

A Pilot Study of the Eloquence of President Obama from an Acoustic Point of View: “World without Nuclear Weapons” Speech in Prague, Czech Republic

Kazuaki ICHIZAKI
Miyazaki Gakuen Jr. College, Japan
email: ichizaki@mwjc.ac.jp

Obama 大統領の「核なき世界」演説の発音分析 —予備研究—

市 崎 一 章

要旨

Obama 大統領は米国の歴代大統領の中でもとりわけ雄弁という評判である。世界を感動させた「核なき世界」演説にその特長を求めた。スピーチ自体は、政治家の演説としては標準的な発話速度で、使用するポーズの長さも頻度も標準値に近かったが、テレプロンプターの活用もあって言い淀みはほとんどなく、使用ポーズのすべてが silent pause となった。他方、氣息群間の発話時間には差が生じた。つまり氣息群には、その長さおよび構成音節数にかなりの幅があり、それが伝達情報量の圧縮率にメリハリを生じさせる結果に至っている。また、特に強調したい氣息群に先行する氣息群の末尾項では持続時間を伸長させることで一気に発話速度を落とすという特徴が認められた。まったく言い淀むことなく日常会話をする調子で、短いフレーズと長いフレーズを混ぜながら、情報の焦点を直前項の持続時間伸長によって一気に強調する。これこそ Obama 大統領の雄弁さの秘訣であろう。

1. Introduction

The United States has produced forty-four presidents since it was founded and it has been said that the current President, Barak H. Obama, is one of the most eloquent speakers in her history. To be called a great speaker, the speaker is desired to be elegant, lofty, careful, friendly, reassuring, persuasive, energetic, dynamic, logical, strict – the list of modifiers would be a long one. Excluding the contents of speech, most of which seem to be made by speech writers, such a positive impression can be made by speech manners of the speaker including gross and fine motor skills, eye attention, ad lib remarks, voice quality, and some

other phonetic features. Some people point out that President Obama is an expert user of the teleprompter which appeared in the middle of the sixties. Why is he outstanding among quite a few speakers who have utilized it for half a century? As long as his eye attention was observed during the speech, which lasted twenty-seven and a half minutes, there were few cases that he surely glanced at the teleprompter and he always looked at the audience right and left. In that sense he would be an expert speaker.

Apart from such behavioral manners, what other skills does President Obama have when making speeches? From an acoustic point of view, this paper is going to deal with objective features which could be measured by sound analyzing software such as rate of speech, pause, pitch, pitch change, sound pressure level, and possibly, tone, and tries to point out what phonetic features President Obama has in his speech. Amateur speakers could be somewhat eloquent by imitating and adopting the phonetic features of his speech.

2. Investigation and Analysis

2.1 speech material

Rather than the inaugural address in Washington D. C. with an audience of two million, the “World without Nuclear Weapons” speech carried out in Prague, Czech Republic, with an audience of twenty thousand must have been more significant for world peace. The latter resulted in a Nobel prize for peace although President Obama had put nothing into practice at that moment. Some of the most impressive paragraphs were extracted out of the speech on nuclear disarmament and adopted as the speech material for this paper. They were separated into tone groups and breath groups as follows for the convenience of making tables listed at the end of this paper.

Just as we stood for freedom || in the 20th century, || we must stand together for the right of people everywhere to live free from fear | in the 21st century. # And as nuclear power; || and as a nuclear power, as the only nuclear power | to have used a nuclear weapon, || the United States has a moral responsibility to act. || We cannot succeed in this endeavor alone, || but we can lead it, | we can start it. ||

So today, || I state clearly | and with conviction || America’s commitment to seek || the peace and security of a world without nuclear weapons. # I’m not naïve. || This goal will not be reached quickly — || perhaps not in my lifetime. || It will take patience and persistence. || But now we, too, | must ignore the voices || who tell us that the world cannot change. || We have to insist, || “ Yes, we can.”

| : the end of a tone group without pause

|| : the end of a tone group with pause

: the end of a tone group with cheers and applause

___: wrong utterance

2.2 procedure

The speech material was extracted from a “You tube” item titled ‘Obama Prague Speech: A World Without Nuclear Weapons¹⁾,’ which originally seemed to have been broadcast through “Sky News,” on the internet. The material, which was played by Adobe flash player, was processed with sound analyzing software called *SUGI Speech Analyzer* (Sugitō, 2000). When analyzing the speech material, it was divided, after each tone group with pause or with cheers and applause, into twenty breath groups as shown in Table 1. The items measured and identified were: the duration of a breath group and that of each pause; number of syllables within a breath group, rate of speech (syllables/second), prominent item(s) in a breath group, the item with the highest pitch within a breath group, and the item(s) with maximum sound pressure level within a breath group.

3. Results and Remarks

The results of the measurement were shown in Table 1 and 2, in the appendix at the end of this paper. The results of the items identified and measured will be discussed here.

3.1 division of tone groups and tone of nucleus

Watanabe (1994) listed a fragment of speech done by N. Kinnock, the leader of British Labour Party, in 1992 as a successful speech and one of the typical styles for public speech. Kinnock divided his speech into as many tone groups as possible and so they became shorter ones without any chaff and with only carefully selected words. Twenty-nine out of thirty tone groups ended with falling tone, many of which were very high falling, with decreasing voice quantity within a tone group for a modulation for effect. Watanabe reported that as a result Kinnock’s speech sounded rhythmical and reassuring. In the case of the speech material here of President Obama, there seems to be only one place, i. e. after “*America’s commitment to seek*,” that looks clearly a redundant division of tone group for a regular conversation. He used nine non-falling tones at the end of twenty-five tone groups. They were six levels, two rise-falls and one rising. In that sense Obama’s speech would have sounded fluent and the fluency might have yielded an impression of his eloquence in the mind of the audience. If the repetition of (high) falling tone has a reassuring effect, the audience could not have helped concentrating themselves on what the speaker would say thereafter with non-falling tones.

Nuclear syllables were identified by the author and listed as the items in the leftmost column of Table 1. As for tones of twenty-five nuclear syllables, there were eighteen fallings, three risings (*only*, *clearly*, *patience*), two rise-falls (*fear*, *too*), and two levels (*used*, *seek*). Although it is difficult to establish a standard ratio of tones for public speech without any samples to compare with, Obama’s usage of tone for nucleus was not so partial as Kinnock’s plethora of falling tone for nuclei. Thus, ideal public speech seems to have some styles in the

use of tone groups, their ultimate tones, and tones for nuclei.

3.2 duration of utterance (pause and speech)

The total duration of the speech material was 85,212 msec. The items of the leftmost column in Table 1 are the first two words of the breath groups interrupted by a pause. President Obama used no fillers such as “ah-” or “erm-” in the speech material and all pauses he used became silent pauses. Table 2 shows the duration of pauses inserted after each breath group listed in Table 1. Cheers and applause of the excited audience were yielded in pause 3 and pause 12 and President Obama waited for a while until the ripple disappeared, and so the two pauses were much longer than the other pauses. The total duration of pauses reached up to 43,052 msec. As the two pauses including cheers and applause were recognized as outliers in statistics, averaged duration of the other pauses, which was 1,123 msec., was tentatively adopted as a pause at the two places, which resulted in 21,337 msec. altogether for all pauses. As a result the total duration of the speech material was 63,497 msec. which consisted of 42,160 msec. breath groups and 21,337 msec. pauses, which means President Obama’s speech consisted of 66.4% utterance and 33.6% pause.

The preceding studies on silent pause, *Goldman-Eisler (1961)* and *Crystal & Davy (1975)*, reported that most silent pauses used in conversation, discussion or speech were within two seconds and there were few whose duration lasted for much longer than two seconds. Excluding the two longest pauses with cheers and applause, fifteen pauses out of seventeen lasted less than two seconds. *Mori, Higgins, and Kiritani (2005)* examined the frequency of silent pause using London Lund Corpus of Spoken English (which included a wide variety of monologues and dialogues such as TV and radio broadcasts, political and legal statements, telephone and face-to-face conversations) as materials and reported that a silent pause appeared after every 6.6 words on the average. If 140, the total number of words of the speech material of this paper, is divided by 6.6, the average number of words which follows after a pause, it is 21.2, from which we can see that, if President Obama used pauses with the average frequency, the extract from his speech would contain 20.2 (21.2-1) pauses. The number of pauses actually used by Obama was 19, which was almost the same as the averaged value in various types of speech. Apart from his value, however, it should be noticed that there was a great difference in the duration between the longest breath group and the shortest breath group and in the number of syllables between the breath group with the most syllables and the breath groups with the fewest syllables. He uttered for more than 5.2 sec. at the longest and for 0.7 sec. at the shortest with one breath. The breath group with the most syllables was “*we must stand together for the right of people everywhere to live free from fear in the 21st century*” consisting of twenty-eight syllables and the breath of people everywhere to live free from fear in the 21st century” consisting of twenty-eight syllables and the breath groups with the fewest were “So today” and “Yes, we can” consisting of

three syllables each. Comparing the former with “Yes, we can,” the former was uttered with a pace of less than 1.3 times as fast as the latter in the rate of speech although the former had 9.3 times as many syllables as the latter had.

3.3 rate of speech

There are not a few preceding studies on the rate of speech. Focusing on U. S. presidents’ speeches, *Heffner (1950)* reported the values of two one-minute samples including pauses of President Franklin D. Roosevelt and those of President Harry S. Truman more than half a century ago. The averaged values of the former were 97.5 words/min, 2.3 syllables/sec, and 6.2 phonemes/sec and those of the latter were 155 words/min, 4.0 syllables/sec, and 10.3 phonemes/sec.

Taniguchi (1988) measured some extracted passages (including pauses) of inaugural addresses of three U. S. presidents, John F. Kennedy, James E. Carter, and Ronald W. Reagan. The averaged values²⁾ of Kennedy were 109.5 words/min, 2.8 syllables/sec, and 6.8 phonemes/sec, the values³⁾ of Carter were 105.6 words/min, 2.4 syllables/sec, and 6.1 phonemes/sec, and the values of Reagan⁴⁾ were 168.7 words/min, 3.6 syllables/sec, and 9.5 phonemes/sec. *Taniguchi* adopted a speech in a debate, a few public speeches, a few speeches in TV commercials, and some speeches in movies besides the presidents’ inaugural addresses mentioned above, and at the same time carried out a perceptual experiment to know how fast such speeches were perceived by listeners. In the conclusion, it was reported that public addresses were mostly spoken at a slower rate than the normal conversational speed, ranging from 90 words/min to 170 words/min, or from 2.3 syllables/sec to 4.2 syllables/sec, or 5.2 phonemes/sec to 9.7 phonemes/sec, while the acceptable range of normal conversational speech was from 110 words/min to 230 words/min, centering 180 words/min, or from 2.7 syllables/sec to 5.8 syllables/sec, centering 4.5 syllables/sec, or from 6.0 phonemes/sec to 12.8 phonemes/sec, centering 10.0 phonemes/sec⁵⁾. Also it was reported as a part of the conclusion that President Reagan spoke at the perfectly normal conversational rate from a viewpoint of perception. That could be one of the reasons why he was said to be a great speaker.

The results of the present paper were obtained with pauses as processed in the previous section 3.1. The total duration of President Obama was 63,497 msec. with 140 words, 198 syllables, and 494 phonemes altogether. The rates of his speech were 132.3 words/min, or 3.12 syllables/sec, or 7.78 phonemes/sec., which means that Obama spoke faster than Roosevelt, Carter, or Kennedy and slower than Truman or Reagan. Considering the range of public addresses reported in *Taniguchi*, President Obama seems to have used the ideally normal speed for them.

Looking at the column of syllables/sec in Table 1, most of the values ranged from 4.0 to 6.0. It is remarkable that the lowest value was found at the penultimate breath group, “*We*

have to insist,” preceding the last breath group “*Yes, we can,*” which is surely the most known catchphrase of President Obama and is the information focus of this utterance. This very low rate of the penultimate breath group was realized by lengthening the last item “insist.” The catchphrase actually sounded so impressive and the response of the audience peaked at this moment. Slowdown just before the message he wants to emphasize seems to be a skillful speech technique of President Obama.

3.4 prominence, pitch, and sound pressure

In the right-hand side of the items having prominent syllables in Table 1, the items having highest pitch, the items having greatest pitch change, and the items having the maximum sound pressure level were also listed to examine the relationship between nucleus and such phonetic features. Concerning sound pressure level shown in *SUGI Speech Analyzer*, as it is a unit used for convenience when referring to intensity, it will be described as sound pressure hereafter.

There were some places where fundamental frequency was not extracted enough to be examined because the vocal cords are apt to vibrate aperiodically especially at the end of a breath group as indicated in *Maekawa (1996)*. Therefore, some cells in the column titled “highest pitch” or “greatest pitch change” in Table 1 were blank (indicated by ϕ).

Twenty-five items having nucleus were recognized for twenty breath groups and nineteen of the items were situated at the end of tone groups. Seventeen items out of the twenty-five were substantiated as the items having nucleus by either the greatest value of pitch, pitch change, or sound pressure: four items were supported by all three values, four items by the value of pitch and pitch change, three items by pitch change and sound pressure, three items by pitch change only, two items by pitch only, and one item by pitch and sound pressure. Five items out of the other eight items seemed to have acquired their prominence from a comparatively longer duration of their prominent syllables although the durations were not listed in Table 1. And the remaining three items seemed to have acquired their prominence from great values of pitch change which were not listed, either. These results clearly show that prominence is made from high pitch, wide pitch change, great sound pressure, or long duration.

Apart from such phonetic features, it was remarkable that President Obama inserted a pause of 105msec⁶⁾ after “*without*” in “*the peace and security of a world without nuclear weapon*” in the second paragraph. The pause, together with the maximum sound pressure on “*(with)out*”, made “*without*” the most prominent item within the breath group.

4. Conclusion

Concerning the frequency and the duration of pause, President Obama made a speech in Prague just like people’s daily conversation, except for a slightly slower rate than normal

conversation (which was a standard value for public speeches). Also, few fillers were used throughout his speech lasting for nearly half an hour, which may have been partially supported with the aid of a teleprompter set on either side of him. His division of tone groups was also neither more nor less than that in ordinary conversation. As a result his utterance without faltering or stammering seems to have sounded fluent and natural.

On the other hand there were a few particular features of President Obama. The breath groups he used varied widely in duration and number of syllables although there was not much difference in the rate of speech. He appropriately put large information or small information into a breath group. Could it be said that he is good at modulating the quantity of information within a breath group according to the circumstances? Another feature is the rate reduction at the breath group preceding the breath group to emphasize, or the sound lengthening at the item just before the emphasized breath group. Skillful usages of such phonetic features seem to be President Obama at his best.

The target material of this paper was just a fragment of a speech made by President Obama. More analysis with more materials would offer us some other features of his utterance and that is surely an assignment for the author to have as the next step.

Notes:

1) The speech material was extracted from

<http://www.youtube.com/watch?v=QKSnlSXjj2s&feature=related>

2) As *Taniguchi* adopted three extracts, whose duration times were 24.54sec, 18.26sec, and 7.52sec, as the material to measure and the averaged values listed here were calculated according to the ratio of duration..

3) The duration of the extract was 28.99 sec.

4) The duration of the extract was 19.56 sec.

5) *Taniguchi* showed rate of speech using syllables/min and phonemes/min and they were shifted to syllables/sec and phonemes/sec.

6) The pause was too short to be recognized as a boundary of breath group.

References

[1] D. Crystal and D. Davy. *Advanced Conversational English*. London: Longman, 1975.

[2] F. Goldman-Eisler. "The rate of changes in the rate of articulation," *Language and Speech*, no. 4, pp. 171-174, 1961.

[3] R-M. S. Heffner. *General Phonetics*. Univ. of Wisconsin Press, 1950.

[4] K. Maekawa. "Onnkyoonnseigaku no tennbo," *The Bulletin of the Phonetic Society of Japan*, no. 211, pp.12-19, 1996.

- [5] Y. Mori, A. Higgins, and S. Kiritani. "Functions and locations of long silent pauses in English discourse," *JACET Bulletin*, no. 41, 91-110, 2005.
- [6] M. Sugitō. *SUGI Speech Analyzer* CD-ROM, Yokohama: Animo, 2000.
- [7] M. Taniguchi. "Average range and limitations of English speech velocity," *The Bulletin of Chikushi Jogakuen University Junior College*, vol. 23, pp. 55-82, 1988.
- [8] K. Watanabe. *Eigo intonation ron*. Kenkyusha, 1994.

APPENDIX

Table 1 Data of phonetic features of the utterance excluding pauses

	duration	no of syll	syll/sec	prominent item(s)	highest pitch	greatest pitch change	max pressure
<i>Just as...</i>	1,475	7	4.75	freedom	freedom	ϕ	Just, stood, freedom
<i>in the...</i>	1,007	7	6.95	20th	20th	20th	20th
<i>we must...</i>	5,218	28	5.37	fear, 21st	21st	21st	right
<i>And as...</i>	1,331	6	4.51	power	nuclear	power	nuclear
<i>and as a...</i>	5,240	21	4.01	only, used	only	only	as, (1st)nuclear, weapon
<i>the United...</i>	2,596	16	6.16	States	United	States	United, States
<i>We cannot...</i>	2,320	12	5.17	alone	alone	alone	alone
<i>but we...</i>	1,739	9	5.18	lead, start	(lead ϕ), start	(lead ϕ), start	start
<i>So today...</i>	709	3	4.23	today	today	today	today
<i>I state...</i>	2,193	9	4.10	clearly, conviction	clearly	clearly, (conviction ϕ)	state
<i>America's commitment...</i>	1,873	9	4.81	seek	commitment	America's	America's
<i>the peace...</i>	3,515	16	4.55	without	security	without	without
<i>I'm not...</i>	980	4	4.08	native	native	(native ϕ)	not
<i>This goal...</i>	1,584	8	5.05	quickly	goal	quickly	not
<i>perhaps not...</i>	1,401	7	5.00	lifetime	perhaps	perhaps	not
<i>It will...</i>	2,224	9	4.05	patience, persistence	patience	ϕ	take
<i>But now...</i>	2,802	10	3.57	too	ϕ	ϕ	now
<i>who tell...</i>	1,655	9	5.44	change	tell	ϕ	tell
<i>We have...</i>	1,589	5	3.15	insist	insist	insist	We
<i>"Yes, we..."</i>	709	3	4.23	can	we	can	Yes

note: ϕ means there obtained no data to be discussed.

Table 2 Duration of pauses (ms)

pause 1	468
pause 2	967
pause 3 (cheers & applause)	10,534
pause 4	2,294
pause 5	1,116
pause 6	1,602
pause 7	192
pause 8	2,342
pause 9	1,243
pause 10	713
pause 11	525
pause 12 (cheers & applause)	13,423
pause 13	1,309
pause 14	866
pause 15	1,462
pause 16	1,331
pause 17	1,239
pause 18	993
pause 19	433