

GE3104 Data is Beautiful: Visualization in the Humanities

Term project: paper

Visualization of Apple product keynotes before and after 10 years

Name: CHAU Yin Shing, Tony
(yschau8)

I. Introduction

This term project aims to compare two speeches presented on-stage keynotes from one of the largest electrical brand in the world, Apple inc. by visualizing the data of two transcripts of the keynotes in Apple inc. product conference at 2007 and 2017.

The outcome of the project is to dig out similarities and differences between two keynotes in Apple inc. product conference at 2007 and 2017. By the visualization, we are expected to see the differences of the speech style from two different keynotes, one from Steve Jobs, ‘the soul of Apple inc.’ and one from the new CEO Tim Cook with his team leaders.

Three main objectives will be stated for the distinction between two keynotes. In this project, two methods plus an online tool will be applied for analyzing the data in different ways. The advantages and disadvantages of using the methods to project different aspects of findings will be discussed in the paper.

II. Data

In this project, two datasets which are from different year of Apple inc. product conference are selected for visualization.

The first dataset is the transcript presented by Steve Jobs, which is from iPhone release keynote at the MacWorld 2007. As the only speaker of the event, Steve Jobs introduces new functions of the first iPhone. The number of word count of his transcript is about 13000 words.

The second dataset is the transcript presented by Tim cook and 11 team leaders, which is from Apple inc. product conference of WWDC 2017 keynote. The speakers were managed to introduce different parts of the new Apple products’ functions. The number of word count of the transcript is about 20400 words.

III. Methodology

By adopting three methods, three main objectives to be visualized are as follow:

1. Modal verbs (can, could, may, might, must, shall, should, will, would)
2. Adjectives
3. Words related to Apple products/services

These three objectives are the word choices of the speakers in the keynotes 2007 and 2017, which are useful for attaining certain outcome of the project.

Three methods for the analysis the transcripts with two different keynotes are as follow:

- A. Graphical Comparison
- B. Keywords & Collocation
- C. Online tool ‘Link’: Word network

A. Graphical Comparison

One of the best method to visualize mass quantity of data is converting numbers into graphs. For each objective, the specified words would be categorized and compounded into useful charts. Tendency and pattern of graphs are easily to be observed and compared systematically. Microsoft Excel will be used as an excellent software for the converting process.

To optimize the effectiveness of visualization, both transcripts of the datasets will be evenly spilt into 10 parts. The information from part-to-part comparison between two keynotes could be obtained easier.

B. Keywords & Collocation

To visualize word-to-word and text-to-text relation, it requires a systematic method to compute the data in terms of frequency as the basis of analysis. For a further step, the assisting tool AntConc. allow users to rank the words in different categories and collect information of word association with the users' preferences.

AntConc. only requires the plain transcript for the analysis. Steps for accurate analysis would be explained further in Discussion part.

C. Online tool 'Link': Word network

The online tool 'Link' provide visualization of word network. Unlike visualizing social network, the word network shows relation between words. Method B and C also visualize relation between words, and in method C, it takes all relation between words as counted.

(see Appendix graph 1.a & 1.b for result)

This method can contribute to information on sum of word relation in selected scope of context. The graphs could provide an overview of relation between all distinct words.

IV. Discussion

A. Graphical Analysis

1. *Modal verbs*

Table 1.a and 1.b (see Appendix for all tables) present the graphs the frequencies of the first objective – modal verbs appear in the keynotes. In these two graphs, since the transcripts of the datasets do not have enough sample of two modal verbs (shall, must), these two modal verbs are not considered to be the compare items.

Findings:

Table 1.a.1 and 1.b.1 show the similar patterns of using the modal verb 'can' between two keynotes. In the middle part of both keynotes, the speakers tend to use more modal verbs 'can', hence there is a rise for the bars from part 3 or 4 of the keynotes.

From part 6 to 8 of the keynotes, a gap indicated a small or big drop in using the modal verb in their presentations.

In Table 1.a.2, the red circle indicates the slight increase of modal verbs (except ‘can’ and ‘might’) at the end of keynote 2007. The end of keynote in 2017 does not follow in the same way (shown in table 1.b part 10), shaping a unique feature of keynote 2007 from the difference between the end of two keynotes.

2. Adjectives

To investigate the preference of using different adjectives by the speakers from two keynotes, the graphs have to be in the scope of most frequent appeared adjectives they have been using in the presentation, which of one of the method to visualize the speech styles of two keynotes. To efficiently select five most frequent appeared adjectives, the function of wordlist by AntConc. has been applied. The following list is the result to be analyzed:

	iPhone Keynote at 2007	WWDC 2017 Keynote
Rank of the most frequent appeared adjectives	1. Little 2. Great 3. Cool 4. New 5. Best	1. New 2. Great 3. Incredible 4. Excited 5. Best

Findings:

Table 2.a and 2.b show the frequencies of five most frequent appeared adjectives for each of the two keynotes. In table 2.a, the adjectives appeared in keynote 2007 do not have big differences in terms of the frequencies of the words appeared. However, in table 2.b, the graph demonstrates the dominance of the adjective ‘new’ throughout the whole keynote at 2017. As the same adjective ‘new’ appears in both keynotes, the word only ranked the fourth at keynote 2007. Comparing two tables in overall, the adjectives ‘new, great, best’ are very likely to be appeared in the Apple product keynotes.

3. Words related to Apple products/services

To visualize particular words that are related to Apple products and services, main products and services that are introduced in the keynotes have been selected for the graphs. The keywords for the words related to Apple products and services in certain year are as follow:

iPhone keynote at 2007: iTunes, iPod, tv, phone(s)/iPhone

WWDC 2017 keynote: iTunes, iPod, tv, phone(s)/iPhone, iMac, iPad

iMac, iPad had not been introduced at 2007 yet, so these two words will not be included in the table 3.a.

Findings:

Table 3.a and 3.b shows the frequency of keywords appeared in the keynotes.

From the beginning part of keynote 2007, the speaker had kept using the keywords of existing products and services. Since the keynote is the Apple inc. product conference of the first iPhone, starting from part 3, the keyword 'phone(s)/iPhone' has a sudden rise. It indicates that the speaker is introducing the new product. The appearance of the keyword 'phone(s)/iPhone' lasted until the end of the keynote 2007, and most of the other keywords stopped at part 5.

Table 3.b shows similarity and difference comparing table 3.a. Both keynotes had introduced the new products not starting from the beginning of the keynotes. However, the speakers in keynote 2017 had broken down the new series of iMac and iPad into separately two parts from part 3 to 4 and part 7 to part 10. Moreover, the keynote in 2017 seldom mentioned the old products and services throughout the whole keynotes, which implies the company may had changed the marketing strategy in these ten years.

Strengths and Weaknesses of method A:

The result of graphical comparison between two subjects directly reflects quantitative data in graphs. The rise and decline of the frequencies of words have been shown through the bars in different parts of the keynotes. It helps how we interpret the speaker altering his/her speech strategy in different parts of the presentation. By comparing the tables in the same objective, the similarities and the differences of pattern between two tables can be obtained easily from how they performed.

The limitation of the method will be overtly countered in the process of analysis and part of the result. The size of datasets in this project requires much of personal manual work. The method may not suit for larger dataset analysis due to low efficiency. Furthermore, if the selected objective words are not specified words, the graphs could not be done by relying on single software. For example, in second objective, the information about most frequent appeared adjectives are obtained by another software. The result of the graphs could not indicate the relations between words to words in the single subject, so it could not reveal more about the reason of altering the word choices in different part of the speakers' speech by solely looking at rise or decline of frequency of a word.

B. Keywords and Collocation

1. Modal Verbs 'can'

In those modal verbs that had appeared in both keynotes, 'can' had been selected for this part's analysis. It is due to the dominance of the modal word shown in method A.

By method B, we look further of the relation between the modal word ‘can’ and other words in the keynotes.

Picture 1.a and 1.b (see appendix) show the result of the words that are most significantly followed by the modal verb ‘can’ appeared in the keynotes 2007 and 2017. In both keynotes, the word ‘see’ ranked the first following by the modal word ‘can’. The combination of the term ‘can see’ may implied that the speakers were demonstrating something by face-to-face interaction with the audiences so that they kept using the term in the keynotes.

2. Adjectives

Base on the above-mentioned result of 5 most frequent appeared adjectives in method A.2 adjectives, ‘great’ and ‘best’ are at the same rank. This part will include the investigation of how ‘great’, ‘best’ differ in two keynotes in terms of word collocation with noun phrase.

iPhone Keynote at 2007		WWDC 2017 Keynote
Word Collocations		
Stereo	Great	Wealth
Quotes		Uplift
Photographer		Organizations
Features		Unification
Sounds		Shares
Vocal	Best	Stylus
Performance		Reflectivity
Solo		Player
Continuum		Scenes
Player		Displays

The table shows how ‘great’ and ‘best’ differed collocating with words in the keynotes. In keynote 2007, the adjectives are appeared with words that a half related to sound ‘Stereo, Sounds, Vocal, Solo, Player’, while in keynote 2017, the adjectives are bounded with a set of words that are unrelated (except ‘player’ appear again and the relation of ‘scenes’ and ‘displays’. By the transcript of keynote 2017, it is believed that the speakers are using the adjectives based on the speech context, such as introducing the products through stories.

3. Words related to Apple products/services

Through the function of computing the keyness of the words by the assisting tool AntConc., word distinction between two texts can be computed. Keywords of two

keynotes has been listed out with the reference file of another dataset, the list is as follow:

iPhone Keynote at 2007	Keyness	WWDC 2017 Keynote	Keyness
phone	+88.86	iPad	+58.56
iPod	+68.92	ios	+54.51
mail	+61.88	siri	+51.48
tv	+53.23	iMac	+45.42
iTunes	+42.43	App	+38.71

The table above are words that related to Apple products and services selected from the list of keywords. After a ten-year interval, the list shows new products and functions from Apple inc..

Strengths and Weaknesses of method B:

The software AntConc. can visualize relation between words and highlight most important words and frequent words. The multi-function software contributes to visualize a text in different aspects.

For obtaining all the results of analysis, the users only require inserting the plain text. The software automatically computes the data with support of statistics. The users can easily collect the information, since all the data are ranked from highest frequency or most significant words.

Although the software is convenient and friendly for users, it has some weaknesses. The software could not visualize mass quantitative data as in method A. Even though it has the function of concordance plot, the function could not visualize statistic of multi items and exact number of rising. Moreover, the data obtained from the assisting tool AntConc. need to be analyzed which could be too subjective for filtering the data.

V. Conclusion

Through visualization of two keynotes, many similarities could be discovered from the findings. The keynotes mainly varied by how they arrange the sorting of introducing the products, and it affects the result of how the words collocated with the products. With more alternations to visualize the data, we could dig out more information for supporting all kinds of researches.

VI. References

European Rhetoric. (N.A.). *Transcript of the iPhone Keynote by Steve Jobs (MacWorld 2007)*. Retrieved 16th April 2018, from <http://www.european-rhetoric.com/analyses/ikeynote-analysis-iphone/transcript-2007/>.

THE SINGJU POST. (N.A.). *Apple CEO Tim Cook Keynote at WWDC 2017 (Full Transcript)*. Retrieved 16th April 2018, from <https://singjupost.com/apple-ceo-tim-cook-keynote-at-wwdc-2017-full-transcript/>.

Voyant tools. (N.A.). *LINKS*. Retrieved 24th April 2018, from <http://docs.voyant-tools.org/tools/links/>

VII. Appendix



Graph 1.a Result of the word network, iPhone keynote at 2007



Graph 1.b Result of the word network, WWDC keynote 2017

Table 1.a Modal verbs in iPhone keynote at 2007

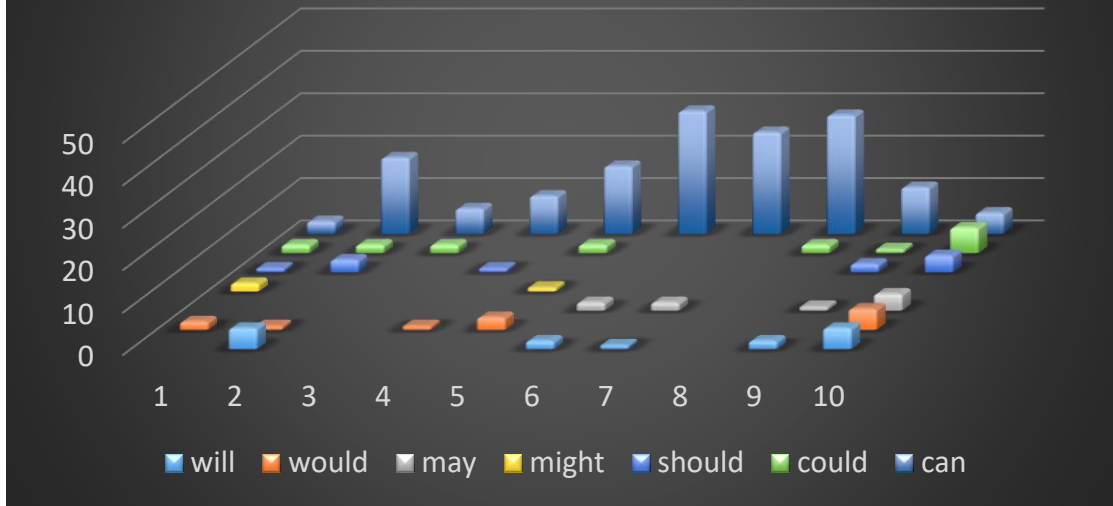


Table 1.a

Table 1.b Modal verbs in WWDC 2017 keynote

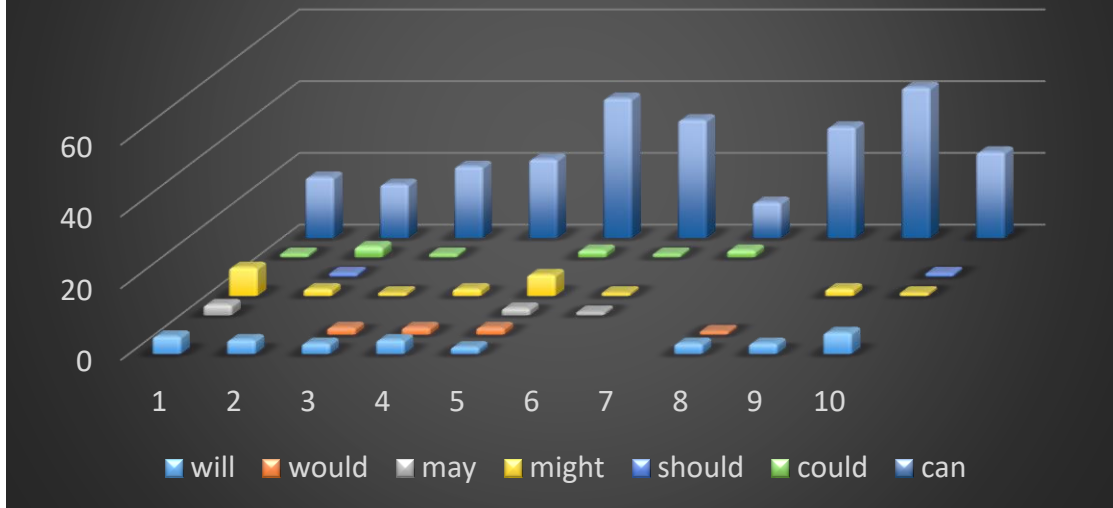


Table 1.b

Table 1.a.1 Modal verbs in iPhone Keynote at 2007

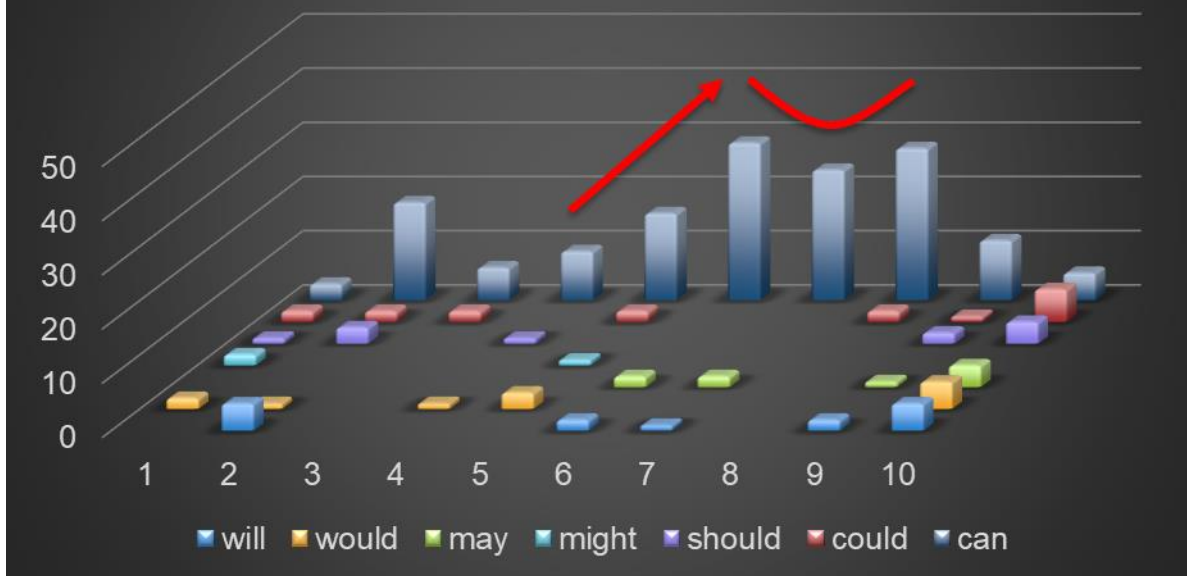


Table 1.a.1

Table 1.b.1 Modal verbs in WWDC 2017 Keynote

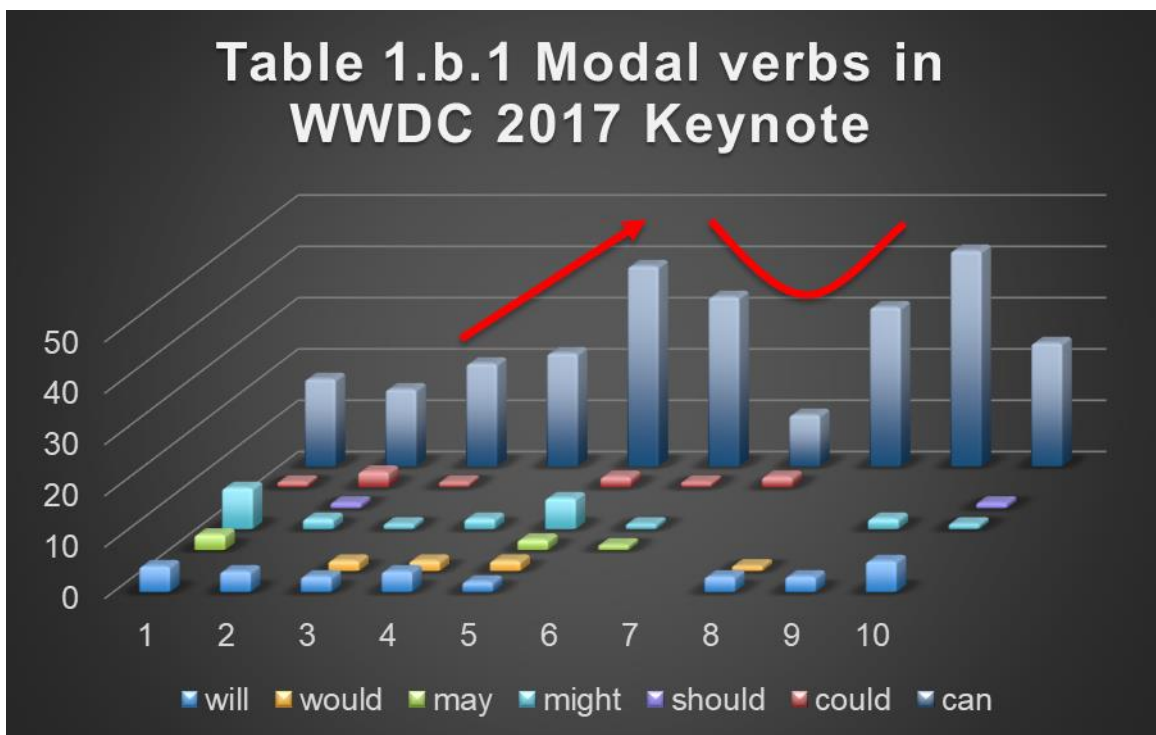


Table 1.b.1

Table 1.a.2 Modal verb in iPhone keynote at 2007

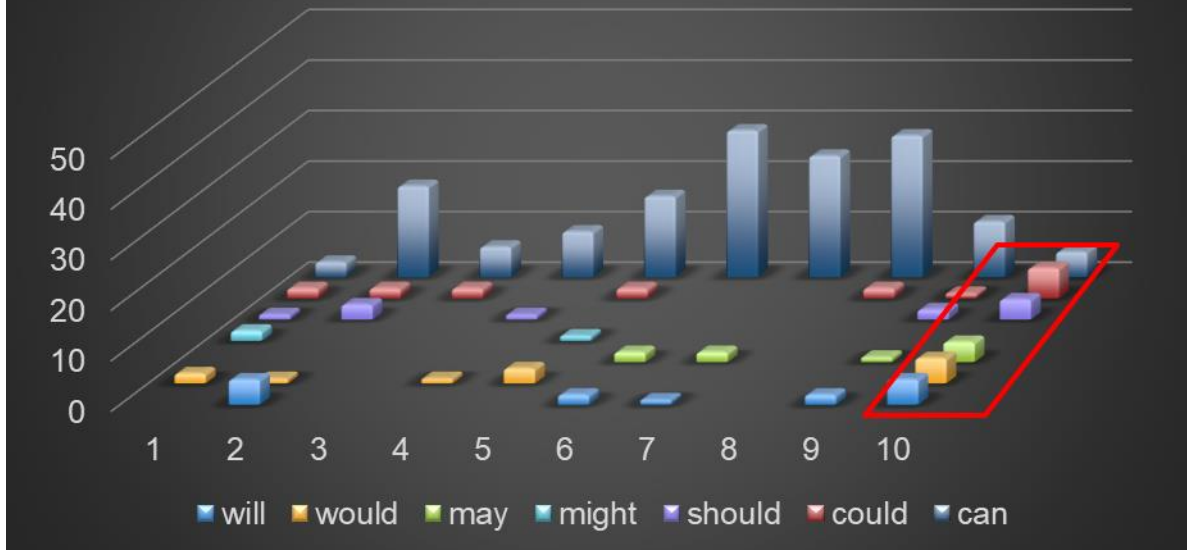


Table 1.a.2

2.a Adjectives in iPhone keynote at 2007

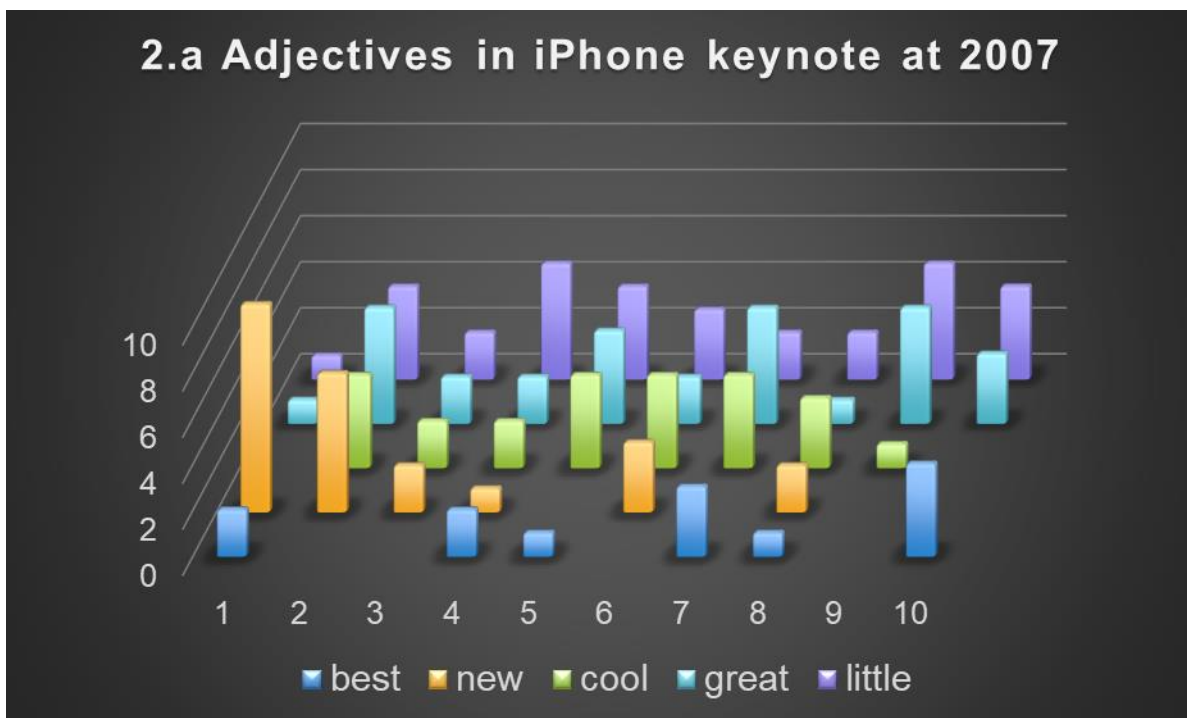


Table 2.a

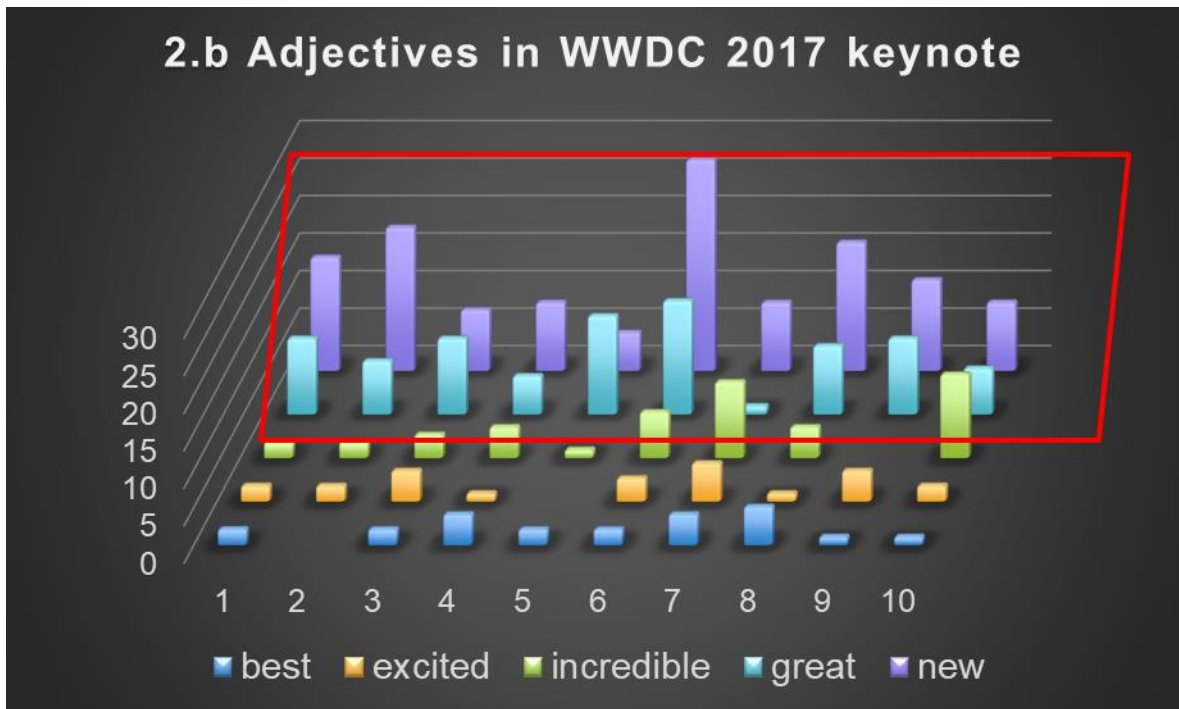


Table 2.b

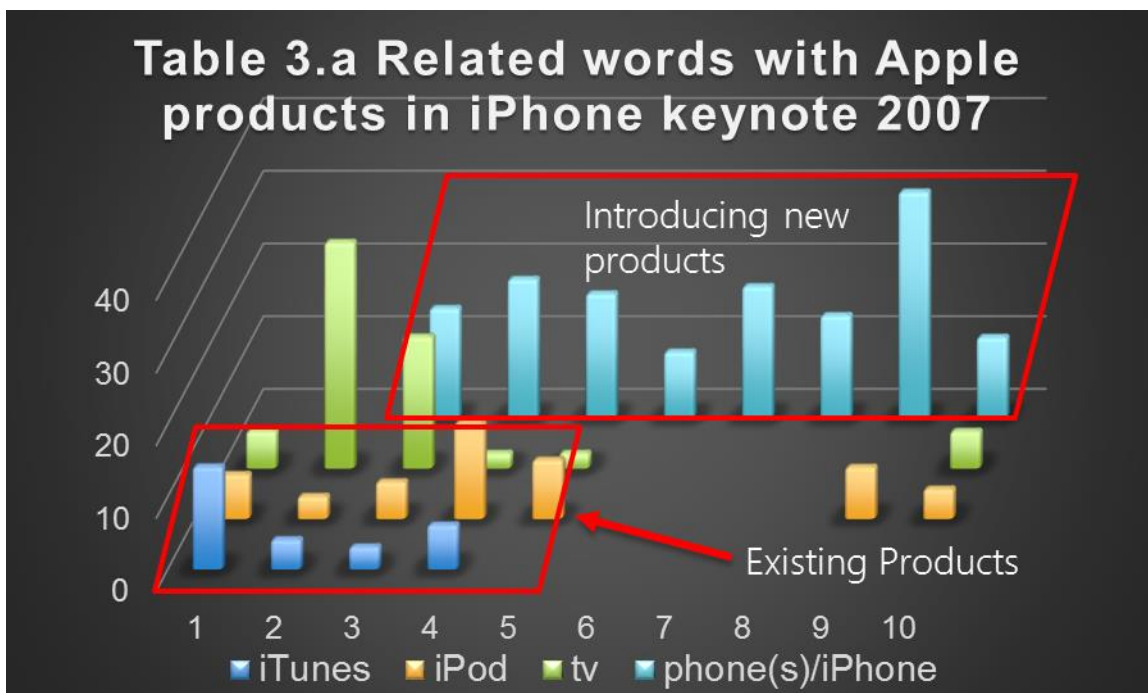


Table 3.a

Table 3.b Related words with Apple products in WWDC 2017 keynote

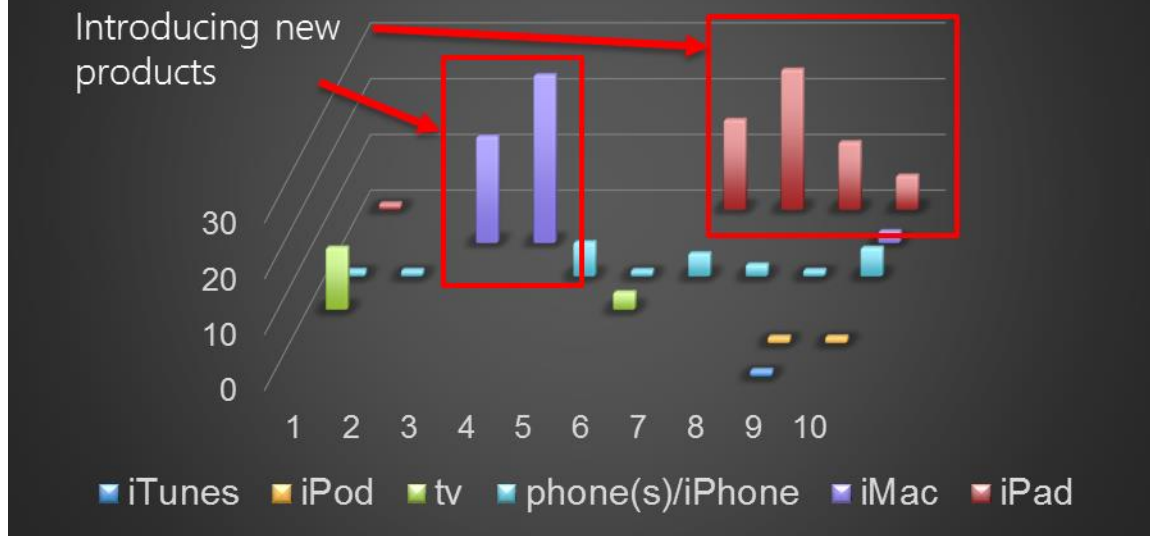


Table 3.b

Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate
1	14	0	14	4.69974	see
2	30	0	30	4.23368	just
3	13	0	13	4.06676	t
4	8	0	8	3.95128	get
5	5	0	5	2.97553	do
6	9	0	9	2.95128	go

Picture 1.a Result of word collocation with the modal verb 'can' in iPhone keynote 2007

Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate
1	22	0	22	5.01729	see
2	11	0	11	4.83087	actually
3	11	0	11	4.32068	get
4	15	0	15	4.11327	do
5	10	0	10	3.73776	also
6	13	0	13	2.10505	now

Picture 1.b Result of word collocation with the modal verb 'can' in WWDC keynote 2017