

영어의 억양 유형화를 이용한 발화 속도와 남녀 화자에 따른 음향 분석

An acoustical analysis of speech of different speaking rates and genders using intonation curve stylization of English

이 서 배¹⁾

Yi, So Pae

ABSTRACT

An intonation curve stylization was used for an acoustical analysis of English speech. For the analysis, acoustical feature values were extracted from 1,848 utterances produced with normal and fast speech rate by 28 (12 women and 16 men) native speakers of English. Men are found to speak faster than women at normal speech rate but no difference is found between genders at fast speech rate. Analysis of pitch point features has it that fast speech has greater Pt (pitch point movement time), Pr (pitch point pitch range), and Pd (pitch point distance) but smaller Ps (pitch point slope) than normal speech. Men show greater Pt, Pr, and Pd than women. Analysis of sentence level features reveals that fast speech has smaller Sr (sentence level pitch range), Sd (sentence duration), and Max (maximum pitch) but greater Ss (sentence slope) than normal speech. Women show greater Sr, Ss, Sp (pitch difference between the first pitch point and the last), Sd, MaxNr (normalized Max), and MinNr (normalized Min) than men. As speech rate increases, women speak with greater Ss and Sr than men.

Keywords: speaking rate, intonation curve, pitch range, pitch slope, pitch distance, stylization of pitch

1. 서 론

발화 속도 증가에 따른 음향 특성의 변화는 음성 신호의 왜곡을 초래하며 음성 신호 처리 시스템의 성능 저하를 야기 시키는 것으로 알려져 있다(Pallett, et al., 1994). 한편, 음성 합성 분야에서는 자연음에서 사람의 발화 속도가 달라질 때마다 음향 자질들의 변화 양상이 달라지기 때문에 단순히 시간 축을 조정하는 것으로는 달라지는 발화 속도에 따라 변하는 음성을 제대로 표현 할 수 없다는 것을 고려해야 한다. 그래서 시간 축과 피치 축을 독립적으로 조정하여 합성된 음성의 발화 속도를 올리거나 내려도 자연성과 명료도를 떨어뜨리지 않는 것이 연구 이슈가 되고 있다(Ninness & Henriksen, 2008).

이런 맥락에서 발화 속도에 따른 억양의 음향적 특성 변화

를 연구하는 것은 의미 있는 일이라 하겠다. 발화 속도에 따라 달라지는 억양의 음향 특성에 대한 연구는 활발하게 이루어져 왔다(Kohler, 1983; Caspers & Van Heuven, 1991; Fougeron & Jun, 1998; Shih & Kochanski, 2003). 그러나 이러한 연구들은 적은 양의 발화 자료를 가지고 사람의 주관적 판단에 기초해 가변적으로 억양곡선의 모양을 유형화함으로써 음향 특징 값들을 뽑았기 때문에 일관성과 정확성에서 한계를 가질 수 밖에 없다.

대용량 음성 코퍼스의 억양 전사를 위한 표준안으로 제안된 ToBI(Tone and Break Indices)(Silverman, et al., 1992)가 억양 연구에 많이 사용되어져 왔지만 이것 역시 장시간 훈련받은 전문가에 의해 레이블링 작업이 이루어져야 하고 전문가들 사이에도 이견이 있을 수 있다. 물리적으로 동일한 억양일지라도 다르게 기술 될 수 있는 한계가 있는 것이다. 그리고 이러한 표현들만 가지고 실제 음성신호의 억양을 복원해 내는 것에는 많은 제약이 따른다. 또한 음성 발화에서 흔히 발견되는 무성 자음들이나 무성화 된 모음들은 F0값이 나타나지 않아 억양 목표점을 찾기 힘든 것도 문제가 되고 있다.

1) 창원대학교, 영어영문학과, sopye@pusan.ac.kr

이러한 한계를 극복하기 위해 개발된 자동 억양 유형화 알고리즘인 Momel(Hirst & Robert, 1993; Hirst, et al., 2000)은 주목할 만하다. Momel은 억양에 나타나는 개별 분절음의 특성이 반영된 미시적인 부분(microprosodic component)과 전체 억양 패턴을 보여주는 거시적인 부분(macroprosodic component)을 잘 표현해 주는 것으로 알려져 있다(Hirst & Robert, 1993; Wang & Narayanan, 2005). 자동화된 Momel은 영어, 프랑스어, 독일어, 스페인어 등의 언어를 대상으로 한 평가에서 높은 성능을 보여 주었다(Campione, 2001). 본 연구는 영어 음성 코퍼스 발화로부터 억양곡선을 추출하여 영어 원어민 남녀 화자의 보통 발화와 빠른 발화에 나타나는 억양의 음향적 특성을 분석하였다. 억양곡선은 Momel을 사용하여 추출하였고 억양의 근접 복사 유형화(close-copy stylization)를 이용해 억양의 정규화가 가능할 수 있다는 것을 보여준 연구(오재혁, 2014a; 오재혁, 2014b)에서 사용한 억양 자질들을 보완하여 분석에 적용하였다.

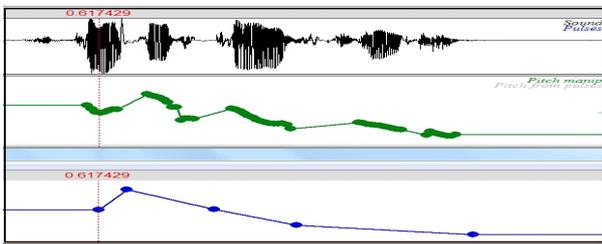


그림 1. Momel을 이용한 자동 억양 유형화
Figure 1. Automatic stylization of intonation curve by Momel

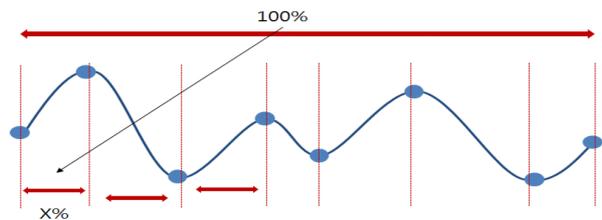


그림 2. 피치포인트 이동시간(Pt)의 계산(오재혁, 2014a)
Figure 2. Measurements of pitch point movement time (Oh, 2014a)

<그림 1>은 본 연구에서 Praat(Boersma, 2001)과 Momel을 사용해 구현한 억양 유형화의 한 예이다. <그림 1> 상단에는 음성파형이 나타나 있고 이 음성에서 추출한 피치 곡선이 바로 중간에 있으며 맨 아래에 Momel에서 추출한 피치포인트들이 나타나 있다. 이것은 원 억양 곡선에 있는 많은 피치 값들을 대표성을 가지는 몇 개의 값으로 단순화한 것인데 위의 예에서 5개의 피치포인트만 가지고 음성 합성을 통해 원래 음

성으로 복원한 음성과 피치 곡선 유형화 이전의 원음성과 비교해서 들어보는 간단한 청취실험 결과 억양의 차이가 느껴지지 않는 것으로 나타났다.

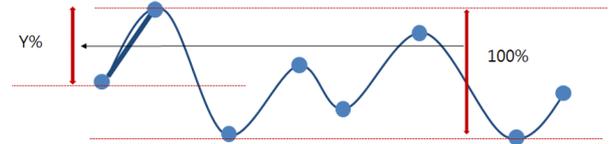


그림 3. 피치포인트 범위(Pr)의 계산(오재혁, 2014a)
Figure 3. Measurements of pitch point range (Oh, 2014a)

억양 자질들을 정규화한 연구(오재혁, 2014a; 오재혁, 2014b)에서는 <그림 2>에서와 같이 각 발화문장의 전체 발화시간을 100%로 했을 때 하나의 피치포인트가 다음 피치포인트의 위치까지 이동한 시간인 피치포인트 이동시간(Pr)을 백분율(%)로 환산(Pt)하고 <그림 3>에서와 같이 각 발화문장별 최소 피치포인트 값에서 최대 피치포인트 값까지의 범위폭을 100%로 했을 때 피치포인트들 간의 피치 값의 차이를 백분율(%)로 환산하여 계산하였다.

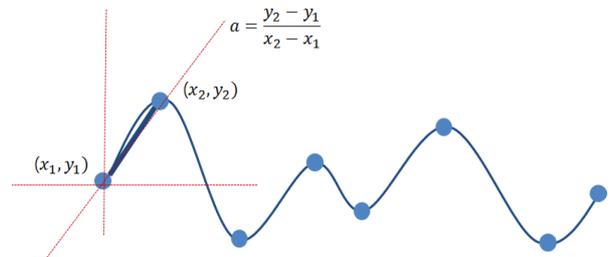


그림 4. 피치포인트들 간의 기울기(Ps) 계산(오재혁, 2014a)
Figure 4. Measurement of the slope between pitch points (Oh, 2014a)

백분율로 환산된 값을 대상으로 두 피치포인트 간의 기울기(Ps)를 측정된 것이 <그림 4>이다. 또한 환산된 값을 대상으로 <그림 5>와 같이 유클리디언 거리계산 방식으로 피치포인트 이동거리(Pd)를 구하였다. 이러한 자질들을 가지고 한국어 자유발화를 연구한 결과 남녀의 차이와 화자들 간의 차이 그리고 발화 맥락상의 여러 가지 가변적 요소들이 있었음에도 불구하고 정규화에 필요한 억양 자질들의 항상성이 발견되었다(오재혁, 2014a; 오재혁, 2014b). 이렇게 정규화 된 억양 자질들은 최근 감정음성의 억양 분석에도 사용될 정도로 적용의 폭이 넓어지고 있다(이서배, 2014).

본 연구는 이러한 억양곡선 정규화기법을 영어 발화의 음향

2) 본 연구에서는 억양 목표점(pitch target)이라는 용어를 피치포인트로 대체하였다.

3) 원문에는 ‘음높이 움직임의 이동시간’이라고 표현한 것을 본 연구에서는 편의상 피치포인트 이동시간으로 바꾸어 사용하였다.

적 분석에 적합하도록 보완하고 Momel로 추출한 억양 곡선에 적용하여 분석하였다. 적은 양의 음성 자료를 대상으로 주관적으로 억양 곡선을 추정한 기존 연구들(Kohler, 1983; Caspers & Van Heuven, 1991; Fougeron & Jun, 1998; Shih & Kochanski, 2003)과 본 연구의 차이는 많은 양의 음성 자료를 대상으로 자동 억양 유형화 기법을 이용해 억양 곡선을 일관성 있게 추출한 뒤 정규화 된 기법으로 분석했다는 점 일 것이다.

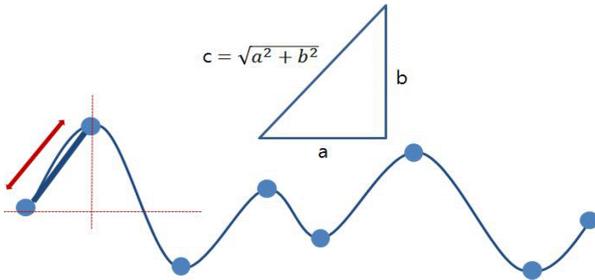


그림 5. 피치포인트 이동거리(Pd) 계산(오재혁, 2014a)
 그림 5. Measurement of the distance between pitch points (Oh, 2014a)

2. 실험

2.1 음성 코퍼스

본 연구에서 사용된 음성 코퍼스(CHAINS Speech Corpus)는 36명의 영어 원어민 화자가 다양한 방식(보통 발화, 동시 발화, 자유 발화, 속삭임, 빠른 발화)으로 발화한 것을 녹음한 것으로서 화자의 특성을 연구하고 화자인식의 성능을 개선하기 위해 개발되었다(Cummins et al., 2006). 조용한 환경에서 Shure SM50 헤드셋 콘텐서 마이크를 Marantz PMD 670 Compact Flash 녹음기에 연결해 16bit 44.1kHz PCM 인코딩 WAV파일 형태로 녹음하였다(Cummins et al., 2006). 본 연구는 이 중 본 연구의 부록에 나온 33문장을 28명의 영어 화자(남성 16명 + 여성 12명)가 보통 발화와 빠른 발화의 형태로 발화한 1,848(33x28x2)개의 발화 문장을 분석대상으로 삼았다. 빠른 발화를 위해 화자들은 샘플 발화를 듣고 샘플 발화 속도와 같은 속도로 발화하도록 지시받았다.

2.2 분석 방법

먼저 피치 측정 값의 물리적인 단위로 기존연구(오재혁, 2014a; 오재혁, 2014b)에서 사용한 Hz 대신 청각적 스케일 단위인 mel을 사용하였다⁴⁾. 자동 음성 신호처리 시스템에서도 성능향상을 이유로 mel 스케일로 필터간격을 디자인하고 있는

것을 감안 할 때 청각적 스케일을 쓰는 것은 공학시스템으로의 적용 가능성을 위해서도 의미 있는 일이라 생각한다.

본 연구는 Momel을 이용해 피치포인트들을 찾은 후 연결시키는 방식으로 억양을 유형화 시켰다. 기존 연구(오재혁, 2014a; 오재혁, 2014b)에서 살펴본 발화 문장 내 억양자질들(피치포인트 이동시간, 피치포인트 이동거리, 피치포인트 피치변위, 피치포인트 기울기)뿐만 아니라 발화 문장 전체에서 구해지는 자질들 즉, 한 문장 당 하나씩 구할 수 있는 억양자질들(<그림 6> 참조)을 추가하여 기존 연구를 보완하였다. 모든 억양 자질들의 설명은 아래와 같다.

- Pt: 피치포인트 이동시간(피치포인트들 간 이동시간의 백분율)
- Pr: 피치포인트 피치변위(피치포인트들 간 피치 차이 절댓값의 백분율)
- Ps: 피치포인트 기울기의 절댓값
- Pd: 피치포인트들 간의 거리
- Sr: 문장 피치변위(문장 내 최대 피치포인트 값과 최소 피치포인트 값의 피치 차이 절댓값)
- Sp: 첫째 피치포인트(P1)와 마지막 피치포인트(Pn)의 피치 차이 절댓값
- Sd: 첫째 피치포인트에서 마지막 피치포인트까지의 시간(Tn - T1)
- Ss: 문장 피치기울기(Ss = Sp / Sd)
- Max: 한 문장 내 최대 피치포인트 값(피치)
- Min: 한 문장 내 최소 피치포인트 값(피치)
- MaxNr: Max와 그 문장 중앙값과의 차이(절댓값)
- MinNr: Min과 그 문장 중앙값과의 차이(절댓값)

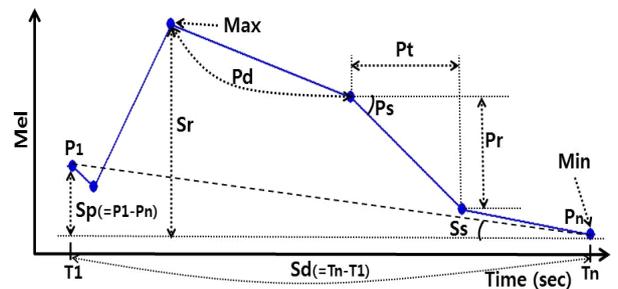


그림 6. 억양 유형화에 따른 억양 자질들
 Figure 6. Intonation features from a stylized pitch curve

위의 억양 자질들 중 Max와 Min의 경우, 두 화자 집단(남녀)의 억양을 비교할 때 남녀의 생리적인 차이(성대의 크기, 성도의 길이, 방패연골의 각도)로 인한 F0의 차이를 정규화 할 필요가 있다. 그래서 Max와 Min을 한 문장의 피치 중앙값(Median pitch value)과의 차이로 정규화한 MaxNr과 MinNr을 계산했다. 중앙값은 일반적으로 사용되는 평균값보다 피치의

4) 모두 아일랜드의 더블린과 그 근방에서 Eastern Hiberno-English를 구사하는 화자들로 구성되었다.
 5) 변환식 $mel = 1127.01048 \times \ln(1 + f/700)$ 을 사용함. 여기서 f0는 기본주파수이고 mel은 청각스케일의 음높이 단위이다.

분석에서와 같이 오류와 정확도의 문제로 인해 생기기 쉬운 극단적 값들(outliers)의 영향을 덜 받기 때문에 정규화가 필요한 억양 분석의 연구에 쓰여 지고 있다(Forsell, 2007, 이서배 & 김수정, 2011).

3. 결과 및 분석

3.1 발화 속도에 따른 분석

발화 속도는 한 문장의 음절수를 시간으로 나누어 초당 음절수(syl/sec)로 계산하여 측정했다. 발화 방식(빠른 발화, 보통 발화)과 두 화자 집단(남녀)을 모수 요인으로 하고 초당 음절수를 종속변수로 하는 분산분석(ANOVA)을 시행했다.

발화 방식에 따라 발화 속도를 분석하면 빠른 발화의 평균은 5.81syl/sec이고 보통 발화의 평균은 4.90syl/sec였는데 발화 방식은 통계적으로 유의한 차이가 있는 것으로 나타났다[F(1, 378.009), p<0.001, $\eta_p^2=0.170$](<그림 7> <표 1>). 기존 연구에 따르면 발화 속도 차이를 느끼게 하는 최소 발화 속도 차이(Just Noticeable Difference: JND)가 5%로 보고되고 있다(Quené, 2007). 본 연구에 사용된 빠른 발화와 보통 발화의 차이는 0.91syl/sec로서 보통 발화를 기준으로 5%인 0.25syl/sec보다 큰 차이를 보이고 있으므로 인지적 관점에서도 두 발화속도는 차이가 있다고 말할 수 있다.

두 화자 집단(남녀)에 따라 발화 속도를 분석하면 여성의 평균은 5.30syl/sec, 남성의 평균은 5.41syl/sec로서 두 집단은 통계적으로 유의한 차이가 있는 것으로 나타났다[F(1, 5.647), p<0.05, $\eta_p^2=0.003$]. 그러나 효과크기(effect size) 면에서 보면, 발화방식에 따른 발화 속도의 차이($\eta_p^2=0.170$)보다 남녀에 따른 발화 속도의 차이($\eta_p^2=0.003$)가 훨씬 작다고 말할 수 있다. 인지적 관점에서도 두 화자 집단 간의 차이인 0.11syl/sec는 여성 발화 속도를 기준으로 5%인 0.27syl/sec보다 작으므로 발화 속도 차이를 느끼게 하는 최소 발화 속도 차이(JND)에 못 미친다.

발화 방식과 두 화자 집단의 상호작용은 통계적으로 유의했다[F(1, 21.510), p<0.001, $\eta_p^2=0.012$]. 이것은 효과크기 면에서도 전반적인 남녀 발화의 차이보다 크다. 그러므로 다음과 같이 발화방식별로 나누어서 남녀의 차이를 분석할 필요가 있다.

보통 발화에 한정해서 남녀 발화 속도의 차이를 보면 여성의 평균은 4.73syl/sec, 남성의 평균은 5.06syl/sec로서 통계적으로 유의한 차이가 있는 것으로 나타났다[F(1, 29.719), p<0.001, $\eta_p^2=0.031$]. 이 때 두 집단의 차이인 0.33syl/sec는 발화속도가 더 큰 남성발화를 기준으로 5%인 0.25syl/sec보다 크므로 인지적 관점에서도 보통 속도일 때 남녀의 발화 속도 차이가 발생한다고 말할 수 있다. 분석 결과 나타난 여성의 발화가 남성의 발화 보다 더 긴(느린) 현상은 기존 연구에서도 보고된 바 있다(Byrd, 1992; Whiteside, 1995; Sandra & Whiteside, 1996;

Whiteside & Hodgson, 2000, Fitzsimons et al., 2001). 또한 이러한 현상은 여성이 남성보다 발화 중간 중간에 긴 휴지를 가지는 경향이 있고 어말 장음화를 하기 때문에 전반적인 발화의 길이가 길어지는 것으로 설명되었다(Sandra & Whiteside, 1996).

한편, 빠른 발화에 한정해서 남녀 발화 속도의 차이를 보면 여성의 평균은 5.87syl/sec, 남성의 평균은 5.76syl/sec로서 근소한 차이인데다 통계적으로 의미가 없는 것으로 나타났다[F(1, 2.181), p=0.140]. 그리고 두 집단의 차이인 0.11syl/sec는 남성을 기준으로 5%인 0.29보다 작은 것으로 나타나 발화 속도 차이를 느끼게 하는 최소 발화 속도 차이(JND)에 못 미치는 것을 알 수 있다.

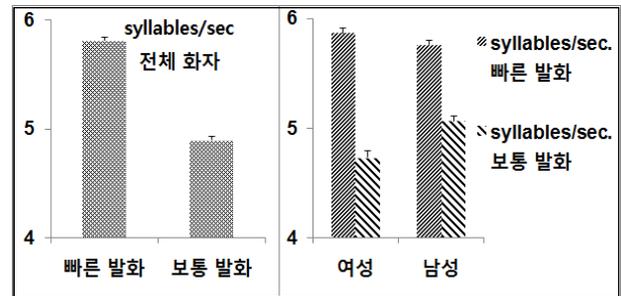


그림 7. 발화 방식에 따른 모든 화자 및 남녀 화자의 평균 발화 속도(syl/sec)

Figure 7. Average speech rate (syl/sec) of all speakers and male & female speakers according to speaking mode (fast vs. normal)

표 1. 발화 방식에 따른 전체 화자 및 남녀 화자의 평균 발화 속도(syl/sec)

Table 1. Average speech rate (syl/sec) of all speakers and male & female speakers according to speaking mode (fast vs. normal)

syl/sec	전체 화자	여성	남성
빠른 발화	5.81	5.87	5.76
표준오차	0.03	0.05	0.04
보통 발화	4.90	4.73	5.06
표준오차	0.03	0.06	0.05

3.2 발화 문장 내 피치포인트 간 억양 자질 분석

발화 방식(빠른 발화, 보통 발화)과 두 화자 집단(남녀)을 모수 요인으로 하고 억양 유형화로부터 얻은 피치포인트 자질들(Pt, Pr, Ps, Pd)을 종속 변수로 하는 MANOVA(다변량 분산분석)를 시행했다.

3.2.1 발화 방식에 따른 분석

발화 방식의 차이는 Pt[F(1, 646.716), p<0.001, $\eta_p^2=0.041$],

Pr[F(1, 41.526), $p < 0.001$, $\eta_p^2 = 0.003$], Pd[F(1, 137.952), $p < 0.001$, $\eta_p^2 = 0.009$], Ps[F(1, 22.021), $p < 0.001$, $\eta_p^2 = 0.001$]에서 모두 유의미했다(<그림 9>, <표 2>).

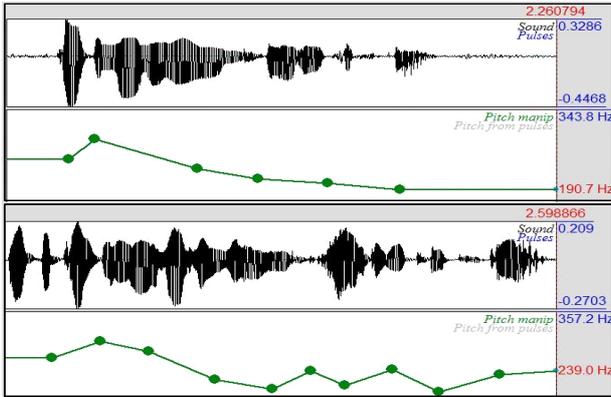


그림 8. 한 화자가 한 문장(If it doesn't matter who wins, why do we keep score?)을 빠른 발화(위)와 보통 발화(아래)로 발화한 것의 파형과 피치포인트들

Figure 9. Waveform and pitch points of fast (top) and normal (bottom) utterance of a sentence read by a speaker.

보통 발화보다 빠른 발화에서 Pt, Pr, Pd의 값이 더 큰 이유는 보통 발화에서 9,575개 이던 피치포인트의 개수가 빠른 발화에서는 7,581개로 줄어드는 것과 관계가 있다. 피치포인트 개수가 준다는 것은 억양 곡선의 단순화를 의미한다(<그림 8> 참조). 이것은 의미 있는 변화를 나타내는 변곡점의 수가 줄어드는 것으로 인해 억양곡선상의 한 피치포인트에서부터 그 다음의 의미 있는 변화가 있는 피치포인트가 나타날 때까지의 시간 비율(Pt)이 길어지고 이들 간의 주파수 차이의 비율(Pr)도 커지게 되는 것을 뜻한다. 그러므로 Pt와 Pr의 제곱합으로 구해진 유클리디언 거리에 해당하는 Pd(피치포인트 이동거리)도 빠른 발화에서 커지게 된다.

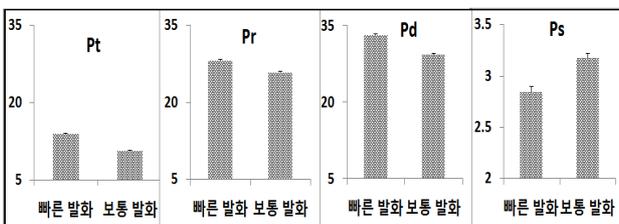


그림 9. 발화 방식에 따른 피치포인트 이동시간, 피치포인트 피치변위의 백분율(%)과 피치포인트 이동거리, 피치포인트 기울기의 평균값

Figure 8. Average percentage of pitch point movement time, pitch point pitch range and average of pitch point distance & pitch point slope according to speaking mode (fast vs. normal)

마찬가지로 보통 발화보다 빠른 발화에서 Ps(피치포인트 기울기)가 더 작아지는 것은 빠른 발화에서 나타나는 억양 곡선

의 단순화가 피치포인트 움직임의 역동성(dynamics)을 둔화시키기 때문이다(<그림 8>참조). 이것은 발화 속도가 느릴수록 피치 움직임이 많아지지만 발화속도가 빠르면 피치 움직임이 단순하게 되어 억양곡선이 부드러워지고 피치 움직임이 적어진다는 기존 연구와 맥락을 같이 한다(Vaissière, 1983; Caspers & Van Heuven, 1991; Jun, 1993; Trouvain & Grice, 1999; Shih & Kochanski, 2003).

표 2. 발화 방식에 따른 피치포인트 이동시간, 피치포인트 피치변위, 피치포인트 이동거리의 백분율(%)과 피치포인트 기울기의 평균

Table 2. Average percentage of pitch point movement time, pitch point pitch range, pitch point distance and average of pitch point slope according to speaking mode (fast vs. normal)

	Pt	Pr	Pd	Ps
빠른 발화	13.90	28.12	33.16	2.84
표준오차	0.10	0.26	0.25	0.05
보통 발화	10.67	25.90	29.27	3.18
표준오차	0.08	0.23	0.22	0.05

3.2.2 두 화자 집단(남녀)에 따른 분석

남녀의 차이는 Pt[F(1, 16.527), $p < 0.001$, $\eta_p^2 = 0.001$], Pr[F(1, 59.199), $p < 0.001$, $\eta_p^2 = 0.004$], Pd[F(1, 55.538), $p < 0.001$, $\eta_p^2 = 0.004$]에서 유의미했지만 Ps[F(1, 2.817), $p = 0.093$]에서는 유의미한 차이가 없었다. 여성보다 남성에서 Pt, Pr, Pd의 값이 더 큰 것은 남성의 발화가 여성보다 빠르므로 앞에서 언급한 빠른 발화에 나타나는 억양 곡선의 단순화에 기인하기 때문으로 보인다.

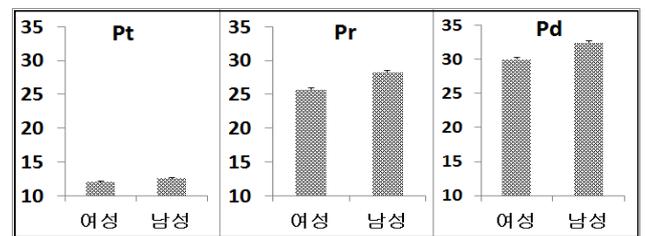


그림 10. 남녀 화자의 피치포인트 이동시간, 피치포인트 피치변위 백분율과 그에 따른 피치포인트 이동거리의 평균
Figure 10. Average percentage of pitch point movement time, pitch point pitch range and average of pitch point distance

여성보다 빠른 남성 발화의 억양 곡선이 더 단순해지면서 피치포인트들 간의 Pt가 길어지고 Pr이 커지고 이로 인해 이들의 제곱합인 Pd도 커지게 된다(<그림 10>, <표 3> 참조). 발화 방식과 두 화자 집단의 상호작용은 Pt[F(1, 3.699), $p = 0.054$], Pr[F(1, 1.394), $p = 0.238$], Pd[F(1, 2.837), $p = 0.092$], Ps[F(1, 0.069), $p = 0.792$] 모두에서 유의미하지 않았다.

표 3. 남녀 화자의 피치포인트 이동시간, 피치포인트 피치변위 백분율과 그에 따른 피치포인트 이동거리의 평균

Table 3. Average percentage of pitch point movement time, pitch point pitch range and average of pitch point distance

	Pt	Pr	Pd
여성	12.02	25.69	29.98
표준오차	0.10	0.26	0.25
남성	12.54	28.33	32.45
표준오차	0.08	0.23	0.22

3.3 발화 문장 전체 억양 자질 분석

발화 방식(빠른 발화, 보통 발화)과 두 화자 집단(남녀)을 모수 요인으로 하고 문장 레벨의 억양 자질들(Sr, Ss, Sp, Sd, Max, Min, MaxNr, MinNr)을 종속 변수로 하는 MANOVA(다변량 분산분석)를 시행했다.

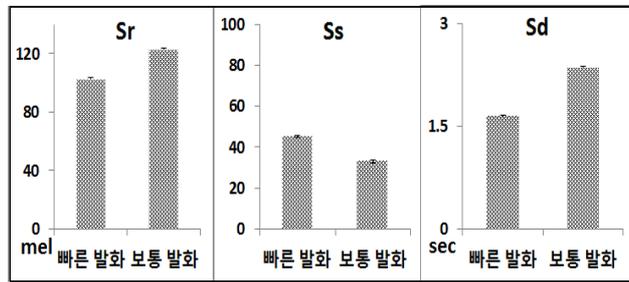


그림 11. 발화 방식에 따른 문장 피치변위, 문장 피치기울기, 첫째 피치포인트와 마지막 피치포인트 간 시간의 평균

Figure 11. Average of sentential pitch range, sentential pitch slope, time between the first pitch point and the last pitch point according to different speaking modes

3.3.1 발화 방식에 따른 분석

발화방식의 차이는 Sr[F(1, 30.301), p<0.001, $\eta_p^2=0.016$], Ss[F(1, 58.521), p<0.001, $\eta_p^2=0.031$], Sd[F(1, 539.577), p<0.001, $\eta_p^2=0.226$], Max[F(1, 34.402), p<0.001, $\eta_p^2=0.018$], MaxNr[F(1, 39.87), p<0.001, $\eta_p^2=0.021$]에서 유의미한 것으로 나타났다. 그러나 Sp[F(1, 0.279), p=0.598], Min[F(1, 0.413), p=0.521], MinNr[F(1, 1.21), p=0.271]에서는 유의미한 차이가 없었다.

그러므로 정의상 'Ss=Sp/Sd'인 것을 고려할 때 발화 방식에 따라 Ss가 유의미한 차이를 보이는 것은 Sp가 아닌 Sd의 차이가 유의미하기 때문이라 할 수 있다. 즉, 발화가 빨라짐에 따라 첫째 피치포인트와 마지막 피치포인트 간의 시간(Sd)이 단축되고 이로 인해 발화문장의 억양기울기(Ss)가 커진 것이다(<그림 11>, <표 4>).

표 4. 발화 방식에 따른 문장 피치변위, 문장 피치기울기, 첫째 피치포인트와 마지막 피치포인트 간 시간의 평균

Table 4. Average of sentential pitch range, sentential pitch slope, time between the first pitch point and the last pitch point according to different speaking modes

	Sr	Ss	Sd
빠른 발화	102.67	45.44	1.64
표준오차	0.79	0.45	0.02
보통 발화	122.66	33.31	2.36
표준오차	0.70	0.39	0.02

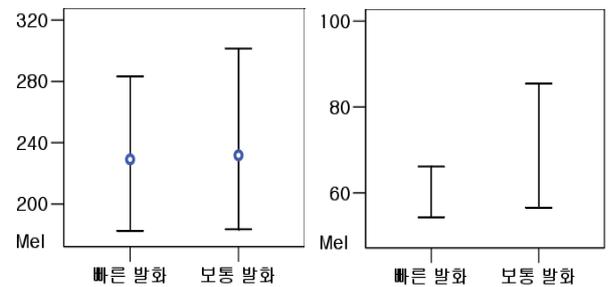


그림 12. 발화 방식에 따른 Max, Mid(중앙값), Min의 평균(좌)과 정규화 된 MaxNr, MinNr의 평균(우)

Figure 12. Average of Max, Mid, Min (left) normalized MaxNr and MinNr (right)

표 5. 발화 방식에 따른 Max, Mid(중앙값), Min의 평균(좌)과 정규화 된 MaxNr, MinNr의 평균(우)

Table 5. Average of Max, Mid, Min (left) normalized MaxNr and MinNr (right)

평균 mel	빠른 발화	보통 발화	빠른 발화 (정규화)	보통 발화 (정규화)
Max	292.25	310.51	68.58	87.97
Mid	236.37	239.12		
Min	188.65	189.63	55.90	58.18

발화 방식에 따라 Sr(발화 문장의 피치변위)이 달라지는 양상을 각 문장별로 구해지는 최댓값, 중앙값, 최솟값들의 평균으로 나타내면 <그림 12><표 5>와 같다. 전반적으로 빠른 발화에서 피치변위의 폭이 줄어드는 것을 볼 수 있는데 이것은 기존 연구의 결과(Kohler, K. 1983; Kohler, 1986; Fougeron & Jun, 1998)와 같다. 그리고 발화 속도가 증가해도 변화가 거의 없는 Min[F(1, 0.413), p=0.521]과는 달리 Max[F(1, 34.402), p<0.001, $\eta_p^2=0.018$]는 빠른 발화에서 더 작아지는 것으로 나타났다. 이러한 현상은 정규화 된 MinNr(=Mid-Min), MaxNr(=Max-Mid)에서도 마찬가지이다. 중앙값과의 거리가 클수록 문장 피치변위가 커지고 중앙값과의 거리가 작을수록 문장 피치변위가 작아지는 것을 고려할 때 빠른 발화의 Max가 보통 발화의 Max보다 중앙값으로부터의 거리가 더 가까워지

고(작아지고) MinNr은 거의 변화가 없기 때문에 빠른 발화에서 Sr이 더 작아지는 현상이 나타난다고 말할 수 있다. 이것은 프랑스어를 대상으로 한 기존 연구(Fougeron & Jun, 1998)에서 발화속도가 증가할 때 화자들이 최댓값을 낮추고 최솟값을 일정하게 유지함으로써 피치변위 폭을 줄이는 경향을 보고 한 것과 부합하는 결과이다.

이러한 Max값의 민감함은 <그림 13>에도 잘 나타나 있다. <그림 13>은 빠른 발화의 Max와 보통 발화의 Max와의 차이를 보통 발화의 Max와 비교해서 그린 것이다. 남녀 둘 다 보통 발화에서 큰 Max일수록 빠른 발화에서 Max가 더 많이 감소하는(보통 발화의 Max를 기준으로) 경향이 있는 것을 보여준다. 그리고 이러한 경향은 여성 보다는 남성에서 더 일관성 있게 나타나고 있다(<표 6>).

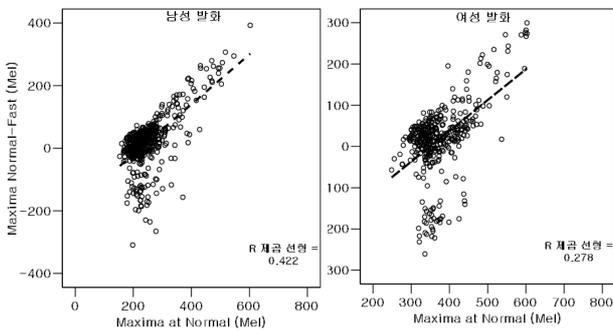


그림 13. 발화 방식에 따른 남녀 Max변화(‘보통 발화 Max’가 가로축, ‘보통 발화 Max-빠른 발화 Max’가 세로축)
Figure 13. Change of Max of men and women according to different speaking mode (horizontal axis is for 'Max of normal speech', vertical axis is for 'Max of normal speech minus Max of fast speech')

이와 같이 보통 발화에서의 Max 크기와 비례해서 보통 발화의 Max와 빠른 발화의 Max 간에 차이가 커지는 선형적인 관계를 포화효과(Saturation Effect)로 설명한 기존 연구(Fougeron & Jun, 1998)에서는 Max가 발화 속도 증가에 따라 도달해야 하는 목표 Max가 일정 수준으로 정해져 있어서 보통 발화에서 큰 Max는 빠른 발화에서 많이 감소하지만 이미 목표 값에 근접한 작은 Max(보통 발화의)는 빠른 발화에서 작게 감소할 수 밖에 없다고 말하고 있다. 이 연구는 프랑스어를 대상으로 했는데 2명의 화자로부터 구한 발화로부터 포화효과를 도출하고 있기 때문에 28명의 영어화자를 대상으로 한 본 연구에서 같은 경향을 확인한 것은 의미 있는 일이라 생각된다. 주목할 만 한 점은 Max가 중앙값으로부터 떨어진 정도가 Min이 중앙값에서 떨어진 정도보다 항상 더 크다는 점이다. 즉, MaxNr이 MinNr보다 항상 큰 것이다(<표 5> 참조). 따라서 발화 속도변화에 Max가 Min보다 더 민감하게 반응하는 이유를 포화효과에 근거해서 분석하면 Min보다 Max가 이루어야

할 감소의 폭이 더 크기 때문에 목표 값에 근접한 Min은 변화가 작은 반면 감소의 여지가 큰 Max는 목표 값과의 거리가 커지게 되므로 많은 변화(감소)를 보여준다는 추정이 가능해진다.

표 6. 남녀에 따른 ‘보통 발화 Max’와 ‘보통 발화 Max-빠른 발화 Max’간의 상관관계수

Table 6. Correlation coefficients between 'Max of normal speech' and 'Max of normal speech minus Max of fast speech' for men and women

	여성	남성
Pearson 상관계수	0.527(**)	0.649(**)
유의확률 (양쪽)	p<0.001	p<0.001
발화 문장수	396	528

3.3.2 두 화자 집단(남녀)에 따른 분석

두 화자 집단의 억양음향을 비교할 때 Max와 Min의 경우 남녀의 생리적인 차이로 인한 F0의 차이를 정규화 할 필요가 있다. 그래서 Max, Min을 종속 변수에서 제외시키고 대신에 Max와 Min을 중앙값과의 차이로 정규화한 MaxNr과 MinNr을 포함해서 MANOVA를 시행한 결과, 두 화자 집단(남녀)은 Sr[F(1, 161.163), p<0.001, $\eta_p^2=0.080$], Ss[F(1, 34.755), p<0.001, $\eta_p^2=0.018$], Sp[F(1, 42.104), p<0.001, $\eta_p^2=0.022$], Sd[F(1, 6.491), p<0.001, $\eta_p^2=0.008$], MaxNr[F(1, 86.188), p<0.001, $\eta_p^2=0.045$], MinNr[F(1, 78.162), p<0.001, $\eta_p^2=0.041$] 모두에서 유의한 차이가 있는 것으로 나타났다.

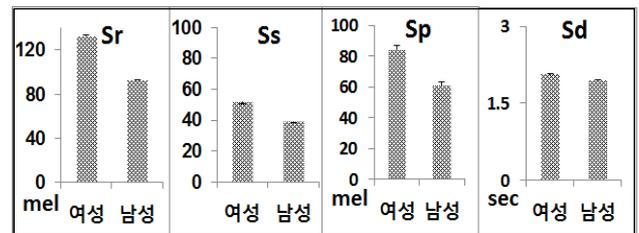


그림 14. 두 화자 집단 간의 발화문장 피치변위, 발화문장 기울기, 첫째와 마지막 피치포인트 간 피치차이, 첫째와 마지막 피치포인트 간 시간의 평균

Figure 14. Average of sentential pitch range, sentential pitch slope, pitch difference between the first and the last pitch point and time difference between the first and the last pitch point

<그림 14>는 여성의 Sr, Ss, Sp, Sd가 모두 남성보다 큰 것을 보여준다. 남성의 Sd가 여성 보다 더 작은 것은 앞 절에서 언급한 바와 같이 남성의 발화가 더 빠르다는 것을 의미한다. 발화 문장의 기울기를 뜻하는 Ss가 여성 발화에서 커지는 이유는 ‘Ss=Sp/Sd’의 계산에서 분모에 해당하는 Sd($\eta_p^2=0.008$)보다 분자에 해당하는 Sp($\eta_p^2=0.022$)의 효과크기가 Sd보다 현저히

크므로 Sp에 의해 Ss가 결정되기 때문이라 할 수 있는데 이 때, 여성의 Sp가 남성의 Sp보다 크기 때문에 여성의 Ss가 남성보다 커진 것으로 해석될 수 있다(<그림 14>, <표 7> 참조). 즉, 여성 발화에서 나타나는 억양 곡선의 시작과 끝부분의 피치 차이가 남성보다 더 큰 점이 문장의 전반적인 억양 기울기가 커지게 하는 요인이었다고 말 할 수 있다.

표 7. 두 화자 집단 간의 발화문장 피치변위, 발화문장 피치기울기, 첫째와 마지막 피치포인트 간 피치 차이, 첫째와 마지막 피치포인트 간 시간의 평균

Table 7. Average of sentential pitch range, sentential pitch slope, pitch difference between the first and the last pitch point and time difference between the first and the last pitch point

성별 \ 발화방식	Sr (mel)	Ss	Sp (mel)	Sd (sec)
여성 평균	132.74	45.42	83.82	2.06
표준오차	0.79	0.44	2.72	0.02
남성 평균	92.59	33.33	60.50	1.94
표준오차	0.70	0.39	2.35	0.02

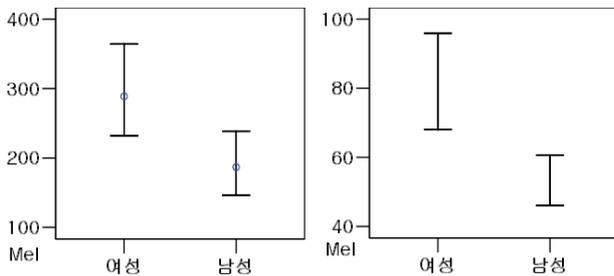


그림 15. 남녀 화자의 Max, Mid(중앙값), Min의 평균(좌)과 정규화 된 MaxNr, MinNr의 평균(우)

Figure 15. Average of Max, Mid, Min (left) and normalized MaxNr & MinNr (right) obtained from male and female speakers

<그림 15>는 Max, Min으로 남녀 화자의 피치변위를 보여주고 있다. 앞 절에서와 같이 남녀 모두에게서 MaxNr이 MinNr보다 항상 큰 것을 볼 수 있는데 이것은 한 문장에서 피치의 최댓값(Max)이 최솟값(Min)보다 중앙값(Mid)에서 항상 더 떨어져 있다는 것을 의미한다(<표 8> 참조). 또한 여성의 피치변위가 남성보다 크다는 것을 알 수 있다. 특히, 중앙값과의 차이로 정규화한 후에도 여성의 MaxNr과 MinNr 둘 다 남성의 MaxNr과 MinNr보다 큰 것을 볼 수 있다. 여성의 Sr과 Ss가 남성보다 큰 현상은 영어를 대상으로 한 기존 연구에서 여성 발화의 피치변위와 역동성이 남성 발화보다 더 크다는 보고와 맥락을 같이 하는 결과이다(Fitch & Holbrook, 1970; Hudson & Holbrook, 1981; Stoicheff, 1981; Graddol & Swann, 1983; Daly & Warren, 2001).

표 8. 남녀 화자의 Max, Mid(중앙값), Min의 평균(좌)과 정규화 된 MaxNr, MinNr의 평균(우)

Table 8. Average of Max, Mid, Min (left) and normalized MaxNr & MinNr (right) obtained from male and female speakers

평균 mel	여성	남성	여성 (정규화)	남성 (정규화)
Max	364.26	238.51	95.86	60.69
Mid	288.76	186.73		
Min	232.08	146.20	68.09	45.99

3.3.3 발화 방식과 두 화자 집단의 상호작용 분석

발화 방식과 두 화자 집단의 상호작용은 Ss[F(1, 6.928), $p < 0.01$, $\eta_p^2 = 0.004$], Sp[F(1, 5.519), $p < 0.05$, $\eta_p^2 = 0.003$], Sd[F(1, 4.107), $p < 0.05$, $\eta_p^2 = 0.002$]에서 유의미한 차이가 있었지만 Sr[F(1, 0.032), $p = 0.859$], MaxNr[F(1, 0.082), $p = 0.774$], MinNr[F(1, 0.009), $p = 0.924$]에서는 유의미한 차이가 없었다. 그래서 빠른 발화와 보통 발화를 따로따로 나누어서 남녀의 차이를 모수요인으로 하고 Ss, Sp, Sd를 종속변수로 하는 MANOVA를 시행했다.

보통 발화에 한정해서 남녀 발화를 보면 Ss[F(1, 8.817), $p < 0.01$, $\eta_p^2 = 0.009$], Sp[F(1, 8.665), $p < 0.01$, $\eta_p^2 = 0.009$], Sd[F(1, 14.320), $p < 0.001$, $\eta_p^2 = 0.015$] 모두에서 유의미한 결과가 나왔다. 그러나 빠른 발화에 한정해서 남녀 발화를 비교하면 Ss[F(1, 26.043), $p < 0.001$, $\eta_p^2 = 0.027$], Sp[F(1, 38.624), $p < 0.001$, $\eta_p^2 = 0.040$]에서는 유의미한 차이를 보였지만 Sd[F(1, 2.321), $p = 0.128$, $\eta_p^2 = 0.003$]에서는 유의미한 차이가 없었다. 보통 발화에서 남녀 간에 차이를 보인 Sd가 빠른 발화에서 차이를 보여주지 못한 것은 3.1절에서 나타났듯이 발화 속도 면에서 보통 발화에서는 남녀의 차이가 유의미했지만 빠른 발화에서 남녀 간에 차이가 없었던 결과와 같은 맥락이라 할 수 있다. 발화 속도는 초당 음절수(syl/sec)로 측정했으므로 발화 속도 계산에서 분모로 쓰인 시간(sec)와 비례하는 Sd는 발화 속도와 반비례한다. 그러므로 보통 발화에 한정해서 볼 때 여성 발화보다 빠른 발화 속도를 보인 남성 발화에서 여성보다 작은 Sd가 측정되는 것이다. 그리고 빠른 발화에 한정해서 볼 때 남녀 간에 Sd차이가 없었다는 것은 빠른 발화에서 남녀 간에 발화 속도 차이가 없었던 3.1절의 결과와 일맥상통하는 내용이다.

빠른 발화, 보통 발화 모두에서 공통적으로 여성이 남성보다 더 큰 Ss, Sp값을 갖고 있다. 이것은 여성의 피치변위와 역동성이 남성 발화보다 더 크다는 3.3.2절의 해석을 재확인하는 결과이다.

두 화자 집단별로 따로따로 나누어서 발화 방식 차이를 모수요인으로 하고 Ss, Sp, Sd를 종속변수로 하는 MANOVA를 시행했다. 여성 발화에 한정해서 보면 Ss[F(1, 40.624), $p < 0.001$,

$\eta_p^2=0.049$], $Sd[F(1, 305.396), p<0.001, \eta_p^2=0.279]$ 에서는 발화 방식 간에 유의한 차이가 있었지만 $Sp[F(1, 2.842), p=0.092, \eta_p^2=0.004]$ 에서는 없었다. 그리고 남성 발화에 한정해서 보아도 마찬가지로 $Ss[F(1, 16.389), p<0.001, \eta_p^2=0.015]$, $Sd[F(1, 246.303), p<0.001, \eta_p^2=0.189]$ 에서는 발화 방식 간에 유의한 차이가 있었지만 $Sp[F(1, 2.436), p=0.119]$ 에서는 없었다. 이것은 남녀 모두에게 있어 발화 속도가 증가함에 따라 문장의 전반적인 피치기울기가 급하도록 하는 요인은 시작 피치포인트와 마지막 피치포인트 간의 피치 차이가 아니라 문장을 발화하는데 걸리는 시간을 줄이는 것이 더 큰 관건이었다는 의미이다. 이것은 두 화자 집단별로 따로따로 나누지 않고 전체 화자를 대상으로 분석한 3.3.1절의 결과와 같은 내용이다.

주목할 만한 점은 보통 발화에서 빠른 발화로 바뀌면서 남성보다 여성의 Ss, Sd의 값이 더 큰 변화를 보인다는 것이다 (<그림 16> 참조). 효과크기 면에서 여성의 $Ss(\eta_p^2=0.049)$, $Sd(\eta_p^2=0.279)$ 는 남성의 $Ss(\eta_p^2=0.015)$, $Sd(\eta_p^2=0.189)$ 보다 크다. 즉, 발화 속도가 증가하면서 문장 전반의 피치기울기는 증가하고 발화 시간은 감소하는데 그 변화의 비율이 남성보다 여성이 더 크다는 것이다.

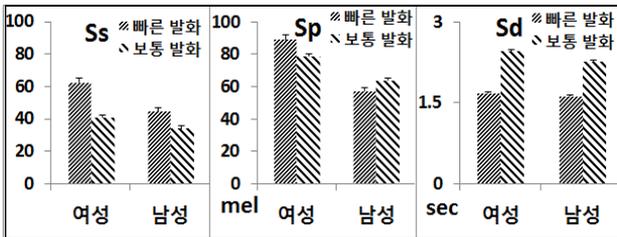


그림 16. 발화 방식과 남녀에 따른 발화문장 피치기울기, 첫째와 마지막 피치포인트 피치 차이, 첫째와 마지막 피치포인트 사이 시간의 평균

Figure 16. Average of sentential pitch slope, pitch difference between the first and the last pitch point, time difference between the first and the last pitch point according to different speaking modes and genders

이상을 종합하면, 전반적으로 여성이 남성보다 발화 문장 피치기울기가 더 급하고 억양 곡선의 처음과 끝의 피치 차이가 더 크다. 그리고 남녀 모두 보통 발화 보다 빠른 발화에서 문장 발화의 길이가 짧아진다. 이 모든 변화의 정도는 발화 속도가 증가하면서 여성이 남성보다 더 심하다. 이것은 여성이 남성보다 더 큰 운율 변화(prosodic variation)를 보였다는 연구와 맥락을 같이 한다(Fitch & Holbrook, 1970; Hudson & Holbrook, 1981; Stoicheff, 1981; Graddol & Swann, 1983; Daly & Warren, 2001).

표 9. 발화방식과 남녀에 따른 발화문장 피치기울기, 첫째와 마지막 피치포인트 피치 차이, 첫째와 마지막 피치포인트 간 시간의 평균

Table 9. Average of sentential pitch slope, pitch difference between the first and the last pitch point, time difference between the first and the last pitch point according to different speaking modes and genders

	Ss 빠름	Ss 보통	Sp 빠름	Sp 보통	Sd 빠름	Sd 보통
여성	54.27	36.55	89.00	78.65	1.67	2.45
표준오차	0.81	0.47	3.84	3.84	0.03	0.03
남성	36.54	30.08	57.22	63.77	1.62	2.26
표준오차	0.71	0.42	3.33	3.33	0.03	0.03

이러한 현상은 앞 절에서 언급한 바와 같이, 중앙값과의 차이로 정규화한 후에도 여성의 MaxNr과 MinNr이 둘 다 남성의 MaxNr과 MinNr보다 큰 것(<그림 15>, <표 8>)과 관련이 있어 보인다. 여성이 남성보다 더 큰 운율 변화를 보이는 이유를 포화 효과로 설명하면, 중앙값과의 거리가 더 큰 여성의 발화가 남성의 발화 보다는 목표 값과의 간격이 더 크고 이로 인해 발화 속도에 더 민감하게 반응할 여지가 더 크기 때문으로 추정된다.

5. 결론

지금까지 영어의 억양 유형화를 이용해 발화 속도가 다르고 화자의 성별이 다른 발화들의 음향 분석을 수행 했다. 보통 속도의 발화에서는 남성이 여성보다 발화 속도가 컸지만 빠른 발화에서는 남녀 간에 차이가 없었다.

발화 문장 내 피치포인트 간 억양 자질 분석에서 빠른 발화가 보통 발화보다 Pt(피치포인트 이동시간), Pr(피치포인트 피치변위), Pd(피치포인트 이동거리)의 값이 더 컸지만 Ps(피치포인트 기울기)는 더 작았다. 이러한 현상은 발화 속도 증가에 따른 억양 곡선의 단순화가 시간 축과 피치 축 그리고 이 두 축으로 이루어진 유클리디언 공간 각각에서 피치포인트들 간의 간격을 커지게 하고 피치포인트 움직임의 역동성을 둔화시키기 때문으로 해석된다. 그러므로 남녀의 비교에서 여성보다 남성의 Pt, Pr, Pd가 컸던 것도 남성이 여성보다 발화 속도가 빨라서 생긴 억양 단순화에 기인한다고 말할 수 있다.

발화 문장 전체 억양 자질 분석에서 빠른 발화는 보통 발화에 비해 Sr(문장 전체 피치변위)과 Sd(문장 발화 시간)와 Max(피치 최댓값)가 더 작았고 Ss(문장 피치기울기)는 더 크게 나타났다. 빠른 발화에 있어서 Ss의 증가는 Sd 감소의 영향이, Sr의 감소는 Max의 감소가 주된 요인으로 분석되었다. 남녀의 비교에서는 Sr, Ss, Sp(첫째와 마지막 피치포인트 간의 피치 차이), Sd, MaxNr(정규화 된 최댓값)과 MinNr(정규화 된 최솟값)

모두에서 여성이 남성보다 더 큰 값을 보였다. 효과크기 면에서, Ss는 Sd보다는 Sp에 의해 더 큰 영향을 받는 것으로 나타났다. 그러므로 여성 억양 변화의 폭이 남성보다 크기 때문에 여성의 Sp가 남성보다 크고 이로 인해 여성의 Ss가 남성보다 큰 것으로 볼 수 있다.

발화 방식(빠른 발화, 보통 발화)과 두 화자 집단(남녀)의 상호작용을 정리하면 여성이 남성보다 Ss가 더 급하고 Sr이 더 큰데 이러한 변화의 정도는 발화 속도가 증가하면서 여성이 남성보다 더 심해진다. 여성이 남성보다 더 큰 운율 변화를 보이는 이유를 포화효과로 설명하면, 중앙값과의 변위가 더 큰 여성의 발화가 남성의 발화 보다는 목표 값과의 간격이 더 크고 이로 인해 발화 속도에 더 민감하게 반응할 여지가 더 크기 때문으로 추정된다.

본 연구에서 나타난 바와 같이, 발화 속도 증가에 따라 문장 별로 피치 최솟값보다는 피치 최댓값이 그리고 남자보다는 여자가 운율변화의 여지가 더 크다는 결과는 음성 신호처리에서 고려할 만한 점 일 것이다. 예를 들어 음성 합성에서는 합성음의 속도를 변화시킬 때 피치 최댓값과 최솟값 그리고 남녀 피치 변화의 가중치를 차별화해서 적용하는 것을 고려해 보자. 그리고 이러한 결과는 화자 인식, 화자 검증, 감정 인식 등과 같이 음성의 피치가 입력 벡터로 자주 쓰이는 인식 시스템의 구축에도 참고가 될 것이다. 이상과 같은 결과가 발화 속도를 자연스럽게 명료하게 변화시키려는 음성 합성의 연구와 발화 속도 증가에 따른 음성 신호 왜곡에 어려움을 겪는 음성 신호 처리 등등의 분야에 조금이나마 도움이 되길 희망한다.

감사의 글

I'd like to express my gratitude to Dr. Beach and Mr. Beach who generously allowed me to use their language lab. I am thankful for their aspiring support and sharing their resources and valuable expertise during this work. 그리고 이 연구가 가능하도록 격려해주고 물심양면으로 뒷받침해 준 아내 김현정에게도 고마움의 마음을 전합니다.

참고문헌

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5:9/10, 341-345.

Byrd, D. (1992). Preliminary results on speaker-dependent variation in the TIMIT database. *Journal of the Acoustical Society of America*, 92(1), 593 - 596.

Campione, E. (2001). *Etiquetage prosodique semi-automatique de corpus oraux : algorithmes et méthodologie*. Thèse de doctorat. Aix-en-Provence: Université de Provence.

Caspers, J. & Van Heuven, V. (1991). Phonetic and linguistic aspects of pitch movements in fast speech in Dutch. *Proceedings of the International Congress of Phonetic Sciences*, Aix-en-Provence, 5, 174-177.

Cummins, F., Grimaldi, M., Leonard, T., & Simko, J. (2006). The CHAINS corpus: Characterizing individual speakers. *Proceedings of SPECOM 2006*, 431 - 435, St. Petersburg, RU.

Daly, D. & Warren, P. (2001). Pitching it differently in New Zealand English: Speaker sex and intonation patterns. *Journal of Sociolinguistics*, 5(1), 85-96

Fitch, J., & Holbrook, A. (1970). Modal vocal fundamental frequency of young adults. *Archives of Otolaryngology*, 92, 379-382.

Fitzsimons M, Sheahan N, & Staunton H. (2001) Gender and the integration of acoustic dimensions of prosody: implications for clinical studies. *Brain Language*, 78, 94 - 108.

Forsell, M. (2007). *Acoustic correlates of perceived emotions in speech*. MS Thesis, KTH, Royal Institute of Technology, Stockholm, Sweden.

Fougeron, C. & Jun, S. (1998). Rate effects on French intonation: prosodic organization and phonetic realization. *Journal of Phonetics*, 26, 45-69.

Graddol, D., & Swann, J. (1983). Speaking fundamental frequency: Some physical and social correlates. *Language & Speech*, 26, 351-366.

Hirst, D. & Robert E. (1993). Automatic modelling of fundamental frequency using a quadratic spline function. *Travaux de l'Institut de Phonétique d'Aix*, 15, 71-85.

Hirst, D., Cristo, A. & Espesser, R. (2000). Levels of representation and levels of analysis for intonation. in M. Horne (ed) *Prosody : Theory and Experiment*. Kluwer Academic Publishers, Dordrecht. 51-87.

Hudson, A., & Holbrook, A. (1981). A study of the reading fundamental vocal frequency of young black adults. *Journal of Speech & Hearing Research*, 24, 197-200.

Jun, S. (1993) *The phonetics and phonology of Korean prosody*. Ph.D. dissertation, The Ohio State University.

Kohler, K. (1983) F0 in speech timing. *Arbeitsberichte des Instituts für Phonetik des Universität Kiel*, 55-97.

Kohler, K. (1986). Parameters of speech rate perception in German words and sentences: Duration, F0 movement and F0 level. *Language & Speech*, 29, 115-139.

Ninness, B. & Henriksen, S. (2008). Time-scale modification of speech signals, *Transactions on Signal Processing, IEEE*, 56(4), 1479-1488.

Oh, J. H. (2014a). A Study of methods of standardization for Korean intonation curve. *Korean Linguistics*, 62, 395-420.
(오재혁 (2014a). 한국어 억양 곡선의 정규화 방안에 대한 연구. *한국어학*, 62, 395-420.)

Oh, J. H. (2014b). A study of intonation curve slopes in Korean spontaneous speech. *Phonetics and Speech Sciences*, 6(1), 21-30.
(오재혁, (2014b). 자유 발화 자료에서 나타나는 한국어 억양 곡선의 기울기 특성에 대한 연구. *말소리와 음성과학*, 6(1), 21-30.)

Pallett, D., Fiscus, J., Fisher, W., Garofolo, J., Lund, B. & Przybocki, M. (1994). WSJ-CSR Benchmark Test Results, *ARPA Spoken Language Systems and Technology Workshop*, March 1994.

Quene, H. (2007). On the just noticeable difference for tempo in speech. *Journal of Phonetics*. 35, 353-362.

Sandra P. & Whiteside, S. (1996). Temporal-based acoustic-phonetic patterns in read speech: Some evidence for speaker sex differences. *Journal of the International Phonetic Association*, 26(1), 23-40.

Shih, C., & Kochanski, G. (2003). Modeling intonation: Asking for confirmation in English, *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 551 - 554.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. & Hirschberg, J. (1992). ToBI: A standard for labeling English prosody. *Proceedings of the International Conference on Spoken Language Processing*, 867-870.

Stoicheff, M. (1981). Speaking fundamental frequency characteristics of nonsmoking female adults. *Journal of Speech & Hearing Research*, 24, 437-441.

Trouvain, J. & Grice, M. (1999). The effect of tempo on prosodic structure. *Proceedings of 14th International Congress of Phonetic Sciences*, San Francisco, 2, 1067-1070.

Vaisseire, J. (1983) Language-independent prosodic features. *In Prosody: Models and measurements*. Berlin: Springer, 53-66.

Wang, D. & Narayanan, S. (2005). Piecewise linear stylization of pitch via wavelet analysis. *Proceedings of Eurospeech 2005*, Lisbon, 3277 - 3280.

Whiteside, S. & Hodgson, C. (2000). Speech patterns of children and adults elicited via a picture-naming task: An acoustic study. *Speech Communication*, 32(4), 267-285.

Whiteside, S. (1995). Temporal-based speaker sex differences in read speech: A sociophonetic approach. *Proceedings of the XIIIth International Congress of Phonetic Sciences*, 3, 516 -

519.

Yi, S., & Kim, S. (2011). A study on low pitch accent produced in different locations in English sentences. *Phonetics and Speech Sciences*, 3(4), 63-70.
(이서배, 김수정 (2011). 영어 문장 내 상이한 위치에 나타난 저성조 피치 액센트 연구, *말소리와 음성과학*, 3(4), 63-70.)

Yi, S. (2014). An acoustical analysis of emotional speech using close-copy stylization of intonation curve, *Phonetics and Speech Sciences*, 6(3), 131-138.
(이서배, (2014). 억양의 근접복사 유형화를 이용한 감정음성의 음향분석, *말소리와 음성과학*, 6(3), 131-138.)

• 이서배 (Yi, So Pae)

창원대학교 영어영문학과
(641-773) 경남 창원시 의창구 사림동
Tel.: 055-540-5466 Fax: 055-540-5465
Cel.: 010-5555-6305
Email: sopaeyi@pusan.ac.kr

부록

CSLU's Phonetically Rich Phrases - 9 sentences in total

- (s01) If it doesn't matter who wins, why do we keep score?
- (s02) Stop each car if it's little.
- (s03) Play in the street up ahead.
- (s04) A fifth wheel caught speeding.
- (s05) It's been about two years since Davey kept shotguns.
- (s06) Charlie, did you think to measure the tree?
- (s07) Tina got cued to make a quicker escape.
- (s08) Joe books very few judges.
- (s09) Here I was in Miami and Illinois.

TIMIT Sentences - 24 sentences in total

- (s10) She had your dark suit in greasy wash water all year.
- (s11) Don't ask me to carry an oily rag like that.
- (s12) A boring novel is a superb sleeping pill.
- (s13) Call an ambulance for medical assistance.
- (s14) We saw eight tiny icicles below our roof.
- (s15) Each untimely income loss coincided with the breakdown of a heating system part.
- (s16) Je thought you argued in favor of a centrifuge purchase.
- (s17) The sermon emphasized the need for armative action.
- (s18) Kindergarten children decorate their classrooms for all holidays.

- (s19) Cory and Trish played tag with beach balls for hours.
- (s20) The frightened child was gently subdued by his big brother.
- (s21) The tooth fairy forgot to come when Roger's tooth fell out.
- (s22) Alice's ability to work without supervision is noteworthy.
- (s23) Special task forces rescue hostages from kidnapers.
- (s24) If Carol comes tomorrow, have her arrange for a meeting at two.
- (s25) Military personnel are expected to obey government orders.
- (s26) Laugh, dance, and sing if fortune smiles upon you.
- (s27) The fish began to leap frantically on the surface of the small lake.
- (s28) The easygoing zoologist relaxed throughout the voyage.
- (s29) Brush res are common in the dry underbrush of Nevada.
- (s30) How much will it cost to do any necessary modernizing and redecorating?
- (s31) Was she just naturally sloppy about everything but her physical appearance?
- (s32) Is a relaxed home atmosphere enough to help her outgrow these traits?
- (s33) The same shelter could be built into an embankment or below ground level.