

부산 지역 청소년 음성의 연령별 특징 변화 분석

백 승 관¹ 노 용 주¹ 윤 종 락¹
¹ 부경대학교 정보통신공학과

Acoustics of Young People's In Busan : Developmental Changes of Spectral Parameters

Back, Sung-Kwan¹ Ro, Yong-Ju¹ Yoon, Jong-Rak¹

¹ Pukyong National University Dept. of Telematics Engineering

E-mail : telepath@mail1.pknu.ac.kr royju@mail1.pknu.ac.k, jryoon@mail.pknu.ac.kr

요약

부산지역 청소년 음성의 지속시간, 피치주파수 포먼트 주파수 특성을 연령별, 성별로 분석하였다. 실제 발음 환경에서의 음성 패턴은 발성화자 개인 및 화자별로 다양하게 변화한다. 이를 모델화 하기 위해서는 다량의 음성 데이터로부터 통계적 방법에 의한 변화 요인별 파라미터 분석이 선행되어야 할 것이다. 실험에 사용된 데이터는 부산지역에 거주하는 청소년(초등학생, 중학생, 고등학생)들이 연령별로 3회 발성한 우화의 일부와 단모음(/아//이//우//에//오/)이다. 실험 결과로부터 얻어진 지속시간, 주파수 특성 변화 패턴을 연령별, 성별로 구분하여 통계적으로 분석한 뒤 이를 정량화 하였다. 실험 결과로부터 부산 지역 청소년 음성의 지속시간, 주파수 특성은 예측된 바와 같이 기 연구된 성인 음성과 많은 차이를 보였으며 이는 부산 지역 방언의 DB 구축 시 설계자가 고려해야 할 기초자료로 활용 될 수 있을 것이다.

1. 서 론

최근의 음성인식 및 합성기술은 인간과 컴퓨터 사이의 의사전달(Human-Computer-Interface)을 위한 주요 기술로서 그 중요성이 크게 부각되고 있으며 이러한 기술로 개발된 시스템의 성능은 시스템이 사용되는 지역의 음운 환경 및 지역 특성을 잘 반영한 음성 DB를 통해 그 성능이 향상 될 수 있을 것이다. 음성 관련 연구에서 음성의 지속시간, 피치 주파수, 포먼트 주파수 등의 음향학적 파라미터는 음성인식시스템이나 음성 합성시스템의 성능을 향상시킬 수 있는 주요 파라미터로 알려져 있다. 이전 연구에서 어린이 음성은 성인음성과 비교하여 높은 피치주파수와 포먼트주파수,

긴 지속시간을 가지며 큰 시간적, 스펙트럼적 변위를 가진다고 연구되어 있다.[1,2,3] 그러나 이러한 연구 결과는 연구 과제에 포함된 전체 음성의 샘플 수 및 발성자의 범위(지역, 연령)가 제한되어 있어 지역 방언 환경에서 성장한 발성자 음성의 전체적인 특성 패턴을 이해하기에는 부족하다. 특히 최근의 음성관련 연구에서 사용하는 음성 DB는 그 주 대상이 성인이므로 이를 탑재한 시스템의 사용자가 지역적 영향을 받는 환경에서 성장한 청소년일 경우에는 시스템의 성능을 저하시키는 요인이 될 수 있다. 본 연구의 목적은 최근에 수집된 음성데이터로부터 발성자의 음운 환경과 관련된 부산 지역 방언을 사용하는 청소년이 발성한 우화 일부에서 분절한 자, 모음의 지속시간 특성과 단모음의 연령별, 성별 주파수 특징 변화를 분석하여 부산 지역 음성 DB 작성에 필요한 기초 자료로 삼기 위함이다. 또한 청소년 음성의 주파수 특징 변화를 연령별, 성별로 관찰함으로써 성인음성으로 변화하는 과정에서의 개인차의 정규화 가능성을 조사하여 DB 작성에 포함될 대상 범위를 선정하는 자료로 활용 하고자 한다.

본 논문에서 다루는 주요 분석 파라미터는 다음과 같다. 부산 지역 청소년음성의 연령별 자, 모음의 지속시간 특징, 연령별, 성별 단모음의 피치 주파수, 포먼트주파수 변화 특성 이다. 분석된 주파수 특성 결과를 확인 하기 위해 연령별, 성별 함수로 현하여 변화 특성을 관찰하였으며 이를 토써 정량화 하였다. 본 논문의 구성은 서론에 이어 2장에서 실험 데이터 획득 및 분석 방법에 대해서 살펴보고 3장에서 실험결과 및 고찰, 4장 결론 및 향후 연구방향 순으로 기술되어 있다.

2. 실험 데이터 및 분석방법

2.1 실험 데이터

본 논문에 사용된 음성 시료는 부산지역에서 태어나 거주하는 청소년들(8세~19세)의 연령별, 성별 음성으로서 전체 240명(남:120,여:120)의 음성을 수집하여 사용하였다. 데이터 수집은 초,중,고등학교 교실 및 상담실에서 행하여 졌으며 데이터 수집은 두 달(60일)간에 걸쳐 녹음하였다. 실험 대상자들이 청소년인 관계로 획득할 시료의 정확한 녹음을 위해 본 실험의 취지 및 발생법에 대해 사전 교육함으로써 평상시 발생 환경을 요구하였다. 지속시간 특징을 분석하기 위한 낭독 문장은 14세~19세 연령의 청소년이 발성한 “바람과 햇님”이라는 우화의 일부분으로 녹음 문장의 전체 음절수는 168개 총 5개의 문장으로 구성되어 있고 주파수 변화 특징 분석을 위한 단모음은 /아//이//우//이//오/ 를 3회 낭독 시켜 녹음하였다. 사용된 마이크는 B&K의 Type 4130이며 DAT를 사용하여 실시간으로 녹음하였다. 전체 음성 중 녹음의 질이 떨어지거나 잘 못 녹음되거나 잡음환경에 과도하게 노출된 자료 (낭독문장 : 14명, 단모음 : 34명의 음성(남:11,여:23))은 연구 대상에서 제외시켰다.

2.2 분석방법

2.1의 과정을 통해 얻어진 녹음 자료 중 2번째 낭독한 음성을 주로 하여 20KHz 샘플링 주파수, 16-BIT로 양자화 하여 분석에 사용하였다. 양자화 된 화자별 실험 데이터의 분절을 위해 사용한 S/V는 Praat[4]으로 시간 파형, 에너지 파형, 영-교차율, 포먼트 특성, 피치 등 변화를 고려하여 자동으로 레이블링 되게 프로그램하였다. 지속시간을 얻기 위한 데이터의 경우 수동 분절된 결과와 자동 분절된 결과와의 차이가 화자에 따라 15%~43%의 큰 차이를 보여 수동 분절된 결과로 분석을 행하였다. 음성의 주파수 특성은 발생시간 내에서 다양한 형태로 변화하는데 제안한 분석 파라미터의 정확한 값을 얻기 위해서는 음성 데이터의 시작점과 끝점의 설정이 중요한 요인이 된다. 자동 분절된 음성데이터의 시작점과 끝점 위치와 수동으로 분절된 위치와의 차이를 확인 하기 위해 자동분절된 결과 중 연령별로 2개의 샘플을 선정해 수동 분절한 음성과 비교한 차이는 평균 36ms,준편차 42 ms로 나타났다. 이는 연령별, 성별에 따른 경향보다는 녹음 시 부과된 잡음의 영향으로 인한 개인적 편차가 큰 것으로 비교되었다. 또한 분절 위치의 차이가 본 연구에 미치는 영향은 기본주파수

F0에는 차이점을 확인 할 수 없었고 포먼트주파수 F2, F3에 큰 영향을 미쳤다. 본 연구에서 제시된 F2,F3 주파수는 수작업으로 진행한 결과이다. 추후 음운학적 지식에 기반한 자동분절법에 관한 연구가 뒤따라야 할 것으로 생각된다. 실험데이터의 기본주파수(F0)와 포먼트 주파수(F1-F3)를 추정하기 위해 암스테르담 대학의 음성학 연구소에서 제작한 Praat 프로그램과 Matlab 코드로 작성된 프로그램을 사용하였다. 각각의 음성 파형은 16KHz로 다운샘플링 후 10ms의 Hamming 윈도우를 5ms 간격으로 중첩 시켜 처리하였고, Pre-emphasis factor(dB/Octave) 0.94, 16차수 LPC로 분석 하였다. 분석된 기본주파수와 포먼트 주파수의 수치 결과는 각각 5개의 단모음*12년연령*성별로 인덱싱 한 후 파일로 저장 하였다.

3. 실험결과 및 고찰

3.1 연령별 자, 모음의 지속시간

음소의 고유 지속시간을 측정하기 위해 수집된 데이터로부터 각 단어의 선두 음절에 위치한 모음 중 인접 음소의 영향을 최소화하기 위해 /자음+모음/ 환경으로 이루어진 음절을 분절하여 사용하였다. 모음의 고유 지속시간 특징에서 성별에 따른 지속시간은 특이점을 찾지 못해 논의에서 제외 시켰다. 그림. 1은 대,음소 /아/의 전체 데이터에 대한 연령별 평균 지속시간이고, 1은 분석 모음 전체에 대한 연령별 평균 지속시간이다. 음소의 분석 대상에서 제외된 음소는 분절 조건에 합당한 음소를 획득 할 수 없어 제외 시켰다.

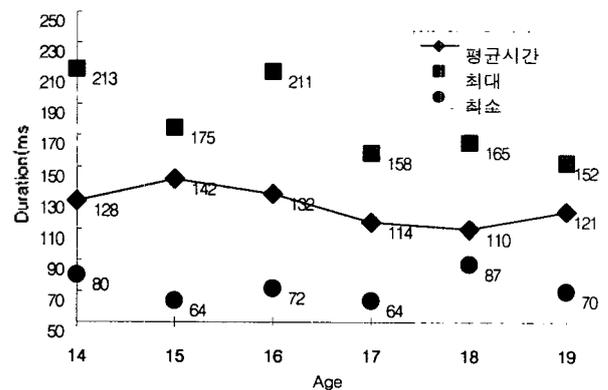


Fig 1. Averaged-vowel(/a/) duration

그림. 1에서 볼 수 있듯이 연령별 모음의 지속시간은 개인적으로 많은 편차를 보이며 그 편차는 15세 이후에 줄어드는 경향을 보인다. 음운학에서 분류한 긴모음(/아/, /애/) / 짧은모음(/이/, /우/) 들의 평균지속시간은

[단위 : ms]

Age \ Vowel	14	15	16	17	18	19	Avg.
아	128	142	132	114	110	121	124
에	123	128	112	121	108	114	117
이	104	118	107	113	131	111	114
오	125	124	122	106	121	117	119
우	113	129	117	118	107	101	114
애	125	113	102	131	106	122	116
의	137	143	141	135	147	131	139

Table 1. Averaged-vowel duration all vowels in each age

120ms/114ms로 분석되었다. 이는 기존 지속 시간 관련 연구에서 제시한 한국어 성인의 평균 모음 지속시간인 158ms/112ms 와 비교해 그 차이가 38ms/2ms로 긴 모음에서 더 큰 차이를 보였다. 이 결과는 긴 모음이 문맥 속에서 짧아지는 특성에 기인하는 것으로 생각된다.[6] 또한 분석 데이터 내의 긴모음과 짧은모음간의 평균 지속시간 차이는 서울 지역 성인의 차이(46ms)와 비교해 볼 때 현재 부산에 거주하는 청소년에게서는 그 차이가 6ms로 긴모음과 짧은모음이 비슷한 지속시간을 갖는 것으로 분석되었다. 2는 자음에 대한 연령별 평균 지속시간이다. 부산 지역 청소년 음성의 자음은 모음과 달리 개인적인 편차는(/ㄱ/ 최대 : 31ms)로 크게 않았으며 연령별에 따른 특성도 볼 수 없었다.

[단위 : ms]

Age \ Con	14	15	16	17	18	19	Avg.
ㄱ	63	59	54	57	62	59	59
ㄴ	78	61	58	80	66	75	70
ㄷ	49	52	51	67	46	55	53
ㄹ	74	84	77	92	81	82	81
ㅁ	61	73	60	78	74	84	71
ㅂ	52	67	54	49	46	58	54
ㅅ	85	59	62	54	52	67	63
ㅆ	81	78	83	61	74	55	72
ㅈ	42	66	63	67	59	70	61
ㅊ	65	84	64	82	78	79	75

Table 2. Averaged-consonant duration in each age

3.2 기본 주파수 변동 특성

음성 데이터로부터 분석된 기본 주파수의 평균값과 편차를 Fig. 3에 나타내었다. 부산 지역 청소년들의 연령별 평균주파수 분석 결과는 남자인 경우에 어린이(13세 이전)음성에서는 기본 주파수의 별다른 변화를 보이지 않다가 13세(F0=235Hz)에서 17세(F0=137Hz) 사이에 두드러진 변화를 보였고 17세 이상에서는 성인 음성

에 근접한 주파수를 보였다. 이는 큰 변화 폭을 보이는 연령층이 변성기 시기에 있음을 알 수 있게 해 주는 결과이다. 또한 어린이(13세 이전) 음성에서의 성별에 따른 차이는 크지 않은 것으로 분석되었다. 여자인 경우에는 전체 연령별 변화 패턴에서의 특이점은 없었고 전체적으로 낮아 졌다. 연령별 분산치는 전체적으로 낮아지는 특성을 보였는데 여자의 경우가 남자 보다 큰 편차를 보였다.

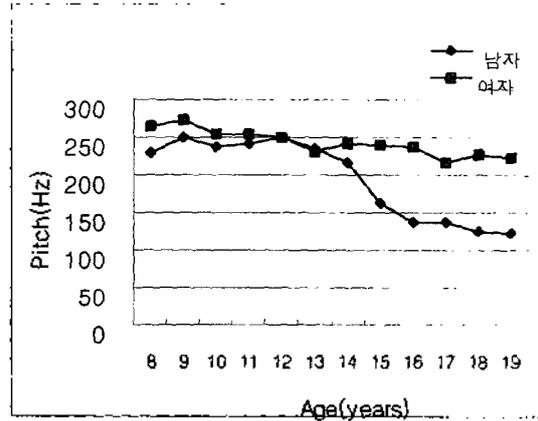


Fig 3. Averaged Fundamental Frequency

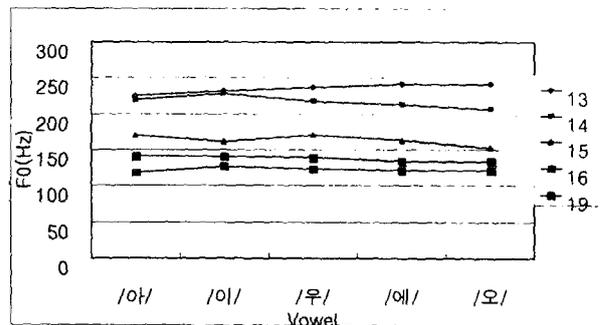


Fig 4. Mean pitch of vowels for male

Fig 4. 은 분석대상 모음별 남자의 평균 주파수이다. 전체 모음에서도 13세~15세 사이의 모음이 큰 주파수 차이를 보였고 그 이후에는 비교적 안정적인 특성을 보였다. 또한 변화 패턴은 초등학교의 경우가 주파수도 높고 변화 폭이 크며 중 고등학교의 경우에는 모음별 차이가 큰 변화 없이 일정하였다.

이상과 같은 분석 결과로부터 부산 지역 청소년들의 기본주파수는 예상된 바와 같이 전체적으로 감소하였고 남자의 경우 변성기에 접어드는 13세~15세에서 큰 변화를 보인다는 것을 알 수 있었고 여자의 경우에는 남자의 경우에서와 같은 큰 변화는 없는 것으로 분석되었다. 표. 3은 분석 데이터 전체에 대한 성별, 연령별 기본주파수 및 포먼트 주파수의 평균값이다.

남자	/이/				/이/				/우/				/에/				/오/			
	F0	F1	F2	F3	F0	F1	F2	F3	F0	F1	F2	F3	F0	F1	F2	F3	F0	F1	F2	F3
M-08	237	952	1695	3621	242	457	1708	3264	231	472	1039	1866	224	644	697	2537	217	561	1060	2089
M-09	245	1081	1260	1860	255	505	1427	2660	258	518	1108	2100	265	792	2202	2930	282	535	846	2476
M-10	253	791	1601	3197	290	504	2461	3146	246	579	1670	2863	137	652	2158	2911	285	531	1004	2188
M-11	240	702	1419	3177	245	430	1630	2998	254	503	955	2474	238	717	2259	3244	232	564	1263	2663
M-12	241	810	1506	3222	247	329	1784	3038	247	545	1634	3017	249	713	1637	2724	267	504	1722	2810
M-13	225	754	1573	3154	232	445	2241	3358	237	664	1789	2983	241	721	2311	2765	241	692	1743	2891
M-14	219	637	1351	2940	228	427	2151	3032	217	517	1413	2844	213	623	2039	2959	206	536	1126	2677
M-15	170	780	1308	2837	160	373	2102	2807	170	434	1058	2628	168	564	1867	2731	152	555	1217	2888
M-16	141	735	1211	2745	141	330	2185	3015	140	610	1398	2668	134	540	1635	2627	134	565	971	2655
M-17	142	760	1213	2573	139	384	2016	2650	137	548	2535	2922	137	584	1784	2560	130	581	1715	2915
M-18	125	750	1325	2706	127	346	2161	2831	125	397	1075	2504	123	555	1610	2601	122	478	913	2700
M-19	118	695	1197	2667	127	335	2195	3002	124	455	1251	2833	121	480	1810	2604	121	460	986	2748

여자	/이/				/이/				/우/				/에/				/오/			
	F0	F1	F2	F3	F0	F1	F2	F3	F0	F1	F2	F3	F0	F1	F2	F3	F0	F1	F2	F3
F-08	260	1187	1940	3678	271	435	1857	3454	267	546	1090	2724	266	629	2331	3575	256	606	1340	2463
F-09	270	887	1591	1734	271	587	750	3003	275	548	688	2262	275	620	1006	2613	272	516	1013	2433
F-10	255	921	1325	3321	279	551	2156	3421	242	661	932	2421	251	754	1451	2587	240	662	1326	2321
F-11	261	823	996	1853	274	517	2034	3170	265	544	976	2301	236	610	1006	2538	232	569	1038	2397
F-12	235	771	1540	2653	243	512	2487	3321	245	580	1730	2910	251	632	2134	2930	252	687	1687	3152
F-13	236	1172	1454	2948	235	443	2236	3098	233	485	1293	2620	232	701	2101	2592	210	668	1044	2940
F-14	238	882	1572	2946	235	439	2084	3000	257	490	1427	2907	234	721	2155	3056	221	475	1304	2709
F-15	243	915	1654	2787	259	472	1811	2788	260	600	1578	2810	230	698	2026	2762	211	718	1187	2908
F-16	240	961	1692	3094	250	503	1511	2800	254	504	1467	2667	227	798	2192	3260	215	719	1479	2984
F-17	218	719	1478	2585	219	437	2256	3032	209	478	1582	2821	215	609	1775	2841	219	544	1588	2458
F-18	220	848	1629	2973	241	450	2196	3068	239	431	1452	2701	225	585	2206	3113	210	578	1228	2690
F-19	222	845	1568	2716	233	446	2075	2959	231	478	1447	2708	220	655	2056	2851	203	516	1093	2598

Table 3. Mean of fundamental frequency and formant frequency Values

3.3 포먼트 주파수 변동 특성

부산지역 청소년들의 포먼트주파수의 변화 특성은 전체적으로 큰 변화 없이 안정적으로 감소하였는데 남자인 경우 14세, 여자의 경우는 13세까지의 포먼트 주파수의 변이가 다소 크게 나타났고 남, 여 모두 15세 이상에서는 그 줄어듦의 경향이 줄었다. 이는 지역 청소년들의 성도 길이의 성장 패턴을 엿 볼 수 있는 결과로 성인 남성과 비교해 부산 지역 청소년들의 성도 길이는 15세 까지 성장하여 성인 길이에 근접하는 것으로 판단된다. 또한 남자의 경우에는 포먼트주파수의 분산이 연령에 따라 크지 않음을 알 수 있었고, 여자의 경우에는 13세까지는 동일 연령 그룹 내 변화가 매우 불규칙하게 변화하는 특성을 보이다가 14세 이후에 다른 연령층과 비슷한 특성을 보였다.

4. 결론 및 향후 연구 방향

부산지역 청소년 음성학의 연령별 모음의 평균지속시간은 개인적으로 많은 편차를 보이며 그 편차는 15세 이후에 줄어들음을 확인하였고 자음의 경우는 연령별 변화 특성을 확인 할 수 없었다. 부산 지역 전체 모음의 전반적인 지속시간은 서울지역 성인과 비교해 볼 때 약 17ms 정도 짧게 지속됨을 확인하였다. 기본 주파수 F0의 경우 분석 데이터 중 남자의 경우 13세(F0=235Hz)에서 17세(F0=137Hz) 사이에 두드러진 변화를 보였고 17세 이상에서는 성인 음성에 가까운 패턴을 보였다.

이는 분석 음성이 변성기에 접어든 시기에 채집된 결과로 판단된다. 여자인 경우에는 전체 연령별 변화 패턴에서의 특이점은 없었고 전체적으로 낮아지는 경향을 보였다. 포먼트 주파수는 전체 변화 패턴이 연령에 따라 크지 않고 전체적으로 낮아짐을 확인하였다.

이상의 결과는 부산 지역 방언의 DB 구축 시 설계자가 기본적으로 고려해야 할 사항인 특징 파라미터의 종류, 특징 추출의 위치, 지속시간 정보 등을 결정하는데 중요한 기초자료로써 활용 될 수 있을 것이다. 향후에는 발성자 개인차의 정규화 방법과 성인음성으로의 보정 기술에 관한 연구를 진행해 나갈 계획이다.

참고문헌

1. 이용주 "한국어 단모음의 성별, 연령별 특징변화 및 인식에 관한 연구", 고려대학교, 석사학위논문, 1987
2. 양병근, "남여 화자가 발성한 한국어 단모음의 음향학적 연구", Journal of Acoustical Society of America, 1992.
3. Sungbok Lee, "Acoustics of children's speech :Developmental changes of temporal and spectral parameters", JASA, Vol 105, No. 3, March, 1999.
4. Paul Boersma, "Optimality-Theoretic learning with the Praat program", IFA Proceedings 23: xx-xx.
5. 성철재, "한국어 리듬의 실험음성학적 연구-시간구조와 관련하여", 서울대학교 박사학위 논문, 1995.
6. 김진영, "한국어음성합성을 위한 운율제어규칙, 서울대학교", 박사학위논문, 1994.