

# 한국어 문음성 변환기의 음운지속시간 제어에 관한 연구

## A Study on Segmental Duration Control for the Korean TTS

김인영, 이양희

동덕여자대학교 전자계산학과

E-mail : inyoung@dongduk.ac.kr, yhlee@www.dongduk.ac.kr

### 요 약

자연스러운 한국어의 음성합성을 위해서는 음운의 지속시간의 제어가 매우 중요하다. 본 연구에서는 POW3848 어절에 대한 음성 데이터에 대해 음운 세그먼트, 음운 라벨링, 품사 태깅을 행한 음성 데이터베이스를 구축하여 한국어 음운의 지속시간을 변화시키는 시간 특징을 통계적으로 분석하였다. 이 시간 특징들 중 변화 폭이 큰 요인들을 제어요소로 각 음운의 고유길이를 최대한 배제하고 단지 음운 발생 환경의 영향에 의한 지속시간 변화만을 고려하는 정규화 지속시간에 대한 회귀트리로 한국어 음운 지속시간을 모델화 하였다. 제안된 음운 지속시간 모델을 실시간 제어 알고리즘으로 구현하여 평가한 결과, 음운 지속시간 예측오차의 88% 정도가 25ms 이내 이었고 예측치와 관측치 간의 다중 상관계수는 0.92 정도로 평가되어, 제안된 모델의 타당성이 입증되었다.

### 1. 서론

본 논문에서는 일반적인 음운지속시간의 변화 요인을 분석하고 정도 높은 지속시간 제어 규칙을 생성하기에 충분한 POW(Phonetically Optimized Word) 3848 어절 음성데이터에 대해 음운 세그먼트, 음운 라벨링, 품사 태깅을 행한 음성데이터 베이스를 구축하고, 이 데이터베이스를 사용하여 일반화된 음운 지속 시간을 통계적 방법에 의해 자동적으로 모델화한다. 여기에서 사용되는 데이터베이스에는 음운이 편중되어 있지 않고, 또한 음운 환경 및 문법정보를 포함하고 있어 일반화된 음운 지속시간

모델화에 충분한 데이터라고 생각된다. 이러한 음성 데이터를 통해서 한국어 음운의 지속시간을 변화시키는 시간 특징을 통계적으로 살펴보고, 분석된 시간 특징들 중 변화 요인이 큰 특징들을 제어요소로 사용하여 한국어 음운 지속 시간을 모델화한다. 이 때 규칙의 최적화 및 규칙의 자동생성을 위해 회귀 트리 모델을 사용하였다. 이 방법에서 통계적 방법으로는 제어요소 간의 의존 관계를 표현할 수 있다. 2절에서는 음운지속시간 제어규칙을 생성하기 위한 음성 데이터베이스에 대해 기술하고, 음운지속시간의 변화요인에 대해 분석결과를 기술한다. 또한 각 음운의 고유 길이를 최대한 배제하여 음운 환경에 의해서만 영향을 받는 지속시간 변화요인을 적용하기 위해 각 음운의 Zscore에 대한 회귀 트리 모델화에 대해 설명한다. 3절에서는 세그먼트지속시간 예측 알고리즘을 구현하며, 그 알고리즘 타당성 평가에 대해 기술한다.

### 2. 음운 지속시간 제어규칙 생성

#### 2.1 음성 DB 구축

통계적인 방법으로 일반화된 규칙을 생성하기 위해서는 다양한 경우를 포함하는 많은 양의 데이터가 요구된다. 보다 일반적이며 정교한 음운지속시간 제어 모델을 생성하기 위하여, 다양한 음운 환경을 고려하는 충분히 많은 자연음성을 분석하여 음운지속시간 변화에 영향을 미치는 요인을 추출하여야 한다. 음운 지속시간을 변화시키는 요인을 크게 음운 환경과 문법적인 요인으로 나누어 생각할 수 있다. 음성DB는 음운환경에 의한 변화를 분석하기

위해 음운 단위로 나누어 세그먼트 되어야 하고 문법적인 요인에 의한 변화를 분석하기 위해 품사 태깅이 필요하다. 본 연구에서는 POW(Phonetically Optimized Word) 3848 어절을 남성, 여성화자 각각 1명이 발성한 음성 데이터를 음운별로 세그먼트, 음운 라벨링 및 음운별 품사 태깅하였다. 음성 데이터 베이스는 다음과 같은 정보를 포함한다.

- 음운레벨 세그먼트(시간정보)
- 음운 라벨링(음운 기호)
- 어절에 대한 음운레벨 품사 태깅(품사)

## 2.2 음운지속시간 변화에 대한 통계적 분석

구축된 음성 데이터베이스를 이용하여 세그먼트의 지속시간 변화에 크게 영향을 미치는 요인을 분석하기 위하여 통계적인 방법을 사용한다. 각 음운의 고유지속 시간의 영향을 배제시킨 순수한 음운환경에 의한 세그먼트 지속시간 변화 요인을 분석하기 위하여 본 연구에서는 식 (1)과 같은 Zscore를 사용하여 음성DB내 세그먼트들의 지속시간을 정규화 하였다.

$$Z_{ip} = (X_{ip} - M_p) / SD_p \quad \text{----- 식 (1)}$$

$Z_{ip}$  : 음운 p의 지속시간에 대한 i번째 세그먼트의 관측치

$M_p$  : 음운 p의 지속시간에 대한 평균치

$SD_p$  : 음운 p의 지속시간에 대한 표준편차

### 1) 조음 양식에 의한 영향

조음양식이나 조음 위치가 유사한 음운들의 정규화 분포를 고려하여 정규화 음운 지속시간이 유사한 분포를 갖는 음운들을 22종류로 분류하였다. 이렇게 분류된 조음 양식이나 조음 위치에 따른 음운들의 정규화 지속시간이 유사하기 때문에 음운의 지속시간 변화에 유사하게 영향을 미친다.

### 2) 음절 유형에 의한 영향

우리말의 음절유형을 다음의 8가지로 분류할 수 있다. (CV, CVC, V, VC, SV, SVC, CSV, CSVC). 음절유형에 따라 분석한 결과 모음의 지속시간이 변화한다. 종성이 있는 CVC, VC, SVC, CSVC의 음절 유형에서 모음의 음운 지속시간이 짧아지는 경향이 있고 초성 없는 음절유형

에서는 모음지속시간이 평균보다 길어지는 경향이 있다. 또한 초성을 갖으며 종성이 없는 음절 유형에서는 평균 모음 지속시간을 유지하는 경향이 있다.

### 3) 어절 내 음절수에 의한 영향

음운수에 의한 모음의 지속시간 변화는 음운의 수가 1-5개 까지는 음운수에 따라 음운지속시간 변화가 크게 감소하며, 6-21개까지는 음운수에 따른 지속시간이 작게 변화하는 경향이 있다. 이러한 현상은 어절 내 음운 수가 음운 지속시간의 변화에 큰 요인이 되고 있음을 나타낸다.

### 4) 어절 내 음절 위치에 의한 영향

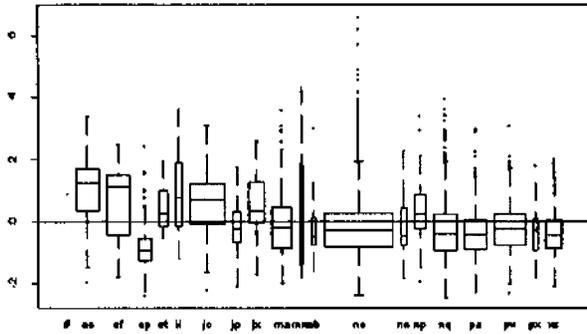
7개 음운을 갖는 어절에 있어서 모음의 위치에 따라 모음의 지속시간이 규칙적으로 변화하지 않았다. 따라서 이러한 요인은 지속시간 제어 요소로는 부적합하다.

### 5) 앞, 뒤 인접음운에 의한 영향

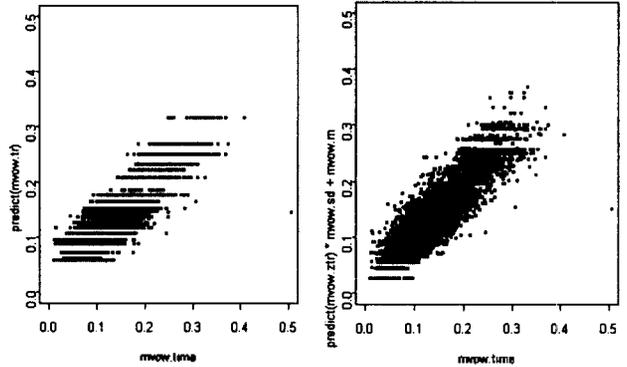
음운 발생 환경에 따른 지속시간 변화를 분석하기 위하여 고려하는 앞 음운과 뒤 음운의 영향을 분석한다. 이때 전후 음운은 조음 양식과 위치로 분류하여 분석한다. 앞 음운에 의한 모음의 정규화 지속시간의 분포와 뒤 음운에 의한 모음의 정규화 지속시간의 분포를 비교해보면 뒤 음운의 영향이 앞 음운의 영향보다 크게 나타남을 알 수 있었다. 따라서 뒤 음운의 영향은 음운지속시간 제어 규칙을 생성하는데 가장 중요한 특징 요소가 된다. 특히 V\_C, U\_C, BDG, ktpc들의 앞에 오는 음운(모음)의 지속시간이 짧아지고, 또한 PAUSE(#)의 앞에 오는 음운(모음) 즉 마지막음운지속시간은 길어짐을 알 수 있다.

### 6) 문법적 요소에 의한 영향

문법적인 요인에 대한 음운지속시간을 분석하기 위하여 한국어의 품사를 다음 [그림 1]과 같이 19품사로 분류하였다. [그림 1]에 나타난 것과 같이 품사별 모음의 지속시간 변화를 분석한 결과는 다음과 같다. 품사가 ec (연결어미), ef (종결어미), jc (격조사), np (지시대명사)의 경우 음운지속 시간이 길어지는 반면, ep (선어말어미), 고유명사를 제외한 내용어인 경우에는 음운 지속시간이 짧아지는 경향이 있다.



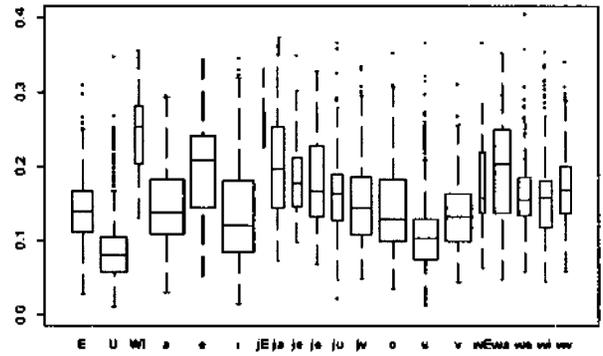
[그림 1] 품사별 정규화 음운지속시간



[그림 2] 예측치와 관측치간의 분포

### 2.3 정규화 회귀트리 구현

음운의 고유지속 시간의 영향을 배제시키고 순수한 음운환경에 의한 세그먼트의 지속시간을 예측하기 위하여 각 세그먼트의 지속시간을 Zscore로 정규화 하였다. 각 세그먼트의 정규화 지속시간은 음운의 고유지속시간을 제외한 음운환경에만 의존하여 변화하게 된다. 따라서 지속시간 변화 요인에 의해서만 분류되기 때문에 보다 정교하게 예측이 가능하도록 정규화 지속시간에 대해 회귀트리로 모델화 하였다.



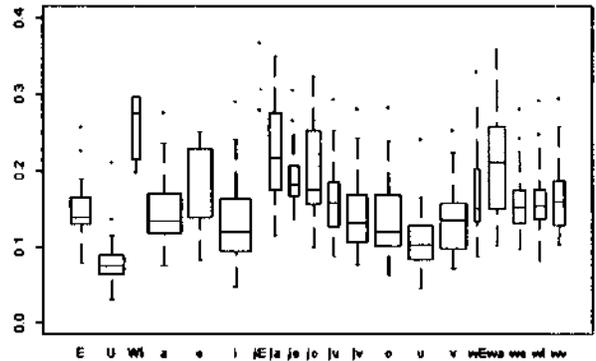
[그림 3] 모음의 음운지속시간 분포

### 2.4 생성된 규칙의 평가

생성된 규칙의 타당성을 확인하기 위하여 관측치와 예측지간의 오류정도를 평가하고 오류 분석을 행한다. 이 때 관측치와 예측지간의 오류정도를 다중상관 계수로 평가한 결과는 [표 1]과 같다.

[표 1]

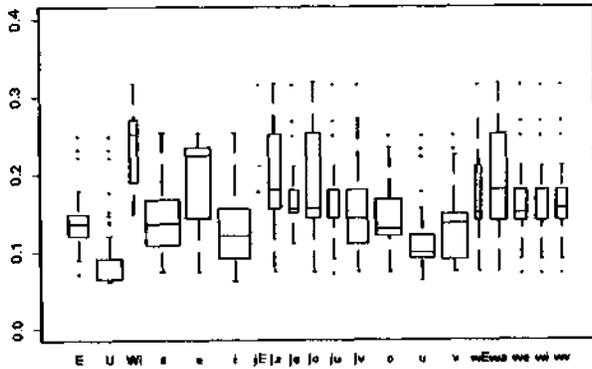
	지속시간 회귀트리	정규화된 회귀트리
다중상관 계수	0.921(남성) 0.904(여성)	0.929(남성) 0.906(여성)
예측오차 25ms 이내	87.0%(남성) 82.8%(여성)	88.2%(남성) 84.6%(여성)



[그림 4] 정규화 지속시간 회귀트리에 의한 예측 모음지속시간 분포

정규화된 지속시간에 의한 예측과 지속시간에 의한 예측의 오류 정도를 평가하기 위하여 다음[그림 2]를 살펴 보면 정규화된 지속시간의 예측(오른쪽)이 일반 지속시간에 의한 예측(왼쪽)보다 정교하고 오류 정도의 차가 더 작게 나타난다.

모음의 세그먼트의 관측치의 분포와 일반 회귀트리에 의해 예측치의 분포, 정규화된 세그먼트 지속시간에 의한 지속시간 예측치의 분포는 이들 [그림 3,4,5]에서 보듯이 정규화 트리에 의한 예측이 관측치 분포와 유사하게 나타남을 알 수 있다.



[그림 5] 비 정규화 음운지속시간 트리에 의한 예측 모음 지속시간 분포

### 3. 세그먼트 지속시간 예측 알고리즘

세그먼트 지속시간 예측 시스템은 다음과 같이 크게 3 부분으로 구성된다.

- 1) 특징 파라미터 변환 시스템 : 입력으로 부터 음소열과 각 음소의 문법적 정보 즉 품사정보를 받아서 각 음소의 조음 양식 및 위치, 그 음소가 포함된 어절내 음운의 수, 어절내 위치, 음절 유형, 인접 음운 정보, 품사 정보등을 출력한다.
- 2) 음운 지속시간 제어규칙 : 1)로부터 출력된 정보들을 회귀트리의 특징 제어 요소로 사용하여 음운 지속 시간 제어 규칙을 생성한다. 즉 음소의 발생 환경에 대한 특징 열을 입력으로 하여 회귀트리로부터 정규화 지속시간을 예측한다.
- 3) 세그먼트 지속 시간 예측 : 예측된 정규화 지속시간으로부터 세그먼트 지속시간으로 변환에서는 정규화 지속시간을 다음 식 (2)와 같은 방법에 의하여 세그먼트 지속시간을 구한다.

$$DURip = Mp + (Zip \times SDp) \text{ ---- 식 (2)}$$

여기에서,

DURip : p음운의 i번째 세그먼트의 예측 지속시간

Mp : 음운p의 평균지속시간

Zip : p음운의 i번째 세그먼트의 예측 정규화 지속시간

SDp : 음운p의 지속시간의 표준편차

세그먼트 지속 시간 예측 알고리즘 구현의 프로그램 실행 예는 다음 [그림 6]과 같고 처리 시간도 실시간 처리에 충분한 정도로 판정 되었다. 청취 실험 결과, 어절 단위에서는 자연스러운 음성이 합성 되었으나 문장 단위 음성 합성에서는 다소 부자연스러운 부분이 나타났다. 이는 어절 단위 음성 DB로부터 구한 제어 규칙이기 때문이라 생각된다.

Consonant	Segment	Duration																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
U	VC	1	24	48-C	2	24	V_C	3	24	48-C	4	24	V_C	5	24	48-C	6	24	V_C	7	24	48-C	8	24	V_C	9	24	48-C	10	24	V_C	11	24	48-C	12	24	V_C	13	24	48-C	14	24	V_C	15	24	48-C	16	24	V_C	17	24	48-C	18	24	V_C	19	24	48-C	20	24	V_C	21	24	48-C	22	24	V_C	23	24	48-C	24	24	V_C	25	24	48-C	26	24	V_C	27	24	48-C	28	24	V_C	29	24	48-C	30	24	V_C	31	24	48-C	32	24	V_C	33	24	48-C	34	24	V_C	35	24	48-C	36	24	V_C	37	24	48-C	38	24	V_C	39	24	48-C	40	24	V_C	41	24	48-C	42	24	V_C	43	24	48-C	44	24	V_C	45	24	48-C	46	24	V_C	47	24	48-C	48	24	V_C	49	24	48-C	50	24	V_C	51	24	48-C	52	24	V_C	53	24	48-C	54	24	V_C	55	24	48-C	56	24	V_C	57	24	48-C	58	24	V_C	59	24	48-C	60	24	V_C	61	24	48-C	62	24	V_C	63	24	48-C	64	24	V_C	65	24	48-C	66	24	V_C	67	24	48-C	68	24	V_C	69	24	48-C	70	24	V_C	71	24	48-C	72	24	V_C	73	24	48-C	74	24	V_C	75	24	48-C	76	24	V_C	77	24	48-C	78	24	V_C	79	24	48-C	80	24	V_C	81	24	48-C	82	24	V_C	83	24	48-C	84	24	V_C	85	24	48-C	86	24	V_C	87	24	48-C	88	24	V_C	89	24	48-C	90	24	V_C	91	24	48-C	92	24	V_C	93	24	48-C	94	24	V_C	95	24	48-C	96	24	V_C	97	24	48-C	98	24	V_C	99	24	48-C	100	24	V_C	101	24	48-C	102	24	V_C	103	24	48-C	104	24	V_C	105	24	48-C	106	24	V_C	107	24	48-C	108	24	V_C	109	24	48-C	110	24	V_C	111	24	48-C	112	24	V_C	113	24	48-C	114	24	V_C	115	24	48-C	116	24	V_C	117	24	48-C	118	24	V_C	119	24	48-C	120	24	V_C	121	24	48-C	122	24	V_C	123	24	48-C	124	24	V_C	125	24	48-C	126	24	V_C	127	24	48-C	128	24	V_C	129	24	48-C	130	24	V_C	131	24	48-C	132	24	V_C	133	24	48-C	134	24	V_C	135	24	48-C	136	24	V_C	137	24	48-C	138	24	V_C	139	24	48-C	140	24	V_C	141	24	48-C	142	24	V_C	143	24	48-C	144	24	V_C	145	24	48-C	146	24	V_C	147	24	48-C	148	24	V_C	149	24	48-C	150	24	V_C	151	24	48-C	152	24	V_C	153	24	48-C	154	24	V_C	155	24	48-C	156	24	V_C	157	24	48-C	158	24	V_C	159	24	48-C	160	24	V_C	161	24	48-C	162	24	V_C	163	24	48-C	164	24	V_C	165	24	48-C	166	24	V_C	167	24	48-C	168	24	V_C	169	24	48-C	170	24	V_C	171	24	48-C	172	24	V_C	173	24	48-C	174	24	V_C	175	24	48-C	176	24	V_C	177	24	48-C	178	24	V_C	179	24	48-C	180	24	V_C	181	24	48-C	182	24	V_C	183	24	48-C	184	24	V_C	185	24	48-C	186	24	V_C	187	24	48-C	188	24	V_C	189	24	48-C	190	24	V_C	191	24	48-C	192	24	V_C	193	24	48-C	194	24	V_C	195	24	48-C	196	24	V_C	197	24	48-C	198	24	V_C	199	24	48-C	200	24	V_C	201	24	48-C	202	24	V_C	203	24	48-C	204	24	V_C	205	24	48-C	206	24	V_C	207	24	48-C	208	24	V_C	209	24	48-C	210	24	V_C	211	24	48-C	212	24	V_C	213	24	48-C	214	24	V_C	215	24	48-C	216	24	V_C	217	24	48-C	218	24	V_C	219	24	48-C	220	24	V_C	221	24	48-C	222	24	V_C	223	24	48-C	224	24	V_C	225	24	48-C	226	24	V_C	227	24	48-C	228	24	V_C	229	24	48-C	230	24	V_C	231	24	48-C	232	24	V_C	233	24	48-C	234	24	V_C	235	24	48-C	236	24	V_C	237	24	48-C	238	24	V_C	239	24	48-C	240	24	V_C	241	24	48-C	242	24	V_C	243	24	48-C	244	24	V_C	245	24	48-C	246	24	V_C	247	24	48-C	248	24	V_C	249	24	48-C	250	24	V_C	251	24	48-C	252	24	V_C	253	24	48-C	254	24	V_C	255	24	48-C	256	24	V_C	257	24	48-C	258	24	V_C	259	24	48-C	260	24	V_C	261	24	48-C	262	24	V_C	263	24	48-C	264	24	V_C	265	24	48-C	266	24	V_C	267	24	48-C	268	24	V_C	269	24	48-C	270	24	V_C	271	24	48-C	272	24	V_C	273	24	48-C	274	24	V_C	275	24	48-C	276	24	V_C	277	24	48-C	278	24	V_C	279	24	48-C	280	24	V_C	281	24	48-C	282	24	V_C	283	24	48-C	284	24	V_C	285	24	48-C	286	24	V_C	287	24	48-C	288	24	V_C	289	24	48-C	290	24	V_C	291	24	48-C	292	24	V_C	293	24	48-C	294	24	V_C	295	24	48-C	296	24	V_C	297	24	48-C	298	24	V_C	299	24	48-C	300	24	V_C	301	24	48-C	302	24	V_C	303	24	48-C	304	24	V_C	305	24	48-C	306	24	V_C	307	24	48-C	308	24	V_C	309	24	48-C	310	24	V_C	311	24	48-C	312	24	V_C	313	24	48-C	314	24	V_C	315	24	48-C	316	24	V_C	317	24	48-C	318	24	V_C	319	24	48-C	320	24	V_C	321	24	48-C	322	24	V_C	323	24	48-C	324	24	V_C	325	24	48-C	326	24	V_C	327	24	48-C	328	24	V_C	329	24	48-C	330	24	V_C	331	24	48-C	332	24	V_C	333	24	48-C	334	24	V_C	335	24	48-C	336	24	V_C	337	24	48-C	338	24	V_C	339	24	48-C	340	24	V_C	341	24	48-C	342	24	V_C	343	24	48-C	344	24	V_C	345	24	48-C	346	24	V_C	347	24	48-C	348	24	V_C	349	24	48-C	350	24	V_C	351	24	48-C	352	24	V_C	353	24	48-C	354	24	V_C	355	24	48-C	356	24	V_C	357	24	48-C	358	24	V_C	359	24	48-C	360	24	V_C	361	24	48-C	362	24	V_C	363	24	48-C	364	24	V_C	365	24	48-C	366	24	V_C	367	24	48-C	368	24	V_C	369	24	48-C	370	24	V_C	371	24	48-C	372	24	V_C	373	24	48-C	374	24	V_C	375	24	48-C	376	24	V_C	377	24	48-C	378	24	V_C	379	24	48-C	380	24	V_C	381	24	48-C	382	24	V_C	383	24	48-C	384	24	V_C	385	24	48-C	386	24	V_C	387	24	48-C	388	24	V_C	389	24	48-C	390	24	V_C	391	24	48-C	392	24	V_C	393	24	48-C	394	24	V_C	395	24	48-C	396	24	V_C	397	24	48-C	398	24	V_C	399	24	48-C	400	24	V_C	401	24	48-C	402	24	V_C	403	24	48-C	404	24	V_C	405	24	48-C	406	24	V_C	407	24	48-C	408	24	V_C	409	24	48-C	410	24	V_C	411	24	48-C	412	24	V_C	413	24	48-C	414	24	V_C	415	24	48-C	416	24	V_C	417	24	48-C	418	24	V_C	419	24	48-C	420	24	V_C	421	24	48-C	422	24	V_C	423	24	48-C	424	24	V_C	425	24	48-C	426	24	V_C	427	24	48-C	428	24	V_C	429	24	48-C	430	24	V_C	431	24	48-C	432	24	V_C	433	24	48-C	434	24	V_C	435	24	48-C	436	24	V_C	437	24	48-C	438	24	V_C	439	24	48-C	440	24	V_C	441	24	48-C	442	24	V_C	443	24	48-C	444	24	V_C	445	24	48-C	446	24	V_C	447	24	48-C	448	24	V_C	449	24	48-C	450	24	V_C	451	24	48-C	452	24	V_C	453	24	48-C	454	24	V_C	455	24	48-C	456	24	V_C	457	24	48-C	458	24	V_C	459	24	48-C	460	24	V_C	461	24	48-C	462	24	V_C	463	24	48-C	464	24	V_C	465	24	48-C	466	24	V_C	467	24	48-C	468	24	V_C	469	24	48-C	470	24	V_C	471	24	48-C	472	24	V_C	473	24	48-C	474	24	V_C	475	24	48-C	476	24	V_C	477	24	48-C	478	24	V_C	479	24	48-C	480	24	V_C	481	24	48-C	482	24	V_C	483	24	48-C	484	24	V_C	485	24	48-C	486	24	V_C	487	24	48-C	488	24	V_C	489	24	48-C	490	24	V_C	491	24	48-C	492	24	V_C	493	24	48-C	494	24	V_C	495	24	48-C	496	24	V_C	497	24	48-C	498	24	V_C	499	24	48-C	500	24	V_C	501	24	48-C	502	24	V_C	503	24	48-C	504	24	V_C	505	24	48-C	506	24	V_C	507	24	48-C	508	24	V_C	509	24	48-C	510	24	V_C	511	24	48-C	512	24	V_C	513	24	48-C	514	24	V_C	515	24	48-C	516	24	V_C	517	24	48-C	518	24	V_C	519