

포만트 정보의 동적 변화특성 조사에 관한 연구

Investigation on Dynamic Behavior of Formant Information

조 철 우)
Jo, Cheolwoo

ABSTRACT

This study reports on the effective way of displaying dynamic formant information on F1-F2 space. Conventional ways of F1-F2 space (different name of vowel triangle or vowel rectangle) have been used for investigating vowel characteristics of a speaker or a language based on statistics of the F1 and F2 values, which were computed by spectral envelope search method. Those methods were dealing mainly with the static information of the formants, not the changes of the formant values (i.e. dynamic information). So a better way of investigating dynamic informations from the formant values of speech signal is suggested so that more convenient and detailed investigation of the dynamic changes can be achieved on F1-F2 space. Suggested method used visualization of static and dynamic information in overlapped way to be able to observe the change of the formant information easily. Finally some examples of the implemented display on some cases of the continuous vowels are shown to prove the usefulness of suggested method.

Keywords: formant, dynamic behavior, F1-F2 space

1. 서론

본 논문에서는 음성의 기본적인 조음 특성을 나타내 주는 포만트 정보는 음성신호의 기본적인 조음 기관의 운동과 밀접하게 연관이 되어 있기 때문에 음성학에서는 화자별 음성의 특징 분석, 특정 언어의 음성 특징 분석 등에 활용되어 왔고 [1][2], 음성공학에서는 음성인식을 위한 기본적인 파라미터로도 활용되어 왔다[3][4]. 말소리장에 분야에서는 장애음성의 측정, 교정, 재활 등의 목적으로도 활용되어 왔다[5]. 그러나 기존에는 F1-F2를 중심으로 정적인 값을 위주로 통계값을 구한 뒤에 이를 도표 상에 그리고 화자 혹은 언어의 특성을 측정하는 자료로 삼아왔다. 전반적인 특성을 측정하는 데는 문제가 없지만 화자의 발성 과정에서의 동적인 특성을 측정하는 데는 어려움이 있다. 음성학에서 정적인 도표를 변형하여 F1-F2공간에서의 움직임을 나타내는 데 활용한 예는 있지만 발성기관의 움직임을 반영할 수 없고 다만 장기적 관점에서 정적인 특

성의 변화를 이용하여 표시하고 있을 뿐이다. 말소리 장애 분야의 경우 실시간 동적 특성의 변화를 통하여 발성교정, 재활 등에 활용한 사례가 있으나 표시방법에서의 문제가 있어서 특정 소프트웨어의 경우에 한정하여 사용되고 있다[6].

음성 신호의 정보는 모음과 자음으로 구성된다. 그 중 모음은 인간의 발성기관의 위치에 따라 결정되고 다양한 공진현상에 따라 서로 다른 모음을 생성하게 된다. 음성인식 혹은 음성합성 등의 응용 혹은 음성의 원리를 규명하려는 여러가지 시도들에서 우리는 모음의 변화를 관찰하려는 여러가지 시도를 해 왔다. 모음의 공진현상을 나타내는 정보 중 가장 기본적인 것은 포만트 값들이다. 대개 두 개의 포만트 조합을 통해 우리는 모음간의 구분을 할 수 있다. 이 값은 조음현상과 연관되어 있기 때문에 포만트 분포를 통해 조음현상을 추정하는 도구로 사용하고 있다. 그러나 대부분의 경우 포만트 정보는 정적인 상태의 모음 음성을 대상으로 하기 때문에 연속 발생된 문장의 경우와 같이 움직이는 형태의 음성의 경우 파라미터의 변화를 용이하게 관찰할 수 없다. 이에 모음의 동적인 현상을 보다 더 잘 관찰할 수 있는 방안을 제시하고자 한다.

본 연구에서는 기존의 F1-F2공간 표시법을 확장하여 포만트 정보를 효과적으로 처리하고 표시할 수 있는 방법을 제안하고

1) 창원대학교, cwjo@changwon.ac.kr

자 한다. 이를 통해 포먼트 정보를 동적으로 표시하는 방법에 대하여 고찰한다.

2. F1-F2공간 표시법

F1-F2공간 표시법은 음성학 분야에서 조음특성 조사를 위하여 전통적으로 오랫동안 사용되어온 방식이다. 가로 축을 F1 주파수, 세로축을 F2 주파수로 하여 그래프로 그리고 이 그래프 상에 각 음소의 주파수 특성의 통계 값을 표시하여 나타내어 왔다. <그림1>은 전형적인 F1-F2그래프의 사례이다.

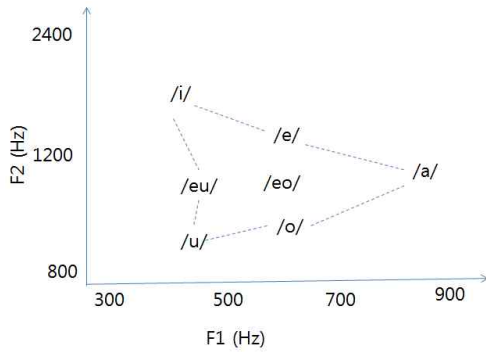


그림1. F1-F2공간 표시법 사례
Figure1. Formant Information on F1-F2 Space

<그림1>에서는 정적인 F1-F2공간에 각 음소의 분포를 표시한 것으로서 각 모음의 대표 위치를 확인할 수 있지만 동적인 변화에 관한 정보는 포함되어 있지 않다. <그림2>의 경우는 동일 모음에 대해 여러 번 발성한 경우의 포먼트 값을 표시한 것이다. 이러한 방식을 통해 한 언어 또는 화자에 대해 여러가지 모음의 포먼트 분포를 통해 특정 화자 또는 언어의 모음 분포를 알 수 있다[7]. 그러나 이 방식 역시 통계적 분포를 통해 모음의 포먼트 분포의 영역을 표시할 수 있을 뿐이지 개별 모음간 변화를 나타내는 방식으로는 적절하지 않다.

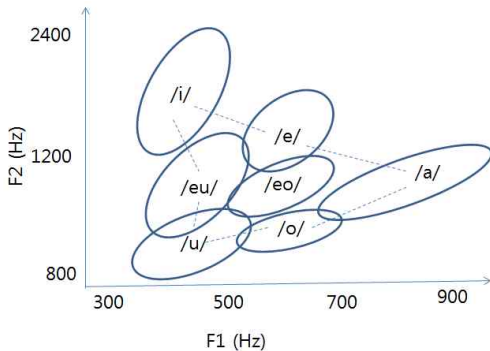


그림2. 한 언어 혹은 화자에 대한 F1-F2공간 표시법 사례
Figure2. Formant Statistics example for a speaker of a language

<그림3>은 기존의 문헌에서 변동하는 모음의 포먼트 분포를 도시한 사례이다. F1-F2 공간에서 특정 음소로부터 다른 음소로 이동하는 경우를 표시한 경우이다. 시작부분과 끝 부분이 화살표로 표시되어 있다. 그러나 세부적인 변화는 표시되어 있지 않고 단지 시작과 끝 부분의 위치만 볼 수 있을 뿐이다[8].

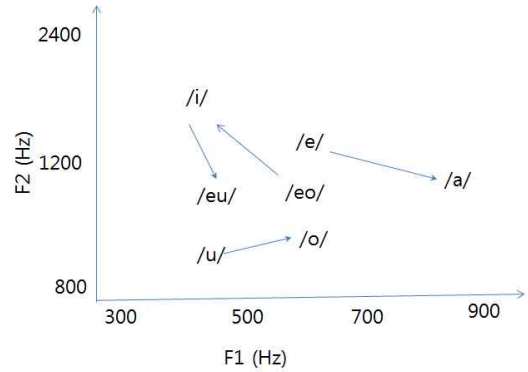


그림3. 기존 문헌에서의 동적 포먼트 표시의 사례
Figure3. An example of dynamic display of formants from the reference

F1-F2공간에 의한 모음의 포먼트 정보 표시방법은 다양한 음성학, 음성공학, 말소리장애 등의 분야에서 활용할 수 있는 유용한 방법으로 널리 사용되고 있다. 그러나 사례에서 살펴 보듯이 다양한 동적인 변화 요소가 있는 음성의 특성을 상대적으로 정적인 표시법에 의존하고 있으므로 기존의 표시방법에 의한 정보 전달에는 한계가 있다.

3. 시각화 방법의 제안

3.1 일반적 고려사항

기존의 포먼트 가시화 방법을 통해서는 다음과 같은 정보들을 얻을 수 있다.

- (1) 각 모음의 정적인 상태의 포먼트 분포
- (2) 특정 언어의 모음 포먼트 분포
- (3) 화자의 발음의 정적인 특성
- (4) 포먼트 변화의 정적인 경향

그러나 기존의 표시방법을 통해서는 다음과 같은 정보는 얻을 수 없다.

- (1) 특정 화자의 모음의 포먼트의 세세한 이동정보
- (2) 특정 화자의 발성 파라미터의 변화하는 형태

본 논문에서는 이러한 정보를 포함하여 하나의 F1-F2공간에 표시할 수 있도록 하기 위해 새로운 가시화 방법을 제안한다. 본 논문에서는 인접한 모음간의 포먼트 값의 변화를 나타내기 위해 연속적으로 분석한 포먼트 값의 궤적을 이용하는 방식을

제안한다. 일반적으로 자연스러운 발성의 경우는 모음간의 변동 구간의 길이가 짧기 때문에 여러 개의 포먼트 값을 구하는 것이 힘들다. 이를 극복하기 위하여 포먼트 값을 분석하는 구간을 정할 때 20ms에서 25ms정도의 분석구간을 택하고 2ms정도의 짧은 구간을 이동하며 분석하도록 한다. 이렇게 함으로써 이동하는 구간에서의 포먼트 값의 갯수를 늘여서 빠른 발성에서의 짧은 전이구간의 경우에도 변화를 표시할 수 있을 만큼의 값을 얻을 수 있다.

이러한 동적인 정보는 정적인 정보와 겹쳐 보이도록 표시한다. 이렇게 하면 정적인 정보와 동적인 정보를 동시에 하나의 그래프에서 관찰할 수 있게 된다. 정적인 정보를 통하여 발성자의 정상상태(혹은 단독발성상태)에서의 조음기관을 대표하는 포먼트 값을 표시할 수 있다. 정상상태의 경우는 특정 화자의 지속적으로 발성한 모음으로부터 여러 개의 프레임을 분석하여 모음의 분포도를 생성한다. 이렇게 하여 모음별로 그 화자의 포먼트 분포를 볼 수 있다. 발성자의 특성을 충분히 반영하기 위하여 정상상태의 모음을 5회이상 발성한 자료를 이용하여 포먼트의 통계값을 구한 뒤 표시한다.

<그림4>에서는 F1-F2공간에서의 개선된 동적 포먼트 표시 과정을 설명한다.

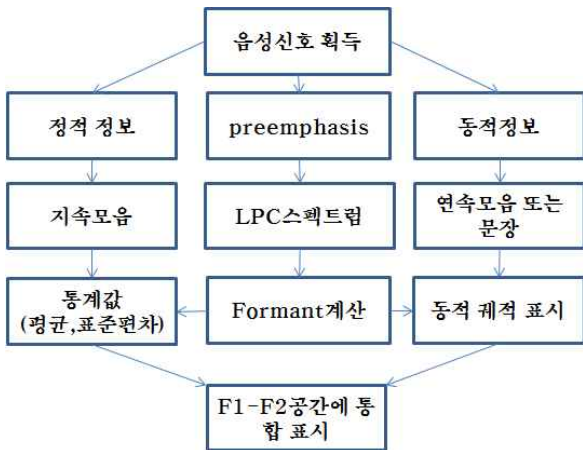


그림4. 제안된 포먼트 궤적 표시방법

Figure4. Suggested way of formant information visualization

3.2 정적 정보의 표시방법

<그림5>는 통계값을 구하기 이전의 원래 포먼트 값들의 분포도이다. <그림6>은 정적인 포먼트 분포도를 평균과 표준편차 값을 통해 가우시안 모델을 통해 나타낸 것이다. <그림7>은 동일한 내용을 3차원 모델을 통해 표시한 것이다. 가우시안 모델 형태로 통계값을 표시한 이유는 다음과 같다. 일반적으로 모음의 포먼트 분포를 나타낼 때 앞서의 <그림2>와 같이 표시하는 경향이 있다. 이 경우는 분석에 포함된 발성자의 모든 발

성음성으로부터 측정된 특정 음소의 포먼트 값의 통계값을 의미하게 되는데 본 논문에서는 동일한 발성자라고 하더라도 정적인 상태의 발성에 대한 부분만을 통해 통계 값을 구하고 이를 그래프로 표시하는 방법을 택하였다. 다양한 형태의 정적인 모델 표시를 기반으로 동적인 변화를 표시한다. 이렇게 정적인 부분과 동적인 부분을 분리하여 표시하는 이유는 인간의발성 중 동적인 발성은 정적인 발성 상태의 조음기관을 중심으로 변이된 형태라는 가정을 전제로 한 것이다. 프레임 크기는 25ms, 겹치기는 2ms 간격으로 분석한다.

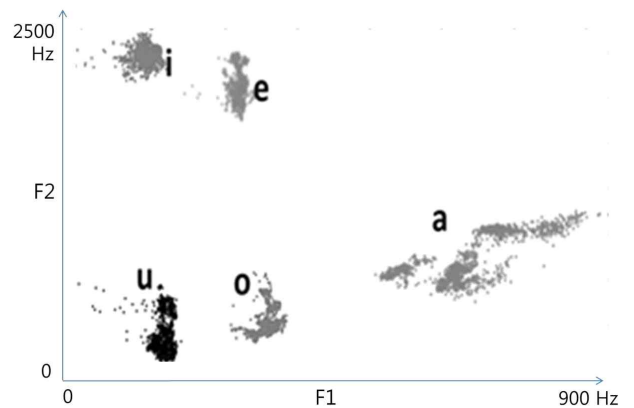


그림5. 모음의 정적인 포먼트 분포도

Figure5. Static formant distribution of 5 vowels

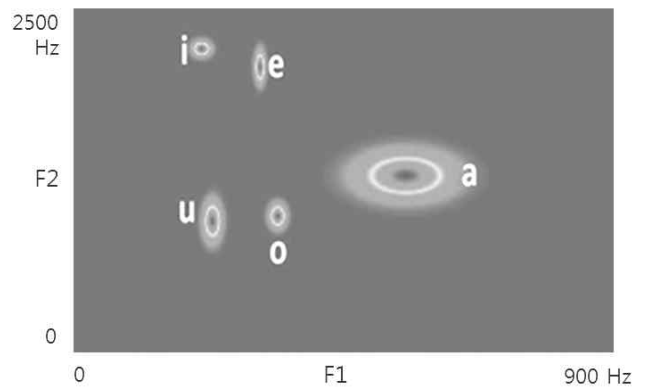


그림6. 정적인 포먼트 분포의 가우시안 통계모델

Figure6. Gaussian statistical model of static formant distribution

이와 같은 정적인 표시를 통해서서는 다음과 같은 정보를 나타낼 수 있다.

- (1) 정상 발성상태에서의 발성자의 모음의 포먼트 분포
- (2) 발성자의 모음의 포먼트 값의 정상상태의 통계값의 범위

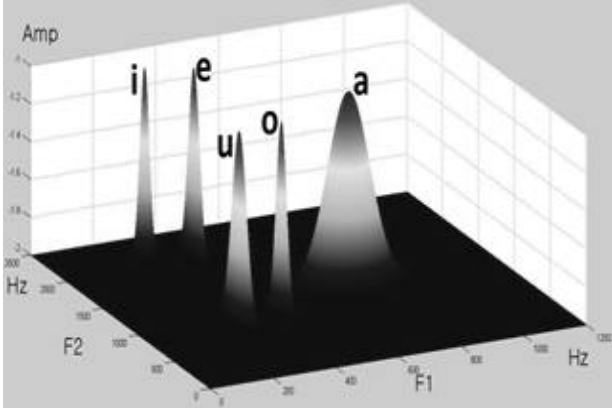


그림7. 정적인 포먼트 분포의 3D가우시안 모델
Figure7. 3D Gaussian model of static formant distribution

3.3 동적인 정보의 표시방법

동적인 발성상태를 표시하기 위해서는 연속 모음 혹은 문장으로부터 연속적인 포먼트 정보를 계산한 다음 단계적으로 계산값을 시간정보에 따라 순서대로 F1-F2공간에 배열하여 표시한다. 실험 목적에 따라 발성자의 특성을 반영하기 위하여 여러 번의 발성을 하나의 그래프에 포함할 수도 있다.

- 1단계: 연속적인 정적인 모음과 발성된 문장으로부터 모음 부분이 구분된 음성 신호를 준비한다.
- 2단계: 각 모음의 정적인 위치를 표시한다.
- 3단계: 연속적인 모음에 대해 F1-F2분포를 표시한다.
- 4단계: 서로 다른 발성 속도에 대해 다른 색으로 변화를 표시한다.(옵션)
- 5단계: 1단계에서 4단계의 정보를 동일한 그래프 상에 표시한다.
- 6단계: 동일한 발성 혹은 서로 다른 발성속도에 대해 위의 단계를 반복한다.
- 7단계: 필요하다면 궤적의 형태를 점 혹은 선으로 다르게 표시한다.

<그림8>은 정적인 포먼트 값의 통계값 위에 겹쳐 그려진 동적 궤적의 예이다. 여기서는 /i/, /o/, /u/가 연속적으로 발성된 경우이다. 그림에서 각 모음을 연결하는 선이 여러 개인 이유는 여러 개의 서로 다른 발성을 함께 표시하였기 때문이다. 동일 화자의 경우에도 발성에 따라 달라지므로 여러 개의 궤적이 만들어 지고 이를 통해 발성자의 발성 경향을 파악할 수 있다. 그리고 이러한 표시법을 통하여 발성자의 정적인 발성과 동적인 발성이 어떻게 차이가 있는지를 시각적으로 쉽게 확인할 수 있다.

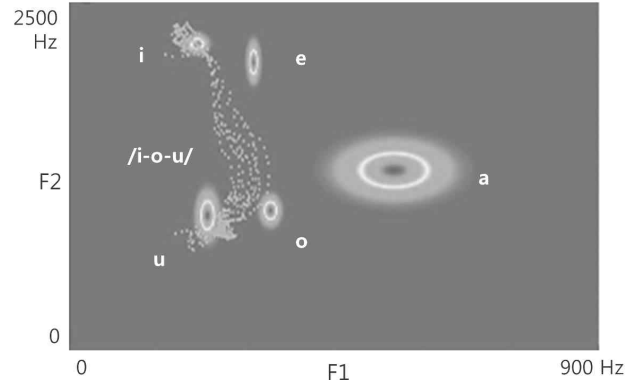


그림8. 정적인 궤적 위에 표시된 동적 궤적 (/i-o-u/의 경우)
Figure8. Overlapped trace of formants on static trace (in case of /i-o-u/)

4. 구현결과

제안된 방법을 한 화자의 6회씩 발성된 연속모음에 대해 적용하여 결과를 제시한다.

<그림9>는 제안한 방법에 의해 그려진 동적인 포먼트 궤적의 예이다. 단일 20대 남성 화자로부터 6번 발성된 음성으로부터 그려진 동적인 궤적이다. 이 그림에서는 정적인 모음의 위치와 함께 동적인 변화를 동시에 관찰할 수 있으므로 발성운동을 관찰하기가 용이하다.

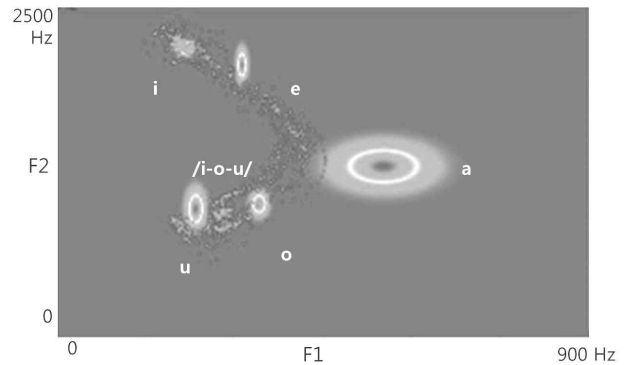


그림9. /i-o-u/의 동적 표시
Figure9. Dynamic trace of /i-o-u/

<그림9>에서 <그림11>까지는 제안한 방법에 의한 또 다른 동적인 궤적 표시의 사례이다. <그림9>는 연속 모음 /i-o-u/의 연속 모음에 대한 궤적이다. 모음은 /i/에서 시작하여 /u/에서 끝난다.

<그림10>에서는 /a-e-o/ (오른쪽 궤적)와 /a-i-o/ (왼쪽궤적)을 비교한 그림이다. 이 그래프에서 서로 다른 두 가지 모음 결합에서 어떻게 포먼트 궤적이 변화하는지를 관찰할 수 있다.

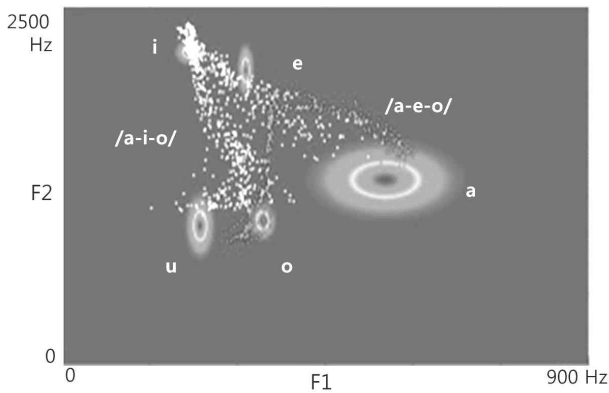


그림10. /a-e-o/와 /a-i-o/의 궤적
Figure10. Dynamic trace of /a-e-o/and /a-i-o/

<그림11>에서는 연속모음 /a-e-u/와 /a-i-u/의 경우를 비교한 것이다. 관찰결과 연속 모음을 발생할 경우 중간에 오는 모음의 경우는 정적인 모음의 중심 값에 도달하지 않고 일정 거리를 두고 궤적이 형성된다는 것을 알 수 있다. 이와 같은 관찰은 기존의 포먼트 표시법에서는 얻을 수 없는 결과였다. 오직 제한한 방법에 의해서만 이런 관찰이 가능하다.

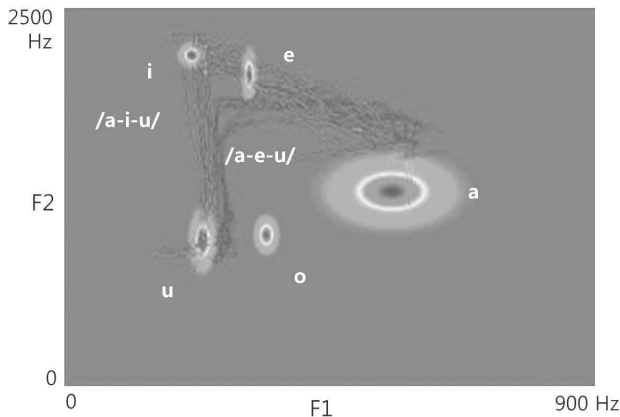


그림11. /a-e-u/와 /a-i-u/의 궤적
Figure11. Dynamic trace of /a-e-u/ and /a-i-u/

제한한 동적표시법의 응용 사례로 서로 다른 발생 속도에 따른 연속 모음이 얼마나 정적인 모음값으로부터 떨어져서 발생 되는가를 알 수 있다. 이러한 정보는 연속 발생시의 연음효과들을 측정하는 데 유용하게 적용될 수 있다.

<그림12>은 서로 다른 속도로 발생한 연속 모음 /i-e-u/의 경우를 발생 속도에 따라 서로 다른 색으로 표시한 경우이다. 그림에서 1번 궤적이 가장 느리게 발생한 경우이고 4번이 가장 빠른 속도로 발생한 것이다. 발생 속도에 따라 서로 다른 색으로 표시하고 궤적이 정적인 지점의 중심과 가장 가까운 지점을 유클리드 거리로 측정하여 표시하였다. 발생 속도는 흑색

이 가장 빠르고, 청색, 적색, 황색, 녹색의 순으로 느리다. 각 궤적에서 O표한 부분이 유클리드 거리 상 중간 모음의 정적인 포먼트 위치와 가장 가까운 부분이 된다. 이를 통하여 연속 발생시 중간 모음의 포먼트 값이 발생 속도의 변화에 따라 얼마나 정적인 포먼트 값으로부터 근접하게 되는지 알 수 있다. 이 점은 인접 모음의 조음기관의 운동 상태에서 움직이는 위치의 상한을 추정하는 데 유용한 정보를 제공해 줄 수 있다.

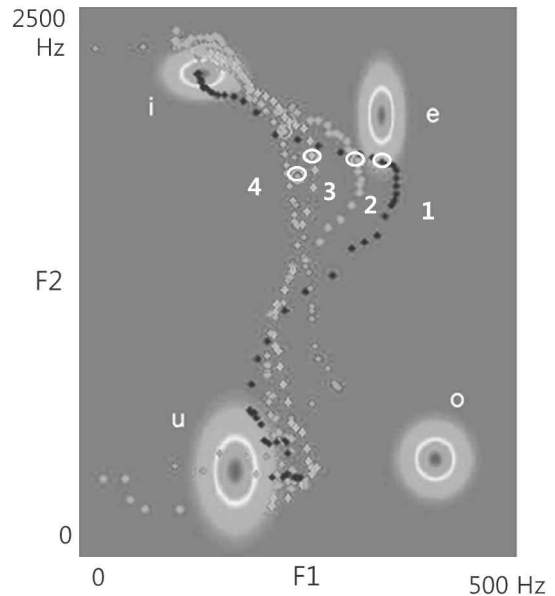


그림12. 서로 다른 발생 속도에 따른 궤적
Figure12. Dynamic trace of F1-F2 for different utterance speed

5. 결론

본 연구에서는 음성이 발생시 동적으로 어떻게 변화하는가를 관찰하는 도구로서 기존의 F1-F2공간 표시 방법을 개선하여 정적인 정보와 함께 동적인 포먼트 정보 표시에도 적합한 표시 방법을 제안하였다.

구현결과 기존의 정적인 포먼트 표시법에 비해 동적인 정보를 표시하는 데 유용하다는 것을 확인하였다. 기존의 F1-F2공간 표시방법에 비해서 정지상태의 발생기관의 포먼트 정보와 동적 상태의 발생기관의 포먼트 정보를 분리하여 이동하는 현상을 가시화 해서 분석할 수 있으므로 음성학, 음성공학, 말소리 장애 등의 분야에서 활용 가능성이 크다고 할 수 있다. 그러나 동적인 분석은 시각적 정보 표시 방법을 활용하는 까닭에 긴 음성이나 복잡한 패턴의 경우는 시각적으로 판독이 쉽지 않은 경우가 있어서 이를 해결할 방법이 추가로 고려되어야 할 것으로 사료된다. 또 한 가지 보완할 점은 실용적으로 활용되도록 하기 위해서는 실시간 분석을 통하여 시각적으로 피드백을 줄 수 있는 방안이 고려되어야 할 것으로 생각된다.

그리고 향후 모음 뿐만 아니라 문장으로부터 포먼트 정보의 변화를 표시하는 데 유용한 방법이 될 수 있도록 확장하는 데 대한 연구가 필요하다.

감사의 글

이 논문은 2013년-2015년도 창원대학교 교내연구비의 지원을 받아 수행되었습니다.

참고문헌

- [1] Li, K. (2006). Formant deflection directions of the voiced alveolar stop consonant in different vowel contexts, Ph.D. dissertation thesis, The university of arizona.
- [2] Schwartz, J.L., Abry, C., Boe, L.J., Mc'hard, L., Valle'e, N. (2005). Asymmetries in vowel perception, in the context of the dispersion-focalizato theory, *Speech Communication*, Vol. 45, 425-434, Elsevier.
- [3] Yan, Q., Saeed, V., Rentzos, D. & Ho, C.H. (2007). "Analysis and Synthesis of Formant Spaces of British, Australian, and American Accents", *IEEE Trans. on Audio, Speech, and Language Processing*, Vol. 15, No.2, pp.683-684.
- [4] Chae, Y.G., Yun, Y.S., Jung, J.M. & Eun, S.B., (2014). "Bilingual Voice Conversion Using Frequency Warping on Formant Space", *Phonetics and Speechsciences*, Vol. 6, No. 4, pp.133-139.
(채의근, 윤영선, 정진만, 은성배, (2014). "포먼트 공간에서의 주파수 변환을 이용한 이중 언어 음성 변환 연구", *말소리와 음성과학*, 6권, 4호, pp. 133-139.)
- [5] Kang, Y.A., Yoon, K.C., Lee, H.S., & Seong, C. (2010). "A Comparison of Parameters of Acoustic Vowel Space in Patients with Parkinson's Disease", *Phonetics and Speechsciences*, Vol. 2, No. 4, pp. 185-192.
(강영애, 윤규철, 이학승, 성철재, (2010). "파킨슨 환자의 음향 모음 공간 파라미터 비교", *말소리와 음성과학*, 2권, 4호, pp. 185-192.)
- [6] Carey, M., (2004). "CALL Visual feedback for pronunciation of vowels: Kay Sona-Match", *CALICO Journal*, Vol. 21, No. 3, pp. 571-601.
- [7] Van Son, R.J.J.H., & Pols, L.C.W., (1992). "Formant movements of Dutch vowels in a text, read at normal and fast rate", *J. Acoust. Soc. Am.*, Vol. 92, pp. 121-127.
- [8] Chitoran, I., Coupe', C., Marsico, E. & Pellegrino, F., (2009). *Approaches to Phonological Complexity*, Mouton de Gruyter, Berlin, NewYork, pp. 59-81.

• 조철우 (Jo, Cheolwoo)

창원대학교 전기전자제어공학부
경남 창원시 의창구 창원대학교 1번지
Tel: 055-213-3662 Fax: 055-262-5064

Email: cwjo@changwon.ac.kr

관심분야: 음성신호처리, 음성신호처리의 의학적 응용
현재 전기전자제어공학부 교수