Online discussions about tinnitus: What can we learn from natural language processing of Reddit posts?

Vinaya Manchaiah,[a,b,c,d] Alain Londero,[e] Aniruddha K. Deshpande,[f,g] Manon Revel,[h] Guillaume Palacios,[i] Ryan L. Boyd,[j,k,l] & Pierre Ratinaud[m]

a. Department of Speech and Hearing Sciences, Lamar University, Beaumont, Texas, USA

b. Virtual Hearing Lab, Collaborative initiative between Lamar University and University of Pretoria, Beaumont, Texas, USA

c. Department of Speech-Language Pathology and Audiology, University of Pretoria, Gauteng, South Africa

d. Department of Speech and Hearing, School of Allied Health Sciences, Manipal, Karnataka, India

e. Hôpital Européen Georges Pompidou, Assistance Publique - Hôpitaux de Paris ; Faculté de Médecine Paris Descartes - Université de Paris, Paris, France.

f. Department of Speech-Language-Hearing Sciences, Hofstra University, Hempstead, New York, USA

g. Long Island Doctor of Audiology Consortium, Garden City, New York, USA

h. Institute for Data, Systems and Society, Massachusetts Institute of Technology, USA

i. PainkillAR, TELECOM ParisTech, Paris, France

j. Department of Psychology, Lancaster University, Lancaster, UK

k. Security Lancaster, Lancaster University, Lancaster, UK

l. Data Science Institute, Lancaster University, Lancaster, UK

m. Laboratory of Applied Studies and Research in Social Sciences, University of Toulous, France

Short Title: Online discussions about tinnitus

Corresponding author:	Dr. Vinaya Manchaiah
Communication address:	Department of Speech and Hearing Sciences, Lamar University, Beaumont, Texas 77710, USA
Email:	vinaya.manchaiah@lamar.edu
Tel:	+1 (409) 880 8927
Fax:	+1 (409) 880 2265

Abstract

Background: The present study was aimed at identifying key topics in online discussions about tinnitus by examining a large dataset extracted from Reddit social media using natural language processing technique.

Method: A corpus of 113,215 posts about tinnitus were extracted from Reddit's application programming interface (API). After cleaning the data for duplications and posts without any text information, the sample was reduced to 101,905 posts which was subjected to cluster analysis using the open-source IRaMuTeQ software to identify main topics based on the cooccurrence of texts. These clusters were named by a panel of tinnitus experts (n=9) by reading typical text segments within each cluster.

Results: The cluster analysis identified 16 unique clusters which belong to two topics which were named as "tinnitus causes and consequences" and "tinnitus management and coping." Based on their characteristics, the clusters were named as: tinnitus timeline (10%), tinnitus perception (9.7%), medical triggers and modulators (8.8%), hearing research (8.8%), attention and silence (8.6%), social media posts about tinnitus (7.4%), hearing protection (7.3%), interaction with hearing healthcare providers (6.7%), mental health and coping (5.8%), music listening (5.7%), hope for a cure (5.6%), interactions with people without tinnitus (5.4%), dietary supplements and alternative therapies (3.2%), sleep (3.9%), dietary effects (1.7%), writing about tinnitus and being thankful to online community (1.4%).

Conclusions: Despite some limitations, tinnitus posts on Reddit provide rich real-world data to identify various issues and complaints that tinnitus patients and their significant others discuss in online communities. Some of the clusters identified here are novel (e.g., tinnitus timeline, interactions with people without tinnitus) and have not been much discussed in the tinnitus

literature. The results suggest that individuals with tinnitus relay on social media for support and highlights the service delivery needs in providing social support through other means (e.g., support groups).

Key Words

Tinnitus, Natural language processing, Text mining, Topic modelling, Computational predictive medicine

Introduction

Tinnitus is a common yet highly heterogenous condition. Approximately 10-15% of adults have tinnitus, with 1-2% experiencing severe, distressing tinnitus (Baguley et al., 2013). The tinnitus population is highly varied in terms of their symptoms and complaints. Many individuals with tinnitus have a normal life other than the occasional nuisance due to the auditory percept; however, some experience deleterious effects of tinnitus (such as anxiety, depression and insomnia) (Manchaiah et al., 2018a) on their professional and personal lives. While various management approaches for tinnitus are available, there is currently no cure (McFerran et al., 2019). Additionally, many individuals with tinnitus lack access to evidence-based tinnitus management such as cognitive behavioral therapy (Bhatt et al., 2016).

In recent years, it has become increasingly common for people to seek health information online (Madrigal & Escoffery, 2019) — this can be particularly true for individuals with chronic, incurable conditions such as tinnitus. The online environment serves as both a social support system as well as a complement to medical professionals, hence it is particularly attractive for

chronic incurable conditions such as tinnitus. Online health information-seeking behaviors may

vary across individuals in terms of why they seek online health information, where they seek

(e.g., internet websites, social media), ability to find reliable information, and how they use this

information. During the early stages of the condition (e.g., during initial tinnitus onset),

individuals may look for possible causes and treatment options. At later stages, they may seek

social support online or alternative therapies to improve coping. Although much of the

information online is unreliable (Swire-Thompson & Lazer, 2020), it is still important for

healthcare professionals to understand the type of information available on different online

outlets as well as how individuals with tinnitus interact with this information as this could impact

the decision-making process as well as how these individuals cope with the condition.

The web-based approach for gathering and sharing ideas on a large-scale is referred to as

"crowdsourcing." This approach may have several advantages including (a) the pool of people

contributing to such discussion are large and diverse, (b) much lower efforts and costs in

collecting data and increases the speed of innovation, and (c) public and professionals are on

equal footing reducing the power imbalance seen in clinical therapeutic relationship (Hall et al.,

2012).

During the last decade several studies have examined online health information about tinnitus on

various outlets as well as on various dimensions such as content, user characteristics, and

misinformation using the crowdsourcing approach to data extraction. Manchaiah et al. (2019a)

examined English language websites about tinnitus and reported that they have poor quality and

high readability. In another study, Manchaiah et al. (2018b) examined tinnitus posts on Facebook

pages using the Hierarchical Cluster Analysis (HCA) and reported that the messages belonged to various topics including (1) diagnosis (21.9%), (2) cause (4.1%), (3) research and development (13.6%), (4) social support (18.8%), (5) challenges (11.1%), (6) symptoms (21.4%), and (7) coping (9.2%). A study on tinnitus tweets suggested that Twitter is mainly used for advocacy, sharing personal information and experiences rather than providing social support or to build relationships (Ni et al., 2020). Examination of the top 100 most-viewed videos on YouTube about tinnitus suggests that the videos were mainly uploaded by individuals with tinnitus sharing their personal experiences (Basch et al., 2018). In another study, Deshpande et al. (2018) examined tinnitus information on various social media platforms and suggested that these social media outlets vary in terms of their use – individuals with tinnitus could learn about tinnitus, seek social support, advocate for tinnitus support, as well as connect with health professionals on different platforms.

More recently, Palacios et al. (2020) examined the symptoms and complaints reported by individuals with tinnitus and hyperacusis in a tinnitus forum using Natural Language Processing (NLP) methods, thereby identifying 15 unique topics of symptoms/complaints based on users'own words. Another recent qualitative study which analyzed the discussions around tinnitus in web-based forums showed that the messages include (a) lived experiences of tinnitus, (b) perspectives on interventions for tinnitus, and (c) experience of participating in web-based discussion forum (Hall et al., 2021). Overall, these studies suggest that internet health information, especially the information about tinnitus on social media, helps identify questions, concerns as well as complaints of individuals with tinnitus. Moreover, online discussions including social media posts are a reflection of the free speech of individuals with tinnitus which

is not biased by any medical interaction. However, despite many recent studies in this area, we believe that there is limited understanding of what kind of discussions happen in social media. This is because, previous studies were limited in the amount of information extracted for analysis. For example, studies by Manchaiah et al. (2018b) and Palacios et al. (2020) only included 1,569 and 3,785 message, respectively in the analysis while thousands of messages are posted in social media platforms each day. Moreover, previous social media platforms were not anonymous due to which some users may not have discussed their opinions freely. Social media platforms such as Reddit provide an anonymous way for users to share their messages which may be more conducive for free speech online. However, no previous studies have examined tinnitus information in Reddit platform.

The present study was aimed at identifying key topics in online discussions about tinnitus by examining a large social media dataset using NLP technique. This study uses a much larger sample size as well as uses a different NLP technique than the study by Palacios et al. (2020). We anticipated that the examination of anonymous free texts would provide us an understanding of user perception of causes and consequences as seen in previous studies (Manchaiah et al., 2018b; Palacios et al., 2020), and most importantly the coping mechanism both in terms of seeking and offering support.

Method

Study Design and Ethical Considerations

The study used a cross-sectional design. Reddit posts about tinnitus were extracted from the Reddit Application Programming Interface (API) via custom-made Python scripts. Ethical

approval (IRB-FY21-89) for the study was obtained from the Institutional Review Board (IRB) at Lamar University. In addition, good practices for analyzing and reporting publicly available data were followed (Eysenbach & Till, 2001).

Data Extraction

While Reddit is a considered as a social media, it is quite different from other social media platforms. Reddit platform is based around communities (i.e., subreddit) which focus on specific topic rather than people. As the Reddit platform ensures anonymity of the user identity no user information (e.g., age, gender, name) was available. As a result of this, it functions more like a forum with social media aspects, which could potentially make it more friendly and easy to use. In addition, in the current study we analyzed text message as a whole text corpus rather than looking at individual messages. Only small segment of individual messages was used as examples in the results section.

Reddit API is publicly accessible and allows researchers to acquire language data directly from the site without using the typical web interface. However, the main limitation to this approach is that Reddit does not provide demographic data on the site's users; hence, characteristics of the sample cannot be described. Therefore, Reddit posts about tinnitus were extracted via Reddit API using a custom-built script. This included both original posts (i.e., a submission that starts a conversation) and associated comments (i.e., a submission that replies to posts or other comments) from the r/tinnitus subcommunity (or "subreddit"). Entire thread histories were extracted and, where possible, all original post-level information (i.e., exact message that was posted) was retained. In cases where comments remained but user-level information (i.e., Reddit

user ID) had been removed, all available data were retained. Data posted from the time Reddit started through March 2020 were extracted chronologically, although the timestamp information was not extracted.

The original corpus included 113,214 posts, which included 12,481 (11%) original posts and 100,733 (89%) comments. However, after cleaning the data for duplications and posts without text information, the text corpus was reduced to 101,905 posts which was used in further analyses. These messages were generated from 12,313 unique user accounts. The mean number of posts was 8.3 (SD 22 posts). Of these, 11,573 (11.4%) were original posts and 90,332 (88.6%) were comments which were generated from 12,303 user accounts. No other preprocessing was performed.

Data Analysis

While there are several types of techniques used in NLP, topic modelling is one of the most commonly used NLP techniques. Moreover, this approach produces results in a similar way to qualitative content analysis or thematic analysis making it easy for hearing healthcare professionals to understand the results as they are familiar with qualitative analysis techniques. Topic modelling is a technique that involves text mining algorithms to identify patterns within the data (Nunez-Mir et al., 2016). This method examines how words commonly occur to produce text with similar meaning. In this study, hierarchical cluster analysis (HCA) was performed using IRaMuTeQ software (IRaMuTeQ, 2021) which is a type of topic modelling. This analysis helped identify key topics within the text corpus. The software produces a dendrogram based on the cluster analysis which shows the most representative words within each cluster as well as the

percentage of texts each of the cluster included from the whole corpus. Latent Dirichlet Allocation (LDA) is another popular topic modelling technique to extract topics from a given text corpus. Unlike LDA which uses words as the unit of analysis, the HCA uses the text segments (i.e., 1-2 sentences) as the unit of analysis. For this reason, the chosen approach helps identify sub-clusters within the main cluster providing both macro and micro view of the data. Table 1 provides description of key steps in the data analysis.

Following the HCA, the top 20 highly representative words for each cluster was provided to a panel of tinnitus experts (n=9) and they were requested to come up with a name for each cluster based on the frequently occurring words. The details were sent to experts in an email and were requested to respond to the email with appropriate names for each of the clusters. The responses were grouped to identify the most appropriate name. Researchers (VM and AL) reviewed the names to come up with the final names. Any disagreements were discussed with the third researcher (AKD) before coming to consensus. Following this step, the typical text segments from each cluster (see Table 1) was compared with the name of the cluster. It is noteworthy that the cluster names themselves are arbitrary, using domain experts in labelling the clusters helped us ensure the cluster names represent the intended meaning of the social media posts and is common topic modelling approach (Boyd et al., 2017; Kennedy et al., 2021).

Table 1: **Steps involved in Hierarchical Cluster Analysis (HCA) of text data**

Steps	Description
Step 1: Text segmentation	Split each unit of text (i.e., single Reddit post) into smaller unit of text segments (generally 1-2 sentences) based on criteria od size and punctuation. This step is important to maintain precision of analysis while maintaining natural segmentation of texts.
Step 2: Lemmatization	Words within the text corpus is sorted by grouping variant or inflated forms of the same word called as "lemmas.". At this stage, "full words" (e.g., verbs, noun, adjectives, and adverbs) and "tool words" (e.g., pronoun, determent, and useful verbs such as—to be and to have) and only the full words are included in the analysis.
Step 3: Generation of sub-corpus	The text segments related to topic of interest (i.e., tinnitus) were extracted to form and a more directed sub-corpus was formed for further analysis.
Step 4: Applying Reinert method for text analysis	The software builds a binary matrix with text segments in rows and full words in columns, and then performs HCA based on a series of bipartitions made with correspondent analysis. At each stage of classification, a larger text segment is divided into two parts based on correspondent analysis until no further distinction is made in terms of information extracted from each cluster or until a maximum set number is reached. The maximum number of allowable clusters was set to 30. This cluster analysis groups the text segments based on co-occurrence of lemmas. Each of the clusters aim to be homogeneous (regrouping text segments with the common pattern of lemmas). Finally, the software produces results in terms of dendrogram and the software also provides profiles typical text segments from each of the clusters. The software uses a Reinert method cluster analysis which is practically developed for analysis of text data. A detailed description of Reinert method cluster analysis of textual data using the IRaMuTeQ software has been provided in our earlier publications (Manchaiah et al., 2018b, 2019b, 2019c).

Results

Main Clusters

The cluster analysis resulted in a 16-cluster solution as illustrated in Figure 1. Here, the

dendrogram presents the words that are most representative of each cluster. The font size of

words within each cluster is proportional to the χ^2 value within the cluster (i.e., higher the χ^2,

larger the font size). The largest cluster contained 10% of the Reddit posts in our sample, and the

smallest cluster contained 1.4% of the posts. Table 2 presents the cluster names, descriptions,

and examples of text segments that are typical representation of these clusters.

The 16 clusters were grouped into two topics relying on experts' assessments: (1) *causes and*

consequences (55.7% of texts), and (2) *management and coping* (44.3% of the texts). The *causes*

and consequences topic was primarily characterized by Reddit users discussing medical triggers

and modulators of their tinnitus, dietary effects of tinnitus, tinnitus perception (e.g., pitch,

loudness, maskability), sleep-related issues, association between tinnitus and

attention/concentration, discussions around the onset of their tinnitus as well as how long it may

last, and also discussions around music listening and hearing protection. The *management and*

coping topic, on the other hand, was largely defined by discussions around Reddit

users''interactions with their hearing healthcare providers, general discussions around social

media posts about tinnitus, hoping for a cure, dietary supplements and alternative therapies,

hearing research, interactions with people who do not experience tinnitus and their limited

understanding of tinnitus, mental health issues and coping, as well some appreciation toward the

online community for answering their questions or concerns and for being supportive.

Table 2: Cluster description and example text segments

Cluster	Description	Examples
Cluster 1: Medical triggers and modulators (8.8%)	Discussions around medical causes, triggers and modulators of tinnitus especially in the head, face, and neck area.	i also seem to have eustachian tube issues ears sometimes feel plugged after blowing my nose and i need to sort of yawn move my jaw around to get them to clear so i guess I'll get checked for any wax buildup and mention the tinnitus congestion issues while I'm at it yeah my left ear feels the same way been having issues with it for a year it could be due to jaw neck muscle putting pressure on eustachian tube which makes tinnitus more noticeable
Cluster 2: Dietary effects (1.7%)	Discussions around dietary items (food and beverages) that may affect tinnitus.	some people have complained smoking weed has had an negative effect on tinnitus so try and avoid that drinking alcohol or caffeine beverages is worse though if yours is getting worse and you don't think its from any of those it might be coming from noise exposure I've been told from doctors that caffeine and alcohol can make it worse and that exercise can help make it better i have experienced the reduced tinnitus from exercise i need to quit drinking coffee and start exercising
Cluster 3: Interaction with hearing healthcare providers (6.7%)	Discussions around hearing health visits including why to seek medical help, what might happen during the visit, as well as the role of audiologists and ENT specialists.	until a cure is found there are tons of medical things you can do to help yourself out firstly book an appointment with an audiologist or an ent doctor they will do an assessment of your symptoms and tell you where to go from there it might be and it might fade i would schedule an appointment with an ent doctor or audiologist as soon as you can because there can be a wait to get in and you want to get it evaluated within two weeks
Cluster 4: Tinnitus perception (9.7%)	Discussions around the perceptual characteristics and masking of tinnitus.	i have a low pitch but high frequency humming sound in my left ear and a high pitch and low frequency ringing in my right ear i also have a static like white noise in both of my ears but it feels like it's coming from the back of my head yes sometimes a white noise will help with the buzzing but will not mask the high pitch tone i hear in the background i can only describe the different tones as constant buzzing throbbing pulsing or electronic high pitch tone
Cluster 5: Writing about tinnitus and being thankful to online community (1.4%)	Writing about tinnitus on social media including the quality of writing, difficulty in writing in English, and being thankful to the online community for their responses.	wow thank you for this in depth reply I'm glad you have found ways to cope and make your situation better i went to the ENT today and they ran all their tests and i don't have any signs of hearing loss or any abnormal ear pressure ok I'll have my doc set up a referral that's awesome how the research is advancing from what I've read tinnitus isn't understood very well so that's great that they are getting somewhere thank you for replying

Cluster 6: Social media posts about tinnitus (7.4%)	Positive or negative views or experiences about tinnitus social media posts.	when i was having a rough time it helps to read all of the success stories on here but be careful because there's a lot of negative posts from the minority also if i were you i would go get a couple different opinions from ENTs stop reading those bad stories lol there are more good ones a lot of people get tinnitus and never post online i wouldn't rely too much on those but habituating really isn't too bad stay off the forums and mask it a lot you should be ok
Cluster 7: Sleep (3.9%)	Discussions around sleep-related disturbances due to tinnitus as well as strategies to have better sleep.	unless I'm super tired it's tough for me to fall asleep try not to disturb my wife by having the tv loud all night i wake up in the middle of the night and the T is the first thing i hear think of and trouble getting back to sleep yesterday i took a xanax when i got home and i took a xanax right before bed because i really needed to sleep that night and i was able to fall asleep pretty easly and i woke up at almost 8 but still before my alarm went off
Cluster 8: Attention and silence (8.6%)	Discussions around effects of tinnitus on attention, concentration, memory, and lack of silence.	I'm obsessive and not focusing on the tinnitus is difficult i am a quiet person who lives alone and enjoys silence and now silence is distressing it is hard to meditate anymore as all i think about is the ringing I'm hearing been having tinnitus for 17 years and smoking weed for 10 years daily as long as there is sound i dont focus on it and never hear it more or less just help not thinking about it the only time I'm bothered by it is in complete silence
Cluster 9: Tinnitus timeline (10%)	Discussions around the onset and duration of tinnitus as well as how long it may last.	just left the doctor i posted a few days ago I've had tinnitus for about 18 months but this past week it increased ten fold i was terrified that it was going to be a permanent change in my tinnitus this all started months ago and it started with my left ear muffled and ringing it went away after a couple days the ringing then became periodic it would come in stay for a few seconds and fade out
Cluster 10: Hope for a cure (5.6%)	Users hoping for a cure in the near future or at least in their lifetime.	thank you man i hope yours gets better i wouldn't wish tinnitus on my worst enemy i pray for a cure in the near future for you'll who are still suffering good luck man all you can do is to follow the things usually prescribed to help t and then hope for a cure in our lifetime man I'm thinking about killing myself daily a lot of it is due to
Cluster 11: Hearing protection (7.3%)	Discussions around importance of hearing protection during recreational activities by minimizing exposure to loud sounds by avoiding loud events or with the use of ear protection.	don't go to concerts or shoot guns even with ear plugs for the time being to see if it starts getting better don't listen to loud music don't wear headphones don't play instruments without earplugs 3 anyone with tinnitus should always keep earplugs on your keychain and yes you should wear them at the movies i recommend custom plugs which are 200 for the best protection and sound quality but some decent ones that are at least 15db will sound better than foam
Cluster 12: Music listening (5.7%)	Discussion around music listening - both in terms of avoiding listening to music for hearing conservation as well as using music to aid distraction and sleep.	in regards of music do not shove it away make use of it i personally stopped listening through headphones and bought myself a pair of speakers which i play at low volume the only method i have of getting to sleep is having speakers next to my bed playing music in the night and i have also recently bought over ear headphones the earmuff kind which i listen to on much lower volume

|		
Cluster 13: Dietary supplements and alternative therapies (3.2%)	Discussion around the benefits, limitations and evidence toward dietary supplements and alternative therapies for tinnitus.	I'd take looking like a doofus over a tinnitus spike or permanent increase in tinnitus anyday of the week i_ll have to look into nac since the onset of my tinnitus i_ve been taking b12 vitamin and a calcium magnesium zinc supplement daily after a few weeks i noticed huge difference also helps with falling asleep supplements and vitamins I've found to work all available on amazon cheap chelated magnesium citrus bioflavonoids taurine zinc
Cluster 14: Hearing research (8.8%)	Discussions around the latest hearing research, especially on tinnitus assessment and management.	the findings have implications for a phase 1 clinical trial now underway that uses gene therapy to restart expression of atoh1 to regenerate hair cells for treatment of hearing loss is it phase 1 of audion therapeutics trial a whole group of companies are racing for that sweet tinnitus money the area of development is mainly hearing restoration by growing cochlea hair cells some of the treatment delivery methods being explored are injection drug pill delivery or perhaps even deep brain stimulation of the auditory cortex
Cluster 15: Interactions with non-tinnitus people (5.4%)	Discussions around interactions with non-tinnitus people, others (e.g., friends, family members, and people in general) not understanding the consequences of tinnitus as well as how tinnitus affects psychology and social interactions.	people sometimes do commit suicide over this and your family friends need to understand that they should not be minimizing what you are experiencing just because they can't understand it few truly understand what it means to lose the will to live tinnitus took that from me and altered my life in a negative way permanently we need a real treatment and real help not blanket dismissal some people suffer in silence we don't even get that
Cluster 16: Mental health and coping (5.8%)	Discussion around consequences of tinnitus on mental health and ways of coping with it.	what do you mean you don't have a choice you can choose to accept the fact of what has happened and learn to deal with it and enjoy life or you can choose to let it rule you and think negative thoughts about it constantly i have gone through depression and anxiety issues because of this and it directly affects my job and many many more aspects of my life i can tell you that you just learn to cope with it the obvious solution to me is a change in lifestyle i probably just need to get out more but depression and anxiety issues tend to make you a recluse so it isn't going to be an instantaneous change
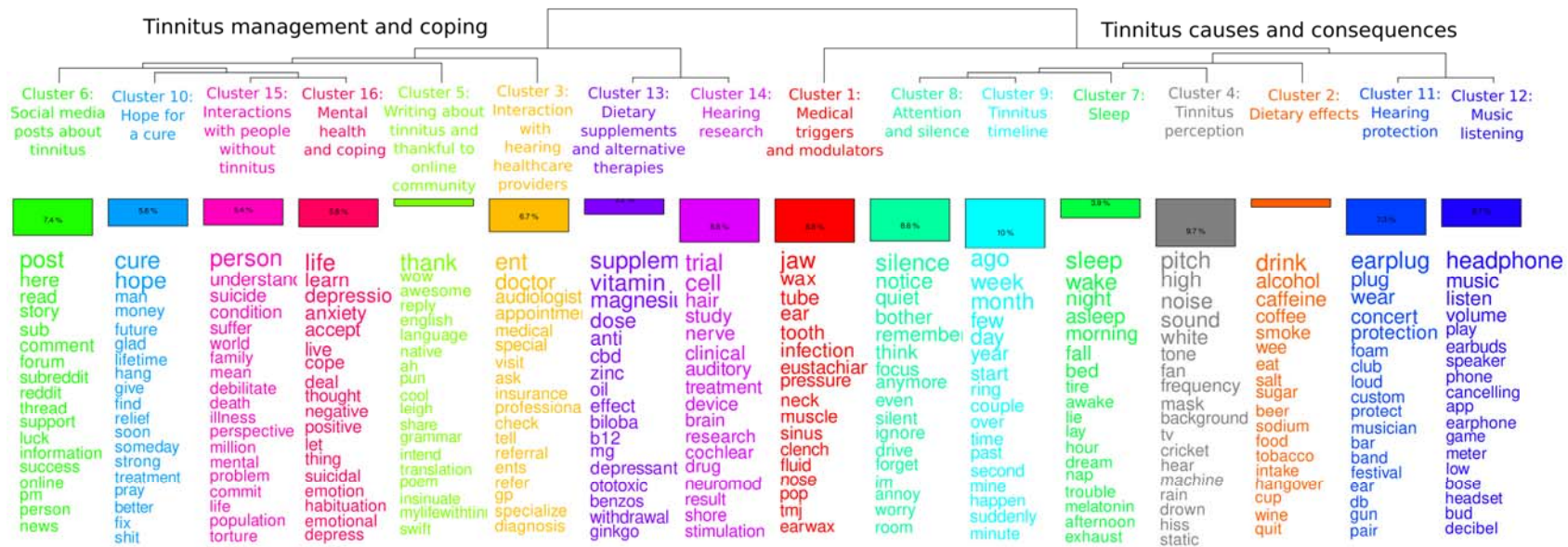
Figure 1: Dendrogram (i.e., classification of clusters), size of clusters as percentage of the text segments, and overrepresented words in each cluster in Reddit posts about tinnitus.

Discussion

The present study examined the main topics in Reddit posts about tinnitus by analyzing users ’ own words using automated text analysis. The study identified 16 key clusters within social media posts about tinnitus, which fell into two main topics (i.e., tinnitus causes and consequences, tinnitus management and coping). These findings help understand the types of discussions taking place on social media, although it may not be possible to examine why individuals use social media for accessing hearing health information. The following sections highlight the key findings and implications.

Most of the discussions (i.e., 55.7% of texts) about tinnitus on Reddit were concerned about the causes and consequences of tinnitus. The dominant cluster within this topic was about tinnitus timeline in which users discussed the onset and duration of their tinnitus and, more importantly, concerns they had about how long they may have to live with tinnitus. This is an especially interesting cluster as several individuals with tinnitus are afraid that they may have to ‘live with tinnitus’ for the rest of their life (Watts et al., 2018). The cluster on medical triggers and modulators of tinnitus highlighted the complex interactions between middle ear, somatosensory tinnitus, stomatological conditions, and sinonasal and eustachian tube dysfunctions that may have triggered or exacerbated tinnitus – topics often discussed in tinnitus literature (Haider et al., 2018). Users’ descriptions of the perceptual attributes of their tinnitus in terms of type of sound, loudness, pitch, location, unpleasantness, and maskability was also a dominant cluster, accounting for nearly 10% of the posts. These descriptions are quite similar to how individuals with tinnitus would describe their problem during clinical interactions (Hall et al., 2018). Another cluster included substantial amount of discussions (13%) around how listening to loud

music could result in tinnitus and ways to mitigate it (i.e., preventative hearing conservation strategies) (Zhao et al., 2020). These posts appeared to come from individuals who developed tinnitus due to exposure to loud music and have had tinnitus for quite some time. The messages are targeted at younger users to prevent hearing damage by taking appropriate precautions. This observation highlights how social media is used for the advocacy of hearing conservation and prevention of tinnitus. Sustaining attention as well as sleep functions are often affected in those experiencing tinnitus (Manchaiah et al., 2018a), and these aspects were also included in the Reddit discussions. Interestingly, users also described how absence of silence became a major burden for them despite the fact that a low level sound stimulation is known to induce tinnitus even in non-clinical subjects (Schaette et al., 2012). Indeed, laboratory experiments have suggested a link between the perception of tinnitus and silence modulated by the auditory attention load (Knobel & Sanchez, 2008; Tucker et al., 2005). However, the discussions on Reddit were predominantly focused on how individuals with tinnitus were bothered by the loss of silence in their life. In contrast, as highlighted by cluster 4, it is interesting to note that avoiding silence and using a variety of sound enrichment techniques is proposed as a therapeutic intervention for tinnitus patients (Pienkowski, 2019). Finally, a smaller proportion of discussions centered around how certain types of food and drinks (e.g., caffeine, alcohol) affected tinnitus perception. Although a few studies have examined the link between diet and tinnitus (e.g., Dawes et al., 2020), this is an area that requires much attention from tinnitus researchers.

The examined Reddit posts also had a substantial amount of discussion around management of and coping with tinnitus (i.e., 44.3% texts). This topic included various clusters ranging from interactions, management options, coping, as well as up-to-date research about tinnitus. The

predominant cluster in this topic was that about hearing research in which users discussed the latest research in diagnosis and treatment of tinnitus. This discussion may have taken place partly due to users having come across tinnitus studies being discussed in newspaper media or other social media posts. These outlets have become popular ways of sharing newly published research (Manchaiah et al., 2018b). There was also substantial amount of discussion about social media posts themselves as many users recognized that posts that get more attention (go 'viral') are often negative, and most people's modest experiences or positive life stories are not shared enough (Vosoughi et al., 2018). Many users discussed their interactions with hearing healthcare providers as well as people without tinnitus. Posts about hearing healthcare providers were mainly about office visits and what to expect during an office visit (e.g., types of investigations that were performed, possible management). However, the posts about interactions with people without tinnitus were mainly about unhappiness that those who do not have tinnitus do not understand how it is to live with tinnitus. This observation may call for inclusion of significant others of tinnitus in the management process as well as public health campaign to raise awareness about tinnitus to general public. The posts also highlight that individuals with tinnitus continue to be hopeful about a possible cure in the near future or in their lifetime, although no cure exists now (McFerran et al., 2019). Several posts discussed relying on dietary supplements and alternative therapies, although the current evidence does not support its use for tinnitus management (Luetzenberg et al., 2020). Many individuals shared the mental health issues they were facing such as anxiety and depression and the mechanisms they use to cope with them. This was an expected finding as many people seek social media information for social support. Lastly, a small number of posts were related to how some users had difficulty writing in English and thanking the online community for being helpful.

Study Implications

Social media has become an integral part of modern human life and has an important role in health. People may use social media for various reasons including seeking information or seeking (or providing) social support. However, previous studies have identified that health information on social media could have misinformation (Deshpande et al., 2018; Deshpande et al., 2019; Swire-Thompson & Lazer, 2020). Identifying and addressing misinformation is important both at an individual level during clinical interactions as well as at public health level through public education campaigns. Because of the negative effects misinformation about health issues may create, our findings further inform the discussion about the need for regulation and moderation of information on social media.

The present study may also help uncover some important topics (i.e., tinnitus timeline, interactions with people without tinnitus, dietary supplements and alternative therapies) that individuals with tinnitus are concerned about, and that are not covered in the current literature. Moreover, examining the patient's own natural language may provide unique insights into their symptoms and complaints that may not be captured easily through standardized patient-reported outcome measures (Spasić et al., 2019). Hearing healthcare professionals should consider discussing if the individuals with tinnitus have engaged in social media discussions as well as offer to help clarify issues that were not clear in the social media. Moreover, such a discussion may also help hearing healthcare professionals to examine if social media has had any influence on their treatment decisions/choices as well as to develop patient-centered strategies for tinnitus management.

Study Limitations and Future Directions

The current study is the first to examine Reddit posts about tinnitus and has a large sample size. Moreover, due to the anonymous nature of Reddit platform, it is likely that users will share their opinions and views more openly and honestly and the data generated from this platform may be more ecologically valid. In the present study, we used a cluster analysis which is one type of NLP. It would be interesting to see if other types of NLP techniques (e.g., LDA) yield similar results (Palacios et al., 2020). Further analysis of some of the clusters using either qualitative methods and/or microanalysis may help provide in-depth understating about specific topic. The microanalysis of the data should also focus on identifying any misinformation about tinnitus in these posts (Deshpande et al., 2018). In addition, the current study was focused on grouping similar messages based on their meaning. However, it would be worthwhile to identify subgroups of users based on the types of messages they post using the Reddit ID as a factor. Such an analysis may help understand how and for what purpose different users may rely on social media platforms for tinnitus information.

Statement of Ethics

The study was approved by the Institutional Review Board (IRB) at Lamar University. As the anonymous data was extracted from social media posts no informed consent was obtained.

Conflict of Interest Statement

There are no financial interests to disclose. There are no conflicts of interest.

Funding Sources

There are no funding sources.

Author Contributions

VM conceptualized the study. RB extracted the data and PR performed the data analysis. AL, AKD, MR, GP, RB and PR provided input to the study design and provided input to draft manuscript. All authors approved the content of the manuscript.

Data Availability Statement

The data that support the findings of this study are not publicly but are available from the corresponding author [VM] upon reasonable request.

References

Baguley, D., McFerran, D., & Hall, D. (2013). Tinnitus. *The Lancet*, *382*(9904), 1600-1607. https://doi.org/10.1016/S0140-6736(13)60142-7

Basch, C. H., Yin, J., Kollia, B., Adedokun, A., Trusty, S., Yeboah, F., & Fung, I. C. (2018). Public online information about tinnitus: A cross-sectional study of YouTube videos. *Noise & Health*, *20*(92), 1–8. https://doi.org/10.4103/nah.NAH_32_17

Bhatt, J. M., Lin, H. W., & Bhattacharyya, N. (2016). Prevalence, severity, exposures, and treatment patterns of tinnitus in the United States. *JAMA Otolaryngology-Head & Neck Surgery*, *142*(10), 959–965. https://doi.org/10.1001/jamaoto.2016.1700

Boyd, R. L. (2017). Psychological text analysis in the digital humanities. In S. Hai-Jew (Ed.), *Data Analytics in Digital Humanities* (pp. 161–189). Springer International Publishing. https://doi.org/10.1007/978-3-319-54499-1_7

Dawes, P., Cruickshanks, K. J., Marsden, A., Moore, D. R., & Munro, K. J. (2020). Relationship between diet, tinnitus, and hearing difficulties. *Ear and Hearing*, *41*(2), 289–299. https://doi.org/10.1097/AUD.0000000000000765

Deshpande, A. K., Deshpande, S. B., & O'Brien, C. (2018). A study of social media utilization by individuals with tinnitus. *American Journal of Audiology*, *27*(4), 559–569. https://doi.org/10.1044/2018_AJA-18-0033

Deshpande, A. K., Deshpande, S. B., & O'Brien, C. A. (2019). Hyperacusis and social media trends. *Hearing, Balance and Communication*, *17*(1), 1-11. https://doi.org/10.1080/21695717.2018.1539321

Eysenbach, G., & Till, J. E. (2001). Ethical issues in qualitative research on internet communities. *BMJ*, *323*(7321), 1103-1105. https://doi.org/10.1136/bmj.323.7321.1103

Haider, H. F., Bojić, T., Ribeiro, S. F., Paço, J., Hall, D. A., & Szczepek, A. J. (2018). Pathophysiology of subjective tinnitus: Triggers and maintenance. *Frontiers in Neuroscience*, *12*, 866. https://doi.org/10.3389/fnins.2018.00866

Hall, D. A., Fackrell, K., Li, A. B., Thavayogan, R., Smith, S., Kennedy, V., Tinoco, C., Rodrigues, E. D., Campelo, P., Martins, T. D., Lourenço, V. M., Ribeiro, D., & Haider, H. F. (2018). A narrative synthesis of research evidence for tinnitus-related complaints as reported by patients and their significant others. *Health and Quality of Life Outcomes*, *16*(1), 61. https://doi.org/10.1186/s12955-018-0888-9

Hall, D. A., Hibbert, A., Vesala, M., Kerr, M., Harrison, S., & Core Outcome Measures in Tinnitus (COMiT) (2021). Web-based discussion forums reveal the person-centered relevance and importance of tinnitus. *Progress in Brain Research*, *260*, 205–221. https://doi.org/10.1016/bs.pbr.2020.12.001

IRaMuTeQ. (2021). IRaMuTeQ software. Retrieved October 13, 2021 from http://www.iramuteq.org/

Kennedy, B., Ashokkumar, A., Boyd, R. L., & Dehghani, M. (2021). Text analysis for psychology: Methods, principles, and practices. https://doi.org/10.31234/osf.io/h2b8t

Knobel, K. A., & Sanchez, T. G. (2008). Influence of silence and attention on tinnitus perception. *Otolaryngology--Head and Neck Surgery*, *138*(1), 18–22. https://doi.org/10.1016/j.otohns.2007.09.023

Luetzenberg, F. S., Babu, S., & Seidman, M. D. (2020). Alternative treatments of tinnitus:

Alternative medicine. *Otolaryngologic clinics of North America*, *53*(4), 637–650.

https://doi.org/10.1016/j.otc.2020.03.011

Madrigal, L., & Escoffery, C. (2019). Electronic health behaviors among US adults with chronic

disease: Cross-sectional survey. *Journal of Medical Internet Research, 21*(3), e11240.

https://doi.org/10.2196/11240

Manchaiah, V., Beukes, E. W., Granberg, S., Durisala, N., Baguley, D. M., Allen, P. M., &

Andersson, G. (2018a). Problems and life effects experienced by tinnitus research study

volunteers: An exploratory study using the ICF classification. *Journal of the American Academy*

of Audiology, 29(10), 936–947. https://doi.org/10.3766/jaaa.17094

Manchaiah, V., Ratinaud, P., & Andersson, G. (2018b). Representation of tinnitus in the US

newspaper media and in Facebook pages: Cross-sectional analysis of secondary data. *Interactive*

Journal of Medical Research, *7*(1), e9. https://doi.org/10.2196/ijmr.9065

Manchaiah, V., Dockens, A. L., Flagge, A., Bellon-Harn, M., Azios, J. H., Kelly-Campbell, R.

J., & Andersson, G. (2019a). Quality and readability of English-language internet information

for tinnitus. *Journal of the American Academy of Audiology*, *30*(1), 31–40.

https://doi.org/10.3766/jaaa.17070

Manchaiah, V., Ratinaud, P., & Beukes, E. W. (2019b). Representation of hearing loss and hearing aids in the U.S. newspaper media: Cross-sectional analysis of secondary data. *American Journal of Audiology*, *28*(1), 11–25. https://doi.org/10.1044/2018_AJA-18-0058

Manchaiah, V., Amlani, A. M., Bricker, C. M., Whitfield, C. T., & Ratinaud, P. (2019c). Benefits and shortcomings of direct-to-consumer hearing devices: Analysis of large secondary data generated from Amazon customer reviews. *Journal of Speech, Language, and Hearing Research: JSLHR*, *62*(5), 1506–1516. https://doi.org/10.1044/2018_JSLHR-H-18-0370

McFerran, D. J., Stockdale, D., Holme, R., Large, C. H., & Baguley, D. M. (2019). Why is there no cure for tinnitus? *Frontiers in Neuroscience, 13*, 802. https://doi.org/10.3389/fnins.2019.00802

Ni, J., Bellon-Harn, M. L., Zhang, J., Li, Y., & Manchaiah, V. (2020). Twitter usage using common reference to tinnitus. *American Journal of Audiology*, *29*(2), 206–217. https://doi.org/10.1044/2020_AJA-19-00055

Nunez‑Mir, G. C., Iannone, B. V., Pijanowski, B. C., Kong, N., & Fei, S. (2016). Automated content analysis: Addressing the big literature challenge in ecology and evolution. *Methods in Ecology and Evolution, 7*(11), 1262–1272. https://doi.org/10.1111/2041-210X.12602

Palacios, G., Noreña, A., & Londero, A. (2020). Assessing the heterogeneity of complaints related to tinnitus and hyperacusis from an unsupervised machine learning approach: An

exploratory study. *Audiology & Neuro-otology*, *25*(4), 174–189.

https://doi.org/10.1159/000504741

Pienkowski, M. (2019). Rationale and efficacy of sound therapies for tinnitus and

hyperacusis. *Neuroscience*, *407*(21), 120–134.

https://doi.org/10.1016/j.neuroscience.2018.09.012

Schaette, R., Turtle, C., & Munro, K. J. (2012). Reversible induction of phantom auditory

sensations through simulated unilateral hearing loss. *PloS one*, *7*(6), e35238.

https://doi.org/10.1371/journal.pone.0035238

Spasić, I., Owen, D., Smith, A., & Button, K. (2019). KLOSURE: Closing in on open-ended

patient questionnaires with text mining. *Journal of Biomedical Semantics*, *10*, 24.

https://doi.org/10.1186/s13326-019-0215-3

Swire-Thompson, B., & Lazer, D. (2020). Public health and online misinformation: Challenges

and recommendations. *Annual Review of Public Health*, *41*, 433–451.

https://doi.org/10.1146/annurev-publhealth-040119-094127

Tucker, D. A., Phillips, S. L., Ruth, R. A., Clayton, W. A., Royster, E., & Todd, A. D. (2005).

The effect of silence on tinnitus perception. *Otolaryngology--Head and Neck Surgery*, *132*(1),

20–24. https://doi.org/10.1016/j.otohns.2005.08.016

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science,
359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Watts, E. J., Fackrell, K., Smith, S., Sheldrake, J., Haider, H., & Hoare, D. J. (2018). Why is
tinnitus a problem? A qualitative analysis of problems reported by tinnitus patients. *Trends in
Hearing*, *22*, 2331216518812250. https://doi.org/10.1177/2331216518812250

Zhao, F., Manchaiah, V. K., French, D., & Price, S. M. (2010). Music exposure and hearing
disorders: an overview. *International Journal of Audiology*, *49*(1), 54–64.
https://doi.org/10.3109/14992020903202520