

한국어에서의 공손함을 나타내는 운율적 특성에 관한 연구*

고현주(ETRI), 김상훈(원광대), 김종진(원광대)

<차 례>

- | | |
|-------------|------------|
| 1. 서론 | 4. 분석 |
| 2. 시료 및 피험자 | 5. 결과 및 해석 |
| 3. 녹음 | 6. 결론 |

<Abstract>

Prosodic Characteristics of Politeness in Korean

Hyun-ju Ko, Sang-Hun Kim, Jong-Jin Kim

This study is a kind of a preliminary study to develop naturalness of dialog TTS system. In this study, as major characteristics of politeness in Korean, temporal(total duration of utterances, speech rate and duration of utterance final syllables) and F0(mean F0, boundary tone pattern, F0 range) features were discussed through acoustic analysis of recorded data of semantically neutral sentences, which were spoken by ten professional voice actors under two conditions of utterance type - namely, normal and polite type.

The results show that temporal characteristics were significantly different according to the utterance type but F0 characteristics were not.

* Keywords: politeness in Korean, duration time, F0

1. 서 론

현재 국내 많은 한국어 합성기들이 명료성 측면에서는 매우 뛰어난 성능을 보이고 있다. 하지만 자연성 측면에서는 여전히 개선의 여지가 많은 것이 사실이며 특히 감정이나 그와 관련된 요소들에 대한 연구는 활발히 이루어지지 않고 있다.

언어를 통한 의사소통은 문장의 통사적, 의미론적 내용 이상의 것을 포함하는 것으로 화자는 다양한 운율적 특성을 통해 자신의 감정이나 발화태도 등을 표현하며 이것은 보다 자연스러운 합성음 구현의 궁극적 목표라고 할 수 있겠다. 본 연구는 이러한 운율적 특성으로 표현되는 여러 언어외적 요소들에 대한 예비 연구로서 그 중 공손함의 운율적 특성을 알아보려고 한다.

모든 언어 사회에서 공손함을 표현한다는 것은 사회적 관계를 보다 원만히 유지하는 데 매우 중요한 요소 중 하나이다(Ofuka et al., 2000). Nakane(1970)는 일본을 수직적 사회(vertical society)로 표현했는데 이처럼 계급적 관계를 중요하게 여기는 언어 사회에서는 나이, 성별, 사회적 지위로 대변되는 사회 계급적 관계를 유지하기 위해 적절한 공손함의 표현이 매우 중요한 역할을 수행한다고 하였다. 한국 사회도 예외는 아니어서 사회적 관계에 맞는 적절한 공손함의 표현이 매우 중요하게 여겨지는 것이 사실이다. 따라서 한국어 대화체 합성기에서 이러한 공손함이 표현될 수 있다면 자연성은 크게 증진될 것으로 기대된다.

본 연구의 기본 가정은 크게 두 가지다. 첫 번째는 만일 공손함을 나타내는 운율적 특성이 예를 들어 FO Range나 발화 속도처럼 발화 전체에 걸쳐 나타나는 것이라면 이러한 특성을 충분히 내포하고 있는 코퍼스가 필요하다는 것이며, 두 번째는 이와 반대로 운율적 특성이 발화의 어느 특정 부분, 예를 들어 발화 마지막 음절의 장음화 정도나 억양구 패턴 등과 관련된 것이라면 대규모 코퍼스보다는 이러한 특성에 관한 규칙의 보장이 유용하리라는 것이다.

2. 시료 및 피험자

발성 실험의 시료는 두 화자가 전화 상으로 대화하는 형식의 의미적으로 중립적인 문장이 사용되었다. 피험자는 전문 성우 남녀 각각 6명 총 12명이며 이 가운데서 두 화자를 뺀 남녀 각각 5명 총 10명의 데이터를 분석에 사용하였는데, 이 2명의 데이터를 제외한 이유는 이들이 대화의 자연성을 강조하기 위해서 주어진 텍스트를 변형하여 발성하거나 텍스트에는 포함되지 않은 간투사를 지나치게 많이 사용하였기 때문이다. 성우들의 연령대는 20대에서 30대이며 출신지역은 대부분 서울이었으나 경기지역이나 경상도 지역도 있었다. 그러나 전문 성우라는 점을 감안하여 출신 지역에 따른 지역 방언의 영향은 고려하지 않았다.

실험에 사용된 시료는 다음과 같다.

- (1) A: 여보세요?
- (2) B: 여보세요? 음성언어팀이죠?
- (3) A: 네. 맞습니다. 누구십니까?
- (4) B: 네, 저는 음성디비팀의 안민영이라고 합니다. 이영애씨 계시면 바꿔 주시겠습니까?
- (5) A: 지금 자리에 안 계시는데요.
- (6) B: 혹시 어디 가셨는지 알 수 있을까요?
- (7) A: 그건 잘 모르겠는데 메시지 전해드릴까요?
- (8) B: 아니요 괜찮습니다. 나중에 다시 전화 드리겠습니다.

3. 녹 음

녹음은 녹음 시설을 갖춘 녹음 전문 업체에서 실시하였고 표본추출율은 16kHz로 하였다. 최대한 자연스러운 발성을 위해 녹음은 피험자와 실험자가 대화하는 형식으로 하였는데 하나의 역할이 끝나면 역할을 바꾸어 처음과 같은 방법으로 한 번 더 녹음하였다. 녹음 전에 피험자로 하여금 두 가지 발화 형태에 대한 설명을 하였고 피험자의 판단에 의해 일반적 형태와 공손한 형태로 녹음하도록 지도하였다. 각 발화 형태별로 반복 없이 한 번만 녹음하였다.

4. 분 석

분석은 피험자의 데이터만을 대상으로 하였는데 분석 항목은 크게 세 가지로 나누었다. 첫째 발화의 지속시간의 특성을 알아보기 위해 발화 전체의 지속시간, 휴지구간의 지속시간 그리고 발화 마지막 음절의 지속시간을 분석하였다. 여기서 발화란 위의 시료에서 볼 수 있듯이 두 화자 간의 대화에서 한 사람의 발성이 완전히 끝나고 다른 화자의 발성이 시작될 때까지의 단위이며 하나의 발화체는 1개 이상의 문장으로 이루어 질 수 있고 위에서처럼 실험에 사용된 시료는 모두 8개의 발화체로 구성되어 있다. 발화체 전체의 지속시간은 하나의 발화체에서 모든 휴지구간을 제외한 지속시간이며 휴지구간의 지속시간은 문장 내의 휴지구간과 문장 간 휴지구간(발화 (4), (8)의 경우 2개 이상의 문장이 하나의 발화체로 구성되어 있기 때문에)을 모두 포함하고 있다. 마지막 음절의 지속시간을 분석한 이유는 한국어에서 억양구의 마지막 음절은 억양구 경계 톤이 없이는 곳이며 따라서 많

은 정보를 포함하고 있는 부분으로서 그 마지막 음절의 지속시간의 변화가 공손함에 대한 어떠한 정보를 가지고 있을 수 있다는 가정에 의한 것이다. 둘째, F0의 특성을 알아보기 위해 발화의 평균 F0와 F0 Range를 살펴보고 억양구 패턴을 비교해 보았다.

먼저 녹음된 데이터를 음소 레이블링한 후 프로그램을 이용해서 지속시간의 특성과 F0 특성 그리고 발화 속도를 자동으로 구했으며 K-ToBI 기준에 따라 경계톤만을 운을 레이블링하였다. 레이블링 프로그램으로는 한국전자통신연구원 음성정보연구센터에서 자체적으로 개발한 윈도우용 프로그램인 Wineat을 이용하였으며 파형과 스펙트로그램을 동시간화시켜 레이블링하였다.

5. 결과 및 해석

<표 1> 발화 형태에 따른 발화 전체 지속시간(ms.), 휴지구간의 지속시간(ms.), 발화 속도(초당음절수): N은 일반적 발화, P-공손한 발화.

피험자		1	2	3	4	5	6	7	8	9	10	평균차이
발화전체 지속시간 (휴지구간제외)	N	2331	2144	2337	2518	2553	1990	1813	2068	1820	1742	201.2
	P	2559	2384	2344	2617	2633	2162	2128	2260	2008	2233	
휴지구간의 지속시간	N	240	129	291	410	159	342	179	160	275	187	184.7
	P	308	313	371	641	197	500	355	572	488	474	
발화속도 (휴지구간 포함)	N	5.86	6.82	5.82	5.35	5.55	6.84	7.82	7.28	7.53	8.17	
	P	5.31	5.79	5.58	4.74	5.36	6.04	6.55	5.82	6.45	6.18	

<표 1>은 일반적 발화 형태와 공손한 발화 형태에 따른 피험자별 발화 전체 지속시간과 휴지구간의 지속시간의 평균값을 나타낸 것이다. 표에서는 평균값만이 제시되어 있는데 표준편차 값을 제시하지 않은 이유는 분석 대상이 여러 발화(8개)이었기 때문에 표준편차 값이 자체적으로 큰 의미를 가지지 않기 때문이다.

표에서 볼 수 있듯이 정도의 차이는 있지만 전체적으로 일관되게 공손한 발화 형태에서의 발화 지속시간이 일반적인 발화 형태보다 평균 201ms 정도 길게 나타났다. 피험자 3의 경우는 두 발화 형태에 따른 차이가 거의 없었다. 또한 휴지구

간의 지속시간도 공손한 발화에서 일관되게 길게 나타났으며 평균 184ms의 차이를 보였다.

휴지구간의 지속시간은 피험자 별로 큰 차이를 보였는데 최소 68ms 차이(화자 1)에서 최대 412ms 차이(화자 8)를 보였다. 또한 발화 속도를 보면 공손한 발화가 일반적인 발화보다 발화 속도가 느린 것으로 나타났다.

<표 2> 발화 형태에 따른 발화 마지막 음절의 지속시간(ms.)

피험자 발화NO.	1	2	3	4	5	6	7	8	9	10	평균차이
1(N) (P)	255	227	286	208	365	134	131	152	161	137	16.7
	317	249	299	220	346	154	120	164	135	219	
2	340	270	276	357	423	282	289	239	244	199	10.6
	361	254	289	422	393	224	300	247	256	279	
3	280	256	339	362	372	241	248	208	195	232	13.6
	343	277	333	379	384	238	291	227	211	176	
4	276	232	248	255	352	207	236	186	161	185	2.7
	303	228	259	262	334	208	239	200	153	179	
5	243	212	271	298	295	230	174	213	118	103	7.5
	298	188	252	276	325	251	160	184	170	128	
6	184	135	282	170	264	125	101	109	73	58	7.4
	267	133	216	201	226	131	91	113	92	105	
7	174	131	148	201	234	145	86	98	69	51	18.4
	222	147	187	189	216	139	101	111	104	105	
8	239	170	223	201	305	146	151	214	109	106	15.6
	305	196	230	210	303	213	115	214	122	112	

<표 2>는 각 발화별 마지막 음절의 지속시간을 피험자별로 나타낸 것이다. 화자별로 차이가 있기는 하지만 전반적으로 공손한 발화에서의 발화 마지막 음절이 일반적인 발화에서보다 길게 나타나는 것을 알 수 있다. 특히 화자 1은 그 차이가 전체적으로 꽤 크게 나타나는 것을 알 수 있고 화자 2, 3, 5는 그 차이가 크지 않으며 오히려 공손한 발화에서 더 짧은 지속시간을 보이기도 하여 일관성이 떨어지는 것을 알 수 있다. 발화별 평균 차이를 보면 발화 4, 6은 다른 발화에 비해 발화 형태에 따른 지속시간의 차이가 크지 않았다.

<표 3> 발화 스타일에 따른 평균 F0, F0 Range(Hz), ()는 표준편차

피험자		1	2	3	4	5	6	7	8	9	10
Mean F0	N	234 (11)	252 (17)	241 (11)	249 (18)	245 (14)	147 (12)	157 (14)	161 (7)	156 (10)	123 (11)
	P	256 (12)	235 (14)	235 (13)	241 (8)	256 (19)	130 (15)	148 (10)	146 (5)	155 (6)	123 (9)
F0 Range	N	142 (142)	200 (20)	160 (27)	186 (29)	189 (38)	104 (9)	90 (22)	92 (14)	86 (16)	84 (27)
	P	174 (31)	207 (29)	162 (21)	192 (33)	206 (43)	99 (8)	96 (14)	97 (15)	87 (18)	95 (14)

<표 3>은 발화 형태에 따른 화자별 평균 F0와 F0 Range를 보여주고 있는데 <표 1>, <표 2>와는 달리 표준편차가 제시되어 있다. 이것은 F0의 특성은 발화체의 다양성, 즉 발화체를 구성하고 있는 단어의 수나 음절의 종류 등에 위의 지속시간의 특성처럼 큰 영향을 받지 않기 때문이다. 표의 분석 결과처럼 화자 평균 F0의 경우 화자간의 어떠한 일관성을 찾기는 어려웠으며 마찬가지로 F0 Range의 경우에도 두 발화 형태에 따라 일관성을 찾기는 어려웠다.

끝으로 발화 형태에 따른 억양구 패턴에 대하여 살펴보았는데 결과적으로 두 발화 형태에 따라 억양구 패턴에는 특별한 차이를 찾아 볼 수 없었다. 두 발화 형태 모두에서 LH%, H% 등이 모두 나타났으며 한 발화 형태가 특정한 억양구 패턴을 선호하는 경향 등은 나타나지 않았다. 강세구나 억양구의 수에서는 공손한 발화 형태에서의 강세구나 억양구의 수가 일반적인 발화 형태보다는 조금 많게 나타났는데 이는 위에서 공손한 발화체의 전체 지속시간이 일반적인 발화체의 전체 지속시간보다 길게 나타난 점과 관련지어 생각해 보면 어느 정도는 설명이 가능하리라 생각된다. 그러나 분석에 사용된 데이터의 양이 매우 한정되어 있고 또한 시료로 사용된 발화체를 구성하고 있는 단어의 수가 적은 것을 고려할 때 이러한 결과에 대해 어떤 단정적인 결론을 내리기는 충분하지 못하였다.

6. 결 론

본 연구에서는 한국어 대화체 합성기의 자연성 증진을 위한 예비연구의 일환으로서 한국어에서의 공손함을 나타내는 운율적 특성에 대해 살펴보았다.

지속시간 특성을 보면 [3]의 결과와도 비슷한 것으로 발화 속도가 공손함의 중

요한 운율적 특성 중 하나라는 것을 알 수 있다. 기본주파수와 관련한 결과에서는 발화 형태에 따른 어떤 특성을 발견하기가 어려웠다. 하지만 기본주파수와 관련해서 다른 특성들에 대한 고려도 있을 수 있기 때문에 본 연구의 결과만으로는 단정적인 결론을 내리기가 어려웠다. 한 가지 분명한 것은 발화의 지속시간 특성이 어느 정도는 공손함의 운율적 특성으로 작용한다는 것이며, 이러한 특성이 실제 인지적인 측면에서는 어떻게 작용하는지 또한 따라서 어떻게 규칙화 할 수 있는지에 대한 연구가 계속되어야 할 것이다. 본 실험의 한계는 발화 형태를 일반적 발화와 공손한 발화로 구분하였으나 실제 피험자들의 일부 일반적 발화의 경우에도 공손한 발화로 느껴지는 경우가 종종 있었고 따라서 두 발화 형태간에 변별성이 그다지 크지 않았다는 점을 들 수가 있다. 이를 위해서 앞으로의 연구에서는 발성된 시료에 대한 인지테스트 등을 통해 객관성을 확보하고 또한 두 가지 발화 형태뿐만 아니라 추가적으로 무례한 형태와 같은, 좀 더 대조를 보일 수 있는 발화 형태와의 비교도 필요할 것으로 생각한다. 본 실험이 가지는 또 다른 한계성은 실험에 사용된 데이터의 양이 매우 부족하다는데 있는데, 앞으로의 연구에서는 좀 더 다양하고 충분한 시료를 통한 고찰이 필요할 것이다.

참 고 문 헌

- [1] <http://www.linguistics.ucla.edu/people/jun/ktobi/K-tobi.html>
 [2] C. Nakane, *Japanese Society*, Univ. of California Press, Berkeley, CA, 1970.
 [3] E. Ofuka et al., "Prosodic cues for rated politeness in Japanese speech", *Speech Communication*, Vol. 32, pp.199-217, 2000.

접수일자: 2002년 5월 30일

수정일자: 2003년 1월 6일

게재결정: 2003년 1월 27일

▶ 고희주(Hyun-ju Ko)

주소: 570-749 전북 익산시 신용동 344-1 원광대학교

소속: 원광대학교 영어영문학과

전화: 063) 850-7452

FAX: 063) 850-7454

E-mail: visualprog@dreamwiz.com

▶ 김상훈(Sang-hun Kim)

주소: 305-350 대전시 유성구 가정동 161 한국전자통신연구원

소속: 한국전자통신연구원 음성언어정보연구센터 음성DB연구팀

전화: 042) 860-5141

FAX: 042) 860-6436

E-mail: ksh@etri.re.kr

▶ 김종진(Jong-jin Kim)

주소: 305-350 대전시 유성구 가정동 161 한국전자통신연구원

소속: 한국전자통신연구원 음성언어정보연구센터 음성DB연구팀

전화: 042) 860-5759

FAX: 042) 860-6436

E-mail: kimjj@etri.re.kr