User Demand Analysis of English Word Learning APP Based on Text Mining
—Taking the APP *Bubeidanci* as an Example

Wenyi Lin, Lingling Lou

School of English Language, Zhejiang Yuexiu University, Shaoxing, China
Email: linwyi@126.com

Abstract

With the wide application of the Internet, as one of the mobile learning methods, English word learning apps are popular among college students. The homogenization between products is becoming more and more obvious. Exploring the new needs of users is becoming the next growth point of such apps. Therefore, it is particularly necessary to study the user needs of such apps. This paper uses the combination of text mining technology and questionnaire survey to collect user comments, preprocess comment data, standardize questionnaire analysis and classify user needs. Taking the app *Bubeidanci* as an example, the research focuses on user needs. The research indicates that, firstly, this kind of apps can develop and improve the functions that are related to English word pronunciation, and enhance the British pronunciation and American pronunciation modes of necessary words; secondly, a simple interface can improve user satisfaction; thirdly, this kind of apps should put the innovation points on the notes and import/export wordbook; fourthly, the apps should add the function of "modifying" the process of memorizing words, such as word interpretation change and picture assisted memory; fifthly, the apps can add functions such as checking-in learning, studying in teams and student sharing; sixthly, the apps can develop English article reading, word spelling and dictation functions; lastly, the apps can innovate the profit model and continue to explore the payment functions such as vocabulary payment and membership system.

Subject Areas

Journalism and Communication, Language Education

Keywords

Text Mining, User Demand, LDA Model, Kano Model, *Bubeidanci*

1. Introduction

1.1. Background of the Research

1.1.1. Classification of Existing English Word Learning Apps

At present, there are many English word learning apps on the market, such as *Baicizhan*, *Bubeidanci* and *Maimemo*. This kind of software can be roughly divided into three categories.

1) Online courses. These apps provide the widest range of services, involving the whole process of English learning. The service they provide is mainly a combination of courses for users with different needs. English word learning is only a part of the whole learning module. Users' learning behavior on such apps is mainly listening to English courses. A typical example of this kind of APP is *Hujiang Online Class*.

2) Dictionary. The service provided by this kind of APP is mainly the query of English words. They usually contain different versions of authoritative English dictionaries, which can give users accurate definitions of words. At the same time, some of these apps also offer English learning online courses. A typical example of this kind of APP is *Youdao Dictionary*.

3) Word memory. The services provided by this kind of APP focus on the segmentation field of word memory. They focus on scientific guidance for users to memorize words, and will scientifically set up a word recitation plan according to the Ebbinghaus forgetting curve. When it comes to memorizing words, such apps will be the first choice for users. A typical representative of this kind of APP is *Bubeidanci*.

1.1.2. Current Situation of the Use of Existing English Word Learning Apps

Take the data of the OPPO software store as an example. As of May 13, 2021, the downloads of English word learning apps are shown in Figure 1. It obviously reflects that the app *Baicizhan* is still the English word learning APP with the most users and takes the lead among the same kind of apps, followed closely by the apps *Shanbay* and *Bubeidanci*. Besides, the apps *Easy English*, *Youdaobeidanci*
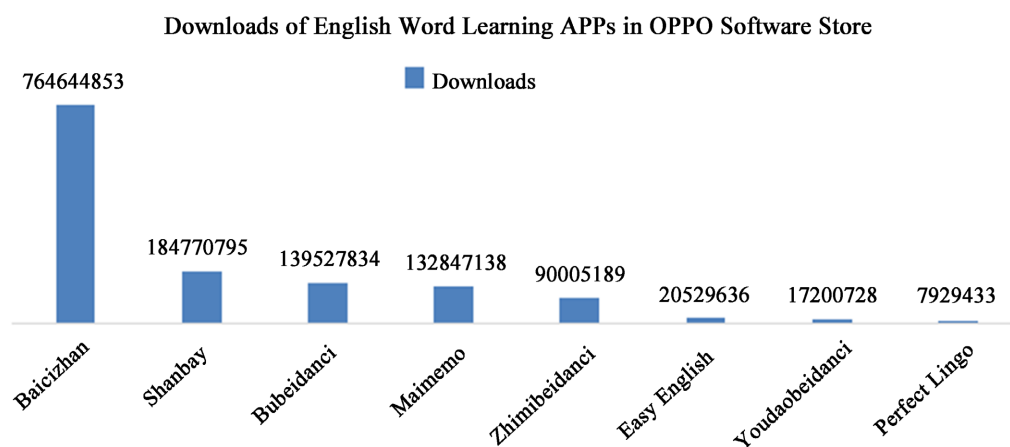

Figure 1. Downloads of English word learning apps in OPPO software store.

and *Perfect Lingo* have lowest downloads.

A review of user reviews in the App Store revealed that many users were unhappy with the feature of relying on images to remember words, saying that images were distracting and that they only remembered the pictures when they memorized the words. Other apps have their innovation points. For example, the APP *Bubeidanci* attracts users with a simple design style, while *Maimemo* attracts a large number of users with a scientific and reasonable memory curve review method. They improve user satisfaction from different aspects respectively.

According to the survey results of Wang Xiaolu [1] on college students, the top six mobile word-memorizing apps used by college students are *Shanbay, Hujiang Kaixincichang, Bubeidanci, Baicizhan, Youdaobeidanci and Zhimibeidanci*. Among them, the APP *Bubeidanci* has a high level of appearance and the example of the original movie sound, which has also attracted a lot of users.

1.1.3. Technical Route

Based on text mining of APP store comments, data processing, word cloud drawing, word frequency statistics, LDA Model keyword extraction, questionnaire survey and Kano Model demand analysis are carried out to analyze the user needs of English word learning apps, study the current situation and future growth points of English word learning apps, and give suggestions based on the guidance of text mining.

1.2. Significance of the Research

In the era of digital information, the Internet affects the behavior and patterns of thinking in our daily life. People are increasingly aware of the importance of language learning and gradually integrate it into daily life. Online learning is a new learning method that has been developing continuously in recent ten years. Almost everyone in the intelligent era will contact it, and the field is still expanding. It has become an indispensable part of the education system.

As a mobile learning method, the English word learning application is widely popular among college students because of its portability, complexity and fragmentation. This makes the technology, marketing and users of this category tend to be stable. The same type of products is emerging one after another in the market, and the homogenization phenomenon is becoming more and more obvious.

Therefore, how to seek a new direction under such market saturation is a problem to be worthy of consideration and research at present. This paper finds that the next growth point of this kind of application is to explore the new needs of users, further improve the application functions and attract more users. Only by grasping the direction of user demand in application market research can we increase user stickiness and usage. The perfection of application function is the key to realize human-computer interaction scientifically and enrich the media of human and information interaction.

Hence, after a rigorous and comprehensive study of the user needs of such applications, this paper analyzes the user needs of English word learning appli-

cations from the perspective of user comment text mining, and adds a theoretical perspective of demand analysis, which can grasp a fairly perfect direction. The combination of text mining technology and questionnaire survey makes the research develop in depth and breadth, so as to draw a scientific conclusion. This has practical significance for the function improvement of such applications and the further improvement of user satisfaction, by helping such apps to guide and plan the direction of the application development, and show its commerciality and purpose.

2. Literature Review

2.1. Studies Abroad on User Demand Analysis Based on Text Mining Technology

The main supporting technologies of text mining are natural language processing and machine learning. Foreign studies on text mining were carried out earlier. Hu Bing *et al.* [2] sorted out and found that H.P. Luhn made pioneering research in this field as early as the late 1950s and proposed the idea of word frequency statistics for automatic classification. Text mining in the modern sense was put forward by scholar Feldman in 1995.

Today, with the rapid development of Internet technology, the analysis of user needs in the form of the traditional questionnaire has fallen behind business needs. Therefore, text mining technology has been widely used in the field of user demand analysis. Kreutler [3] mainly studies the individuation of consumer needs from two aspects: the relationship between consumers and product development designers, and whether product development designers can find a design scheme that adapts to user needs. Chen Lianghsuan [4] studied the perspective of how to maximize user satisfaction, and believed that user demands and design requirements in the process of product transformation and innovation could be quantified through a nonlinear fuzzy number model. Jungmok M [5] proposed the technology of snowball trend mining, which can carry out a detailed and accurate design of the whole life cycle of the product and predict related risks before the actual production design of the product. This technology has been verified in practical operation and proved to be highly feasible.

2.2. Studies at Home on User Demand Analysis Based on Text Mining Technology

Chen Jianlin [6] believes that with the rapid development of information technology, traditional foreign language teaching elements (textbooks, content, methods, etc.) are replaced by new teaching elements (computer, Internet, information resources, information technology methods, etc.). At present, the relevant research on the combination of English learning software and teaching mainly focuses on three aspects: 1) The promotion effect of mobile software on English teaching reform. Liu Xiaoli *et al.* [7] discussed the prospect of the combination of the Internet, mobile devices and English teaching and its implications for col-

lege English teaching reform. 2) The influence of modern educational technology on the reform of teaching mode. Li Qianqu [8] proposed the combination of modern educational technology development and college English teaching reform to build a new model of college English teaching. 3) The educational security of learning software.

The research on text data mining is relatively late in China, but in recent years, text mining has been applied more and more widely in all walks of life. Liu Qingtang *et al.* [9] concluded that text mining generally includes data collection, text preprocessing (data selection and cleaning, document representation, feature selection, etc.), text mining (classification, clustering, association rule mining, etc.), and text post-processing (model evaluation and feedback, knowledge interpretation and visualization, etc.).

In terms of policy research, Yang Chengqi *et al.* [10] took government work reports as samples and conducted quantitative evaluation of energy and environmental protection policies based on the PMC-index model, establishing a set of scientific evaluation index system. In terms of public opinion analysis, Qiu Zeguo, He Baiyan [11] used text mining technology to analyze high-value public opinion topics in online public opinion, and used data visualization method to study the emotional tendency of netizens.

Yang Deqing [12] proposed a dynamic Kano Model construction method of user demand and built a fine-grained automatic marking model of product problems, which can accurately and effectively extract the problems existing in products. However, it is not enough to locate the charismatic needs of users, and the demand analysis is not comprehensive enough. Tan Qiang [13] took the H brand mobile phone as a case study and proposed a user demand identification method for online comment mining, which solves the problem that enterprises cannot effectively obtain user demand in the Internet era and enriches the sources for enterprises to obtain user demand information. Hu Yunbao [14] analyzed the data of users' online comments, with the help of text mining technology and the Kano Model of analyzing users' demands, so as to help enterprises accurately and effectively grasp users' demands.

Jiang Ruyu [15] proposed that the interaction design principles of adult self-learning apps are: pertinence, education, consistency, interaction, emotion, motivation and individuation. Wang Xun *et al.* [16] proposed three functional requirements for the TCSL mobile learning platform: page design of learning resources, content design of learning resources, and APP function design, including 16 APP design principles such as the consistency principle, the principle of human-computer interaction and design of information management. Tang Shu [17] summarized the user experience design principles of knowledge payment APP from the perspective of micro-learning into five categories: matching user usage scenarios, efficiently accessing learning content, promoting multi-dimensional group co-creation, guiding the establishment of learning habits, and improving the perception of fragmented results.

2.3. Comment on Relevant Studies

After collecting and sorting out the literature, it is found that most of the current analysis of user needs is through questionnaires distribution, sentiment analysis, and demand classification of user comments. Only a few users analyze user needs by combining online review data with the questionnaire. Therefore, the combination of text mining and questionnaire survey is chosen to analyze the user needs of English word learning apps. At the same time, through the study of literature related to the design principles of online learning APP and the combination of specific principles proposed by scholars, the following demand categories are sorted out as the criteria for the classification of questionnaire indicators, as shown in Table 1.

3. Research Methodology

3.1. Research Design

3.1.1. Research Ideas and Procedure

Based on sorting out domestic and foreign studies, combining the current situation of several English word learning apps in the market, the text mining of APP store reviews is carried out by taking the APP *Bubeidanci* as an example. To begin with, collect app store reviews. And then, by using text mining technology, the user demand index is obtained and a standardized questionnaire is set up. Combined with Kano Model, user demand classification is obtained, and suggestions for the development of English word learning apps are put forward.

The specific research ideas of this paper are as follows:

1) With the help of the platform KUCHUAN, https://www.kuchuan.com/, the user comment information about the app *Bubeidanci* in Huawei App Market, Xiaomi App Store, OPPO App Store and VIVO App Store are collected.

2) Preprocess the text data. Remove repeated comments and short comments, distinguish between Chinese and English comments, and remove blank lines.

3) Use Jieba to make word segmentation and word frequency statistics, and use wordcloud to draw wordcloud.

4) Use LDA Model to extract keywords from the comment information, and combine them with the top 200 words of word frequency to design questionnaire indicators.

Table 1. Principles of APP design.

Demand Category	Instructions
Interactive Demand	Relevant functions that allow users to interact
Functional Requirements	Basic functions that should have as an English learning app
Design Requirements	APP interface design style
Motivational Needs	Functions that motivate users to learn
Personalized Needs	Users can customize related functions

5) Make standardized questionnaires and conduct questionnaires.

6) Combine the questionnaire survey results with Kano Model, and then the characteristics and priorities of user needs are obtained.

7) Put forward specific suggestions for English word learning apps based on the conclusions.

The research process of this paper is shown in **Figure 2**.

3.1.2. Data Collection

1) App Store selection

According to the total download data of the APP *Bubeidanci* on May 5, 2021 on the platform KUCHUAN, as shown in **Figure 3**. Huawei APP Market, Xiaomi App Store, OPPO App Store and VIVO APP Store ranked the top four in terms of APP downloads, accounting for 35.42 percent, 24.95 percent, 11.43 percent and
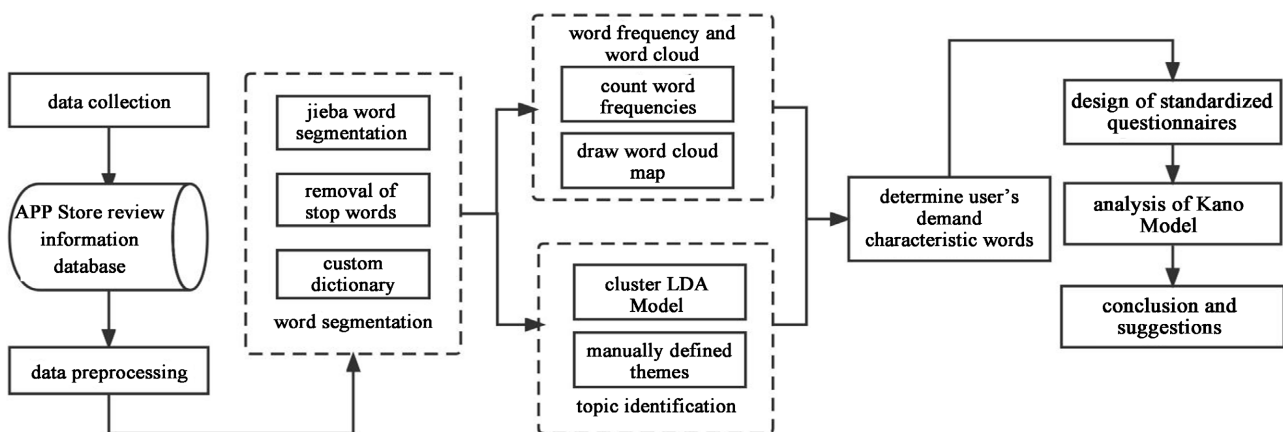

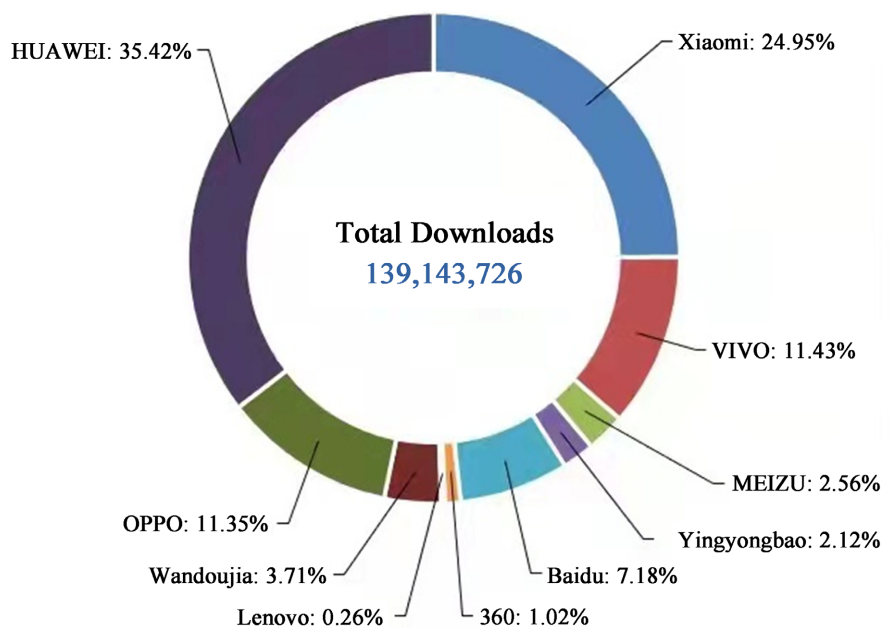**Figure 2.** Research flow chart.

Figure 3. The app *Bubeidanci*: total downloads and the share of each app market on May 5, 2001.

11.35 percent respectively. Accordingly, user reviews in these four app markets were selected as data for text mining.

2) Selection of comment text

To analyze user needs, the review information selected must be time-sensitive so that the analysis results can make effective recommendations for application development in the present moment. Therefore, with the help of the platform KUCHUAN, the above four app stores were selected to collect 65,864 comments from December 6, 2020 to December 5, 2021 as the data source of text analysis.

After preprocessing, 52,919 useful comments were obtained as the data of text mining.

3.1.3. Data Preprocessing

The raw data of user reviews is shown in Table 2. In order to screen out effective comments, the results of text mining are more representative. The user comment information is preprocessed as follows:

1) Data deduplication. Deduplication deletes duplicate data.

2) Remove English comments. According to the difference between English characters and Chinese characters, Chinese and English comments are distinguished.

3) Get rid of short comments. Remove short comments of less than five words to make the text mining results more representative. This is because the preliminary browsing of the comment information, most of the comments with words less than 5 words are meaningless comments like "very good", without describing their feelings and needs in detail.

4) Remove blank lines. After the above processing, blank lines should be removed from the data to prevent blank lines in the data from affecting subsequent operations.

Table 2. User review raw data.

Comment Attributes	Reviews	Processing Mode
Effective Comments	It's very useful to remember words, roots, affixes, word extensions, learn one and remember other forms.	Remove
Effective Comments	I have used the best software to recite words, according to the number of errors and memory curve to choose the right time to review, but also through the root affixes and the original example sentences to speed up memory, capital Nice!	Remove
English Comments	It's very useful.	Get Rid of
English Comments	It is an application which is wonderful and free to learn English.	Get Rid of
A Short Review	Very nice.	Get Rid of
A Short Review	Like to use.	Get Rid of

3.1.4. Process Analysis

1) Generation of Questionnaire Indicators

On the basis of the LDA Model analysis, the user demand-related feature words presented by word frequency statistics were added to form the final user demand index, and the questionnaire was designed by me. The design and interpretation of questionnaire indicators are shown in Table 3.

Table 3. Questionnaire indicators.

Demand Category	Indicators	Indicators Source	Explanation of Index
Interactivity	1. Friends PK Learning	High-Frequency Words	Users can compete with their friends to learn.
	2. Checking-in Learning	The LDA Model	Users can clock in and share after learning.
	3. Study in Teams	The LDA Model	Users can do team challenges and get rewards.
	4. Note Sharing	High-Frequency Words	Notes and other aspects can be shared between students.
Functionality	5. Definition of Words	The LDA Model	Provide detailed definitions when reciting words.
	6. Smart Refrigerator	The LDA Model	When reciting a word, you can change the meaning of the word.
	7. Notes	The LDA Model	You can take notes on the APP when reciting words.
	8. Examples	The LDA Model	Provide example sentences when reciting words to show how they are used.
	9. Root Affix Association	The LDA Model	Words that provide the same root affixes when reciting words.
	10. Picture-Aided Memory	The LDA Model	Remember words with pictures to help you remember them.
	11. Video-Aided Memory	The LDA Model	When reciting words, there are related videos to help you remember them.
	12. Original Movie Sound	The LDA Model	When reciting words, there are original movie sound example sentence readings.
	13. Write Words by Heart	The LDA Model	When reciting words, there are word spelling or silent writing modules.
	14. British Pronunciation	The LDA Model	Recite the words with British pronunciation.
	15. American Pronunciation	The LDA Model	Recite the words with American pronunciation.
	16. Forgetting Curve	High-Frequency Words	Vocabulary review time is arranged according to the forgetting curve.
	17. Derivative	High-Frequency Words	There are derivative hints when reciting words.
	18. Phrase	The LDA Model	When reciting words, there are related phrase prompts.
	19. Word Group	The LDA Model	Show the related phrases when reciting the words.
	20. Meaning Discrimination	The LDA Model	When reciting words, it is shown by meaning discrimination.
	21. Words Downloaded Offline	The LDA Model	Words can be downloaded offline.

Continued

Functionality	22. English Listening	The LDA Model	You can listen to words on a walkman.
	23. Vocabulary Words for Free	The LDA Model	Vocabulary is free when memorizing words.
	24. Search for New Words	The LDA Model	You can search for words you don't know.
	25. Listening Exercises	The LDA Model	You can practice listening.
	26. Text Translation	High-Frequency Words	Text translation is available.
	27. English Articles Reading	High-Frequency Words	English reading can be trained.
	28. English Learning Videos	High-Frequency Words	There are instructional videos for learning English.
	29. Past Exam Papers	The LDA Model	There are old exam questions on display.
Designability	30. Simple Interface	The LDA Model	APP interface is simple and clean, and easy to operate.
	31. Word Lock Screen Wallpaper	The LDA Model	Provide word lock screen service.
	32. APP Background Image	The LDA Model	APP background pictures have a sense of design.
Motility	33. Calendar Reminders	High-Frequency Words	You can use a calendar to remind you of your vocabulary learning schedule.
	34. Gold/Points Rewards	The LDA Model	After completing the study, the user will be rewarded.
Personalization	35. Import/Export Vocabulary Books	The LDA Model	You can import/export word books.
	36. Payment for Vocabulary	The LDA Model	Vocabulary pay for custom purchase when memorizing words.
	37. Membership System	High-Frequency Words	Membership services.
	38. Word Book	The LDA Model	You can choose your own word book to recite.
	39. Dictionary	The LDA Model	You can choose your own dictionary definition when memorizing words.

2) Distribution and Collection of Questionnaires

On the basis of research and reference, the questionnaire is designed into two parts. The first part is to investigate the user's situation, mainly including the current grade and the purposes of using English learning apps. The second part is to investigate users' demand for English word learning apps. A total of 39 services are summarized from the 5 aspects mentioned above, and standardized questionnaires are designed based on the Kano Model, and are distributed online on the professional questionnaire survey website "Wenjuanxing".

The questionnaire was released online, and the respondents were college students. Among the collected questionnaires, 165 were valid, accounting for 100% of the total questionnaires.

3) Questionnaire Evaluation

To evaluate the rationality of the questionnaire design, reliability and validity of analysis was conducted on the questionnaire, and the results are shown in Table 4.

Cronbach α coefficient was 0.934 higher than 0.8, indicating that the overall reliability of the questionnaire was high. KMO value of 0.811 was higher than

Table 4. Reliability and validity evaluation of questionnaires.

Number of Terms	Sample Size	Cronbach a Coefficient	KMO Value
84	165	0.934	0.811

0.8, showing that the overall validity of the questionnaire was high. The questionnaire passed reliability and validity tests, and the questionnaire design was reasonable.

3.2. Exploration of User Requirements

3.2.1. Chinese Word Segmentation and Word Frequency Visualization

1) Chinese Word Segmentation

a) Set the stop word

On the basis of the Baidu stop words table, the stop words which were related to comment information are added, such as "for", "indeed" and so on, to form the final stop words table. According to this stop word list, the comment information is de-stopped word processing.

b) Set a custom dictionary

The comment information of the English word learning APP has its own special words, so a custom dictionary should be set to avoid the Jieba default word segmentation method separating it. Adding "root and affix", "memory curve" and other characteristic words to the custom dictionary makes word segmentation more ideal.

c) Divide the words

After setting the stop words and user-defined dictionary, the Jieba tool is used to divide the words in users' comments.

2) Word Cloud Map Drawing

Based on Jieba word segmentation, the Wordcloud library is used to draw Wordcloud. The cloud map of comment information words is shown in Figure 4.

As can be seen from the figure, the words describing the function are "checking-in", "wallpaper", "dictation", "walkman", "root and affix" and so on. These different words reflect the diversity of users' preferences for functions and services. The size of words in the word cloud also shows the specific preferences of users. For example, the size of "example sentences" is larger than that of "dictation", which shows that users pay more attention to the functions related to example sentences than to the functions related to dictation.

3) Word Frequencies Counting

Word frequency statistics are based on Jieba segmentation, which counts the frequency of occurrence of each word in Jieba segmentation. According to word frequency, the feature word are ranked, and then output the top 200 features. Most of the high-frequency words after the removal of stop words are function-related words. On the one hand, these high-frequency words can be seen in the specific types of functions and services that are highly discussed by users. On the other hand, questionnaire indexes obtained from LDA Model results can be

Figure 4. Total word cloud.

Table 5. Top50 words.

Feature Word (word frequency)	Feature Word (word frequency)	Feature Word (word frequency)	Feature Word (word frequency)	Feature Word (word frequency)
Review (4876)	Checking-in (1216)	Mode (620)	Pronunciation (484)	Dictionary (362)
Example Sentences	*Bubei* (1197)	Picture (612)	Activities (473)	Recite Words (355)
Interface (3136)	Good Looking (1043)	Cool Currency (601)	Context (466)	Content (351)
Memory (3031)	Simple (950)	Clean (599)	Film (461)	Walkman (336)
Learning (2800)	Affix (906)	Sentence (551)	Picture (454)	Definition (329)
Page (2040)	Comfortable (838)	Team up (541)	Word Book (432)	Reminder (325)
Spelling (2019)	Advertising (756)	Wallpaper (521)	New Words (422)	Add (315)
English (1816)	Simplicity (742)	Free (512)	Recite (412)	Curve (308)
Root (1443)	Phrase (652)	Dictation (507)	Style (373)	Word Meaning (303)
Download (1227)	Design (644)	Vocabulary (487)	Friends (367)	Forgetting Curve (294)

supplemented to analyze user needs more comprehensively.

50 of them are shown in Table 5.

3.2.2. LDA Model Modeling

1) Visual interaction

The LDA Model results were visualized using the pyLDAvis library in Python, as shown in Figure 5. This diagram is interactive. Place your mouse on a theme to see the words associated with that theme. Each theme displays 30 feature words, and the correlation between feature words and the theme can be obtained by the following formula:

$$\text{relevance}\left(\text{term } w \middle| \text{topic } t\right) = \lambda * p\left(w \middle| t\right) + \left(1 - \lambda\right) * p\left(w \middle| t\right) / p\left(w\right)$$

Taking theme 1 as an example, the 30 feature words with a high degree of relevance to theme 1 are shown in Table 6.

2) Manually define themes

Analyze the characteristic words of each theme, and determine the name of each theme, as shown in Table 7.

Specific analysis shows that most of the featured words are words related to the requirements of English word learning APP users, such as spelling, cool currency, and so on. After the repeated description words were removed, 30 questionnaire indicators could be summarized in the LDA Model results. However, there are
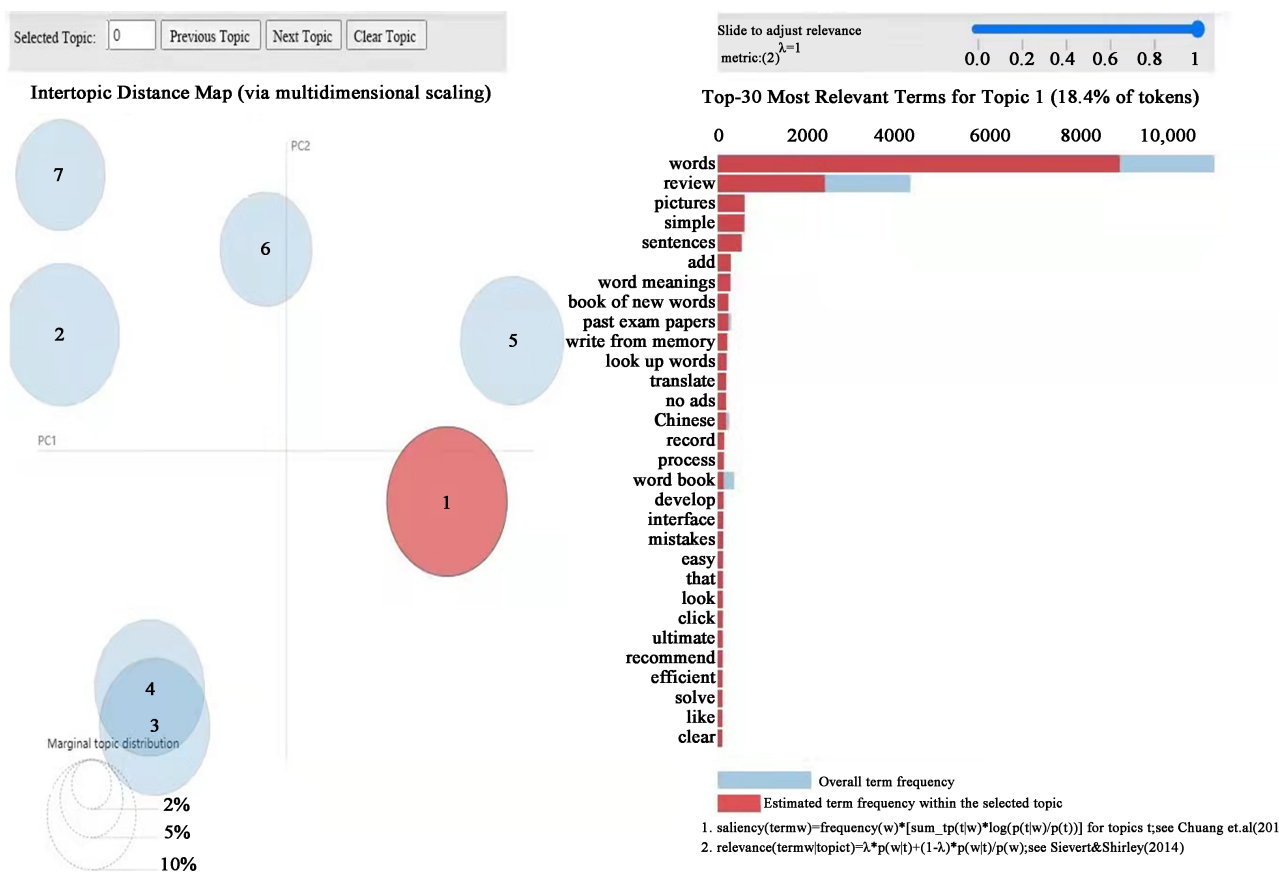

Figure 5. Theme 1 bubble diagram.

Table 6. Related words to theme 1.

30 Words that Are Highly Relevant to Theme 1					
Word	Example Sentences	*Bubeidanci*	Concise	Interface	Learning
Memory	*Baicizhan*	*Bubei*	Spelling	Page	English
Memorize Words	Check-in	Review	Download	Good-Looking	Simple
Phrase	Design	Contracted	Comfortable	Study in Team	Model
Picture	Phrase	Root	Activity	Stem Affix	Picture

some limitations in the user requirements described by these indicators. So the indicators should be supplemented.

4. Results and Discussion

4.1. Results

The findings from the survey of users' personal information are shown in Table 8. Among the 165 questionnaires collected, senior students have a larger proportion, reaching 50.3%. The proportion of freshmen is 10.91%, sophomores 15.76% and juniors 23.03%. Among the purposes of using English word learning apps, students preparing for CET-4 and CET-6 accounted for the highest proportion, reaching 76.96%.

Table 7. LDA Model results.

The Theme	The Number of Theme	Subject Words
Interactive Demand Motivational Needs	I	Spelling, checking-in, team, vocabulary, activity, style, reciting vocabulary, cool currency, phrases, band four and six
Functional Requirements	II	Learn, *Bubei*, English, memorize vocabulary, simple, review, free, word book, dictionary, cool currency
Functional Requirements	III	Example sentences, memory, *Baicizhan*, review, root and affix, root, movie, context, recite, *Maimemo*
Design Requirements	IV	Brevity, interface, page, download, good-looking, comfortable, review, wallpaper, dictation, acoustic example sentences
Functional Requirements	V	*Bubeidanci*, words, simplicity, mode, spelling, pronunciation, Walkman, background, listening, review words
Personalized Needs	VI	Words, review, pictures, clean, sentences, add, word meanings, book of new words, past exam papers, write from memory
Design Requirements	VII	Phrase, design, picture, simplicity, advertisement, beautiful, the number of vocabulary, vocabulary book, payments, new words

Table 8. Basic information of users.

Questions	Situation	Population	Proportion (%)
Grade	Freshman	18	10.91%
	Sophomore	26	15.76%
	Junior	38	23.03%
	Senior	83	50.3%
Purpose (multiple choice)	CET4&6	127	76.96%
	IELTS	31	18.79%
	TOEFL	8	4.85%
	Kaoyan (take the postgraduate entrance exams)	12	7.27%
	Other	12	7.27%

Carry on cross-analysis. The result is shown in **Figure 6**.

College students use English word learning apps for the main purposes of CET-4 and CET-6 and postgraduate exams, and their emphasis gradually changes with the increase of grades. For the freshman, the importance of CET-4 and CET-6 is the highest. As the grade increases, the importance of CET-4 and CET-6 gradually decreases, but it is still the highest among the four tests listed. The trend of the degree of postgraduate entrance examination is opposite to that of CET-4 and CET-6. It becomes more and more important as the grade increases, but it still does not exceed the importance of CET-4 and CET-6.

It should be noted that after the statistics of the students who choose "other", the purpose of using an English word learning APP for students who choose this option is just to learn and accumulate English words daily.

4.2. Discussion

According to the Kano evaluation standard table, the number of users of A, O, M, I, R and Q are counted respectively, and the calculation formula is as follows:

$$\text{Better Index} = \frac{A+O}{A+O+M+I} \qquad (1)$$

$$\text{Worse Index} = \frac{O+M}{A+O+M+I} \times (-1) \qquad (2)$$

A Better index refers to the improvement in user satisfaction when a service or function is provided; A Worse index refers to the decline in user satisfaction when a product does not provide a service or feature. The calculation results of the Kano Model are shown in **Table 9**.

Fill the calculated Better index and Worse index into the Better-Worse index quadrant graph, as shown in **Figure 7**.
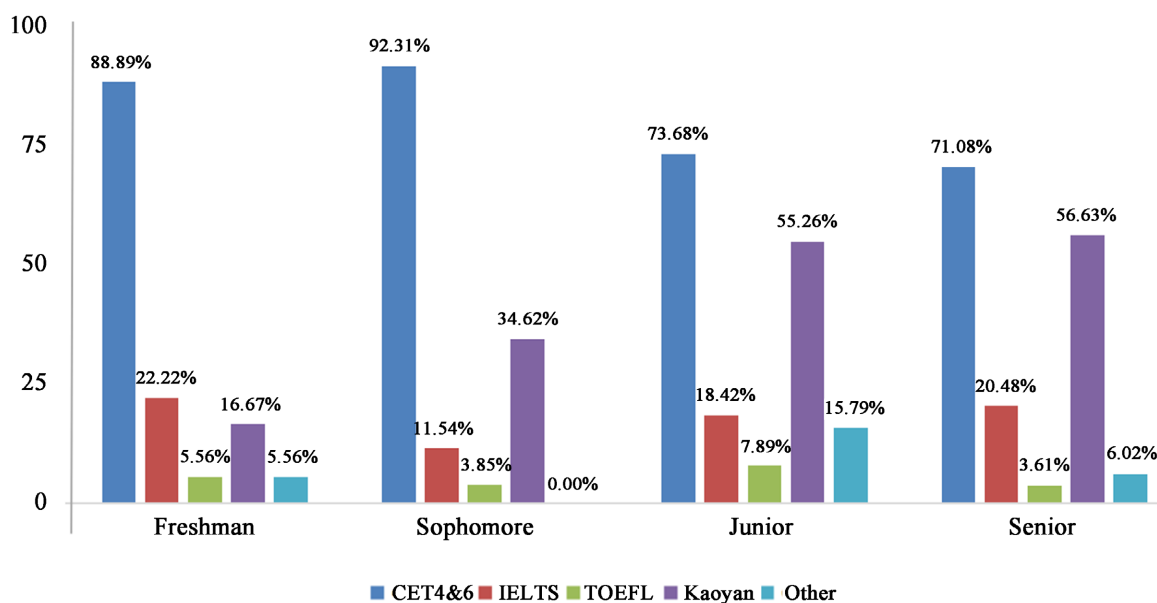

Figure 6. Cross-analysis results.

Table 9. Calculation results of Kano model.

Demand Category	Indicators	A	O	M	I	R	Q	Better Index	Worse Index
Interactivity	1. Friends PK Learning	17	5	4	131	7	1	14.01%	−5.73%
	2. Checking-in Learning	29	20	8	103	3	2	30.63%	−17.50%
	3. Study in Teams	24	5	7	119	7	3	18.71%	−7.74%
	4. Note Sharing	26	9	7	120	2	1	21.60%	−9.88%
Functionality	5. Definition of Words	29	47	15	69	2	3	47.50%	−38.75%
	6. Change of Word Meaning	27	11	6	114	5	2	24.05%	−10.76%
	7. Notes	41	21	6	93	2	2	38.51%	−16.77%
	8. Example Sentences	35	41	10	74	2	3	47.50%	−31.87%
	9. Root Affix Association	37	29	12	80	4	3	41.77%	−25.95%
	10. Picture Aided Memory	37	13	5	103	4	3	31.65%	−11.39%
	11. Video Aided Memory	25	7	6	118	5	4	20.51%	−8.33%
	12. Original Movie Sound	36	12	5	106	5	1	30.19%	−10.69%
	13. Write Words by Memory	36	19	5	101	3	1	34.16%	−14.91%
	14. British Pronunciation	28	24	8	103	1	1	31.90%	−19.63%
	15. American Pronunciation	29	22	10	100	3	1	31.68%	−19.88%
	16. Forgetting Curve	47	26	8	80	3	1	45.34%	−21.12%
	17. Derivative	44	21	9	89	0	2	39.88%	−18.40%
	18. Phrase	46	21	9	84	3	2	41.88%	−18.75%
	19. Word Group	43	23	10	88	0	1	40.24%	−20.12%
	20. Meaning Discrimination	41	21	11	90	1	1	38.04%	−19.63%
	21.Words Downloaded Offline	42	22	13	85	2	1	39.51%	−21.60%
	22. English Listening	36	20	8	96	2	3	35.00%	−17.50%
	23. Vocabulary Words for Free	41	47	7	66	1	3	54.66%	−33.54%
	24. Search for New Words	35	37	13	78	1	1	44.17%	−30.67%
	25. Listening Exercises	42	30	11	81	0	1	43.90%	−25.00%
	26. Text Translation	44	32	13	74	1	1	46.63%	−27.61%
	27. English Articles Reading	36	22	9	95	1	2	35.80%	−19.14%
	28. English Learning Videos	27	17	7	110	3	1	27.33%	−14.91%
	29. Past Exam Papers	51	30	8	74	1	1	49.69%	−23.31%
Designability	30. Simple Interface	47	32	11	72	2	1	48.77%	−26.54%
	31. Word Lock Screen Wallpaper	29	8	9	114	4	1	23.13%	−10.63%
	32. APP Background Image	26	6	8	119	5	1	20.13%	−8.81%
Motility	33. Calendar Reminders	33	18	8	103	1	2	31.48%	−16.05%
	34. Gold/Points Rewards	27	10	7	115	4	2	23.27%	−10.69%

Continued

Personalization	35. Import/Export Vocabulary Books	42	20	8	93	1	1	38.04%	−17.18%
	36. Payment for Vocabulary	3	2	0	95	62	3	5.00%	−2.00%
	37. Membership System	5	5	5	118	31	1	7.52%	−7.52%
	38. Word Book	36	39	9	77	2	2	46.58%	−29.81%
	39. Dictionary	45	25	9	83	2	1	43.21%	−20.99%
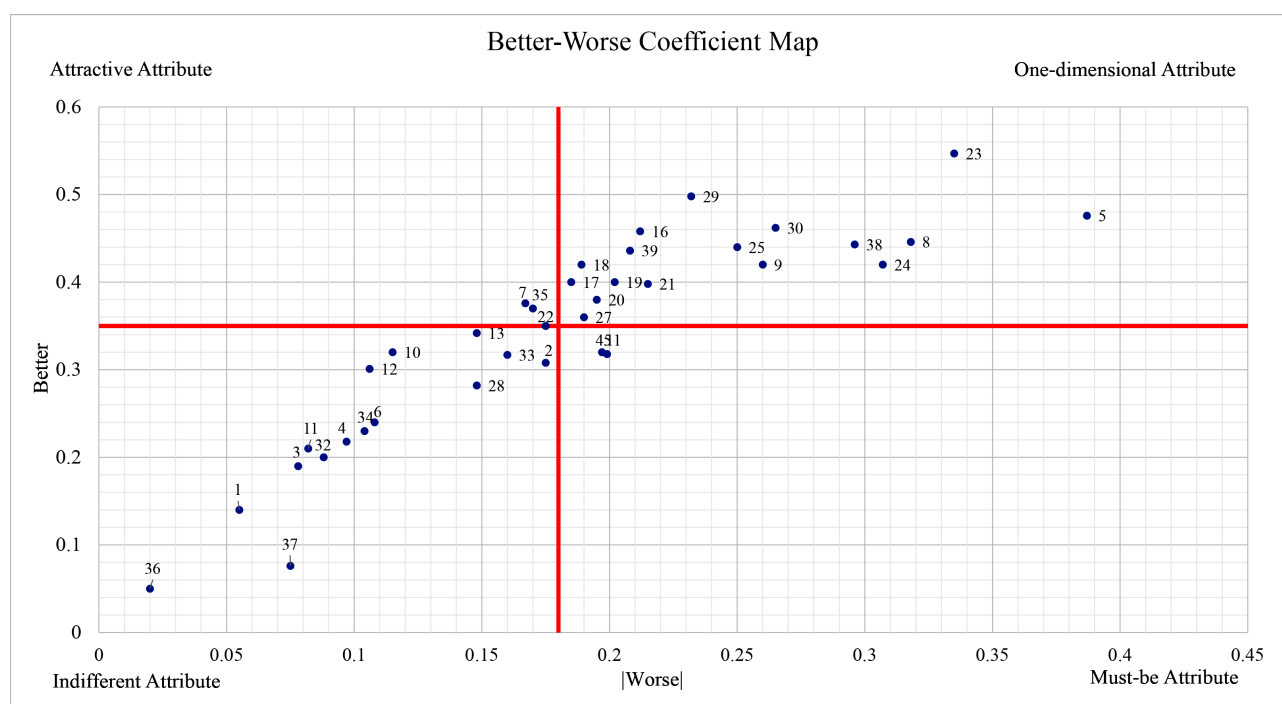
Figure 7. Quadrants of the better-worse index.

The first quadrant is the one-dimensional attribute. The better value is high, and the worse value is a high absolute value. The functions/services in this quadrant should be met first. Definition of words (5), example sentences (8), root affix association (9), forgetting curve (16), derivative (17), phrase (18) and word group (19), meaning discrimination (20), words downloaded offline (21), vocabulary words for free (23), search for new words (24), listening exercises(25), text translation(26), English articles reading(27), past exam papers (29), simple interface (30), word book (38), dictionary (39), these 18 services are expected attributes. Taking the index serial number as the object, the degree of satisfaction with the function and service is ranked as: 23 > 29 > 30 > 5 = 8 > 26 > 38 > 16 > 24 > 25 > 39 > 18 > 9 > 19 > 17 > 21 > 20 > 27. The more services provided by English word learning apps fall into this quadrant, the higher the user satisfaction will be. From the specific situation, most of the requirements in this quadrant belong to functional requirements. It can be seen that as an English word learning APP, function-related requirements are the requirements that users care most about. Free vocabulary, the display of real questions over the years,

simple interface design, etc., will greatly improve user satisfaction and make users more loyal to the product.

The second quadrant is the attractive attribute. The better value is high, and the worse value is a low absolute value. The functions/services in this quadrant should be met first; Notes (7), import/export word book (35), these two services belong to the charm attribute. Taking the index serial number as the object, the order of function and service satisfaction is 7 > 35. There are 1 functional requirement and 1 personalized requirement in this quadrant. In addition to functional requirements, if an English word learning APP can better meet users' personalized needs, users will be more satisfied with it. The note-taking function and word-book setting function greatly meet the needs of users. Users can customize according to different learning purposes. If word learning apps can make their services fall into this quadrant as much as possible, the appeal of the APP to users will be greatly enhanced. English listening (22) is between the attractive attribute and the indifferent attribute. This shows that users' evaluation of this function is quite polarized. Some users believe that this function is the innovation of the APP *Bubeidanci*. The other part of users has no sense of this function. And whether this function exists or not, it will not change their satisfaction with the APP.

The third quadrant is the indifferent attribute. The better value is low, and the worse value is also low in absolute value. The indifferent attribute means that customer satisfaction does not change when the product provides or does not provide services falling into this quadrant. The 16 services include friends PK learning (1), checking-in learning (2), studying in teams (3), notes sharing (4), change of word meaning (6), picture aided memory (10), video aided memory (11), original movie sound (12), write words by heart (13), English learning videos (28), word lock screen wallpaper (31), APP background image (32), calendar reminders (33), gold/points rewards (34), payment for vocabulary (36), and membership system (37) are indifferent attribute. Taking the index serial number as the object, the order of function and service satisfaction is: 13 > 10 > 33 > 2 > 12 > 28 > 6 > 34 > 31 > 4 > 11 > 32 > 3 > 1 > 37 > 36. Requirements that fall into this quadrant are covered in all five types. It is worth noting that the interaction requirements all fall in this quadrant. It can be seen that when college students use English word learning apps to learn, the personal learning process is the most important, while whether they can study with friends is not important to most of them. Functional requirements in this quadrant are more "additional requirements", that is, the need to decorate the process of memorizing words, such as image and video-assisted memory, original movie sound, etc. There is no difference for users in the services similar to the "decorating" word memory process. However, some common word memory apps often treat these services as gimmicks to gain more users. This takes time and effort, but has little impact. In terms of design requirements, the lock screen wallpaper function falls into this quadrant. In the incentive demand, the reward function of gold coins/

points falls in this quadrant, indicating that it is necessary to pay attention to the method of motivating users, and not all incentives will play a positive role. What kind of incentive method should be adopted is still to be further studied by APP developers. In the personalized demand, the function of vocabulary payment and membership system falls in this quadrant, which is enough to see that freedom is the most popular. At this point, developers need to innovate their monetization models.

The fourth quadrant is the must-be attribute. The better value is low, and the worse value is the high absolute value. The must-be attribute refers to that when the product provides services falling into this quadrant, user satisfaction will not be improved; On the contrary, if the service is not provided, user satisfaction will be greatly reduced. The functions/services in this quadrant must be satisfied. The two services, British pronunciation (14) and American pronunciation (15) are the must-be attributes. Taking the index serial number as the object, the order of function and service satisfaction is 14 > 15. The two services in this quadrant are functional requirements related to word pronunciation. This shows that users attach great importance to standard English pronunciation. English word learning APP can tilt the service in this aspect as much as possible to obtain user satisfaction.

According to the general theory of Kano Model, the priority of function/ service provision usually needs to be: must-be attribute > one-dimensional attribute > attractive attribute > indifferent attribute. By observing the overall results, it can be found that the correct pronunciation of words, the definition and category of words, the independent degree of reciting words and the functions derived from the whole aspect of English learning, such as listening, speaking, reading and writing, should be provided to users first. The "decoration" of the word memory process, such as picture/video-assisted memory, checking-in learning, gold rewards, etc., belongs to indifferent attribute. Users' expectations are more rested on the process of "reciting words" service. English word learning apps should return the focus of providing services to word memory itself.

Demand attributes and demand categories of indicators are summarized, as shown in Table 10.

Better values of the LDA Model were calculated and the average value was 35.47%. The Better value of high-frequency words is calculated and the average value is 29.95%. It can be seen that the quality of the index obtained by using the LDA theme model to cluster is higher than that obtained by simple word frequency ranking. It is feasible to apply text mining technology to the design of a questionnaire index. At the same time, the indexes obtained by using text mining technology are obtained from scientific analysis of user comment information, which are the functions and services that users care most, and require APP developers to pay more attention to these functions and services.

Table 10. Summary of requirement attributes and requirement categories.

Requirement Attributes	Indicators	Demand Category
Must-Be Attribute	14. British Pronunciation	Functional Requirements
	15. American Pronunciation	Functional Requirements
One-Dimensional Attribute	23. Vocabulary Words for Free	Functional Requirements
	29. Past Exam Papers	Functional Requirements
	30. Simple Interface	Design Requirements
	5. Definition of Words	Functional Requirements
	7. Example Sentences	Functional Requirements
	26. Text Translation	Functional Requirements
	38. Word Book	Personalized Needs
	16. Forgetting Curve	Functional Requirements
	24. Search for New Words	Functional Requirements
	25. Listening Exercises	Functional Requirements
	39. Dictionary	Personalized Needs
	18. Phrase	Functional Requirements
	9. Root Affix Association	Functional Requirements
	19. Word Group	Functional Requirements
	17. Derivative	Functional Requirements
	21. Words Downloaded Offline	Functional Requirements
	20. Meaning Discrimination	Functional Requirements
	27. English Articles Reading	Functional Requirements
Attractive Attribute	7. Notes	Functional Requirements
	35. Import/Export Vocabulary Books	Personalized Needs
Attractive Attribute/Indifferent Attribute	22. English Listening	Functional Requirements
Indifferent Attribute	13. Write Words by Heart	Functional Requirements
	10. Picture Aided Memory	Functional Requirements
	33. Calendar Reminders	Motivational Needs
	2. Checking-in Learning	Interactive Demand
Indifferent Attribute	12. Original Movie Sound	Functional Requirements
	28. English Learning Videos	Functional Requirements
	6. Change of Word Meaning	Functional Requirements
	34. Gold/Points Rewards	Motivational Needs
	31. Word Lock Screen Wallpaper	Design Requirements
	4. Note Sharing	Interactive Demand
	11. Video Aided Memory	Functional Requirements
	32. APP Background Image	Personalized Needs
	3. Study in Teams	Interactive Demand
	1. Friends PK Learning	Interactive Demand
	37. Membership System	Design Requirements
	36. Payment for Vocabulary	Personalized Needs

5. Conclusions

College students are the main user group of word-memorizing apps. Students on campus are all 18 - 23 years old. They are enthusiastic about new things, full of creativity, and like to reveal their personality. They also have different needs for English word learning apps. But among these differences, there is also something in common. Their demands for such apps are simplicity and efficiency. The core competitiveness of English word learning apps lies in "memorizing words". Therefore, services centering on this core demand are effective services.

This paper uses the methods of text mining and questionnaire survey to analyze the needs of users of English word learning apps. Firstly, text mining of user comments can be more efficient and intuitive to understand the main areas of users' concerns. Secondly, based on the combination of demand feature words mined in this paper and existing demand feature words of the word memorizing apps, a questionnaire survey can make more comprehensive and extensive statistics of user needs. Finally, combining the traditional Kano Model and using a quantitative analysis method, the user needs for English word learning apps are classified and prioritized.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Wang, X. (2020) A Survey of College Students' Use of APP for Memorizing English Words on Mobile Phones. *Science and Education Guide* (*Electronic Edition*), No. 8, 241-242.

[2] Hu, B., Hu, D. and Ma, W. (2008) Study and Research in Text Mining. *Computer Knowledge and Technology: Academic Exchange*, **vol**, p.

[3] Kreutler, G. and Jannach, D. (2006) Personalized Needs Elicitation in Web Based Configuration Systems. In: Blecker, T. and Friedrich, G., Eds., *Mass Customization: Challenges and Solutions*, Vol. 87, Springer, Boston, 27-42. https://doi.org/10.1007/0-387-32224-8_2

[4] Chen, L. and Ko, W. (2008) A Fuzzy Nonlinear Model for Quality Function Deployment Considering Kano's Concept. *Mathematical and Computer Science Modelling*, **48**, 581-593. https://doi.org/10.1016/j.mcm.2007.06.029

[5] Jungmok, M., Minjung, K. and Harrison, M.K. (2014) Demand Trend Mining for Predictive Life Cycle Design. *Journal of Cleaner Production*, **68**, 189-199. https://doi.org/10.1016/j.jclepro.2014.01.026

[6] Chen, J. (2010) On the Integration of Computer and Internet into Foreign Language Teaching Program—From an Ecological Point of View. Shanghai Foreign Language Education Press, Shanghai, 42.

[7] Liu, X. and Jing, G. (2016) College English Teaching Reform Based on Mobile Learning. *Chinese Vocational and Technical Education*, No. 29, 90-96.

[8] Li, Q. (2020) On the Development of Modern Educational Technology and Reform Strategies of College English Teaching. *Journal of Hubei Open Vocational College*, **33**, 161-162+165.

[9] Liu, Q., He, L., Wu, L., Yang, W. and Li, J. (2020) Educational Text Mining Model and Its Application in the Age of Intelligence. *Modern Distance Education Research*, **32**, 95-103.

[10] Yang, C. and Zhuang, X. (2021) Research on Big Data of Energy and Environmental Protection Policy Based on Text Mining. *Accounting Learning*, No. 7, 146-149.

[11] Qiu, Z. and He, B. (2021) Topic Discovery and Sentiment Analysis of Internet Public Opinion Based on Source Data Mining. *Foreign Trade*, No. 2, 76-79.

[12] Yang, D. (2017) Studying of User Requirements Based on Online Product Community. Tianjin University, Tianjin.

[13] Tan, Q. (2018) Research on User Reqirement Identification and Classification of H Brand Phone Based on Online Reviews. Jinan University, Jinan.

[14] Hu, Y. (2020) User Demand Analysis Based on Text Mining. Anhui University of Finance and Economics, Bengbu.

[15] Jiang, R. (2016) Research on Interactive Design of Autonomous Learning class APP for Adult. Harbin Engineering University, Harbin.

[16] Wang, X., Wang, F. and Shi, B. (2019) Research on the Design Principles of the Mobile Learning Platform for Teaching Chinese as a Foreign Language. *Software Guide*, **18**, 213-217.

[17] Tang, S. (2019) Research on User Experience Design of Paying-for-Knowledge Applications from the Perspective of Micro-learning. Jiangnan University, Wuxi, China.