때문에 피치의 에러가 발생할 확률이 큰 반면, 잔차신호를 이용하면 inverse filtering이 되어 포먼트의 영향이 제거되기 때문에 잔차신호를 이용하였다.

그림 2.2는 피치검출과정을 나타냈다. 피치 검출에 사용되는 잔차신호는 LPC해석에서 예측계수 $\{a_i\}$ 를 이용하여 구한다.

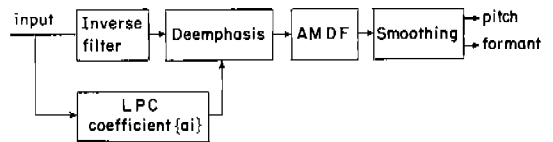


그림 2.2 피치검출의 블럭선도

Fig. 2.2 Block diagram of pitch detection

3. 실험대상 및 방법

3.1 실험대상

평안방언을 쓰는 제보자를 찾아 같은 내용을 녹음케하여 FM테이프에 녹음하였다. 제보자들은 평안 남·북도 36개 시군(광복전 기준)에서 각도 10개 지점에서 월남한 60세 이상의 남자 제보자를 선정하여 119개어에 각각 어미에 조사가 연결된 어형이 쓰인 문장을 얻도록 조사자의 질문에 대해 녹음을 하는 과정을 거쳤고, 표준어에 대한 같은 내용의 어음재료는 출생후 서울지역에서 계속 거주해 온 정상남자 5명에 대한 발성음을 녹음하였다.

3.2 실험방법

FM 카세트 테이프에 기록된 어음재료를 16비트 A/D변환한 후 예측차수 12차의 LPC스펙트럼을 얻어서 각 어음재료의 포먼트 주파수와 그 대역을 산출하고 각 어음재료의 각 256샘플에 대하여 AMDF

를 얻어서 그 구간에서 피치의 평균주파수를 산출하여 서울 표준어와 평안 방언과의 통계적인 특성을 비교하였다.

4. 연구결과 및 분석

4.1 단모음

표 4.1은 서울지역 표준어 단모음의 포먼트 주파수 및 표준편차를 정리한 데이터이며, 그림 4.1은 이에 대한 분포도를 나타낸다.

표 4.1 서울지역 표준어 단모음의 제 1, 2 포먼트 주파수 및 표준편차

Table 4.1 1st and 2nd formant frequency & standard deviation of seoul standard vowels

	제 1 포먼트	제 2 포먼트
아	736±45 [Hz]	1000±112 [Hz]
에	500±29	2050±193
오	405±43	1300±96
애	512±33	2040±165
우	289±17	875±78
어	500±59	2272±203
으	320±37	1000±74
이	285±23	2074±132

제 2 포먼트

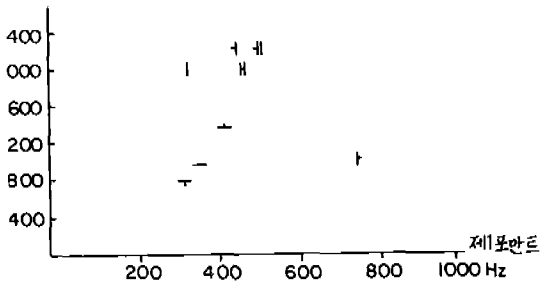


그림 4.1 서울지역 표준어 단모음의 제 1, 2 포먼트의 분포

Fig. 4.1 1st and 2nd formant distribution of Seoul standard vowels

4.2 조사가 연결된 어음자료

단어와 조사를 분리하여 포먼트와 대역폭을 살

펴보면 표준어의 ‘그림’에 대한 제 1, 2, 3 포먼트 주파수와 대역폭은 322[56]Hz, 1064[100], 2237[184]이고 표준어 ‘ㅣ’에 대한 제 1, 2 포먼트 및 대역폭은 285[23], 2074[132]이다. 이때 어미와 조사를 연결한 표준어에서의 ‘그림ㅣ’에서 제 1, 2, 3 포먼트 및 대역폭은 651[84]Hz, 1110[381], 2178[368]이고, 평안도 자성 지점의 ‘그림ㅣ’에 대한 제 1, 2, 3 포먼트 및 대역폭은 672[84], 1542[153], 2191[124]로 나타났는데 자성 지점의 ‘그림ㅣ’가 표준어 ‘그림ㅣ’에서 보다 제 1, 2, 3 포먼트가 조금씩 높게 나타났다. 따라서 단어에 조사가 연결될 때에 연결되는 조사의 포먼트와 포먼트 대역폭에 의한 영향이 있다는 것을 알 수 있는데 ‘그림’의 ‘ㅣ’의 제 1, 2 포먼트가 320, 3000(Hz)인데 ‘ㅣ’의 제 1, 2 포먼트가 285, 2074(Hz)로서 조사 ‘ㅣ’의 제 2 포먼트의 주파수와 대역폭이 ‘ㅣ’의 두배이다. 강동의 ‘가래가’의 제 1, 2, 3 포먼트도 728[278], 1364[153], 2179[124]로서 표준어 ‘가래가’의 606[384], 1075[381], 2086[368]보다 크게 나타났다.

표준어 ‘두드러기 보다’에서 제 1, 2, 3 포먼트는 263, 1412, 2191이고 강동의 262, 1408, 2183이다. 이는 ‘ㅣ’의 제 1, 2 포먼트가 289, 875이므로 제 1 포먼트가 낮다는 것을 알 수 있다.

그 외의 모든 어음자료에서 방언과 표준어의 통계학적인 특성은 각 지역 모든 방언의 제 1 포먼트가 표준어의 제 1 포먼트 보다는 작고, 제 2, 3 포먼트는 각 지역의 방언이 표준어 보다 모든 큰 값들을 보였다. 또한 평안도 각 지점간의 각 어음자료에서는 서로 유사한 포먼트 분포를 보였다. 그림 4.2는 표준어 ‘는데’의 LPC스펙트럼이고 제 1, 2, 3 포먼트는 각각 429, 2129, 2826Hz의 분포를 갖는다.

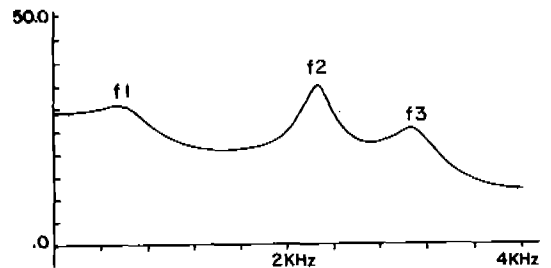


그림 4.2 표준어 ‘는데’의 LPC스펙트럼

Fig. 4.2 LPC spectrum of Seoul standard “neundae”

그림 4.3은 '코끼리보다'에 대한 양덕 지점(a)과 서울 표준어(b)에 대한 제 1, 2, 3 포먼트 위치를 나타내고 있다. 제 1 포먼트는 표준어가 양덕 지점보다 주파수가 낮으나 제 2, 3 포먼트는 표준어에서보다 더 낮다.

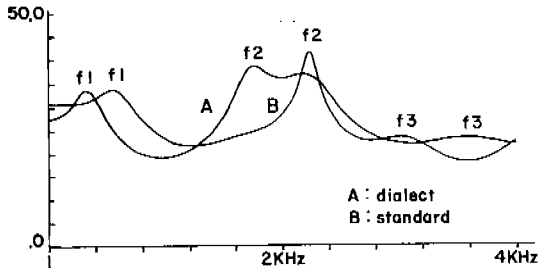


그림 4.3 “코끼리보다”의 양덕 지점과 서울 표준어에 대한 제 1, 2, 3 포먼트 위치

Fig. 4.4 1st, 2nd and 3rd formant position about yangdeok and Seoul standard language of “kokiriboda”

그림 4.4는 ‘밥먹고’에 대한 영원지점(a)과 서울 표준어(b)의 제 1, 2, 3 포먼트 위치를 나타냈고, 비슷한 포먼트 분포를 나타낸다.

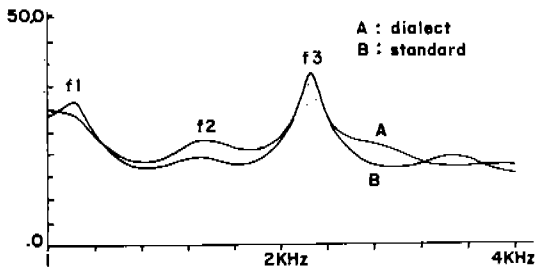


그림 4.4 ‘밥먹고’에 대한 영원 지점과 서울 표준어에 대한 제 1, 2, 3 포먼트 위치

Fig. 4.4 1st, 2nd, and 3rd formant position about youngwon and Seoul standard language of “babmukgo”

시계열 자료 처리에 유용한 통계 패키지인 MINITAB을 이용하여 방언 및 표준어에 대해서 얻어진 각각의 제 1, 2 포먼트 주파수에 대한 분포를 분석하였다.

그림 4.5는 제 1 포먼트에 대하여 표준어와 방언

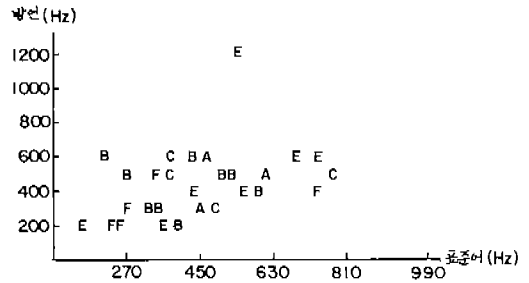


그림 4.5 표준어와 방언의 제 1 포먼트의 분포

Fig. 4.5 1st formant distribution of Seoul standard language and dialects

의 분포를 도시하였다. 동일한 어음 재료로서 A는 표준어와 양덕지점 어음의 제 1 포먼트, B는 표준어와 영원, C는 표준어와 용강, D는 표준어와 평양, E는 표준어와 자성, F는 표준어와 강동지점의 어음에 대한 제 1 포먼트의 비교이다. X-축은 표준어이고, Y-축은 각 지점의 방언이다. 그림의 B, C, F에서 알 수 있듯이 표준어와 방언의 제 1 포먼트가 낮은 주파수를 갖고, E(표준어와 자성의 비교)의 표준편차가 가장 크다.

5. 결과 및 고찰

모음의 가장 중요한 음향 지표는 포먼트의 주파수 상의 분포상태로 알려져 있는데, Peterson[9]은 제 1, 2 포먼트만으로서 분별하기 힘든 모음도 구별되어 들을 수 있는 것은 제 3 포먼트의 역할에 의존하는 것이라 하여 제 3 포먼트설을 주장하였으나, Fischer-Jørgensen[10]은 인접한 두 포먼트는 그 사이의 주파수대역에 있는 한개의 포먼트와 동일한 청음효과를 주장한 바 있는데 본 연구의 결과, 모든 어미와 조사가 연결된 어음에서 제 1, 2 포먼트는 100%의 출현도를 보이고 폭이나 강도도 넓고 강하여 이를 뒷받침할 수 있다. Fant[11]도 제 1, 2 포먼트가 모음의 성질을 결정하는데 중요한 역할을 한다는 동일한 견해를 보였으며 제 3, 4 포먼트는 전설모음에서 중요한 역할을 한다고 하였는데, 본 연구결과에서도 전설모음이 포함된 ‘가래가’, ‘해개오리가’ 등에서 제 3 포먼트의 주파수 분포가 높은 이유를 설명해 준다. 또한 Ylppö 및

Sovijävi는 포만트와 공명실과의 관계를 연구한 결과 제 1 포만트는 인두강과 관계가 있고 제 2 포만트는 구강, 그리고 제 3 포만트는 전구강과의 관련이 있음을 설명한다.

6. 결 론

LPC를 이용한 표준어와 평안도 각 지점의 방언의 어미와 조사가 연결된 어음재료를 분석한 결과 다음과 같은 결론을 얻었다.

1. 표준어 단모음의 포만트 분포는 ‘i’ ‘æ’ ‘e’에서 제 1, 2 포만트가 넓은 간격을, 제 2, 3 포만트가 좁은 간격을 유지하고 있었고 ‘u’ ‘o’ ‘a’에서는 반대로 제 1, 2 포만트가 좁은 간격을, 제 2, 3 포만트가 넓은 간격을 유지하고 있다.
2. 평안도 각 지점의 방언과 표준어의 어미와 조사가 연결된 어음재료에서 방언의 제 1, 2, 3 포만트가 표준어의 제 1, 2, 3 포만트 보다도 약간씩 높은 포만트 주파수 분포를 보였다. 이러한 현상은 악센트에 의한 포만트 주파수의 이동 현상을 말해준다.
3. 그외의 모든 어음재료에서 방언과 표준어의 통계학적인 특성은 각 지역 모든 방언의 제 1 포만트가 표준어의 제 1 포만트 보다는 작고, 제 2, 3 포만트는 각 지역의 방언이 표준어 보다 모두 큰 값들을 보였다. 따라서 방언은 성도의 인두강의 전달함수와 깊은 관계를 갖는다는 것을 유추할 수 있다. 또한 평안도 각 지점간의 각 어음재료에서는 서로 유사한 포만트 분포를 보였다.
4. AMDF를 구한 후 악센트에 의한 피치 주기의 상승과 하강을 평균한 결과 140-180Hz 사이의 값을 가졌고 표준어의 피치주기의 평균과 같거나 약간 높은 주파수 값들을 가졌다. 즉 방언과 표준어에서 억양이 피치의 상승과 하강에 대한 평균주파수에는 큰 영향을 미치지 않는다는 것을 알 수 있었다.
5. 평안도의 각 지점의 어음들의 서울 지점보다 제 1 포만트는 낮으나 제 2 포만트 주파수는 비교적 높게 나타나고 있다. 따라서 이러한

결과는 제 1 포만트는 인두강과 제 2 포만트는 구강과 관계가 깊다는 Ylppö 등의 연구결과를 따르면, 평안방언이 음성해석시 전달함수수에 해당하는 성도에서 특히 인두강과 구강에 의한 영향이 밀접한 관계가 있음을 지적할 수 있다.

6. 악센트 유형 분류기준 설정을 위해서 Discrimination 판별법을 적용하기 위한 판정함수의 설정에 대한 연구가 진행되어야 한다.

참 고 문 헌

- 1) 허용, : 국어음운학. 정음사, 1979.
- 2) A. Ylppö, A. Sovijärvi, "Sonographic and palatographic studies of full denture, half denture and edentulous cases, Acta Odont Scand 20 : 257, 1962.
- 3) W. A. Lea, : Trends in speech Recognition. Wngglewood cliffs, NJ, Prentice-hall, 1980.
- 4) J. D. Markel, A. H. Gray, : Linear prediction of speech. springer verlag, New York, 1976.
- 5) Le Roux, C. Gueguen, "A fixed point computation of partial correlation coefficients", IEEE Trans. Acoust., Speech, Signal Processing, ASSP-25, 3, 1977.
- 6) L. R. Rabiner, "A comparative performance study of several pitch detection algorithm", IEEE Trans. Acoust., Speech, Signal Processing ASSP-24, 5, 1976.
- 7) C. K. Un, S. C. Yang, "A pitch extraction algorithm based on LPC inverse filtering and AMDF", IEEE Trans, Acoustic, Speech, Signal Processing, ASSP-25, 6, 1977.
- 8) G. E. Peterson, H. L. Barney, "Control methods used in a study of the vowels", J. Acoust., Soc., Am., 24, 175-184, 1952.
- 9) David A. Krubsack, Russell J. Niederjohn, "An autocorrelation pitch detector and voicing decision with confidence measures developed for noise corrupted speech", IEEE Trans. signal processing, vol. 39, No. 2, pp. 319-329, Feb,

1991.

- 10) E. Fischer-Jorgensen, "What can the techniques of acoustic phonetics contribute to linguistics?", Proceedings of the VIII International Congress of Linguists, Oslo, pp. 433-478, 1958.
- 11) G. Fant, G. Henningsson, and U. Stalhmmer, "Formant frequencies of Swedish Vowels", STL-QPSR Speech Transmission Lab, Quarterly Progress and Status Report, KTH Stockholm, No. 4 : 26-31, 1964.