Topic Modeling for Analyzing Patients’ Perceptions and Concerns of Hearing Loss on Social Q&A Sites: Incorporating Patients’ Perspective

Junghwa Bahng 1,* and Chang Heon Lee 2,*

1 Department of Audiology and Speech Language Pathology, Hallym University of Graduate Studies, HUGS Center for Hearing and Speech Research, Seoul 06157, Korea
2 College of Business and Economics, United Arab Emirates University, Al Ain 15551, UAE
* Correspondence: bahng.jh@hallym.ac.kr (J.B.); changlee@uaeu.ac.ae (C.H.L.)

Received: 27 June 2020; Accepted: 24 August 2020; Published: 27 August 2020

Abstract: Hearing loss is the most common human sensory deficit, affecting normal communication. Recently, patients with hearing loss or at risk of hearing loss are increasingly turning to the online health community for health information and support. Information on health-related topics exchanged on the Internet is a useful resource to examine patients’ informational needs. The ability to understand the patients’ perspectives on hearing loss is critical for health professionals to develop a patient-centered intervention. In this paper, we apply Latent Dirichlet Allocation (LDA) on electronic patient-authored questions on social question-and-answer (Q&A) sites to identify patients’ perceptions, concerns, and needs on hearing loss. Our results reveal 21 topics, which are both representative and meaningful, and mostly correspond to sub-fields established in hearing science research. The latent topics are classified into five themes, which include “sudden hearing loss”, “tinnitus”, “noise-induced hearing loss”, “hearing aids”, “dizziness”, “curiosity about hearing loss”, “otitis media” and “complications of disease”. Our topic analysis of patients’ questions on the topic of hearing loss allows achieving a thorough understanding of patients’ perspectives, thereby leading to better development of the patient-centered intervention.

Keywords: hearing loss; latent topic; LDA; topic modeling; social Q&A

1. Introduction

Hearing loss is the most common human sensory deficit, affecting normal communication. Hearing loss is a multifactorial disorder caused by both genetic and environmental factors [1]. According to recent reports from the World Health Organization (WHO) [2], the global prevalence of hearing loss has increased rapidly over the last decade. The WHO estimated that, in 2018, 6.1% of the world’s population or 466 million people experienced some degree of hearing loss. Furthermore, it is expected that the number of people with hearing loss will rise to 900 million by 2050 [2]. Hearing loss can negatively affect the quality of life as it has adverse effects on communication performance as well as emotional and social functions [3,4]. Thus, understanding the patients’ perspective of their hearing loss is an essential step toward patient improvement.

Developing adequate healthcare and public health promotion interventions requires not only an in-depth knowledge of diseases and traits, but also a comprehensive understanding of patients’ perceptions, misconceptions, concerns, and needs about the diseases. Clinical and behavioral interventions are more effective when seeking to improve outcomes that are central to patients’ experiences and perspectives [5]. A review of prior literature showed that patient-centered interventions are responsive to patients’ needs and thus, patients are encouraged to actively participate in the...
research process, especially regarding the identification of salient issues and concerns. By recognizing the importance of patient engagement, scholars have increasingly paid attention to electronic patient-authored texts as sources of valuable information [6].

With the advances in Internet and communication technologies, healthcare professionals and patients can easily communicate about patient-provided data. In addition, patients not only can easily generate and share their health data and concerns with others, but also look for relevant medical information from peers who have experienced a similar diagnosis, set of symptoms or treatments. Fox and Duggan [7] investigated the use of online resources and found that 72% of Internet users used the Internet to search for health-related information, and 39% of these online health seekers looked for health information related to their own health or medical situation [7]. Research about health communication has demonstrated that health information exchanged on the Internet is a useful resource to examine patients’ perceptions, concerns, and needs. Among Internet-based platforms, community-based question-and-answer forums are increasingly becoming popular as a medium for exchanging health information [8,9].

It is not surprising that patients with hearing loss or hearing difficulties use social media channels and social Q&A, such as Yahoo! Answers in English, Baidu Knows in Chinese, and Naver Knowledge-IN in Korean, to exchange information to address their concerns about hearing problems. Prior research has demonstrated that a collection of questions that patients post on social Q&A provides not only interesting but also important information to health professionals [9–11]. Topic analysis of patients’ questions on the topic of hearing loss allows for achieving a thorough understanding of patients’ perspectives, thereby leading to better development of the patient-centered intervention.

In this paper, we investigate the electronic patient-authored questions on social Q&A sites by applying the natural language analytics. This study adopts a multi-component semantic and computational linguistics method to discover and analyze themes or topics from hearing loss-related health conversations in the social Q&A sphere. Computational linguistics analysis is suitable to process unstructured textual data and identify hidden patterns in the data. Notably, a probabilistic topic modeling is applied to answer our research questions—topic prevalence, topic correlation, and topic evolution: what are hearing loss-related topics being asked or communicated in the social Q&A sphere? What are the prevalent topics in the hearing loss health community? How are those topics interrelated?

First, to the best of our knowledge, no prior studies have examined questions and conversations on social media Q&A as a viable source of understanding patients’ prevalent concerns on hearing loss. By applying LDA topic modeling for text analysis techniques on electronic patient-authored hearing loss questions, our present study extracts latent topics and their dominant words, which might provide useful insights for understanding patients’ perspectives and concerns on hearing loss. Second, our results reveal that patients post about the relationship between hearing loss and other diseases such as otitis media, tinnitus, and chronic kidney disease. These topic relations between categories have implications for the development of a patient-centered intervention. Lastly, our present research demonstrates that topic model analysis techniques applied to electronic patient-authored questions are effective in studying patient-focused engagement strategy. The extant literature on hearing loss indicates that survey-based methods have been a primary study design to assess patients’ concerns and perceptions about hearing loss and related issues. Using this traditional engagement strategy is costly to capture and analyze a massive volume of patients’ concerns. Data analytics on patient-authored texts on social Q&A have the potential to improve patient-centered outcome research.

2. Related Literature

2.1. Hearing Loss

Hearing loss is the most common sensory deficit. Among 432 million people with hearing loss, 93% of them are adults, and 54% of them are male—over 5% of the world’s population have hearing loss. Furthermore, it is estimated that the number of people with hearing loss will rise to
over 900 million by 2050. Untreated hearing loss can negatively affect individuals’ communication performance and thus the quality of life in individuals and their families. Hearing loss can be reduced speech understanding, declined acoustic information, and impaired localization of sound sources [12]. Hearing loss is associated with comorbidities such as social isolation, loneliness [13], depression [14–16], balance problem [17], acoustic neuroma (vestibular schwannoma), multiple sclerosis, cardiovascular disease [18], and diabetes [19,20]. Recently, a growing body of research has shown that hearing loss and dementia are related [21–24]. Hearing loss in later life is one of the factors that play an important role in decreasing cognitive ability and developing dementia.

Hearing loss can be caused by damage to any portion of the peripheral and central auditory systems. The main causes of sensorineural hearing loss are degenerative processes associated with aging, genetic mutations, noise exposure, exposure to therapeutic drugs that have ototoxic side effects, and chronic conditions [25]. The most common cause of hearing loss is aging [17]. Age-related hearing loss, generally referred to as presbycusis, typically arises from gradual changes in the inner ear, affecting the sum of sensory, neural, and metabolic causes. Additionally, other factors such as ear diseases and the effects of noise exposure may affect people at all ages and stages in life [26]. Noise-induced hearing loss is caused by loud noise exposure for more extended periods. It has been suggested that more than 12% of the global population is at risk for hearing loss from noise [27]. The WHO estimates that one-third of all cases of hearing loss can be attributed to noise exposure.

Several options are available for hearing loss, ranging from medical treatment to listening devices such as hearing aids and cochlear implants. Treatment depends on the cause and severity of hearing loss. For age-related and noise-induced hearing loss, hearing cannot be treated, but hearing can be restored after using hearing aids or cochlear implantation [25].

2.2. Social Q&A Community

Social Q&A is an online question-and-answer platform enabled by Internet and Web technologies. It is a community-driven platform that allows online users to exchange information by asking questions and providing answers [28,29]. It is open to the public, where interested parties submit questions to be answered by other fellow online users around the world. Over the last decade, social Q&A has gained popularity, and the number of visits to the top Q&A sites such as Yahoo! Answers has increased dramatically [30]. For example, Yahoo! Answer includes more than 300 million questions and 90 million unique users worldwide as of 2012 since the service launched in 2005 [10]. Another popular social Q&A site launched in 2006, Wiki Answers, has 17 million answers posted. Questions and answers on topics ranging from education to diet become a source of rich experience and opinion for anyone with similar concerns or problems. Naver is the largest online platform in South Korea, referred to as the “Google of South Korea”, and Naver provides the Social Q&A platform, Knowledge-iN. Users can post any topics on Knowledge-iN, and professionals or people who know the issues make comments and provide information and solutions, in content-centered platforms, and then users select the most valuable answers, and respondents earn awards or points.

2.3. Social Q&A Log Analysis

According to the Pew Internet and American Life Project data, more than 70% of Internet users use the Internet to search for health or medical information [7]. Increased use of such online platforms such as social Q&A sites leads to the generation of unprecedented volumes of information about symptoms, treatments, and health directly from patients, which is generally referred to as electronic patient-authored text [6,31]. As the volume of potentially valuable patient-authored text on social Q&A is growing, more researchers have paid attention to identifying the potential of online data sharing platforms for education and health service. Online patient narratives are a reliable data source for detecting disease trends and identifying medical terms [31]. Moreover, novel insights into patients’ treatment decisions and drug-treatment effects were discovered on PatientsLikeMe [32].
A review of the literature on the electronic patient-authored text on the social Q&A community indicates that the existing research streams can be divided into content-centered (e.g., question and answer narratives) and user-centered (e.g., questioners, answerer, and the community) studies. The content-centered studies have mainly focused on three areas: (1) detection of diverse types of health-related questions and answers [9,33,34]; (2) identification of medical concepts in the patient-authored text [31]; and (3) evaluation of the quality of questions and answers with a distinct set of criteria [35,36].

The first type of research has examined electronic patient-authored questions and answers from social Q&A sites to detect health-related hot topics [33,34,37]. Lu et al. (2013) applied text clustering techniques to detect disease topics such as lung cancer, breast cancer, and diabetes, and related symptoms, medical tests, drugs, procedures, and complications. Sadah et al. [33] identified a set of popular topics and associated sentiments based on the patients’ demographics. The second type has focused on identifying medically relevant terms and mapping words from the patient-authored text to medical concepts [9,31]. A language gap between patients and health care professionals is known to hinder effective communication between the two groups, so identifying and bridging the vocabulary gap is crucial [31]. Park et al. [8] applied the named entity recognition method to identify medical terms in their collected diabetes dataset and then map the identified terms to the formal medical vocabularies in the Unified Medical Language Systems (UMLS). Lastly, with concerns about the quality of both health-related questions and answers, researchers have proposed a diverse set of quality criteria and empirically examined them [36,38]. Harper et al. [37] employed supervised machine learning algorithms to distinguish information and conversational intent questions automatically. Their findings show evidence that conversational questions yield a lower archival value than informational questions.

3. Methods

Our approach carries our semantic and linguistic analysis to reveal the health characteristics of patients’ questions in online textual questions containing hearing loss-related words. The present study consists of three phases: data collection, topic discovery, and topic extraction. Figure 1 shows the overall procedure of the research analysis.

![Figure 1. Overall research flow.](image_url)

3.1. Data Collection

Electronic patient-authored texts on the topic of hearing loss were collected from the social Q&A, namely Naver Knowledge-iN. Launched in 2002, Naver Knowledge-iN is the largest social Q&A community platform in South Korea, where online users can post and share questions related to various topics ranging from insurance policy to medical treatment. Health topics are popular among the questioners. To collect research data, we developed a software program to access and gather the questions posted from 2009 to 2019 on Naver Knowledge-iN. We collected 68,327 questions using the
key word of “hearing loss”. Repeated or duplicated questions posted were excluded. In addition, questions were excluded if a question contained less than 10 words. As a result, our final sample dataset consists of 65,842 questions that were analyzed for this study.

3.2. Topic Discovery via Latent Dirichlet Allocation (LDA)

To discover the topics from the collected textual questions, we utilized a topic modeling approach that clusters the semantically associated words with “hearing loss” into subtopics. Topic modeling has been widely applied in health and medical domains such as extracting relevant clinical concepts from patient health records [31], discovering health topics in social media [11,39,40] identifying emerging patterns of clinical events [41], and detecting new disease breakout [42]. Among diverse topic modeling techniques, Latent Dirichlet Allocation (LDA) [43] has gained popularity as a tool for automatic text summarization and visualization. In this study, we apply the LDA model to extract topics from the collected corpus.

The automatic text analysis method is usually divided into supervised and unsupervised methods. The unsupervised method does not classify the text content in advance but reduces the dimension of the text through statistical probability inference and explains the text as a whole by means of the reduced dimension theme. The LDA model is an unsupervised machine learning method that uses a bag-of-words representation method. It utilized a latent variable topic between observed variables document and word to explain the semantic topic distribution of documents. The LDA modeling approach considers each document to be presented as a random mixture over latent topics, where each topic is characterized by a probability distribution over words.

LDA is a generative probabilistic model that assigns sets of words collected from documents to be explained by unobserved topic groups that explain why some parts of the data are similar. Each document consists of a small number of different topics, and each word’s generation is attributional to one of the topics of the document. The plate diagram of the LDA model is shown in Figure 2, which helps to explain the components of the LDA model.

LDA assumes that documents and the words within them are derived from a generative probabilistic model [43]. Here, \( k \) is the number of topics. \( M \) is the number of documents and \( N \) is the number of words within the document. Given a corpus \( D \) consisting of \( M \) documents, with documents having \( N_d \) having \( (d \in 1, \ldots, M) \), LDA models \( D \) according the followings.

- The number of words \( N \) is represented by the vector \( w \) as a bag-of-words.
- The model parameter \( \theta_d \) represents topic proportions for the documents, which is a random variable drawn from a Dirichlet (\( \alpha \)) prior with parameter \( \alpha \).
- For each topic, the model parameter \( \beta_k \) is the multinomial distribution of words and is drawn from the prior Dirichlet distribution using the parameter \( \eta \).
- For each word in the document, the topic of the word \( z_{wj} \) is a random variable drawn from a multinomial (\( \theta_d \)) distribution. Thus, \( z_{wj} \) is the topic that generates \( w_{ij} \) having the \( j \)th word in the \( i \)th document.
\[
P(\theta, z \mid w, \alpha, \beta) = \frac{P(\theta, z \mid w, \alpha, \beta)}{P(w \mid \alpha, \beta)} \tag{1}
\]

3.3. Topic Extraction

The topic extraction is the final step of understanding and understanding the topics found. The meaning of the topic is determined by analyzing the most likely terms along with the most likely related documents. First, a thorough investigation was performed on the relevant documents to verify that the initial interpretation based on word probability was meaningful. Second, the two researchers independently interpret and label 21 topics, and each topic is classified into one of five theme categories, except for minor word differences.

4. Results

The LDA topic extraction results concerning the first question reveal similar dimensions to those addressed in the literature on hearing loss. The twenty-one topics resulting from the LDA computation are shown in Table 1. The lists of words that make up a particular topic are displayed below the topic title for each theme. The finding suggests that the identified themes not only resembled the categorization of hearing loss diagnosis and treatments but also cover issues on topics relating to hearing loss.

We then examined the extracted 21 topics and merged similar topics. Finally, the topics were classified into five themes, which were “sudden hearing loss”, “tinnitus”, “noise-induced hearing loss”, “hearing aids”, “dizziness”, “curiosity about hearing loss”, “otitis media” and “complications of disease”. Topics and words with probabilities for each topic are shown in Table 1.

Theme 1. Sudden Hearing Loss. Four of the identified topics, namely Topics 3, 13, 17, and 22, pertain to sudden sensorineural hearing loss, commonly known as “sudden deafness”, so the main theme derived from the four topics is labelled as sudden hearing loss. Specifically, this theme addresses causes, diagnosis, medication treatments, prescriptions, and effects associated with both sudden hearing loss.

Topic 3 contains the words related to the medication treatments of sudden sensorineural hearing loss: “steroid”, “injection”, “treatment”, “drug”, “prescription”, “shot”, and “sudden deafness”. We found that, based on the selected keywords, the main concern is related to steroid use for the treatment of sudden deafness. Steroids taken orally or given by injection are commonly used to treat patients who experience sudden deafness. Topic 3 shows evidence that people with sudden hearing loss post questions to learn more about the effectiveness of treatment with steroid drugs and direct injections into their ears.

Topic 13 reflects diagnosis and test related to sudden hearing loss with keywords specifically “diagnosis”, “cause”, “hospital”, “test”, “result”, “nose”, and “infection”. It is evident that patients who have been diagnosed with sudden deafness recently posted questions about what causes sudden deafness.

Topic 17 and 21 represent prescriptions and side effects, respectively. Terms such as “side effect”, “treatment”, “prescription”, “drug”, “worry”, and “recovery” explain that patients with sudden deafness express concerns about the side effects of steroid and hearing recovery. A greater portion of patients diagnosed with sudden deafness have expressed concerns about the side effects of steroid use and have posted questions about whether their steroid treatment is appropriate in terms of frequency and dosage.

Theme 2. Tinnitus. Four of the identified topics, namely Topics 1, 6, 15, and 19, pertain to tinnitus, so the central theme derived from the four topics is labeled as tinnitus. Specifically, the tinnitus theme addresses symptoms, causes, and treatments associated with both tinnitus and hearing loss.
### Table 1. The most probable keywords in the topic of LDA with 21 topics.

| Topic No. | Topic Label | Keywords | Theme Category |
|-----------|-------------|----------|----------------|
| 1         | Tinnitus Symptoms | ringing, buzzing, tinnitus, sound, hear, constant, continue, hissing, severe, ear | Tinnitus |
| 2         | Loud Music | loud, music, earphone, listen, eardrum, affect, damage, cause, headset, longtime | Noise-induced Hearing Loss |
| 3         | Steroids for Sudden Deafness | steroid, injection, treatment, drug, pill, prescription, shot, sudden, deafness, effect | Sudden Hearing Loss |
| 4         | Symptoms of Otitis Media | symptoms, otitis media, pain, headache, sore throat, infection, deafen, today, morning | Otitis Media |
| 5         | Hearing Test Results | diagnosis, treatment, cure, curious, communication, hospital, learning, disability, quality, doubt | General Inquiry |
| 6         | Tinnitus Symptoms | tinnitus, loud, noise, hear, sleep, night, worse, stress, ear, feel, tired | Tinnitus |
| 7         | Complications | complication, cause, chronic, diabetes, anemia, cancer, thyroid, hepatitis, kidney, disease | Complications |
| 8         | Degree of Disability | handicap, disability, degree, disabled, support, experience, hearing, aids, cochlear, implant | Hearing Aids |
| 9         | Dizziness | dizzy, cause, awake, morning, severe, pain, balance, nausea, sudden | Dizziness |
| 10        | Otitis Media Treatments | surgery, hospital, medication, otitis media, medicine, doctor, condition, after, hearing, test | Otitis Media |
| 11        | Cochlear Implant | cochlear, implant, surgery, cost, normal, restore, wear, test, support, wonder | Hearing Aids |
| 12        | Health Screening | health, screening, increase, urine, test, radiation, surgery, complication, infection, hospital, | Complications |
| 13        | Diagnosis of Sudden Deafness | sudden, loss, diagnosis, cause, hospital, test, results, nose, ear, infection, damage | Sudden Hearing Loss |
| 14        | Headache | headache, stress, frequent, dizziness, symptom, discomfort, menière, disease, nerve, brain | Dizziness |
| 15        | Tinnitus Treatments | tinnitus, treatment, hospital, recommend, food, symptoms, effects, audiologist, pain, doctor | Tinnitus |
| 16        | Hearing Aids Recommendation | Hearing, aids, recommend, brand, place, good, hospital, uncomfortable, price, impaired | Hearing aids |
| 17        | Treatment and Side Effects | treatment, sudden deafness, medicine, right, ear, left, test, normal, restore | Sudden Hearing Loss |
| 18        | Shooting | shooting, military, gun, explosion, damage, hospital, degree, diagnosis, protection, accident | Noise-induced Hearing Loss |
| 19        | Causes of Tinnitus | tinnitus, infection, drug, earwax, loud, noise, stress, medication, cause, sick | Tinnitus |
| 20        | General Inquiry | stress, headache, serious, nerve, exercise, job, worry, cause, diet, decibel | General Inquiry |
| 21        | Treatment and Side Effects | side effect, steroid, drug, virus, hours, memory, worry, recovery | Sudden Hearing Loss |

Topic 1 and 6 represent issues of hearing loss, with tinnitus symptoms and problems due to tinnitus such as “ringing”, “buzzing”, “sound”, “loud”, “noise”, “sleep”, “hissing”, “night”, “severe”, and “stress”. Interestingly, the keywords identified in the two topics show that questioners or patients in social Q&A detail their hearing problems with specific vibrations and sounds (e.g., “ringing”, “buzzing” and “hissing”), and specify a level of severity when posting a tinnitus symptom. Tinnitus,
commonly referred to as buzzing in the ears, is “often accompanied by hearing loss but not everyone with hearing loss experiences tinnitus” [44]. However, the results indicate that a significantly greater portion of patients who experience tinnitus are involved in the hearing loss topic. Additionally, patients seem to be more stressed at night due to tinnitus and worry that severe or persistent tinnitus is a sign of going deaf.

Topic 15 shows the keywords involved in treatments of tinnitus. This included words such as “treatment”, “hospital”, “doctor”, “food”, “audiologist”, “pain”, and “doctor”. The topic results show that patients are in search of diverse tinnitus treatments, and specifically seek for good hospitals or doctors, and alternative treatments for tinnitus. We also found that, based on the selected keywords, effective food is a frequently occurring word, implying that tinnitus patients also search for foods to help reduce tinnitus.

Topic 19 represents a cause of tinnitus with keywords specifically including “infection”, “drug”, “earwax”, “noise”, “stress”, and “medication”. The topic results show that people want to understand what causes tinnitus. Patients who are suffering from tinnitus want to know more about whether their recent illness, such as ear infection, ear wax, or stress, causes tinnitus. Additionally, patients with tinnitus post inquiries about the relationship between tinnitus and hearing loss, and particularly ask whether tinnitus can lead to hearing loss. Most importantly, the results show that many patients with hearing loss post questions whether exposure to loud noise causes tinnitus.

Theme 3. Noise-Induced Hearing Loss. Topic 2 contains terms such as “loud”, “music”, “earphone”, “headset”, “cause”, “damage”, referred to as “noise-induced hearing loss”. These words imply that users want to ask about the relationship between listening to loud music through earphones or headsets with a personal listening device and hearing loss. In addition, people want to know how long listening to music is safe for not getting hearing loss per day.

The authors can observe the terms “shooting”, “military”, “gun”, “explosion”, “compensation”, “degree”, and “diagnosis”, “hospital” in Topic 18. These terms imply that people who are exposed to excessive and continuous noise wonder and get the information whether and how they can get any compensation with their degree of hearing loss. Notably, South Korea has adopted a conscription system. Noise-induced hearing loss is a severe disease in the military because military personnel remain in a noisy environment for the completion of their missions. Thus, the prevalence of hearing loss and tinnitus in the military is higher than the general public [45].

Theme 4. Hearing Aids. Topics 8, 11, and 16 were identified as hearing aids. Hearing aids are the most popular option of treating hearing loss for improving speech understanding. Topic 8 contains the words, “hearing aids”, “cochlear implant”, “handicap degree”, “disabled”, and “support”. In South Korea, the registered disabled person with hearing loss gets the subsidy of their hearing aid expense every five years. People with hearing loss or their family members want to know whether their hearing loss degree gets financial support.

In Topic 11, the authors also observe the term “cochlear implant” and “surgery”. In a person with severe to profound hearing loss, hearing aids do not give many benefits. So, instead of hearing aids, cochlear implantation can help to improve speech understanding to people with severe to profound hearing loss. People want to ask about alternative interventions, such as cochlear implantation surgery, other than hearing aids.

Topic 16 consists of the keywords: “hearing aids”, “recommend”, “brand”, “hospital”, and “price”. Those keywords indicate that patients post questions by asking for a specific recommendation for an excellent hearing aid with a reasonable price in social Q&A.

Theme 5. Otitis media. Topics 4 and 10 include the terms related to hearing loss with symptoms of otitis media, such as “pain”, “sore throat”, “infection”, “deafen”, “headache” and “eardrum”. Otitis media is an infection of the middle ear, and it is one of the most common disease in children [46]. The symptoms of otitis media are sore throat, night restlessness, fever, and ear pain [47].
Topic 10 also contains keywords related to the treatment of otitis media, “hospital”, “surgery”, “medication”, “test”, “after”. The results indicate that people asked about hearing loss after taking medication or undergoing surgery for otitis media. The term “hospital”, “after”, and “doctor” imply that, right after visiting doctors, users ask questions about otitis media treatment, such as drugs and surgery. Mild hearing loss can come with otitis media, and usually, hearing loss recovers after otitis media is cured. However, if otitis media occurs repeatedly, it causes permanent hearing loss [48].

**Theme 6. Dizziness.** Topic 9 contains the words “dizzy”, “cause”, “headache”, “awake”, “morning”, “severe” and “sudden” referred to as “dizziness”. Topic 14 also contains the keywords “headache”, “stress”, “dizziness” referred to as “dizziness”. One of the most common cause of dizziness is problems of the inner ear. The vestibular system, which is responsible for balance, is located within the inner ear with the auditory system. There are many diseases that affect dizziness and hearing loss, including severe cold, and bacterial or virus infections of the inner ear. In particular, Ménière’s disease is a disorder of the inner ear that leads to dizziness and sudden hearing loss. The affected ear may lead to progressive and/or permanent hearing loss [49]. Other symptoms of Ménière’s disease are tinnitus, migraines, and nausea. Ménière’s disease is very hard to diagnose due to the fact that not all of the symptoms are shown at the early stage [50]. These words relating to topic 9 may represent the symptoms of Ménière’s disease.

**Theme 7. Curiosity or general inquiry about hearing loss.** Topics 5 and 20 contain a higher number of words relating to hearing loss itself. The authors could observe that users want to know about assessment results, such as the units (decibel) which are used for describing hearing loss. Additionally, the terms such as “stress”, “nerve”, “headache”, “worry” imply that users ask about their symptoms relating to hearing loss.

The terms “treatment”, “hospital”, “communication” “learning” imply that people went to the hospital to have assessments of their hearing, but they did not fully understand their hearing loss and how to treat their hearing loss. In social Q&A, they want to get the information about their hearing loss.

**Theme 8: Complications of disease.** Topic 7 and 12 represent the terms “chronic”, “anemia”, “cancer”, “thyroid”, “hepatitis”, “kidney” and “disease”. All the diseases in topic 12 are related to hearing loss. Hearing loss is a side effect of some chemotherapy drugs [51–54]. Additionally, anemia, iron deficiency, and thyroid hormone deficiency lead to hearing loss [55–58]. The hepatitis virus B and C are very strong risk factors of sudden sensorineural hearing loss [59]. Hearing loss is linked to chronic kidney disease [60], too. In social Q&A, patients or their family members seek the information about different complications leading to a hearing loss.

5. Discussion

Hearing loss subtopics indicated that users posted about the relationship between hearing loss and hearing loss-related disease. The themes of sudden deafness, noise-induced hearing loss, and otitis media, and related complications of diseases are covered by hearing loss-related disease. Sudden deafness and otitis media themes show that users searched for social Q&A about the cause of diseases, treatment options, and side effects of the medicine. These topics confirmed that people are not well-informed about their medications. Notably, sudden deafness is considered an otologic emergency. Patients with sudden deafness can recover from the hearing loss problem if they receive appropriate treatments promptly. Prior studies reported that treatment within seven days of the onset of sudden deafness is effective for better hearing recovery [61]. The standard gold treatment of sudden deafness is oral high-dose corticosteroids [62]. Our findings reveal that users want to get the information and are worried about the side effect of the steroids, which means that patients did not get the information about their treatment and medicine.

In the theme complications, we found several keywords are related to diseases, including diabetes, anemia, cancer, thyroid, hepatitis, and kidney disease. The medications for these diseases have side
effects that can cause hearing loss. Based on our findings, health providers, such as physicians, did not give enough information about treatment and medicine. Lee et al. [63] indicated that the average real consultation time for outpatients was 4.2 min in Korea, and patient satisfaction was too low regarding consultations. This consultation time is not long enough to ask questions about their treatments. The physicians need to ensure that they have enough consultation time for effective communication.

On the other hand, noise-induced hearing loss is a slow-occurring hearing loss. In subtopics of noise-induced hearing loss, users post about the causes of hearing loss, such as loud noise exposure and the usage of listening devices with headphones. People wonder how much loud noise exposure causes hearing loss. Exposure to noise increases the risk of tinnitus, as well. Noise-induced hearing loss is the only preventable loss, so education on the usage of a safe listening device is necessary [64]. For preventing noise-induced hearing loss, people are advised to avoid listening to loud music or noise above 85 dB (A) for no longer than 1 h. People need to avoid listening to music because it increases the volume when they listen to it under noisy circumstances. Interestingly, in these subtopics, we found keywords related to the military. Military personnel exposed to significant impulse noise. For this reason, in social Q&A, many people seek information about the relationship between hearing loss and military noise. The hearing conservation program needs to be developed adequately to ensure the health of military personnel in Korea.

Next, users posted the relationship between hearing loss and symptoms of hearing loss, tinnitus, and dizziness. The highest subtopics are related to tinnitus. Tinnitus is a symptom associated with ear disease, including hearing loss, but tinnitus does not cause hearing loss. More than 50 million people, with an estimated prevalence of 10–15% in adults, reported that they experienced tinnitus in the U.S. [65]. The keywords related to hearing loss indicated that people want to know about the cause of tinnitus, medicines, and difficulties caused by tinnitus. The findings show that people want to find a way to cure tinnitus. However, the treatment of tinnitus focuses on minimizing the impact and burden of tinnitus rather than the “cure” of tinnitus [65]. There are several rehabilitation methods for relieving tinnitus, such as counseling and sound therapy. Additionally, there are applications that patients can use to alleviate their tinnitus. Health providers need to inform the options of tinnitus relief.

Tinnitus is a symptom that people feel chronically, while dizziness is a symptom that people feel acutely. Our results reveal several keywords related to the symptoms of Ménière’s disease. Ménière’s disease is a disorder of the inner ear that leads to dizziness and sudden hearing loss. The affected ear may lead to progressive or permanent hearing loss [49]. Other distinct symptoms of Ménière’s disease include tinnitus, migraines, and nausea. Ménière’s disease is hard to diagnose because not all of the signs are shown at the early stage [50]. The patients who feel dizziness need to visit the Ear, Nose and Throat (ENT) doctor and check their conditions.

Lastly, people want to search for information about the treatment options of hearing loss, such as hearing aids and cochlear implants. Hearing aid usage is the most common treatment to reduce the adverse effects, including communication difficulties and cognitive decline, caused by hearing loss [66]. In addition, the use of hearing aids improves productive time use, quality of life, economic circumstances and mental health for listeners with hearing loss [67]. In our results, people searched for information on hearing aid benefits or the specification of hearing aids. Additionally, they searched about brands of hearing aids and reasonable prices for hearing aids. They also asked for recommendations regarding good hearing aid centers. Since the national insurance subsidizes all or part of the cost of purchasing hearing aids or cochlear implants for hearing impairments in Korea, users search the information about disability degree and national insurance benefit. This information should be provided to websites so that those who want to use hearing aids can accurately check the price and specifications of hearing aids and get national insurance benefits.
6. Conclusions

Hearing loss is the most frequently occurring human sensory deficit and has many different causes. Hearing loss is also known to be related to many other health issues. Analyzing patient-authored questions can be a useful approach to better understand patients’ perceptions, concerns and needs. Traditional surveys are limited to small sample sizes. However, social Q&A offers a new environment for patients to easily share various opinions and medical experiences, so a large volume of patient-authored data can accumulate. Exploring how patients with hearing loss use social questions and answers to find health information not only helps to identify a set of critical topics and issues for various types of research but also improves communication between patients and healthcare professionals. To the best of our knowledge, this is the first analysis of patient-authored contents on the topic of hearing loss from the social Q&A community, and these results provide valuable methodological and content insights.

This study provides a computational linguistic approach to perform an in-depth analysis using patient-authored data from sizeable online data sets. Our framework decodes public health views from hearing loss-related questions, which can be useful for developing adequate healthcare and public health promotion interventions. Our results reveal that those characterized topics ranging from sudden deafness to hearing aids are both representative and meaningful, and mostly correspond to sub-fields established in hearing science research.

Author Contributions: Conceptualization, J.B. and C.H.L.; Data curation, C.H.L.; Formal analysis, J.B. and C.H.L.; Funding acquisition, J.B. and C.H.L.; Investigation, J.B.; Project administration, J.B.; Software, C.H.L.; Supervision, J.B.; Writing—original draft, J.B. and C.H.L.; Writing—review & editing, J.B. and C.H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A5A2A01039904). This research was partially funded by the UAE University, grant number G00002617 (Funding No. 31B088).

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Willems, P.J. Genetic causes of hearing loss. N. Engl. J. Med. 2000, 342, 1101–1109. [CrossRef]
2. World Health Organization. Deafness and Hearing Loss. Available online: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss (accessed on 10 March 2020).
3. Dalton, D.S.; Cruickshanks, K.J.; Klein, B.E.; Klein, R.; Wiley, T.L.; Nondahl, D.M. The impact of hearing loss on quality of life in older adults. Gerontologist 2003, 43, 661–668. [CrossRef]
4. Seidman, M.D.; Standring, R.T. Noise and quality of life. Int. J. Environ. Res. Public Health 2010, 7, 3730–3738. [CrossRef]
5. Iezzoni, L.I.; O’Day, B.L.; Killeen, M.; Harker, H. Communicating about health care: Observations from persons who are deaf or hard of hearing. Ann. Intern. Med. 2004, 140, 356–362. [CrossRef]
6. Dreisbach, C.; Koleck, T.A.; Bourne, P.E.; Bakken, S. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. Int. J. Med. Inform. 2019, 125, 37–46. [CrossRef]
7. Fox, S.; Duggan, M. Health online 2013. Health 2013, 2013, 1–55.
8. Oh, S.; Yi, Y.J.; Worrall, A. Quality of health answers in social Q&A. Proc. Am. Soc. Inf. Sci. Technol. 2012, 49, 1–6.
9. Park, M.S.; He, Z.; Chen, Z.; Oh, S.; Bian, J. Consumers’ use of UMLS concepts on social media: Diabetes-related textual data analysis in blog and social Q&A sites. JMIR Med. Inform. 2016, 4, e41. [PubMed]
10. Zhang, J.; Chen, Y.; Zhao, Y.; Wolfram, D.; Ma, F. Public health and social media: A study of Zika virus-related posts on Yahoo! Answers. J. Assoc. Inf. Sci. Technol. 2019, 71, 282–299. [CrossRef]
11. Jones, J.; Pradhan, M.; Hosseini, M.; Kulanthaivel, A.; Hosseini, M. Novel Approach to Cluster Patient-Generated Data Into Actionable Topics: Case Study of a Web-Based Breast Cancer Forum. JMIR Med. Inform. 2018, 6, e45. [CrossRef]
12. Divenyi, P.L.; Stark, P.B.; Haupt, K.M. Decline of speech understanding and auditory thresholds in the elderly. J. Acoust. Soc. Am. 2005, 118, 1089–1100. [CrossRef] [PubMed]
37. Lu, Y.; Zhang, P.; Liu, J.; Li, J.; Deng, S. Health-related hot topic detection in online communities using text clustering. *PLoS ONE* **2013**, *8*, e56221. [CrossRef]
38. Harper, F.M.; Moy, D.; Konstan, J.A. Facts or friends? Distinguishing informational and conversational questions in social Q&A sites. In *Proceedings of the Sigchi Conference on Human Factors in Computing Systems*, Boston, MA, USA, 4–9 April 2009; pp. 759–768.
39. Paul, M.J.; Dredze, M. Discovering health topics in social media using topic models. *PLoS ONE* **2014**, *9*, e103408. [CrossRef]
40. Ghosh, D.; Guha, R. What are we ‘tweeting’ about obesity? Mapping tweets with topic modeling and Geographic Information System. *Cartogr. Geogr. Inf. Sci.* **2013**, *40*, 90–102. [CrossRef]
41. Chandola, V.; Sukumar, S.R.; Schryver, J.C. Knowledge discovery from massive healthcare claims data. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Chicago, IL, USA, 11–13 August 2013; pp. 1312–1320.
42. Charles-Smith, L.E.; Reynolds, T.L.; Cameron, M.A.; Conway, M.; Lau, E.H.; Olsen, J.M.; Pavlin, J.A.; Shigematsu, M.; Streichert, L.C.; Suda, K.J. Using social media for actionable disease surveillance and outbreak management: A systematic literature review. *PLoS ONE* **2015**, *10*, e0139701. [CrossRef]
43. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *6*, 993–1022.
44. Husain, F.T.; Medina, R.E.; Davis, C.W.; Szymko-Bennett, Y.; Simonyan, K.; Pajor, N.M.; Horwitz, B. Neuroanatomical changes due to hearing loss and chronic tinnitus: A combined VBM and DTI study. *Brain Res.* **2011**, 1369, 74–88. [CrossRef] [PubMed]
45. Yong, J.S.; Wang, D.Y. Impact of noise on hearing in the military. *Mil. Med. Res.* **2015**, *2*, 6. [CrossRef] [PubMed]
46. Zielhuis, G.A.; Rach, G.H.; van den Bosch, A.; van den Broek, P. The prevalence of otitis media with effusion: A critical review of the literature. *Clin. Otalaryngol. Allied Sci.* **1990**, *15*, 283–288. [CrossRef]
47. Kontiokari, T.; Koivunen, P.; Niemela, M.; Pokka, T.; Uhari, M. Symptoms of acute otitis media. *Pediatric Infect. Dis. J.* **1998**, *17*, 676–679. [CrossRef]
48. Gravel, J.S.; Wallace, I.F.; Ruben, R.J. Auditory consequences of early mild hearing loss associated with otitis media. *Acta Oto-Laryngol.* **1996**, *116*, 219–221. [CrossRef] [PubMed]
49. Sarna, B.; Abouzari, M.; Lin, H.W.; Djalilian, H.R. A hypothetical proposal for association between migraine and Meniere’s disease. *Med. Hypotheses* **2020**, *134*, 109450. [CrossRef]
50. Sajjadi, H.; Paparella, M.M. Meniere’s disease. *Lancet* **2008**, *372*, 406–414. [CrossRef]
51. Hirose, Y.; Simon, J.A.; Ou, H.C. Hair cell toxicity in anti-cancer drugs: Evaluating an anti-cancer drug and Meniere’s disease. *Adv. Integr. Anat. Evol. Biol.* **2011**, *12*, 719–728. [CrossRef] [PubMed]
52. Ding, D.; Allman, B.L.; Salvi, R. Review: Ototoxic characteristics of platinum antitumor drugs. *Anat. Rec. Adv. Integr. Anat. Evol. Biol.* **2012**, *295*, 1851–1867. [CrossRef]
53. Vermorken, J.B.; Remenar, E.; van Herpen, C.; Gorlia, T.; Mesia, R.; Degardin, M.; Stewart, J.S.; Jelic, S.; Betka, J.; Preiss, J.H.; et al. Cisplatin, fluorouracil, and docetaxel in unresectable head and neck cancer. *N. Engl. J. Med.* **2007**, *357*, 1695–1704. [CrossRef]
54. Oldenburg, J.; Gietema, J.A. The Sound of Silence: A Proxy for Platinum Toxicity. *JAMA Otolaryngol. Head Neck Surg.* **2014**, *140*, 417–422. [CrossRef] [PubMed]
55. Shih, J.-H.; Li, I.; Pan, K.-T.; Wang, C.-H.; Chen, H.-C.; Fann, L.-Y.; Tseng, J.-H.; Kao, L.-T.; Health, P. Association between Anemia and Auditory Threshold Shifts in the US Population: National Health and Nutrition Examination Survey. *Int. J. Environ. Res. Public Health* **2020**, *17*, 3916. [CrossRef] [PubMed]
56. Chung, S.D.; Chen, P.Y.; Lin, H.C.; Hung, S.H. Sudden sensorineural hearing loss associated with iron-deficiency anemia: A population-based study. *JAMA Otolaryngol. Head Neck Surg.* **2014**, *140*, 417–422. [CrossRef] [PubMed]
57. Ben-Tovim, R.; Zohar, Y.; Zohiar, S.; Laurian, N.; Laurian, L. Auditory brain stem response in experimentally induced hypothyroidism in albino rats. *Laryngoscope* **1985**, *95*, 982–986. [CrossRef] [PubMed]
58. Vanasse, M.; Fischer, C.; Berthezene, F.; Roux, Y.; Volman, G.; Mornex, R. Normal brainstem auditory evoked potentials in adult hypothyroidism. *Laryngoscope* **1989**, *99*, 302–306. [CrossRef]
59. Chen, H.-C.; Chung, C.-H.; Wang, C.-H.; Lin, J.-C.; Chang, W.-K.; Lin, F.-H.; Tsao, C.-H.; Wu, Y.-F.; Chien, W.-C. Increased risk of sudden sensorineural hearing loss in patients with hepatitis virus infection. *PLoS ONE* **2017**, *12*, e0175266. [CrossRef]
60. Vilayur, E.; Gopinath, B.; Harris, D.C.; Burlutsky, G.; McMahon, C.M.; Mitchell, P. The association between reduced GFR and hearing loss: A cross-sectional population-based study. *Am. J. Kidney Dis.* 2010, 56, 661–669. [CrossRef]

61. Chen, W.T.; Lee, J.W.; Yuan, C.H.; Chen, R.F. Oral steroid treatment for idiopathic sudden sensorineural hearing loss. *Saudi Med. J.* 2015, 36, 291. [CrossRef]

62. Chandrasekhar, S.S.; Surgery, N. Updates on methods to treat sudden hearing loss. *Oper. Tech. Otolaryngol.-Head Neck Surg.* 2003, 14, 288–292. [CrossRef]

63. Lee, C.H.; Lim, H.; Kim, Y.; Park, A.H.; Park, E.-C.; Kang, J.-G. Analysis of appropriate outpatient consultation time for clinical departments. *Health Policy Manag.* 2014, 24, 254–260. [CrossRef]

64. You, S.; Kwak, C.; Han, W.; Health, P. Use of Personal Listening Devices and Knowledge/Attitude for Greater Hearing Conservation in College Students: Data Analysis and Regression Model Based on 1009 Respondents. *Int. J. Environ. Res. Public Health* 2020, 17, 2934. [CrossRef] [PubMed]

65. Tunkel, D.E.; Bauer, C.A.; Sun, G.H.; Rosenfeld, R.M.; Chandrasekhar, S.S.; Cunningham Jr, E.R.; Archer, S.M.; Blakley, B.W.; Carter, J.M.; Granieri, E.C. Clinical practice guideline: Tinnitus. *Otolaryngol. Head Neck Surg.* 2014, 151, S1–S40. [CrossRef] [PubMed]

66. Arlinger, S. Negative consequences of uncorrected hearing loss—A review. *Int. J. Audiol.* 2003, 42 (Suppl. 2), 2S17–2S20. [CrossRef] [PubMed]

67. Spreckley, M.; Macleod, D.; González Trampe, B.; Smith, A.; Kuper, H.; Health, P. Impact of Hearing Aids on Poverty, Quality of Life and Mental Health in Guatemala: Results of a before and after Study. *Int. J. Environ. Res. Public Health* 2020, 17, 3470. [CrossRef] [PubMed]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).