Readability and quality of online information regarding tinnitus in the Spanish language

Christina M. Greer

A thesis submitted in partial fulfilment

of the requirements for the

Degree of Master of Audiology

at the University of Canterbury

2019

Acknowledgements

I would like to thank my supervisor Dr. Rebecca Kelly-Campbell, without whom this or any other thesis would not have been possible. Your exemplary work ethic, organisation and research skills, alongside your support and understanding have made it a pleasure to work with you this past year. I would also like to thank my family and partner, for all their support during the tougher moments of this journey and especially my Dad for always being willing to proof read my work. Finally, thank-you to all my classmates for all the help, moral support and laughs over the past two years. Good luck to you all!

Abstract

Background: Tinnitus is a common symptom of multiple different audiological impairments which can lead to negative psychological ramifications for some people. Given the link between low health literacy and poor health outcomes, it is important to analyse the information available to the public regarding tinnitus. As more people turn to the internet as their initial source of health information, analyses of quality and readability of tinnitus information available online is a necessary direction for current research.

Although Spanish is the second most spoken global language, few published health studies on readability and quality have investigated online Spanish-language health information, thereby creating a need specifically regarding the topic of Spanish-language tinnitus information.

Method: The terms *tinnitus*, *acúfenos*, *pitido en los oídos*, and *zumbido en los oídos* were typed into the search engine Google, for each of the 24 different country-coded Top-Level Domains (ccTLDs), and the first 10 webpages were saved for analysis. Four readability formulas (RFs) were used to assess readability: SOL, Rate Index, Gilliam-Peña-Mountain Graph and Crawford. The presence of HON code certification was recorded and the DISCERN tool was used to analyse the quality of the individual webpage.

Results: Forty-four Spanish language webpages on tinnitus were analysed. Of this sample, the mean reading grade level (RGL) was 9.58, a score which significantly exceeds the recommended 6th grade level. Only three webpages displayed HON code certification, and in general the quality of the sample was low with a mean DISCERN score of 2.20.

Conclusions: Online Spanish-language information on tinnitus is generally not accessible to a wide audience because of poor readability. The general low quality of the information could mean the information Spanish-speakers are accessing is potentially erroneous, misleading or biased. Clinician and web-developers need to be aware of health

literacy issues such as readability and quality and use validated instruments, such as DISCERN, HON code and readability formulas to try to improve online Spanish tinnitus literature.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ANOVA = Analysis of Variance

CBT = Cognitive Behavioural Therapy

ccTLDs = country-coded Top-Level Domains

EU = European Union

GPM = Gilliam-Peña-Mountain

HON = Health on the Net

ICC = Intra-class correlation coefficient

LIX = *Lasbarhetsindex*

RF = Readability Formula

RGL = Reading Grade Level

RIX = Rate Index

SAM = Suitability Assessment of Materials

SMOG = Simple Measure of Gobbledygook

UNASUR = Union of South American Nations

URL = Uniform Resource Locator

WHO = World Health Organization

1. Introduction

1.1 Overview

Hearing impairment affects roughly 466 million people worldwide, with disproportionate representation in populations over 65 years of age (World Health Organization, 2018). Tinnitus, the perception of sound which is not present in the physical environment, is a common symptom of hearing loss and other audiological impairments (Nondahl et al., 2011). As life expectancy increases globally, the number of people with hearing loss and tinnitus is likely to also increase (World Health Organization, 2018).

Good health literacy is crucial for improving health outcomes and quality of life (DeWalt, Berkman, Sheridan, Lohr, & Pignone, 2004), and this likely holds true in the field of audiology. Improving health literacy can lead to improved audiological outcomes such as the treatment and prevention of hearing loss (Reavis, Tremblay, & Saunders, 2016). However, it is difficult to achieve good health literacy when much of the information that people encounter regarding their audiological needs is either not readable, of poor quality, or both.

The continual growth of the internet and access to information has changed the way people search for and consume information (McInnes & Haglund, 2011). People have become more dependent on internet resources as their primary source of information. This is especially true in the case of healthcare, which is commonly cited as one of the main reasons that people use internet search engines (Ritchie, Tornari, Patel, & Lakhani, 2016).

However, it is difficult to verify much of the information available on the internet, and previous studies have found that the quality and readability of webpages tends to be poor (e.g. Berland et al., 2001; Laplante-Lévesque, Brännström, Andersson, & Lunner, 2012;

Manchaiah et al., 2018). By contrast, little research on texts available in the Spanish language appears to have been published (e.g. Castillo-Ortiz et al., 2017).

This chapter reviews the literature pertaining to the readability and quality of information available online in Spanish regarding tinnitus and the need to critically analyse this information. Readability and quality will be addressed by discussing the importance and relevance of easy to read, high quality information available in Spanish language webpages. Hearing impairment and tinnitus will be explained, as will the relevant topics of health literacy, and access to information. Finally, the literature regarding readability and different readability formulas as well as quality and the DISCERN tool will be reviewed. This information will lead to a rationale of the current study, and the aims and hypotheses will be stated.

1.2 Spanish language

Spanish is a global language with some 480 million native speakers (Instituto Cervantes, 2017). It is the principal language of communication of the majority of the populations in 20 different states or countries. Additionally, it is a significant minority language in four other nations and territories, including the USA and Gibraltar. It is recognized as an official language in multiple international organizations, including but not limited to the United Nations (UN), the European Union (EU) and the Union of South American Nations (UNASUR). In terms of first-language speakers, it is the second most spoken language in the world (Instituto Cervantes, 2017), after Mandarin Chinese.

The number of people using the internet in Spanish has also increased, accounting for 7.7% of internet users (Instituto Cervantes, 2017). Castillo-Ortiz et al. (2017) argued the need to analyse readability and quality of online information on rheumatoid arthritis because

of the 103.8 million Spanish-speaking internet users in Spain and Mexico alone. Similarly, in the case of audiological related issues, the vast majority of research into readability and quality of health-related information has been conducted in and about English. This is despite the high prevalence of the Spanish language globally. Fackrell, Hoare, Smith, McCormack, and Hall (2012) evaluated the content and quality of English language tinnitus websites recommended by general practitioners, while Manchaiah et al. (2018) examined the readability and quality of tinnitus webpages available in English. However, to the best of the author's knowledge there are no comparable published studies that have analysed the readability and quality of information regarding tinnitus in the Spanish language. Given the high incidence (5.3%) of hearing impairment (World Health Organization, 2018) and tinnitus (10-15%, Henry, Dennis, & Schechter, 2005) in the general population, it can be reasonably assumed that there is a large number of Spanish-speaking internet users that search for information regarding tinnitus. It is, therefore, important to analyse the quality and readability of the information available to those people.

1.3 Hearing impairment

Hearing impairment (HI), also termed hearing loss, is a chronic health condition (Dalton et al., 2003; Gates & Mills, 2005), affecting millions of people globally (World Health Organization, 2018). An individual can be affected from birth (congenital HI), or can acquire a HI at any point during their life. Causes can be genetic, environmental (such as excessive noise), due to medications (ototoxicity), or trauma (Eggermont, 2017). HI can affect individuals at any age, but those older than 65 years are affected more than other populations (World Health Organization, 2018). Permanent HI can negatively affect a person's ability to communicate, as well as their quality of life (Convery, Hickson, Meyer, & Keidser, 2018; Dalton et al., 2003). Health literacy was found to be a crucial factor for managing the effects of HI in everyday life (Convery et al., 2018).

1.4 Tinnitus

Tinnitus can be defined as the perception of sound which is not present in the physical environment (Bauer, 2018; Henry et al., 2005). Tinnitus can manifest itself in a variety of ways, some of the more common being whistling, ringing, buzzing, hissing or even cicada-like sounds perceived in the ear (Adjamian, Sereda, & Hall, 2009; Bauer, 2018).

Tinnitus is not considered a disease itself, but rather a symptom of other diseases and health conditions (Hoare & Hall, 2011; Nondahl et al., 2011). It is heterogeneous in nature, and has been attributed to a wide range of audiological and medical conditions (Fackrell et al., 2012). The most frequent is sensorineural hearing loss (Hoffman & Reed, 2004; Savastano, 2008). This is reflected in the increase of prevalence of people experiencing tinnitus with age (De Brito Macedo Ferreira, Novaes Ramos Júnior, & Pereira Mendes, 2009). As people get older and are more likely to experience HI due to ageing (presbycusis), they are also more likely to experience tinnitus. However, tinnitus is not restricted to age-related HI, but also derives from other types of HI. It is commonly reported by those who have experienced excessive noise exposure (Henry et al., 2005; Nondahl et al., 2011), trauma to the head and ears (Henry et al., 2005; Tyler et al., 2008), the use of certain drugs and medication, which can be ototoxic (Sataloff, Sataloff, & Lueneburg, 1987), and even wax impaction in the ear canals (McFadden, 1982). Unilateral tinnitus may also be a clinical indicator of Ménière's disease or vestibular schwannomas – benign tumours that develop on the 8th cranial nerve (Bauer, 2018; Sataloff et al., 1987). Non hearing-related causes of tinnitus can include neck injuries (Henry et al., 2005; Sindhusake et al., 2003) and other medications such as aspirin (Adjamian et al., 2009) as well as cardiovascular disease (Heller, 2003; Nondahl et al., 2002).

Although tinnitus is treated as symptomatic, it can have a significant impact on the quality of life of an individual, due to their perception of its severity. There is great variety in

the ways which tinnitus can affect a person's daily life, mental and physical function (Bartels, Middel, van der Laan, Staal, & Albers, 2008). The degree of an individual's HI does not predict if or how they might experience tinnitus (Savastano, 2008). There is evidence of a multifaceted relationship between the "psychological, psychosocial and environmental factors and personality traits" (Adjamian et al., 2009, pp. 15-16). When tinnitus is severe, comorbidities such as anxiety and depression are common (Bartels et al., 2008) and can also include insomnia (Savastano, 2008). Negative ramifications caused by tinnitus can include disturbed sleep patterns, lack of concentration in both professional and social activities as well as disrupted emotional balance (De Brito Macedo Ferreira et al., 2009; Savastano, 2008).

The cause-effect relationship between tinnitus and these co-morbidities is not well-understood. Evidence has been found supporting different theories. For example, mental health issues may exacerbate tinnitus, and the onset of tinnitus during an emotionally difficult or stressful time of life could contribute to the development of chronic tinnitus (Bartels et al., 2008). Alternatively, the presence of tinnitus may increase the severity of mental health issues such as depression or anxiety (Bartels et al., 2008; Nondahl et al., 2011). Other research suggests that a vicious circle can develop between tinnitus and other common co-morbidities, where one or more of these symptoms exacerbates the others (Folmer, Griest, & Martin, 2001).

Prevalence of tinnitus is difficult to determine due to its subjective nature (Sataloff et al., 1987). However, tinnitus has been estimated to affect 10-15% of the population at some point in their lives (Henry et al., 2005; Yang et al., 2018). Of those affected, around 1 in 5 experience tinnitus in a form severe enough to seek treatment (Adjamian et al., 2009).

Because tinnitus can originate from a wide variety of underlying conditions, there is no treatment or management plan that will be fully effective for all, or even a majority of people experiencing it (Theodoroff, Schuette, Griest, & Henry, 2014). If HI is the cause, many people with tinnitus find that treating the HI, for example with hearing aids, will also help reduce the perception of tinnitus (Del Bo & Ambrosetti, 2007; Folmer et al., 2001). For people who experience tinnitus alongside anxiety or depression but without HI, referral to a psychologist would be appropriate, where the management plan would involve teaching coping strategies through different therapies such as Cognitive Behavioural Therapy (CBT) or sleep management (Adjamian et al., 2009).

Given the complexities and uncertainties that surround tinnitus and its treatment, it is crucial for the information available to those who experience tinnitus to be accurate as well as understandable. Information empowers people to acclimatize and better cope with tinnitus, alleviating some of its severity (Langguth, Kreuzer, Kleinjung, & De Ridder, 2013). However, even tinnitus websites recommended by general practitioners are not comprehensive sources of information (Fackrell et al., 2012). It is, therefore, important to analyse what tinnitus information is available to the general public, to highlight problem areas with the aim of improving and providing easy-to-read, quality information that will ultimately benefit information-seekers. The importance of good information will be discussed in the following sections.

1.5 Health literacy

Health literacy has been defined in New Zealand as “the capacity to obtain, process and understand basic health information and services in order to make informed and appropriate health decisions” (Ministry of Health, 2010, p. 1). The World Health Organization (WHO) defines it as being able to “gain access to, understand and use information in ways which

promote and maintain good health" (World Health Organization, 2009, para 1). Health literacy is not limited to the ability to read health-related information, but utilises all the skills required to "effectively search for, interpret and use health information to navigate the various levels of the health care system" (McInnes & Haglund, 2011, p. 174).

The concept of health literacy is a relatively recent one. Evidence emerged in the USA in the 1990s, indicating the pervasiveness of low literacy skills and the connection to poor health outcomes (DeWalt et al., 2004). This relationship is imperative to also understand the true causes of poor health outcomes, identify any clinical indications of patients at risk of worse health outcomes and inform the development of health-related treatment and prevention strategies (DeWalt et al., 2004). The WHO has identified literacy as a key factor in determining health inequities (2013). A systematic review by DeWalt et al. (2004) confirmed that low literacy is associated with adverse health outcomes. The relationship between literacy skills and health literacy will therefore be discussed.

The literature on health literacy can be divided into two separate schools of thought. The first is that health literacy can be interpreted as a risk factor for poor health outcomes, through analysing the relationship between literacy skills and health (Nutbeam, 2008). Following this line of thinking, achieving high literacy skills across the population becomes the goal by which public health benefits can be realised (Nutbeam, 2008). Improving access to effective schooling and providing further adult education for those who slipped through the cracks are possible responses to health literacy as a risk factor.

The second and alternative conceptualization is that health literacy is an asset (Nutbeam, 2008), rather than a risk factor. In this interpretation, health literacy is a concept distinct from literacy and numeracy. Health literacy is the outcome, instead of a factor that affects the outcome (Nutbeam, 2008). The focus is for the individual to develop knowledge

specific to their age and their health status. In doing so, the individual should develop the self-efficacy to use that knowledge to exert greater control over their health and health-based decisions (Nutbeam, 2008). It is this school of thought that will be adopted in the present study.

At present, there is no gold standard available to measure an individual's health literacy (Berkman, Sheridan, Donahue, Halpern, & Crotty, 2011). Figures will vary, depending on the measure being used. Literacy skills and health literacy are distinct concepts; although a clear link exists between the two. Both schools of thought (risk factor or asset) have merit; society should aim to improve literacy skills to ensure individuals' self-efficacy in the healthcare system. However, the healthcare system also needs to assume some responsibility, and should aim to provide materials that are comprehensible to the widest possible audience. In doing so, health literacy will be supported.

Literacy skills are known to be remarkably poor for the majority of adult populations in many developed countries and estimates of illiteracy are much higher in developing countries (Roser & Ortiz-Ospina, 2013). Greater functional literacy allows members of society to participate more fully, both socially and economically; allowing them to exercise greater control in their everyday activities and decisions (Nutbeam, 2008). Low literacy skills are commonly linked to lower socioeconomic status, as well as reduced use of health information and services (Sudore & Schillinger, 2009). Estimates of health literacy are also low in developed nations, including New Zealand (56.2% adults have low health literacy (Ministry of Health, 2010)), Australia (59% did not satisfy minimum requirements to meet the demands of life and work in the health domain (Australian Bureau of Statistics, 2006)), and USA (36% had basic or below basic health literacy (Kutner, Greenburg, Jin, & Paulsen, 2006)).

Literacy and health literacy cannot be assumed from years of schooling. As Baker et al. (2002) and Baker et al. (2007) discussed, years of schooling is an imprecise measure of education. Baker and colleagues (2007) found that years of schooling was a weak predictor of mortality in older adults, while low health literacy was a strong predictor. Reasons for this could be that progressing through the years of school does not automatically ascertain that educational goals are met, including the ability to read at an appropriate age or grade level (Baker et al., 2007). Moreover, an older adult's ability to deal with the health care system is less related to the amount of time they spent in school than the lifelong learning they have attained since leaving school (Baker et al., 2002; Baker et al., 2007). Measuring years of schooling does not factor in age-related declines in reading fluency, or the impact of reduced cognitive function in older adult (Baker et al., 2007).

Good health literacy is crucial for better health and quality of life. Poor health literacy has been shown to be the single biggest predictor of poor health outcomes, regardless of age, sex, socio-economic status, race, or education level (McInnes & Haglund, 2011; Weiss, 2003). Low health literacy is associated with decreased understanding of health-related information, in turn leading to poorer health behaviours and resulting in worse health outcomes (Donald & Kelly-Campbell, 2016). Examples include the inability to understand and follow medication labels and health messages appropriately (Berkman et al., 2011) or even understanding basic concepts of common diseases (Weiss, 2003).

Lower levels of health literacy correspond to lower levels of quality of life and higher mortality rates (McInnes & Haglund, 2011), which can be twice as high in elderly patients (Baker et al., 2007). Those with lower health literacy are more likely to seek health care in emergency departments (DeWalt et al., 2004), require hospitalization (Berkman et al., 2011; DeWalt et al., 2004) and are less likely to access screening and preventative services such as immunizations (Sudore & Schillinger, 2009). An economic burden results from the additional

health care expenditures linked to low health literacy: a recent report estimated costs in the USA due to low health literacy to be $106-238 billion annually (Vernon, Trujillo, Rosenbaum, & DeBuono, 2007). Given these figures, a strong economic incentive exists to improve the health literacy of all patients, regardless of their reading ability (Baker et al., 2002).

According to the WHO, health inequality is common to all corners of the globe, and rates of disease and poor health are highest in poorer areas (World Health Organization, 2013). Steps to improving health literacy globally require action at various different levels of society, including attention from policy makers, clinicians and individuals (Berkman et al., 2011). Interactions between the health and education sectors at the local, national and international level are also important in the pursuit of improved health literacy (Nutbeam, 2000).

In the interactions between health system and patient, interventions aimed at low health literacy might include well-designed, user-friendly health education materials which are easy to read (Sudore & Schillinger, 2009). It is recommended that information is kept short and succinct, uses a large font, makes use of helpful pictures and diagrams and the writing is set at a 5th – 6th grade reading level or lower (Doak, Doak, & Root, 1996; Sudore & Schillinger, 2009). These simple techniques should be utilised for information for all patients, regardless of their health literacy levels. Studies have shown easy-to-read materials benefit all reading levels, and act as a step in empowering patients to be more active participants in their health care (Baker et al., 2002; Sudore & Schillinger, 2009).

However, as the dominance of and dependence on the internet grows, more and more people are turning to the internet and search engines as their initial source of health inquiries,

and utilising this information to help inform decision-making regarding their own health care (Ritchie et al., 2016).

1.6 Access to online health care information

Internet penetration has been increasing progressively over the last two decades. According to the Internet World Stats: Usage and Population Statistics (2018), 54.4% of the global population had access to the internet as of December 31st 2017. The geographical regions with the highest internet penetration rates are North America (95%) and Europe (85.2%). While all the regions' penetration rates have risen considerably since the year 2000, regions such as Africa, the Middle East and Latin America/Caribbean have shown the most dramatic increases (9941%, 4893%, and 2318%, respectively). As such, the internet has become an important source for health information for both patients and practitioners globally (Bates, Romina, Ahmed, & Hopson, 2006; Cline & Haynes, 2001; Porter & Edirippulige, 2007).

Increased internet access means the internet population is changing. It is no longer restricted predominantly to those with greater wealth or higher levels of education, but is beginning to more closely resemble the general population (McInnes & Haglund, 2011). With greater access to all types of information, internet users can search anonymously for information on sensitive health-related topics (Friedman, Hoffman-Goetz, & Arocha, 2006). Access to internet based health information is also a means for health consumers to advance their equality and self-efficacy, by providing greater input in the specialist-patient relationship (Friedman et al., 2006), and participating in shared decision-making (Convery et al., 2018).

However, a key issue with obtaining health-related information on the internet is that anyone can publish anything without verifying facts, sources of information or reporting potential biases. It is left to the consumer to decide what is trustworthy and what is not. Of

concern to the medical community, medical advice is often freely available through unofficial means, such as blogs or comments sections of webpages (Mayer, Leis, & Sanz, 2009), where information has not been verified.

1.7 Readability

Another caveat to online health-related information is that it is frequently presented as complex information which is effectively inaccessible to those with low literacy skills, because of poor readability (Laplante-Lévesque et al., 2012; McInnes & Haglund, 2011).

Readability can be defined as the ease with which a person can read and understand written materials (Atcherson et al., 2014). Readability is commonly expressed as a reading grade level (RGL), typically the US equivalent, indicative of the standard at which a person should be able to read after *x* years of education (Laplante-Lévesque & Thorén, 2015). However, it is not a direct representation of the number of years of schooling a person has had. Education does not equal literacy (Weiss, 2003). Comparable to health literacy, poor readability is a widespread problem, across different populations.

Given that a substantial percentage (25%) of adults in countries such as the USA or Canada are regarded as functionally illiterate – that is, reading below the 5th grade level, it is commonly accepted that a 5th or 6th grade level should be the standard to which health-related information should be set (Doak et al., 1996). Despite this recommendation, previous research into the readability of online health-related information has shown that websites have consistently been found to be at much higher levels, such as at secondary or tertiary levels, which are too difficult for a lay person to read and comprehend (Friedman et al., 2006).

Many studies have reported similar findings when assessing the readability of online health information: internet health information has poor readability and is written above the

recommended reading level of 6 (Berland et al., 2001; Laplante-Lévesque & Thorén, 2015). This finding is also common to audiology-specific information (Laplante-Lévesque et al., 2012) as well as health information available in the Spanish language. Berland et al. (2001) compared accessibility, quality and readability of online health information in English and Spanish, finding that although the readability was poorer in English, the Spanish information also exceeded the recommended RGL(13.2 RGL English, 9.9 Spanish). Castillo-Ortiz et al. (2017) found only 25% of the websites analysed on rheumatoid arthritis had acceptable readability levels.

One consequence of poor readability is that health inequalities can worsen (McInnes & Haglund, 2011) by isolating readers from the subject matter and deterring them from seeking information (Doak et al., 1996). Simply put, if the reading level required is too high, the reader will stop reading. This can be particularly problematic when lower quality healthcare information is easy to read and as such is more appealing to those with lower literacy skills. Ritchie et al. (2016) found that chat forums and blogs on glue ear had good readability, but low quality. If readability is a barrier, readers with lower literacy skills are more likely to access inaccurate health care information via this format than accurate information verified by medical professionals.

Although health literacy is multifaceted and not determined by literacy skills alone (Baker et al., 2002), improving the readability of healthcare information can also help improve health literacy (Donald & Kelly-Campbell, 2016). Written materials presented at an appropriate RGL are not just important for those with poor literacy skills. Research has shown that all readers, regardless of level and capability, prefer easy-to-read materials (Doak et al., 1996; Donald & Kelly-Campbell, 2016).

Readability has been shown to be an important factor in accessing healthcare information. It is useful to critically analyse the readability in order to highlight barriers to proper understanding of healthcare. For this reason, it has been included as a focus of the present study. Ways of measuring readability are discussed in section 1.7.1.

1.7.1 Readability formulas

Readability is the quantifiable measure of how easy a text is to read (Freda, 2005; Shepperd, Charnock, & Gann, 1999). Developed out of the field of primary education in the 1920s (Doak et al., 1996), readability is typically gauged through the use of a variety of readability formulas (RFs).

RFs are useful for a variety of reasons: they are quick and simple ways to estimate readability (Janan & Wray, 2014). They have become even easier to use with the advent of computer and internet readability programs (Janan & Wray, 2014). However, like anything, there are also caveats.

RFs typically examine a limited number of features of written language, and while there is variability between formulas, most RFs analyse two common indicators of difficulty: the difficulty of the vocabulary and the syntactic complexity (Bailin & Grafstein, 2001; Collins-Thompson & Callan, 2004). Vocabularic difficulty might be measured by how common a word is (Doak et al., 1996), or by how many syllables it has (Bailin & Grafstein, 2001; McInnes & Haglund, 2011). Syntactic complexity is typically measured by the length of the sentence. There are several underlying assumptions to these principles: the use of jargon or uncommon words will reduce readability (Laplante-Lévesque et al., 2012), the longer a word is the harder it is to read, and finally there is a parallel between the length of a sentence and its difficulty (Bailin & Grafstein, 2001).

Several limitations are inherent to these linguistics assumptions. When considering the commonality of words, it is important to remember the ever-changing nature of language.

Words that were common several decades ago no longer remain so today, while new words will have entered the everyday vernacular (Bailin & Grafstein, 2001). Additionally, words common to one socio-economic group will differ from another (Bailin & Grafstein, 2001), and there are many regionalisms and variations between countries that speak the same language.

The length (of words or sentences) does not necessarily increase difficulty. Word length is often increased by affixation. A basic premise of language is knowing how to manipulate words in predictable ways with affixation (Bailin & Grafstein, 2001). Adding *–ing* to *do* creates the longer but no more complex word, *doing*. Similarly, reducing sentence length can in some cases reduce clarity and intelligibility (Redish, 2000). However, those with limited literacy skills tend to favour information presented in shorter words and sentences (Weiss, 2003). As Spaulding (1956) pointed out, the assumptions of RFs are valid on the whole, but should not be followed blindly. To use a RF as a rule would be misuse, as they were never meant to be taken in isolation (Contreras, Garcia-Alonso, Echenique, & Daye-Contreras, 1999).

It is important to note that RFs do not measure comprehension (Atcherson et al., 2014; McInnes & Haglund, 2011). Although written comprehension depends on readability, other factors such as a person's familiarity with the topic, motivation and interest are also significant, and cannot be measured by RFs (Atcherson et al., 2014). Moreover, RFs do not take into account the visual layout, such as font size, long paragraphs versus bulleted lists and the use of diagrams and illustrations, which can enhance readability if used well (Weiss, 2003).

Despite these caveats, RFs remain popular because they are simple and valuable assessment tools (Klare, 1974). Texts can be analysed in a computer program in seconds and

can reliably discriminate between harder and easier texts (Janan & Wray, 2014). The use of RFs in healthcare information is important as they can highlight problematic texts, putting the producer in a better position to understand the shortcomings of their work, improve upon it and make it more accessible to a wider audience (Ritchie et al., 2016).

The majority of literature available on readability formulas is focused on English; however, some studies have looked at their applicability in other languages. Several RFs have been adapted from English into Spanish. Parker, Hasbrouck, and Weaver (2001) found that the features identified as important in English were reasonably important in Spanish as well. Coco, Colina, Atcherson, and Marrone (2017) found that poor readability issues are also common in Spanish healthcare materials: audiological and otolaryngological patient materials were consistently above the recommended 5th – 6th RGL.

Because RFs differ slightly, two formulas may give different RGLs for the same text. For this reason, previous studies (Coco et al., 2017; Friedman et al., 2006; Ley & Florio, 1996) argue for the use of two or more formulas to improve reliability of results. Following this argument, this study will use a total of four Spanish readability formulas: Crawford (Crawford, 1984); Gilliam-Peña-Mountain Graph (Gilliam, Peña, & Mountain, 1980), Rate Index (Anderson, 1983) and SOL (Contreras et al., 1999). These specific formulas are explained in more detail in section 2.4.

Analysing readability is important to ensure that individuals understand the message and can help improve health literacy, but it is not the whole picture. Quality is an equally important aspect of healthcare information, and many studies have analysed both readability and quality in their research (Berland et al., 2001; Castillo-Ortiz et al., 2017; Guo et al., 2018; Laplante-Lévesque et al., 2012; Ritchie et al., 2016). Quality of online health information is discussed in the following section.

1.8 Quality

Quality of information refers to how trustworthy the information is (Cline & Haynes, 2001). More than this, quality can be thought of as the degree to which information can affect the health outcome or quality of life of a user in a positive way (Risk & Petersen, 2002).

The quality of information pertaining to healthcare has traditionally been held to the rigorous standards of the scientific community (Ávila de Tomás, Portillo Boyero, & Pajares Izquierdo, 2001). New discoveries regarding treatments or disease prevention are published in peer-reviewed journals, with acknowledgements of previous and related research and findings. The same cannot be said of the information available on the internet. While scientific publications are available on the internet, they are often highly technical, requiring specific background knowledge, and so are not appropriate for the average internet user. Additionally, many require payment to the journal for access. Internet health information can be considered anything from personal narratives, blogs and illness discussion groups, to scientific journal articles (Purcell, Wilson, & Delamothe, 2002). These different formats cover a wide range of audiences; health consumers prefer more pragmatic, simple explanations and reassurance, while healthcare professionals might be more inclined to read studies in peer-reviewed journals (Purcell et al., 2002).

A lot of free information is available on the internet, but the internet's strength is also its weakness in this context: anyone can contribute (Bates et al., 2006). Publishing information on the internet is a simple and inexpensive process (Price & Hersh, 1999), and while often seen as a democratic way to access information, there is also great potential for erroneous and misleading information to become commonplace (Benotsch, Kalichman, & Weinhardt, 2004). There is no formal process to ascertain the quality of the information written on any given webpage, and in many cases the information is of dubious quality (Price & Hersh, 1999).

The problem of unchecked quality has been known since the beginnings of the internet, especially in the health care community. Low quality information can be harmful to consumers, caused by inaccurate, misleading, fraudulent, or biased information (Risk & Dzenowagis, 2001). It is also difficult to find a standard against which online health information can be held, when it comes in such varied formats. Purcell et al. (2002) suggested one solution might be to use scientific standards for medical knowledge, whereas personal accounts are better suited to literary or journalistic standards.

As a result, different initiatives have emerged to deal with this problem, and to protect and prevent consumers from physical, mental or emotional harm (Risk & Dzenowagis, 2001). Some initiatives put the onus on the creators of websites, those who are producing the information, to uphold quality standards. Examples of this include organizations or institutions creating codes of conduct which attempt to address the quality of online health information. Some of these initiatives manifest purely as guidelines to which websites can choose to adhere, as is the case for the European Commission Quality Criteria (European Commission, 2001). Other codes, such as the HON code (discussed in greater detail in section 1.8.1) and Health Internet Ethics (Hi-Ethics Inc., 2000) provide a quality seal or marker for websites to display if they comply with the code of conduct. This allows consumers who are aware of the quality markers to quickly identify a website that meets quality standards.

The aim of these different codes is to first and foremost provide protection for consumers, but the secondary goal is to protect the “good name” of the company which chooses to adhere to the code, ideally creating market competition based on quality (Risk & Dzenowagis, 2001). Quality guidelines or seals are advantageous in the way that any organization can implement them. They can also be updated easily and can be helpful for implementing ethical standards in corporate environments (Risk & Dzenowagis, 2001).

However, a major deficiency of this type of initiative is that there is no way to enforce implementation. It is the decision of the website providers as to whether they choose to follow the guidelines. It is possible for codes to be abused by web providers, or to misinterpret the principles and guidelines (Risk & Dzenowagis, 2001). Another disadvantage is that it is difficult to measure the utilization and effectiveness of these codes and initiatives (Risk & Dzenowagis, 2001).

An alternate strategy is to teach consumers to effectively judge the quality of the websites they visit. This may present a challenge, as the average internet user is not always a good judge of quality. Bates et al. (2006) found that participants in their study were unable to distinguish high credibility sources (defined as experts, with up-to-date information without any commercial biases) from sources with no credibility. The popularity of websites undoubtedly influences what the average consumer reads. Various studies have found that consumers are most likely to only read information from the first page of search results (Eysenbach & Köhler, 2002; Morahan-Martin & Anderson, 2000). It is also difficult for health specialists to know what websites they can recommend to their patients when the information available is constantly changing and growing. Instead it is more efficient and effective to empower patients to evaluate the content they come across online (Bernstam, Shelton, Walji, & Meric-Bernstam, 2005) through education regarding the importance of quality information and teaching them how to use quality evaluation tools.

Several different tools are available to help consumers do this, such as DISCERN (Charnock, Shepperd, Needham, & Gann, 1999), discussed in section 1.8.2. Again, quality standards outlined in the tools cannot be enforced, and moreover their implementation depends on the consumer's awareness of both problems regarding quality of information and the existence of these tools (Risk & Dzenowagis, 2001). A separate issue is that while many tools are available to analyse quality, few of them have been validated (Purcell et al., 2002).

However, validated tools can provide a means to empower consumers to decide whether the information they find on the internet can be trusted.

Another approach, investigated by Price and Hersh (1999), was to examine the role software could play in assessing quality of webpages based on an algorithm which would then automatically filter search results to provide the best quality websites for the user. Their study found that automatic filtering was helpful in assessing the quality of consumer health pages. Their software was able to rank webpages in order of quality; however they also found that automated analysis did not remove the need for information on webpages to be critically evaluated (Price & Hersh, 1999).

While absolute standards for the quality of health information have not been formally established (Risk & Petersen, 2002), many studies have noted several key criteria that are useful in determining quality. The most common criteria listed in the literature include currency (date of publication or of last update), displaying sources of information, reliability, relevance, and accuracy (Ademiluyi, Rees, & Sheard, 2003; Price & Hersh, 1999; Risk & Dzenowagis, 2001; Shepperd et al., 1999). Eysenbach and Köhler (2002) investigated how consumers search for and assess health information online, used focus groups and an observational study to identify how non-health professionals and non-academics judged the quality of the websites. Quality markers identified in the focus groups included scientific reference citations, understandable and professional writing as well as knowing about the credentials of the author. These were notably different to the standards by which the participants later conducted their healthcare searches on the internet. While it is not known how representative the participants were of the general population, it is worth mentioning that authors described the participants' search strategies as "suboptimal" and "ineffective" (p. 575), despite being able to find the desired answers. The participants failed to actively search for information regarding the author or last update of the webpage, only 35% of their search

queries consisted of more than one search term, and 97% of the webpages accessed by participants were among the first 10 results (Eysenbach & Köhler, 2002).

The results of this study (Eysenbach & Köhler, 2002) are important because they demonstrate the difficulty consumers face in finding and assessing the quality of health information on their own. The current initiatives and strategies available, such as codes of conducts and consumer appraisal tools, have advantages but also have several significant drawbacks, including a lack of awareness on the part of consumers and a lack of enforceability on the part of webpage providers. Two of these, which are focuses of this study, the HON code and DISCERN, will now be discussed in greater detail.

1.8.1 HON code

The Health on the Net code or HON code is a quality marker that was developed by the not-for-profit Health on the Net Foundation in 1995. It aims to "promote the deployment of useful and reliable health information online and to enable its appropriate and efficient use" (Health on the Net, 1995, para 3).

HON allows certified websites to display the HON quality seal, enabling consumers and viewers to easily identify a website as being of an acceptable quality standard. Websites must apply for certification and have to meet the eight integral HON code principles. These principles are designed to give websites credibility as well as support consumer confidence by:

1. Claiming that the information comes from a qualified medical professional (authority)
2. Acting as additional support rather than replacing the role of a specialist (complementarity)
3. Keeping visitors' information and identity confidential (privacy policy)
4. Declaring sources and information regarding the date of publication (attribution and date)

5. Using evidence to support any claims (justifiability)
6. Providing contact information for consumers that require further support (transparency)
7. Indicating sources of funding (financial disclosure)
8. Including advertising (advertising policy).

Certification of websites is re-assessed on an annual basis. While the initial certification is free, websites have to pay an annual fee (ranging between 50-325€) after the first year to maintain certification (Health on the Net, 1995), which may be a deterrent for some companies.

The main caveat of the HON code is its small online presence. While the HON Foundation cites over 8,000 webpages have HON code certification at present, this number is actually relatively low considering the sheer number of health information available on the internet (Laplante-Lévesque et al., 2012). This comes back to the issue of enforceability of quality codes and guidelines. Laplante-Lévesque et al. (2012) assessed quality of online information regarding hearing impairment, finding that only 14% of the articles they analysed displayed HON code certification. Serban (2018) found in his study of 432 French language websites on hearing and audiological matters, 86% did not have HON code certification. In a study of English language information on voice disorders, Dueppen, Bellon-Harn, Radhakrishnan, and Manchaiah (2017) reported only 1.2% of the 85 webpage sample were certified by the HON code. Websites with HON code certification are typically found to have better quality ratings than those without HON code certification (Bompastore, Cisu, & Holoch, 2018). Low rates of certification could be due to a lack of awareness of the part of the online healthcare community, whether that means awareness of the HON organization itself, or the issue of poor quality information (Risk & Dzenowagis, 2001).

Despite the low uptake, the HON code has the largest distribution of voluntary certification worldwide. A major benefit is that HON is available in 35 different languages, giving it global applicability. On the other hand, a code such as Web Médica Acreditada (WMA, 1999) reportedly has more users but a smaller global presence (Mayer et al., 2009). However, HON works collaboratively with different organizations, such as the WHO, the International Organization for Standardization (ISO) and the European Commission to promote high quality health-related information online (Health on the Net, 1995).

Because the HON code is a global quality marker, its presence (or absence) will be examined in the present study. Given that quality seals are not commonly used, it is uncertain whether the presence of the HON code will be a useful indicator of quality in Spanish-language websites regarding tinnitus.

1.8.2 DISCERN

The DISCERN tool developed by Charnock and colleagues (1999), was designed to enable both consumers and health professionals to evaluate the quality of information regarding treatment choices. It is a 16-item tool in which each item is given a rating from one to five, indicating the extent to which the quality criteria is fulfilled, a score of 1 indicates the criterion is not fulfilled, scores of 2 – 4 indicate partial fulfilment and a score of 5 is clear criterion fulfilment. Questions 1 – 8 are designed to help the reader assess whether the publication is a valid source of information, while questions 9 – 15 examine at the specifics of the text. The latter questions have a greater focus on treatment, and as such are not always applicable if a publication does not discuss treatment. The final question (16) asks the reader to consider the ratings in questions 1 – 15, and make a final rating depending on the applicability of the questions to the particular publication (Charnock et al., 1999). It is worth noting that the DISCERN tool assesses quality according to certain criteria regarding the content, rather than in terms of scientific rigour or accuracy (Castillo-Ortiz et al., 2017) .

The somewhat subjective nature of the DISCERN tool has generally not been found to be an issue; Charnock et al. (1999) found good agreement between raters, although higher levels of agreement were found when the criteria were more objective, with specific indicators such question 5 “Is it clear when the information used or reported in the publication was produced?” The results from Rees, Ford, and Sheard (2002) supported this finding, and Ademiluyi et al. (2003) also reported good internal consistency for ratings in their study. High agreement is an indication that DISCERN can reliably discriminate publications of low and high quality (Rees et al., 2002).

Using DISCERN is not dependent on having specialist knowledge, and as such is appropriate for not only healthcare professionals but consumers as well (Charnock et al., 1999). However, training on the use of the tool has been shown to increase reliability of ratings (Rees et al., 2002) and although this is not always possible, the DISCERN handbook can also help users with their ratings.

Given that there are many invalidated instruments to assess quality, the fact that the DISCERN tool has been validated in several different studies (Ademiluyi et al., 2003; Ávila de Tomás et al., 2001; Rees et al., 2002) was an important factor in the consideration of its use in this study. The following studies used DISCERN to analyse the quality of different English language healthcare topics: Laplante-Lévesque et al. (2012) assessed the quality of online information regarding hearing impairment, finding this information was generally of low quality; Ritchie et al. (2016) analysed online information on glue ear (otitis media with effusion), reporting variable quality levels; Dueppen et al. (2017) used DISCERN to analyse quality of online information on voice disorders, finding that websites on voice disorders had acceptable quality levels; while Azios, Bellon-Harn, Dockens, and Manchaiah (2017) and Manchaiah et al. (2018) analysed online information on aphasia and tinnitus respectively, both finding the quality was generally low.

The applicability of DISCERN to other languages has also been demonstrated. Castillo-Ortiz et al. (2017) made use of the DISCERN tool when analysing online health information related to rheumatoid arthritis in Spanish. They found that only 25% of the pages analysed were of high quality according to the DISCERN tool. Quality of online tinnitus information in Spanish will be assessed in this study using the DISCERN tool.

1.9 Rationale

This chapter has reviewed the literature relevant to the topic of readability and quality of online tinnitus information written in Spanish. The causes, prevalence and management strategies of tinnitus have been discussed in detail. Additionally, the literature has shown that health literacy is important to the general population, because it is the single most important factor in predicting health outcomes (McInnes & Haglund, 2011; Weiss, 2003). Literacy and health literacy have become important in an increasingly internet–dependent world, where people are turning to online information more and more. However, several barriers stand in the way of people achieving good health literacy: a lack of literacy skills restricts access to information, and people are dissuaded from making use of information if it is not written at an acceptable reading level. Measuring the readability of health information is important because it is a quantifiable measure of how easy a text is to read, and can be achieved using readability formulas, which typically measure difficulty in terms of syntactic complexity and length of words. Despite some limitations, readability formulas remain effective at detecting if a text is technically difficult. Furthermore, because internet healthcare information can be written by anyone with access to the internet, it is difficult to verify the quality of such information. Again, because people are increasingly turning to online sources, it is important that the information they access is of a high standard, with credible sources.

This chapter has also demonstrated that analysing the quality as well as the readability of online sources is common practice, due to the aforementioned reasons. However, the vast

majority of this work is carried out in the English language (e.g. Berland et al., 2001; Dueppen et al., 2017; Laplante-Lévesque et al., 2012), whereas studies on online health information in Spanish, the world's second most spoken language, are relatively few (e.g. Castillo-Ortiz et al., 2017). Some research has been conducted in English regarding online tinnitus information, finding low quality and poor readability (Manchaiah et al., 2018); however no published studies to date have investigated the readability or quality of online tinnitus information in the Spanish, which is the focus of the present study.

1.10 Aims and hypotheses

The aims of this present study are to evaluate the readability and quality of online tinnitus information written in the Spanish language, using four different Spanish readability formulas, as well as the DISCERN tool and presence of the HON code. The hypotheses are:

1. There will be an even distribution of the place of origin of the webpages (classified by country) that were found using the search criteria.
2. There will be an even distribution of the type of organisation of the webpages (commercial, governmental, non-profit) that were found using the search criteria.
3. There will be an even distribution of the webpages' type of organisation by place of origin.
4. There will be an even distribution of HON based on place of origin.
5. There will be an even distribution of HON code certified webpages based on type of organisation.
6. There will be no significant relationship between the RGLs derived from each formula.
7. The mean RGL of the webpages will not be significantly different from 6.
8. There will be no significant difference in mean RGL based on place of origin.
9. There will be no significant difference in mean RGL based on type of organisation.

10. There will be no significant difference in RGL scores based on HON code certification.
11. There will be no significant difference in DISCERN scores based on place of origin.
12. There will be no significant difference in DISCERN scores based on type of organisation.
13. There will be no significant difference in DISCERN scores based on HON code certification.
14. There will be no significant relationship in DISCERN scores and mean RGL.

2. Methods

2.1 Overview

This study investigated the readability, quality and suitability of online information regarding tinnitus in the Spanish language. This study's methodology was based on that used by Laplante-Lévesque et al. (2012) and was broken into two stages. The first stage consisted of analysing the readability of the selected webpages, using four different readability formulas: Crawford, Gilliam-Peña-Mountain (GPM) graph, Rate Index (RIX) and SOL. In the second stage, the DISCERN tool was used to assess the quality of the information in the webpages.

The main purpose of this study was to assess and describe the readability and quality of online information regarding tinnitus in the Spanish language. The secondary purpose was to determine whether there were differences in the readability and quality based on geographical location or type of organisation. Finally, this study aimed to analyse and identify differences in readability and quality of webpages based on Health on the Net certification (HON).

2.2 Measures

The dependent variables of the study were the mean readability level (measured as an RGL), and quality rating. Quality ratings were established using the DISCERN tool, while RGLs were determined using several different readability formulas, outlined in section 2.4.

2.3 Procedure

2.3.1 Participants

No participants were required for this study, however six native Spanish-speaking informants, known to the researchers, were recruited to nominate search terms that they would use on the internet to seek more information on tinnitus. These informants were first-

language Spanish speakers from Spain, North or Central America. Search terms were collected via personal communications, using the phrase "If you were experiencing noises (such as ringing/buzzing, etc.) in your ears, and you wanted to find out more about it, what would you type into Google to get more information?"

2.3.2 Identification of search terms

Search terms identified by the informants included *tinnitus*, *zumbido en los oídos* (buzzing in the ears), *sonando en los oídos* (ringing in the ears) , *silbidos en los oídos* (whistling in the ears), *pitido en los oídos* (ringing in the ears), *pitido molesto en los oídos* (annoying ringing in the ears), *pitido constante en los oídos* (constant ringing in the ears), *zumbido molesto* (annoying buzzing), *zumbido constant* (constant buzzing), *ruidos en el oído* (noises in the ears) and *acúfenos* (tinnitus). These Spanish search terms were entered into Google.com/trends. Search terms that did not turn up any results or produced results unrelated to the hearing health were disregarded. Terms that were identified by the informants and verified by Google Trends are displayed in Table 1.

Table 1.

Search terms that were identified and used in the present study

Terms identified by informants	**English translation of terms**	**Geographic location identified by Google Trends**
Zumbido en los oídos	Buzzing in the ears	Argentina, Spain
Pitido en los oídos	Ringing in the ears	Colombia, Spain, Mexico
Acúfenos	Tinnitus	Spain, Argentina, Mexico
Tinnitus	Tinnitus	52 localities worldwide, including all 24 Spanish speaking ccTLDs

2.3.3 ccTLDs

Google country-coded Top-Level Domains (ccTLDs) were used as the basis from which the Google search was conducted. The 2 key criteria that identified the ccTLDs were that 1) Spanish is spoken by at least 5% of the locality's population and 2) the locality had its own ccTLD. The 24 ccTLDs identified for use in this study are shown in Table 2. The internet penetration rate of each locality is also shown (the percentage of the population that has access to the internet).

The four search terms, as determined by the informants and Google Trends, were entered into the Google search engine, for each of the 24 ccTLDs. This search was carried out on the 26th of July 2018. Ninety-six searches in total were completed (four search terms, for each of the 24 ccTLDs). Only the first ten webpages displayed for each search were individually inspected (as discussed in section 1.8). A copy of the webpage's URL (uniform resource locator) and its contents were saved separately, provided that the website met the inclusion criteria.

2.3.4 Inclusion and exclusion criteria

The inclusion criteria specified that the website: 1) is predominantly written in the Spanish language, 2) contains information relevant to tinnitus and 3) is freely available to the public. Search results were excluded if the website was less than 100 words in length, the search result was a video, a directory-listing or a Google-advertisement. Duplications of websites from the search were removed. Based on the inclusion and exclusion criteria, 44 unique webpages were identified to be analysed for readability and quality in this study.

Table 2.

Country-coded Top-Level Domains (ccTLDs) for Spanish

Locality	Google Domain	Internet Penetration (%)
Argentina	google.com.ar	59.9
Belize	google.com.bz	31.7
Chile	google.cl	66.5
Colombia	google.com.co	51.7
Costa Rica	google.co.cr	46.0
Cuba	google.com.cu	25.7
Dominican Republic	google.com.do	45.9
Ecuador	Google.ec	54.3
El Salvador	Google.com.sv	29
Gibraltar	google.com.gi	65.0
Guatemala	google.com.gt	19.7
Honduras	google.hn	17.8
Mexico	google.com.mx	43.5
Nicaragua	google.com.ni	15.5
Panama	google.com.pa	42.9
Paraguay	google.com.py	36.9
Peru	google.com.pe	39.2
Puerto Rico	google.com.pr	73.9
Spain	google.es	71.6
Trinidad and Tobago	google.tt	63.8
Uruguay	google.com.uy	58.1
US	google.us[a]	84.2
Venezuela	google.co.ve	54.9
Virgin Islands	google.co.vi	45.3

Note. Reprinted from the The World Bank Group (2018), retrieved from https://data.worldbank.org/indicator/IT.NET.USER.ZS?view=map

[a] Redirects to Google.com

2.3.5 Identifying information

Information regarding the website was collated in Microsoft Excel, including the URL, whether the website was commercial, non-profit or governmental, and whether the website had the HON code certification. The majority of the webpages appeared in multiple searches, depending on which ccTLD was being used. For this reason, the geographical location of the website host was also determined, i.e. the location of the author, rather than the ccTLD in which the webpage appeared.

2.3.6 Determining location

The location was not always immediately evident to the reader on the webpage itself, and in these cases further internet searches were required to establish this information. Several webpages belonged to larger, international corporations (such as news publications), so were based on the local version of the publication where possible. In cases such as Wikipedia, which allow author contributions from across the Web, the pages were deemed to be of international origin. Other websites were blog pages which did not clearly state their author. In such cases, attempts were made to contact the authors to establish the origin, and failing this, the pages were deemed to be of international origin, as they were not clearly specific to one country.

2.3.7 Determining type of organisation

The type of organization, for example whether the webpage was commercial, non-profit or governmental was recorded for each of the webpages. Governmental pages were established if they contained .gov in the URL, or they bore an official seal of any type of government department. Non-profit websites were determined by the presence of .org in the URL and were cross-checked by follow up internet searches of the foundations. If any advertisements were present on a website, it was deemed to be commercial in nature.

2.3.8 HON code

HON code certification was determined by downloading the Google Extension HON code Toolbar and visiting each page. These results were cross-checked by searching each website individually on the HON website.

2.4 Readability

Four readability formulas were selected to analyses the webpages: Crawford, Gilliam-Peña-Mountain graph, SOL (Spanish version of SMOG) and the Rate Index.

2.4.1 Crawford

The Crawford formula (Crawford, 1984) was developed as an alternative to the Gilliam-Peña-Mountain graph (Gilliam et al., 1980), explained in section 2.4.2. Both formulas are based on the Fry Graph (Fry, 1968) a commonly used English readability measure. The Crawford formula assesses both the number of syllables and the sentence length in a text. The formula is a multiple regression analysis of a 100-word sample, written as:

$$Crawford\ RGL = (nSentences * -.205 + nSyllables * .049) - 3.407$$

2.4.2 Gilliam-Peña-Mountain-Graph

The GPM Graph (Gilliam et al., 1980) is an adaptation of the Fry Readability grade for the Spanish language. It takes into consideration that words in the Spanish language typically have more syllables than their English equivalents, and sentences are also typically longer than their translation in English. The formula, assessing the number of syllables and sentence length, is calculated by:

$$GPMG = \frac{\left(3 * \left(\frac{100}{wS}\right)\right)}{(10 + \left(\frac{100}{wS}\right))/Ts}$$

Where *wS* = 100-word sample text and *Ts* = total number of syllables (Gilliam et al., 1980).

2.4.3 SOL

The SOL formula (Contreras et al., 1999) is the Spanish equivalent of the Simple Measure of Gobbledygook, or SMOG (McLaughlin, 1969). The SMOG was developed from an earlier readability index the Gunning fog index, or FOG as it is more commonly known (Gunning, 1952). Rather than being an acronym, the name *SOL* derives from a play on words: after the SMOG and FOG, it was time for the SOL to come out (*sol* in Spanish means "sun") (Contreras et al., 1999). The SOL formula takes into account the complexity of words as well as sentence length when determining the RGL. This formula is calculated by:

$$G = \left(1.0430 * \sqrt{C} + 3.1291\right) * .74 - 2.51$$

Where *C* = the number of 3+ syllable words per 10 sentence sample, and *G* = grade level (Contreras et al., 1999).

2.4.4 RIX

The RIX, developed by Anderson (1983) is a variation of the LIX formula and counts the number of syllables as well as sentence length to determine readability. The *Lasbarhetsindex* (LIX) is a Swedish formula (Björnsson, 1968), designed to measure readability of foreign texts. The formula calculation is:

$$LIX = \left(\frac{total\ words}{total\ sentences}\right) + \left(\frac{long\ words * 100}{total\ words}\right)$$

In this case, long words are defined as having more than 6 letters. The derivation, RIX, is a simpler formula and can be used on any Western European language. It is calculated as follows:

$$RIX = \frac{long\ words}{sentences}$$

2.5 Readability analyses

The content of the webpages was copied into a Word document, because readability programmes cannot interpret *.html* files. The files were run through the computer software Oleander Readability Studio 2012 at the University of Canterbury. Spanish was selected as the analysis language, and the texts were defined as "non-narrative, fragmented text" and as having "sentences split by illustrations or extra spacing". This was important to ensure the composition and layout of the webpage did not cause the text to be analysed incorrectly.

2.6 DISCERN

The DISCERN tool was used in this study as an instrument design to assess the quality of consumer health information. The DISCERN tool is comprised of 15 questions, which can be rated from 1 – 5. A rating of 1 indicates NO, 3 =PARTIAL and 5 =YES, in terms of whether the quality criteria have been met. The user then gives a final rating based on both their responses to the previous 15 questions as well as their overall impression of the article. The author, a fluent second-language Spanish-speaker who completed a four-year university degree in Spanish, analysed the 43 webpages used in this study, in accordance with the DISCERN guidelines to determine the quality of the information available. To ascertain that the author's ratings were consistent, a random 20% sample of the websites was also rated by a second person. The second rater was a native Spanish speaker with four years of university-level Spanish coursework. Consistency and inter-rater reliability were statistically analysed, as described in section 2.7.2.

2.7 Statistical analyses

2.7.1 Planned statistical analyses

Several different statistical tests were used to test the hypotheses identified in this study. The ideal tests and analyses for each hypothesis are outlined in Table 3. Statistical significance was determined in this study by using an alpha level of $p = 0.05$.

2.7.2 DISCERN inter-rater reliability

In addition to the aforementioned tests, two statistical tests were run in IBM SPSS to determine the inter-rater reliability of the DISCERN scores of the webpages analysed in this study: the intra-class correlation coefficient (ICC) and Cronbach's Alpha.

The kappa value (ranging from 0 to 1), is generated from the ICC and gives an indication of inter-rater reliability by indicating "the proportion of agreement corrected for chance." (Fleiss & Cohen, 1973) . Values greater than .75 represent excellent agreement between raters beyond chance, and values between .40 and .75 represent fair agreement beyond chance (Fleiss, 1981).

The ICC assesses the reliability of coding by using an ANOVA. Because the same two raters established the DISCERN scores for each of the webpages, a two-way mixed model was selected for this analysis. The single measures results was used, given that the reliability analysis was for the mean DISCERN scores for each webpage, rather than for each DISCERN item. The kappa value from the ICC using single measures was .94 (95% CI = .88 - .97), which indicated excellent inter-rater agreement.

IBM SPSS also generates a Cronbach's Alpha within the ICC analysis. Cronbach's Alpha is typically used to assess internal consistency within a scale. However, it can also be

Table 3.

Hypotheses and planned statistical analyses for this study

	Hypothesis	Statistical Analyses
$H_0$1	There will be an even distribution of the place of origin of the webpages (classified by their country or as International pages) that were found using the search criteria	Chi-square goodness of fit
$H_0$2	There will be an even distribution of the type of organisation of the webpages (commercial, Governmental, non-profit) that were found using the search criteria	Chi-square goodness of fit
$H_0$3	There will be an even distribution of the websites' type of organisation by place of origin	Cross tabulation
$H_0$4	There will be an even distribution of HON based on place of origin	Cross tabulation
$H_0$5	There will be an even distribution of HON based on type of organisation	Cross tabulation
$H_0$6	There will be no significant relationship between the RGLs derived from each formula	Descriptive statistics, Pearson's correlation
$H_0$7	The mean RGL of the webpages will not be significantly different from 6	t-test
$H_0$8	There will be no significant difference in mean RGL based on place of origin	Descriptive Statistics, One-way ANOVA
$H_0$9	There will be no significant difference in mean RGL based on type of organisation.	Descriptive Statistics, One-way ANOVA
$H_0$10	There will be no significant difference in RGL scores based on HON certification	Descriptive Statistics, One-way ANOVA
$H_0$11	There will be no significant difference in DISCERN scores based on place of origin	Descriptive Statistics, One-way ANOVA
$H_0$12	There will be no significant difference in DISCERN scores based on type of organisation	Descriptive Statistics, One-way ANOVA
$H_0$13	There will be no significant difference in DISCERN scores based on HON certification	Descriptive Statistics, One-way ANOVA
$H_0$14	There will be no significant relationship in DISCERN scores and mean RGL	Descriptive Statistics, One-way ANOVA

Note. ANOVA = Analysis of Variance

used to measure the extent to which a group of values measure a single thing (in this case, DISCERN score). The alpha can range from 0 to 1. A higher value indicates greater internal consistency which increases the likelihood of the scores measuring the same thing. Once

reliability was established, a single DISCERN rating for each webpage was used for subsequent analyses. The Chronbach's alpha was .97, indicating a high degree of internal consistency. In addition, the ratings were not significantly different: $F(28, 28) = 32.97$, $p < .001$.

3. Results

3.1 Overview

When the internet search was performed, 960 webpages were initially identified (the first 10 search results for each of the four search terms, performed for each of the 24 ccTLDs). Many websites did not meet the inclusion criteria, such as all the *tinnitus* search term results which only returned webpages in English. Others appeared multiple times from different ccTLDs searches. Any duplicates or pages not meeting the criteria were removed, leaving 44 unique webpages. Information regarding the geographical origin, type of organisation and HON code certification is described in section 3.2.3. The results from the statistical analyses are shown in section 3.3.

3.2 Descriptive Statistics

3.2.1 Location

The geographical origins of the websites comprised 6 different countries, as well as international sites such as Wikipedia. Half of the 44 unique webpages were from Spain. The breakdown of websites by geographical origin is shown in Figure 1. The geographical origins were later regrouped, as International, European or American (including North, Central and South America). This was carried out to increase the frequency of each group, to allow for further statistical testing. The frequency of these groups is shown in Figure 2.

3.2.2 Type of organisation

Each webpage was categorized as either a commercial, governmental or non-profit page. Of the 44 webpages, 36 were commercial, 2 were governmental and 6 were non-profit.

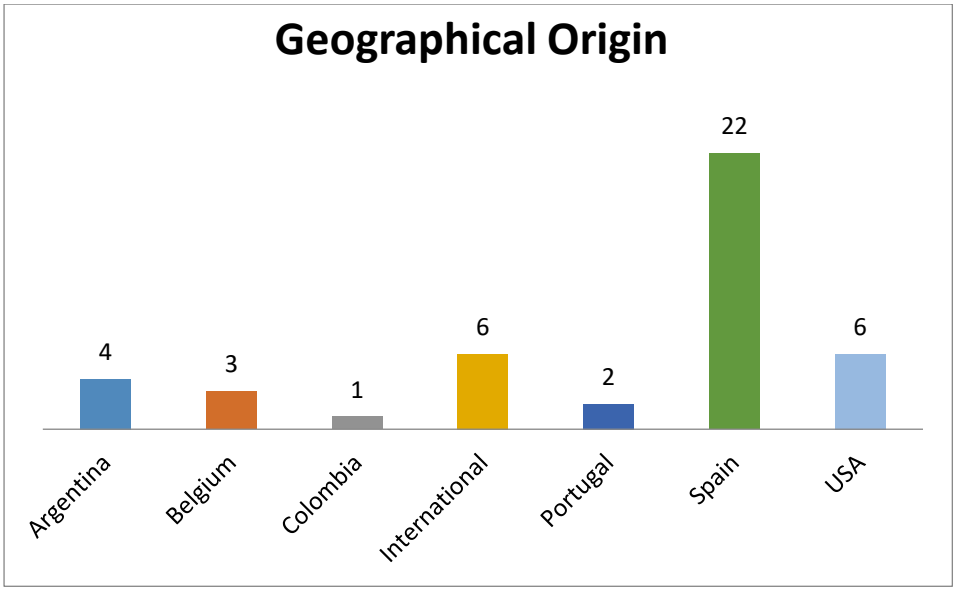

Figure 1. The frequency of websites from each country.

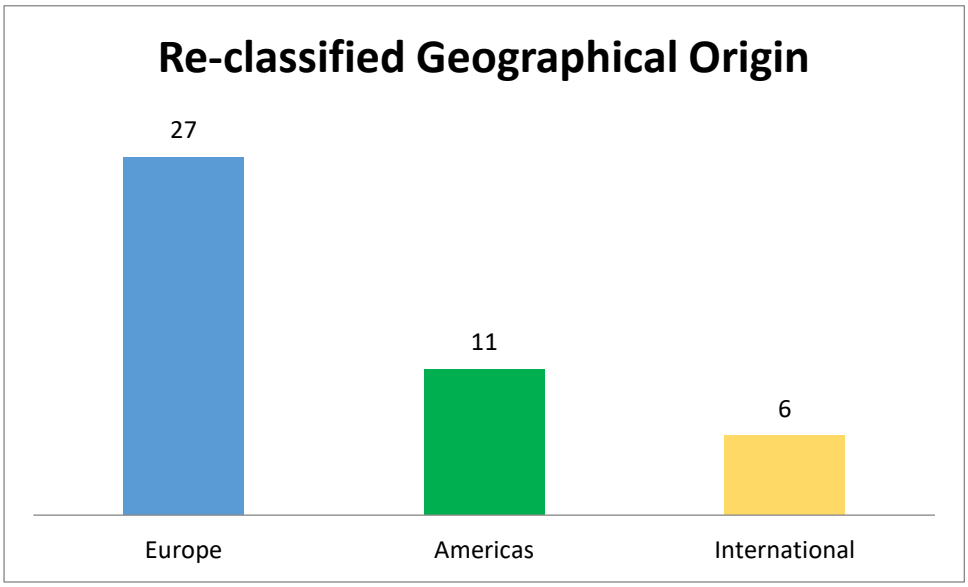

Figure 2. The frequency of websites, using broader geographical classifications.

3.2.3 HON code certification

The majority (41) of the 44 webpages had no HON code certification.

3.2.4 Readability and DISCERN

Readability was analysed using four different readability formulas; the Crawford, GPM Graph, RIX and SOL formulas. The mean RGL of each of these formulas was also

calculated and used for certain analyses, as dictated by the hypotheses. Quality of information presented in the webpages was analysed using the DISCERN tool. The descriptive statistics for both are shown in Table 4; while Table 5 displays the means scores and standard deviations for each of the individual DISCERN items.

Table 4.

Descriptive statistics for each of the readability formulas and the DISCERN tool.

Readability formula	***N***	**Min.**	**Max.**	**Mean**	**SD**
Crawford	44	4.80	8.00	6.19	0.66
Gilliam-Peña-Mountain Graph	39	7.00	17.00	10.26	2.24
Rate	44	8.00	13.00	11.45	1.45
SOL	44	8.30	17.20	11.52	1.86
Mean RGL	39	7.08	13.67	9.58	1.23
DISCERN	44	1.25	3.50	2.20	0.51

Note. RGL - reading grade level
Min. – minimum score
Max. – maximum score
SD – Standard Deviation

The *n* values in Table 4 are different between the formulas, namely the number of webpages analysed was lower for the GPM graph and consequently for the mean RGL. This is due to the inability of the GPM graph to calculate an RGL in cases where the text level is too difficult. A total of five webpages returned a “failed” result in the GPM graph analysis, which meant the *n* values and subsequently the degrees of freedom reported for correlations between RFs (section 3.3.6) are different.

Table 5.
Means and Standard Deviations of individual DISCERN items

DISCERN item	Mean	SD
1. Are the aims clear?	2.64	0.91
2. Does it achieve its aims?	3.38	0.51
3. Is it relevant?	3.36	0.70
4. Is it clear what sources of information were used to compile the publication (other than author or procedure)?	1.70	0.97
5. Is it clear when the information used or reported in the publication was reported?	1.99	0.82
6. Is it balanced and unbiased?	2.88	0.65
7. Does it provide details of additional sources of support and information?	2.15	1.32
8. Does it refer to areas of uncertainty?	1.94	1.04
9. Does it describe how each treatment works?	1.93	0.92
10. Does it describe the benefits of each treatment?	1.44	0.73
11. Does it describe the risks of each treatment?	1.16	0.48
12. Does it describe what would happen if no treatment is used?	1.17	0.55
13. Does it describe how the treatment choices affect overall quality of life?	1.18	0.50
14. Is it clear that there may be more than one possible treatment choice?	2.27	1.11
15. Does it provide support for shared decision making?	1.23	0.52
16. Based on the answers to all of the above questions, rate the overall quality of the publication as a source of information about treatment	2.20	0.50

3.3 Hypothesis testing

3.3.1 Distribution of webpages based on location

$H_0$1: There will be an even distribution of the geographical origins of the webpages

(classified by their country or as international pages) that were found using the search criteria

A Chi-square test of goodness-of-fit was performed to determine whether the geographical origins of the webpages were evenly distributed in the sample. Hypothesis 1 was not supported, as the origins were found to have a significantly uneven distribution, X^2 (6, N = 44) = 49. 23, p <.001.

3.3.2 Distribution of webpages based on type of organisation

$H_0$2: There will be an even distribution of the type of organisation of the webpages (commercial, Governmental, non-profit) that were found using the search criteria

A chi-square test of goodness-of-fit was performed to determine whether the type of organisation of the webpages was evenly distributed in the sample. The types of organisations were found to be significantly uneven in their distribution in the sample of webpages, X^2 (2, N = 44) = 47. 09, p <.001. Therefore, hypothesis 2 was also not supported.

3.3.3 Distribution of location by type of organisation

$H_0$3: There will be an even distribution of the websites' type of organisation by place of origin

Due to the significantly uneven distributions of both the geographical origin and type of organisation of the webpages, the planned statistical analysis for this hypothesis (Chi-square cross tabulation) could not be performed as it would have violated frequency count assumptions.

3.3.4 Distribution of HON certification by location

$H_0$4: There will be an even distribution of the websites' type of organisation by place of origin

Due to the uneven distributions of the geographical origins and low frequency count of websites with the HON code certification, a Chi-square cross tabulation could not be performed as it would have violated frequency count assumptions.

3.3.5 Distribution of HON certification by type of organisation

$H_0$5: There will be an even distribution of HON based on type of organisation

Due to the uneven distributions of the type of organisation and the low frequency count of the HON code certification, a Chi-square cross tabulation could not be performed as it would have violated frequency count assumptions.

3.3.6 Relationship between results from each RFs

$H_0$6: There will be no significant relationship between the RGLs derived from each formula

Pearson's product moment correlations were performed to examine the relationships between each of the RGL formulas:

There was a significant positive correlation between the Crawford and GPM Graph formulas r (37) = .95, $p < .001$.

There was a significant positive correlation between the Crawford and RIX formulas r (42) = .63, $p < .001$.

There was a significant positive correlation between the Crawford and SOL formulas r (42) = .65, $p < .001$.

There was a significant positive correlation between the SOL and RIX formulas r (42) = .76, $p < .001$.

There was a significant positive correlation between the SOL and GPM Graph formulas r (37) = .71, $p < .001$.

There was a significant positive correlation between the RIX and GPM Graph formulas r (37) = .69, $p < .001$.

These results showed that the null hypothesis 6 was not supported, because all of the readability formulas were significantly positively correlated to one another.

3.3.7 Mean RGL compared to recommended RGL

$H_0$7: The mean RGL of the webpages will not be significantly different from 6

Descriptive statistics indicated that scores were above the recommended RGL of 6. A one-sample t-test was run to establish if the mean RGL was significantly different from the recommended RGL. The 44-webpage sample had a mean RGL of 9.58 (SD = 1.23) which was significantly higher than the recommended RGL of 6, t (43) = 19.24, $p < .001$. Therefore, hypothesis 7 was rejected.

3.3.8 Mean RGL and location

$H_0$8: There will be no significant difference in mean RGL based on geographical origin

Hypothesis 2 showed that there was not an even distribution of geographical locations, and descriptive statistics indicated that the frequency count of certain geographical origins was too small for it to be used in further statistical testing. For this reason, the locations were re-classified into broader categories: Europe (N = 27), Americas (N = 11) or International (N = 6). Descriptive statistics showed that these group sizes would have violated parametric assumptions; however, they were acceptable for non-parametric testing.

The Kruskal-Wallis test, a one-way non-parametric ANOVA, was performed on the data to determine significant differences in mean RGL scores based on the redefined locations.

There were no significant differences in mean RGL based on the place of origin of the webpages (H (2) = 2.62, p = .27). Mean RGL for European webpages (M = 9.47, SD = 1.28) were not significantly different to American webpages (M = 9.90, SD = 1.18) d = .35, nor

were either significantly different to the international pages (*M* = 9.47, *SD* = 1.22), *d* = 0 and .36 respectively. Hypothesis 8 was supported by the data.

3.3.9 Mean RGL and type of organisation

$H_0$9: There will be no significant difference in mean RGL based on type of organisation.

The uneven distribution of the type of organisation, as indicated by hypothesis 3, meant that parametric tests could not be used to test this hypothesis. Instead, the non-parametric one-way ANOVA Kruskal-Wallis test was performed. It was found that there were no significant differences in mean RGL based on the type of organisation of the webpages (H (2) = 1.01, *p* = .60).

Scores from commercial websites (*M* = 9.67, *SD* = 1.28) were not significantly different from governmental websites (*M* = 9.44, *SD* = .30) *d* = .25. Neither commercial nor governmental pages were significantly different from non-profit websites (*M* = 9.06, *SD* = 1.06), *d* = .52 and .49 respectively. Hypothesis 9 was supported by the data.

3.3.10 Mean RGL and HON certification

$H_0$10: There will be no significant difference in RGL scores based on HON certification

Neither parametric nor non-parametric ANOVAs could not be performed to test this hypothesis; given the low number of webpages with HON code certification would have violated frequency count assumptions.

3.3.11 DISCERN and location

$H_0$11: There will be no significant difference in DISCERN scores based on place of origin

The uneven distribution of the type of organisation, as indicated by hypothesis 3, meant that parametric tests could not be used to test this hypothesis. Instead, the non-

parametric one-way ANOVA Kruskal-Wallis test was performed. It was found that there were no significant differences in DISCERN scores based on the geographical location of the webpages (H (2) = 3.75, *p* = .15). Scores for the European webpages (*M* = 2.09, *SD* = .43) were not significantly different to the American webpages (*M* = 2.30, *SD* = .43) *d* = .49 or the international webpages (*M* = 2.54, *SD* = .81), *d* =.69. The American and international webpages were also not significantly different, *d* =.37. Hypothesis 11 was supported by the data.

3.3.12 DISCERN and type of organisation

$H_0$12: There will be no significant difference in DISCERN scores based on type of organisation

The uneven distribution of the type of organisation, as indicated by hypothesis 3, meant that parametric tests could not be used to test this hypothesis. Instead, the non-parametric one-way ANOVA Kruskal-Wallis test was performed. There were no significant differences in DISCERN scores based on the type of organisation of the webpages (H (2) = 3.42, *p* = .18). Scores from commercial pages (*M* = 2.15, *SD* = .52) were not significantly different from governmental webpages (M = 2.75, *SD* .35) *d*= 1.35. Neither commercial nor governmental pages were significantly different from non-profit webpages (*M* = 2.33, *SD* = .41), *d* =.39 and 1.09 respectively. Hypothesis 12 was supported by the data.

3.3.13 DISCERN and HON certification

$H_0$13: There will be no significant difference in DISCERN scores based on HON certification

This hypothesis could not be tested given the low number of websites with HON code certification would have violated frequency count assumptions for both parametric and non-parametric testing.

3.3.14 Relationship between DISCERN and mean RGL

H014: There will be no significant relationship in DISCERN scores and mean RGL

Descriptive statistics showed there were no significant outliers for any variable, and as such, a parametric test was used to test this hypothesis. A parametric bivariate correlation using Pearson's product moment correlation was used to examine the relationship between the DISCERN scores and the RFs. It was found that there was a non-significant positive relationship between the DISCERN tool and mean RGL, r (42) = .07, p = .64. Hypothesis 14 was supported by the data.

4. Discussion

4.1 Overview

The aim of this study was to analyse the readability and quality of online tinnitus information written in the Spanish language, using a method based on the Laplante-Lévesque et al. (2012) study. Hypotheses were formulated in this study to examine readability and quality, with reference to the geographical origins of the webpages, the type of organisation, and the presence or absence of HON code, the results of which shall be addressed in this chapter.

Overall, 44 unique webpages were analysed for quality and readability. Webpages were found using the search terms *zumbido en los oídos* (buzzing in the ears), *pitido en los oídos* (ringing in the ears) and *acúfenos* (tinnitus). The search term *tinnitus* did not return any usable results. Searches were carried out on the Google search engine, using each of the 24 different Spanish speaking ccTLDs. The majority of tinnitus information in Spanish was produced in Europe, and in particular, Spain. The type of organisation was also recorded as commercial, governmental or non-profit, with a majority of webpages being of a commercial origin. The following sections address the results of the readability and quality analyses of the webpages.

4.2 Relation to literature

4.2.1 Readability

In general, online information written in Spanish on tinnitus was written at a higher RGL than what is recommended for health information. The mean RGLs of the webpages were similar for the GPM Graph, Rate Index and SOL RFs (10.26, 11.45 and 11.52, respectively). RGLs such as these suggest that a reader requires advanced reading skills, of at least 10 years of schooling to effectively read the information. Such levels exceed the recommended 5th - 6th RGL (Doak et al., 1996). The Crawford formula, however, calculated

that the mean RGL of the webpages to be 6.19, suggesting that in fact the information was written at an appropriate level. The difference in results is likely due to the minor derivations between the formulas (Wang, Miller, Schmitt, & Wen, 2013). This is not the first reported instance of the Crawford formula calculating low RGLs compared to other RFs. Gallego Andrés (2017) found that the Crawford RF reported acceptable RGLs (~6) where other RFs (*Legibilidad* μ and *INFLESZ*) reported readability for the same texts to be "somewhat difficult" or "difficult". Similarly, Votta et al. (2018) compared RGLs of the four RFs used in this study, with the addition of LIX. For two separate groups on burns treatment written in Spanish, the Crawford RF was the only formula to give RGLs ~6 when all other four RFs reported RGLs above 10. This is not to say that the Crawford formula is not an accurate measure of readability. It may reflect certain aspects of readability; in fact, the present study found that all four of the RFs were positively correlated with each other. This suggests that all four RFs do measure aspects of readability, despite the variability between them. However, the results of the Crawford formula in isolation highlight the risks of forming a judgement on readability based off a singular RF, and further support the argument of using multiple RFs (Coco et al., 2017; Friedman et al., 2006; Ley & Florio, 1996).

Precisely due to the likelihood of variation between formulas, the use of all four formulas was chosen in order to produce more reliable results. The mean of the four RFs showed the webpages to have a mean RGL of 9.58, which was determined to be significantly higher than the recommended 6th grade level. These results are consistent with those of comparable studies: Berland et al. (2001) reported a mean RGL of 9.9 for Spanish healthcare webpages and 13.2 for English healthcare webpages and McInnes and Haglund (2011) reported the mean RGL of English online healthcare information was 13.2. Dueppen et al. (2017), who also based their methodology on Laplante-Lévesque et al. (2012), analysed online information on voice disorders using three English RFs, finding the mean RGL to be

between 8.56 and 10.92. The results of the present study further contribute to the trend that online healthcare information is simply too difficult for the average consumer to read.

No differences in RGL were found based on geographical location or type of organisation. These results are fairly consistent with previous literature: Cardelle and Rodriguez (2005) compared the readability of U.S. healthcare websites to non-U.S. websites (in Spanish) and found that all websites were all written at a high RGL, regardless of origin. Laplante-Lévesque et al. (2012) also found that type of organisation was not significantly associated with RGL for online English language hearing impairment information. Conversely, Dueppen et al. (2017) identified differences between non-profit and commercial webpages, concluding that some non-profit websites were more difficult to read.

4.2.2 Quality

The main finding regarding quality in this study was that the analysed webpages consistently were of low quality.

4.2.2.1 HON code

Within the sample of 44 webpages analysed in the present study, only 3 webpages presented with the HON code certification. Laplante-Lévesque et al. (2012) found that government websites were significantly more likely to have HON certification. However, because this sample size was too small to carry out any further statistical testing, the interaction of HON code certification and the type of organisation cannot be compared to previous studies. Alternatively, Laplante-Lévesque et al. (2012) reported 14% of websites had HON code certification in 2012, and more recent studies have reported a wide range of HON code certification penetration in their samples. For instance, (Guo et al., 2018) studied a selection 72 articles of English healthcare information pertaining to spinal surgery, of which

8.3% had HON code certification. In addition, Dueppen et al. (2017) reported that only 1.2% of their 85 webpage sample on voice disorders had certification, whereas 20% of English language websites on Peyronie disease had certification (Bompastore et al., 2018). By comparison, 6.8% of the current webpage sample had HON code certification. Given the variable penetration rate of HON code certification across the aforementioned studies; it is uncertain how much value the HON code certification has in helping consumers find good quality health information on the internet.

4.2.2.2 DISCERN

As was the case for RGLs, neither geographical location nor type of organization was found to significantly affect DISCERN scores of the webpage sample. This contrasts with previous literature. Dueppen et al. (2017) reported that commercial websites were significantly lower quality than both non-profit and governmental websites. Laplante-Lévesque et al. (2012), on the other hand, found that non-profit websites tended to be of better quality than both commercial and governmental websites. The significantly uneven distribution of both geographical locations and types of organisations could account for the difference in results between the literature and the present study. With larger sample sizes, the present results may have resembled those of earlier literature more closely. However, because the sample size was indicative of what was available on the internet at the time, there was no way of knowing what the geographical distribution of sample size might look like going into the study, and therefore could not be controlled for.

The DISCERN tool webpage analyses demonstrated that online tinnitus information in Spanish overall had serious shortcomings regarding the quality of information. DISCERN scores of the assessed webpages ranged from 1.25 to 3.50 out of a possible 5.00. The mean DISCERN score for the 44-webpage sample was 2.20. This result is comparable to previous

studies on various healthcare websites written in English: Hargrave, Hargrave, and Bouffet (2006) reported that 60% of websites on paediatric neuro-oncology were poor or very poor (equivalent to a score of 2.5 and below in the present study); Guo et al. (2018) reported the mean for spinal surgery websites to be 35.26 (2.3 equivalent in present study); and Pusz and Brietzke (2012) found the mean DISCERN score for websites on common otolaryngological problems to be 2.58.

Specifically, in the field of communication disorders Azios et al. (2017) found English-language online information on aphasia was of low quality, with a mean DISCERN score of 2.04, Ting and Hu (2014) found the mean DISCERN scores of thyroplasty information in English was 2.20 and finally, Manchaiah et al. (2018) reported online English-language tinnitus information was of variable quality, with a mean score of 2.39. The DISCERN handbook states that a score of 1-2 indicates a low quality rating, whereas a 3 indicates a publication of fair quality (Charnock & Shepperd, 1999). This type of scoring makes it difficult to analyse the intervals in-between, however for the purpose of simplicity a score of 2.2 is considered low quality in the present study.

The individual questions of the DISCERN tool that scored best overall were question 2 (*Does the article achieve its aims?*) and question 3 (*Is it relevant?*). The DISCERN scores suggest that Spanish language webpages tend to present clear overviews of the subject of tinnitus for consumers. However, questions 4 and 7 *(Is it clear what sources of information were used to compile the publication?* and *Does it provide details of additional sources of support and information?*) had the most variable answers, with scores ranging from 1 – 4.5 and 1 – 5, respectively. The webpages that scored highest in questions 4 and 7 tended to achieve higher overall scores. This finding points to the importance of utilising named sources and giving supplementary evidence when writing online healthcare information. Similarly, Hargrave et al. (2006) argued the main reason that paediatric neuro-oncologic

websites return “poor” or “very poor” scores, was due to the publisher not providing sufficient details regarding the validity of the information and references. It is important for consumers to be able to trust the information they are reading, and they can be more assured of this if there are ways to cross check what they find with other sources, either by providing hyperlinked suggested readings, or complete citations of where the information came from.

Possibly because tinnitus is a symptom of other conditions rather than an illness, treatment was not discussed in many of the webpages. As such, Questions 9 – 15, regarding the specific information on treatment choices, were often not applicable to the webpages. This did not mean those webpages automatically received lower quality ratings, but rather the section was omitted when making a final quality judgement, unless the aims of the webpage led the reader to believe treatment would be discussed. Therefore, disregarding questions 9 – 15, the lowest scoring questions were 4 (aforementioned) and 8 (*Does [this publication] refer to areas of uncertainty*?). It is important for consumers to be aware of uncertainty regarding the research available regarding their particular health condition, so they are able to manage their expectations. The DISCERN scores demonstrate that Spanish language webpages on tinnitus are inclined to either present biased information, such as by omitting gaps in the knowledge base or differences in expert opinion, or by promoting one particular treatment option as if it were the only option. This is problematic, as consumers may not be motivated to investigate the validity of claims further, and as such may base decisions off incomplete or misleading information.

It should be noted that the findings of this study indicated that readability and quality of information are separate issues: no statistical relationship was found to exist between RGL and DISCERN scores. Improving readability does not necessarily equate to an improvement in quality of information, and vice-versa. This finding is consistent with Guo et al. (2018), who also found no significant correlation between readability and DISCERN scores. That

said, they cannot be dealt with in complete isolation either. To improve online information on tinnitus, both quality and readability must be treated as two equally important factors to be improved upon.

4.3 Clinical implications

With around 480 million native speakers, living in least 20 different countries or states, it is likely that many Spanish speakers will use the internet to investigate tinnitus, a condition that affects an estimated 10-15% of the population at some point in their lifetime. The results of this study suggest that Spanish speakers do not currently have access to high quality, easy-to-read information regarding tinnitus on the internet. As such, clinicians, audiological businesses, medical professionals and government organisations all need to make a concerted effort to ensure their consumers and patients are directed to high quality information online that is also accessible in terms of their literacy skills.

While Spain accounts for only 47 of the 480 million of the Spanish-speaking population, it is interesting to note that the great majority of tinnitus related websites were of Spanish origin. This may be related to Spain's aging population, which is more pronounced than that of Central and South American countries (Leeson, 2013). An aging population may mean there is greater demand for audiological services in Spain than other Spanish-speaking countries, potentially leading to more awareness (and publications) around hearing-related issues such as tinnitus. Because of their dominance on the topic of online tinnitus information, the clinical implications of this study are somewhat more pertinent in the context of healthcare in Spain.

Research (Baker et al., 2002; Doak et al., 1996; Sudore & Schillinger, 2009) has revealed that access to health information written at an appropriate reading level is crucial for

understanding and improving the individual's health literacy. By providing healthcare materials with good readability, consumers are in a better position to increase their understanding and in doing so; improve their healthcare outcomes (Donald & Kelly-Campbell, 2016; McInnes & Haglund, 2011). Top-down implementation of readability standards would support individuals with low health literacy; however, this would require greater awareness of health literacy issues. When websites are created by government departments, doctors, hearing aid manufacturers and clinicians, they need to ensure this information is written at a 6th grade level. This specific RGL aims to reach the widest possible audience rather than placing the onus on consumers to have the necessary health literacy skills to keep up with the healthcare system. The results of the present study highlight that the current tinnitus information available to Spanish speakers online is not sufficiently easy to read. Poor readability can act as a deterrent for people with low reading abilities, which mean that online tinnitus healthcare is effectively inaccessible to any of the Spanish-speaking population who read at an RGL lower than 9.58.

As the internet continues to grow and its role as a first point of healthcare information solidifies, web developers need to be aware of the issues that poor readability presents to healthcare consumers. Whether this refers to clinicians, hearing aid manufacturers, governmental institutions or non-profit organisations, they should be aware that poor readability negatively affects all readers and should make it their aim to try and minimise this barrier to improve healthcare outcomes. Several readability formulas are available in Spanish, such as those used in the current study. Although RFs do have flaws (as discussed in chapter 1.7.1) they are extremely quick and simple tools which can be used to direct the publisher towards any major readability problems before it is published (Janan & Wray, 2014; Klare, 1974; Ritchie et al., 2016).

Furthermore, greater awareness of online quality issues could lead to greater motivation to improve the quality standards on online healthcare information. For example, being trained on how to use the DISCERN tool would be beneficial for many clinicians. If clinicians can identify good quality websites on tinnitus, they are able to guide their clients to appropriate materials. They would also be in a better position to teach their clients to use the DISCERN tool. This type of collaborative interaction of client and clinician is valuable, as it can serve to empower clients and achieve greater client satisfaction, as well as giving the clients greater confidence in their attempts to search for online healthcare information (Bernstam et al., 2005).

Despite the increasing prevalence of healthcare searches on the internet, the HON code certification was not frequently found on Spanish-language tinnitus websites. It is uncertain if this is because the quality of the websites is too low to achieve certification, or because website publishers are simply unaware of the HON code. For the HON code to be truly useful, publishers of commercial websites need to be aiming to fulfil the criteria and actively applying for certification. Additionally, HON code certification would ideally be endorsed at a governmental level as well. The purpose of this would be two-fold: 1) high quality standards become policy for governmental healthcare websites and 2) raising the HON code certification profile might lead to greater uptake for commercial websites as well. Given that websites with HON code certification tend to have better quality ratings (Bompastore et al., 2018), it is certainly in the best interest both at a societal level and a personal level for consumers to have access to more HON code certified Spanish-language websites on tinnitus.

4.4 Limitations and future research

4.4.1 Study limitations

While great care was taken to ensure the methodology and study design were both rigorous and practical, as with any study there were limitations. A principal limitation is the constantly-changing nature of the internet. The analysed webpages in the present study are but a snapshot of the Spanish tinnitus literature available online in July of 2018. As the internet continues to expand, it is possible that if the study was replicated the webpage sample could be different.

As was the case for Laplante-Lévesque et al. (2012), attempts in the present study were made to imitate the search strategy of the target population. However, the possibility remains that the search strategy used here might differ from Spanish-speaking adults in search of tinnitus information. In fact, the informants used in the present study were from Spain, North and Central America but no informant was from South America, which might have biased the search terms. Another possibility is that the target population uses different search engines and website selection strategies to those used in the present study. In the present study, only the first page of web results was accessed for each search term (10 results per page), as it is assumed that the average consumer does not look past the first page (Anderson, 1983; Castillo-Ortiz et al., 2017; Eysenbach & Köhler, 2002; Morahan-Martin & Anderson, 2000). To counter these limitations in future research, an alternative could be the use of participants to find webpages, rather than an individual search strategy. Additionally, further research could aim to obtain regionally specific information and webpages for further analysis using South American informants.

Finally, because most of the webpages were both commercial and of Spanish origin, sample sizes of other countries as well as government and non-profit websites were small and

meant non-parametric testing had to be used. Although several hypotheses returned non-significant findings, some of the reported effect sizes were large (such was the case for $H_0$12, which analysed the interaction of DISCERN and type of organisation, finding effect sizes of d = 1.09 and d = 1.35). This most likely was the product of underpowered hypothesis testing. Larger sample sizes permitting parametric testing may have instead found more significant differences when carrying out hypothesis testing. Similarly, because only three webpages had HON code certification, the sample size was too small to carry out statistical analyses investigating the interaction between HON code certification and DISCERN scores or RGL. The search strategy was designed to replicate what consumers would actually encounter on the internet, so any small sample sizes were unavoidable. Future research may wish to reconsider the search strategy utilised to find the webpages in order to prioritise sample sizes for more reliable statistical testing.

4.4.2 Limitations of readability formulas and DISCERN

The inter-rater agreement was considered to be excellent in the present study, an indication that the DISCERN can consistently differentiate between low and high quality publications (Rees et al., 2002), however several items of the DISCERN tool are inherently subjective, leaving them open to the rater's interpretation. For instance, question 4 (*Is it clear what sources of information were used to compile the publication*?) had a very clear scoring guideline, indicating what needed to be present in the publication to score a score of 5, 3 or 1. However, a comparable scoring guideline is not provided for question 6 (*Is it balanced and unbiased*?) , which leaves it open to further interpretation. While any individual rater can create their own scoring checklist to keep their own scoring consistent between articles for such items, interpretation will remain subjective between raters. Charnock et al. (1999) and Rees et al. (2002) also found some questions to be more subjective than others, but they argued that small amounts of subjectivity were not significant to overall ratings.

Moreover, the DISCERN tool does not measure how accurate the information provided is, but rather addresses the reliability of information as well as the presentation and discussion of treatment choices (Castillo-Ortiz et al., 2017). Although a webpage may have a low DISCERN rating, this is not necessarily indicative that the information presented is not true. DISCERN also ignores pertinent factors such as design and use of diagrams or illustrations. These are all areas which remain to be explored. The Suitability Assessment of Materials (SAM) appraisal tool may be useful in analysing the content, graphics and layout of tinnitus webpages written in Spanish in future research (Doak et al., 1996).

Furthermore, it is important to remember that RFs do not measure comprehension or understanding (Atcherson et al., 2014; McInnes & Haglund, 2011). The main caveats of RFs were discussed in section 1.7.1. A specific issue that arose during data collection in the present study was that two of the formulas may have provided incorrect mean RGLs. This was because the GPM Graph was unable to calculate an RGL for five of the websites, stating that this "text is too difficult to be classified to a specific grade level because it contains too many high syllable words." Additionally, when the GPM Graph and the RIX reached the limits of their calculations, instead of assigning a specific RGL, they would write it as 17+ and 13+, respectively. When calculating mean RGLs, these numbers were interpreted as 17 and 13, which do not reflect their true RGL and therefore lowered the mean RGL. Nevertheless, the mean RGL were significantly higher than the recommended level, so this inconsistency in interpretation was unlikely to have had any impact on the study or, at least, the effects would be relatively minor. Future researchers may wish to use different RFs, but it should be emphasized that no one RF is perfect and the combined data produce more reliable results (Coco et al., 2017; Friedman et al., 2006). Alternatively, the use of the Cloze reading test alongside RFs may further strengthen the validity of the results.

4.5 Conclusions

The aims of the current study were to analyse the readability and quality of 44 tinnitus webpages written in Spanish, available on 24 unique ccTLDs versions of Google. Using Crawford, Gilliam-Peña-Mountain Graph, SOL and Rate Index readability formulas, the mean RGL was found to significantly exceed the recommended 5th - 6th grade level. These findings were highly consistent with previous research on readability of other online healthcare information in both English and Spanish. Furthermore, it was concluded that quality was generally low across webpages, based on DISCERN tool ratings, consistent with previous research on internet-based healthcare information, including English-language tinnitus information. HON code certification was not a common marker amongst the 44-webpage sample. Given that there are approximately 480 million Spanish speakers globally and tinnitus is estimated to affect 10-15% of the population, it is important that Spanish speakers have access to good quality, easy-to-read online information. However, this study has shown that this is not the case at present. From a clinical perspective, audiologists and doctors need to be conscious of the health literacy issues faced by their clients or patients; taking readability and quality into account when recommending tinnitus information to consumers. Furthermore, webpage developers should aim to cross-check the readability of the information they are producing using simple readability formulas, as well as using tools or certified quality seals such as DISCERN and HON to improve quality of online tinnitus information.

5. References:

Ademiluyi, G., Rees, C. E., & Sheard, C. E. (2003). Evaluating the reliability and validity of three tools to assess the quality of health information on the Internet. *Patient Education and Counseling*, *50*(2), 151-155.

Adjamian, P., Sereda, M., & Hall, D. A. (2009). The mechanisms of tinnitus: Perspectives from human functional neuroimaging. *Hearing Research*, *253*(1-2), 15-31.

Anderson, J. (1983). Lix and rix: Variations on a little-known readability index. *Journal of Reading*, *26*(6), 490-496.

Atcherson, S. R., DeLaune, A. E., Hadden, K., Zraick, R. I., Kelly-Campbell, R. J., & Minaya, C. P. (2014). A computer-based readability analysis of consumer materials on the American Speech-Language-Hearing Association website. *Contemporary Issues in Communication Science and Disorders*, *41*, 12.

Australian Bureau of Statistics. (2006). *Health literacy, Australia*. Retrieved from Canberra, Australia: www.abs.gov.au

Ávila de Tomás, J., Portillo Boyero, B., & Pajares Izquierdo, J. (2001). Calidad en la información biomédica existente en Internet. *Atención primaria*, *28*(10), 674-679.

Azios, J. H., Bellon-Harn, M., Dockens, A. L., & Manchaiah, V. (2017). Quality and readability of English-language internet information for aphasia. *International Journal of Speech-Language Pathology*, 1-9.

Bailin, A., & Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. *Language & Communication*, *21*(3), 285-301.

Baker, D. W., Gazmararian, J. A., Williams, M. V., Scott, T., Parker, R. M., Green, D., . . . Peel, J. (2002). Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *American Journal of Public Health*, *92*(8), 1278-1283.

Baker, D. W., Wolf, M. S., Feinglass, J., Thompson, J. A., Gazmararian, J. A., & Huang, J. (2007). Health literacy and mortality among elderly persons. *Archives of Internal Medicine, 167*(14), 1503-1509.

Bartels, H., Middel, B. L., van der Laan, B. F., Staal, M. J., & Albers, F. W. (2008). The additive effect of co-occurring anxiety and depression on health status, quality of life and coping strategies in help-seeking tinnitus sufferers. *Ear and Hearing, 29*(6), 947-956.

Bates, B. R., Romina, S., Ahmed, R., & Hopson, D. (2006). The effect of source credibility on consumers' perceptions of the quality of health information on the Internet. *Medical Informatics and the Internet in Medicine, 31*(1), 45-52.

Bauer, C. A. (2018). Tinnitus. *New England Journal of Medicine, 378*(13), 1224-1231.

Benotsch, E. G., Kalichman, S., & Weinhardt, L. S. (2004). HIV-AIDS patients' evaluation of health information on the internet: The digital divide and vulnerability to fraudulent claims. *Journal of Consulting and Clinical Psychology, 72*(6), 1004.

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low health literacy and health outcomes: An updated systematic review. *Annals of Internal Medicine, 155*(2), 97-107.

Berland, G. K., Elliott, M. N., Morales, L. S., Algazy, J. I., Kravitz, R. L., Broder, M. S., . . . Lara, M. (2001). Health information on the Internet: Accessibility, quality, and readability in English and Spanish. *JAMA: The Journal of the American Medical Association, 285*(20), 2612-2621.

Bernstam, E. V., Shelton, D. M., Walji, M., & Meric-Bernstam, F. (2005). Instruments to assess the quality of health information on the World Wide Web: What can our patients actually use? *International Journal of Medical Informatics, 74*(1), 13-19.

Björnsson, C. H. (1968). *Lesbarkeit durch Lix*. Stockholm, Sweden: Pedagogiskt centrum.

Bompastore, N. J., Cisu, T., & Holoch, P. (2018). Separating the Wheat From the Chaff: An Evaluation of Readability, Quality, and Accuracy of Online Health Information for Treatment of Peyronie Disease. *Urology*.

Cardelle, A. J. F., & Rodriguez, E. G. (2005). The quality of Spanish health information websites: An emerging disparity. *Journal of Prevention & Intervention in the Community, 29*(1-2), 85-102.

Castillo-Ortiz, J. D., de Jesus Valdivia-Nuno, J., Ramirez-Gomez, A., Garagarza-Mariscal, H., Gallegos-Rios, C., Flores-Hernandez, G., . . . Barajas-Ochoa, Z. (2017). Readability, relevance and quality of the information in Spanish on the Web for patients with rheumatoid arthritis. *Reumatología Clínica (English Edition), 13*(3), 139-144.

Charnock, D., & Shepperd, S. (1999). DISCERN online: Quality criteria consumer health information. Retrieved from http://www.discern.org.uk/index.php

Charnock, D., Shepperd, S., Needham, G., & Gann, R. (1999). DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology & Community Health, 53*(2), 105-111.

Cline, R. J., & Haynes, K. M. (2001). Consumer health information-seeking on the Internet: The state of the art. *Health Education Research, 16*(6), 671-692.

Coco, L., Colina, S., Atcherson, S. R., & Marrone, N. (2017). Readability level of Spanish-language patient-reported outcome measures in audiology and otolaryngology. *American Journal of Audiology, 26*(3), 309-317.

Collins-Thompson, K., & Callan, J. P. (2004). *A language modeling approach to predicting reading difficulty.* Paper presented at the Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics.

Contreras, A., Garcia-Alonso, R., Echenique, M., & Daye-Contreras, F. (1999). The SOL formulas for converting SMOG readability scores between health education materials written in Spanish, English, and French. *Journal of Health Communication, 4*(1), 21-29.

Convery, E., Hickson, L., Meyer, C., & Keidser, G. (2018). Predictors of hearing loss self-management in older adults. *Disability and Rehabilitation*, 1-10.

Crawford, A. N. (1984) A Spanish language Fry-type readability procedure: Elementary level. In*: Vol. 7. Bilingual Education Paper Series*: Evaluation, Dissemination and Assessment Center, California State University, Los Angeles.

Dalton, D. S., Cruickshanks, K. J., Klein, B. E., Klein, R., Wiley, T. L., & Nondahl, D. M. (2003). The impact of hearing loss on quality of life in older adults. *The Gerontologist, 43*(5), 661-668.

De Brito Macedo Ferreira, L. M., Novaes Ramos Júnior, A., & Pereira Mendes, E. (2009). Characterization of tinnitus in the elderly and its possible related disorders. *Brazilian Journal of Otorhinolaryngology, 75*(2), 249-255.

Del Bo, L., & Ambrosetti, U. (2007). Hearing aids for the treatment of tinnitus. *Progress in Brain Research, 166*, 341-345.

DeWalt, D. A., Berkman, N. D., Sheridan, S., Lohr, K. N., & Pignone, M. P. (2004). Literacy and health outcomes. *Journal of General Internal Medicine, 19*(12), 1228-1239.

Doak, C., Doak, L., & Root, J. (1996). *Teaching patients with low literacy skills* (2nd ed.). Philadelphia, PA: J.B Lippincott Company.

Donald, A. J., & Kelly-Campbell, R. J. (2016). Pediatric audiology report: Assessment and revision of an audiology report written to parents of children with hearing impairment. *Journal of Speech, Language, and Hearing Research, 59*(2), 359-372.

Dueppen, A. J., Bellon-Harn, M. L., Radhakrishnan, N., & Manchaiah, V. (2017). Quality and readability of English-language Internet information for voice disorders. *Journal of Voice*.

Eggermont, J. J. (2017). *Hearing Loss: Causes, Prevention, and Treatment*: Academic Press.

European Commission: Quality Criteria for Health Related Websites, (2001).

Eysenbach, G., & Köhler, C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *The BMJ, 324*(7337), 573-577.

Fackrell, K., Hoare, D. J., Smith, S., McCormack, A., & Hall, D. A. (2012). An evaluation of the content and quality of tinnitus information on websites preferred by General Practitioners. *BMC Medical Informatics Decision Making, 12*(1), 70.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York, NY: John Wiley.

Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correltion coefficient as measures of reliability. *Educational and Psychological Measurement, 33*(3), 613-619. doi:10.1177/001316447303300309

Folmer, R. L., Griest, S. E., & Martin, W. H. (2001). Chronic tinnitus as phantom auditory pain. *Otolaryngology—Head and Neck Surgery, 124*(4), 394-400.

Freda, M. C. (2005). The readability of American Academy of Pediatrics patient education brochures. *Journal of Pediatric Health Care, 19*(3), 151-156.

Friedman, D. B., Hoffman-Goetz, L., & Arocha, J. F. (2006). Health literacy and the World Wide Web: Comparing the readability of leading incident cancers on the Internet. *Medical Informatics and the Internet in Medicine, 31*(1), 67-87.

Fry, E. (1968). A readability formula that saves time. *Journal of Reading, 11*(7), 513-578.

Gallego Andrés, A. (2017). *Análisis de algoritmos para determinar el nivel de complejidad de textos sanitarios y recomendaciones para mejorar el empoderamiento de un paciente.* Universitat Politècnica de València, Valencia, Spain.

Gates, G. A., & Mills, J. H. (2005). Presbycusis. *The Lancet, 366*(9491), 1111-1120.

Gilliam, B., Peña, S. C., & Mountain, L. (1980). The Fry graph applied to Spanish readability. *The Reading Teacher, 33*(4), 426-430.

Guo, W., Wang, W., Xu, D., Qiao, Z., Shi, Y., & Luo, P. (2018). Evaluating the quality, content and readability of online resources for failed back spinal surgery. *Spine*.

Hargrave, D. R., Hargrave, U. A., & Bouffet, E. (2006). Quality of health information on the Internet in pediatric neuro-oncology. *Neuro-oncology, 8*(2), 175-182.

Health on the Net. (1995, 2018). About Health on the Net. Retrieved from https://www.hon.ch/en/about.html

Heller, A. J. (2003). Classification and epidemiology of tinnitus. *Otolaryngologic Clinics of North America, 36*(2), 239-248.

Henry, J. A., Dennis, K. C., & Schechter, M. A. (2005). General review of tinnitus: prevalence, mechanisms, effects, and management. *Journal of Speech, Language, and Hearing Research, 48*(5), 1204-1235.

Hi-Ethics Inc. (2000). Health Internet Ethics. Retrieved from http://www.hiethics.org/

Hoare, D. J., & Hall, D. A. (2011). Clinical guidelines and practice: A commentary on the complexity of tinnitus management. *Evaluation & the Health Professions, 34*(4), 413-420.

Hoffman, H. J., & Reed, G. W. (2004). Epidemiology of tinnitus. In J. B. Snow Jr (Ed.), *Tinnitus: Theory and Management* (Vol. 16, pp. 16-41). Hamilton, Ontario: BC Decker Inc.

Instituto Cervantes. (2017). *El Español: Una Lengua Viva. Informe 2017*. Retrieved from

Internet World Stats: Usage and Population Statistics. (2018, 25/09/2018). World Internet Users and 2018 Population Stats. Retrieved from https://www.internetworldstats.com/stats.htm

Janan, D., & Wray, D. (2014). Reassessing the accuracy and use of readability formulae. *Malaysian Journal of Learning and Instruction, 11*, 127-145.

Klare, G. R. (1974). Assessing readability. *Reading Research Quarterly*, 62-102.

Kutner, M., Greenburg, E., Jin, Y., & Paulsen, C. (2006). *The health literacy of America's adults: Results from the 2003 national assessment of adult literacy.* Washington, DC: National Center for Education Statistics.

Langguth, B., Kreuzer, P. M., Kleinjung, T., & De Ridder, D. (2013). Tinnitus: Causes and clinical management. *The Lancet Neurology, 12*(9), 920-930.

Laplante-Lévesque, A., Brännström, K. J., Andersson, G., & Lunner, T. (2012). Quality and readability of English-language internet information for adults with hearing impairment and their significant others. *International Journal of Audiology, 51*(8), 618-626.

Laplante-Lévesque, A., & Thorén, E. S. (2015). Readability of Internet information on hearing: Systematic literature review. *American Journal of Audiology, 24*(3), 284-288.

Leeson, G. (2013). The demographics of population ageing in Latin America, the Caribbean and the Iberian Peninsula, 1950-2050. *La agenda del envejecimiento y las políticas públicas hoy*, 53-74.

Ley, P., & Florio, T. (1996). The use of readability formulas in health care. *Psychology, Health & Medicine, 1*(1), 7-28.

Manchaiah, V., Dockens, A. L., Flagge, A., Bellon-Harn, M., Azios, J. H., Kelly-Campbell, R. J., & Anderssonk, G. (2018). Quality and readability of English-language Internet information for tinnitus. *Journal of the American Academy of Audiology*.

Mayer, M. Á., Leis, Á., & Sanz, F. (2009). Información sobre salud en internet y sellos de confianza como indicadores de calidad: El caso de las vacunas. *Atención Primaria, 41*(10), 534-542.

McFadden, D. (1982). *Tinnitus: Facts, theories and treatments. Report of working group 89. Committee on Hearing Bioacoustics and Biomechanics. National Research Council.* Washingon, DC: National Academy Press.

McInnes, N., & Haglund, B. J. (2011). Readability of online health information: Implications for health literacy. *Informatics for Health and Social Care, 36*(4), 173-189.

Ministry of Health. (2010). *Kōrero mārama: Health literacy and Māori.* Retrieved from Wellington, New Zealand:

Morahan-Martin, J., & Anderson, C. D. (2000). Information and misinformation online: Recommendations for facilitating accurate mental health information retrieval and evaluation. *CyberPsychology & Behavior, 3*(5), 731-746.

Nondahl, D. M., Cruickshanks, K. J., Huang, G.-H., Klein, B. E., Klein, R., Javier Nieto, F., & Tweed, T. S. (2011). Tinnitus and its risk factors in the Beaver Dam offspring study. *International Journal of Audiology, 50*(5), 313-320.

Nondahl, D. M., Cruickshanks, K. J., Wiley, T. L., Klein, R., Klein, B. E., & Tweed, T. S. (2002). Prevalence and 5-year incidence of tinnitus among older adults: The epidemiology of hearing loss study. *Journal of the American Academy of Audiology, 13*(6), 323-331.

Nutbeam, D. (2000). Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century. *Health Promotion International, 15*(3), 259-267.

Nutbeam, D. (2008). The evolving concept of health literacy. *Social Science & Medicine, 67*(12), 2072-2078.

Parker, R. I., Hasbrouck, J. E., & Weaver, L. (2001). Spanish readability formulas for elementary-level texts: A validation study. *Reading &Writing Quarterly, 17*(4), 307-322.

Porter, A., & Edirippulige, S. (2007). Parents of deaf children seeking hearing loss-related information on the internet: The Australian experience. *Journal of Deaf Studies & Deaf Education, 12*(4), 518-529.

Price, S. L., & Hersh, W. R. (1999). *Filtering web pages for quality indicators: an empirical approach to finding high quality consumer health information on the world wide web*. Paper presented at the Proceedings of the AMIA Symposium.

Purcell, G. P., Wilson, P., & Delamothe, T. (2002). The quality of health information on the internet: As for any other medium it varies widely; regulation is not the answer. *The BMJ*. doi:https://doi.org/10.1136/bmj.324.7337.557

Pusz, M. D., & Brietzke, S. E. (2012). How good is Google? The quality of otolaryngology information on the internet. *Otolaryngology--Head and Neck Surgery, 147*(3), 462-465.

Reavis, K. M., Tremblay, K. L., & Saunders, G. (2016). How can public health approaches and perspectives advance hearing health care? *Ear and Hearing, 37*(4), 376.

Redish, J. (2000). Readability formulas have even more limitations than Klare discusses. *ACM Journal of Computer Documentation (JCD), 24*(3), 132-137.

Rees, C. E., Ford, J. E., & Sheard, C. E. (2002). Evaluating the reliability of DISCERN: A tool for assessing the quality of written patient information on treatment choices. *Patient Education and Counseling, 47*(3), 273-275.

Risk, A., & Dzenowagis, J. (2001). Review of internet health information quality initiatives. *Journal of Medical Internet Research, 3*(4).

Risk, A., & Petersen, C. (2002). Health information on the internet: Quality issues and international initiatives. *JAMA: The Journal of the American Medical Association, 287*(20), 2713-2715.

Ritchie, L., Tornari, C., Patel, P., & Lakhani, R. (2016). Glue ear: How good is the information on the World Wide Web? *The Journal of Laryngology & Otology, 130*(2), 157-161.

Roser, M., & Ortiz-Ospina. (2013, 20th September 2018). Literacy. Retrieved from https://ourworldindata.org/literacy

Sataloff, J., Sataloff, R. T., & Lueneburg, W. (1987). Tinnitus and vertigo in healthy senior citizens without a history of noise exposure. *The American Journal of Otology, 8*(2), 87-89.

Savastano, M. (2008). Tinnitus with or without hearing loss: Are its characteristics different? *European Archives of Oto-Rhino-Laryngology, 265*(11), 1295-1300.

Serban, M. (2018). *A readability analysis of French language online information on hearing related websites.* University of Canterbury,

Shepperd, S., Charnock, D., & Gann, B. (1999). Helping patients access high quality health information. *The BMJ, 319*(7212), 764-766.

Sindhusake, D., Golding, M., Newall, P., Rubin, G., Jakobsen, K., Mitchell, P., & hearing. (2003). Risk factors for tinnitus in a population of older adults: The blue mountains hearing study. *Ear and Hearing, 24*(6), 501-507.

Spaulding, S. (1956). A Spanish readability formula. *The Modern Language Journal*, *40*(8), 433-441.

Sudore, R. L., & Schillinger, D. (2009). Interventions to improve care for patients with limited health literacy. *Journal of Clinical Outcomes Management: JCOM*, *16*(1), 20.

The World Bank Group. (2018). Individuals using the Internet (% of population). *World Telecommunication/ICT Devlopment Report and database.* Retrieved from https://data.worldbank.org/indicator/IT.NET.USER.ZS?view=map

Theodoroff, S. M., Schuette, A., Griest, S., & Henry, J. A. (2014). Individual patient factors associated with effective tinnitus treatment. *Journal of the American Academy of Audiology*, *25*(7), 631-643.

Ting, K., & Hu, A. (2014). Evaluating the quality and readability of thyroplasty information on the internet. *Journal of Voice*, *28*(3), 378-381.

Tyler, R., Coelho, C., Tao, P., Ji, H., Noble, W., Gehringer, A., & Gogel, S. (2008). Identifying tinnitus subgroups with cluster analysis. *American Journal of Audiology*, *17*(2), S176-S184.

Vernon, J., Trujillo, A., Rosenbaum, S., & DeBuono, B. (2007). *Low health literacy: Implications for national health policy*. Retrieved from Washington, DC:

Votta, K., Metivier, M., Romo, S., Garrigan, H., Drexler, A., Nodoushani, A., & Sheridan, R. (2018). Readability of Spanish language online information for the initial treatment of burns. *Burns*, *44*(4), 956-961.

Wang, L. W., Miller, M. J., Schmitt, M. R., & Wen, F. K. (2013). Assessing readability formula differences with written health information materials: Application, results, and recommendations. *Research in Social and Administrative Pharmacy*, *9*(5), 503-516.

Weiss, B. D. (2003). Health literacy. *American Medical Association*.

WMA. (1999). Web Médica Acreditada. Retrieved from http://wma.comb.es/es/home.php

World Health Organization. (2009). *Track 2: Health literacy and health behaviour*. Paper presented at the 7th Global Conference on Health Promotion, Nairobi, Kenya. https://www.who.int/healthpromotion/conferences/7gchp/track2/en/

World Health Organization. (2013). *Closing the health equity gap: Policy options and opportunities for action*. Retrieved from Geneva: http://www.who.int/iris/handle/10665/78335

World Health Organization. (2018). *WHO global estimates on prevalence of hearing loss*. Paper presented at the Prevention of deafness, Geneva, Switzerland.

Yang, H., Cai, Y., Guo, H., Xiong, H., Sun, Y., Huang, X., & Zheng, Y. (2018). Prevalence and factors associated with tinnitus: Data from adult residents in Guangdong province, South of China. *International Journal of Audiology*, 1-8.