The simplicity complex: exploring simplified health messages in a complex world

CHRISTINA ZARCADOOLAS*
Department of Preventive Medicine, Health and Environmental Literacy Initiative, Mount Sinai School of Medicine, Box 1057, One Gustave L Levy Place, New York, NY 10029, USA
*Corresponding author. E-mail: christina.zarcadoolas@mssm.edu

SUMMARY
A challenge in individual and public health at the start of the 21st century is to effectively communicate health and science information about disease and complex emergencies. The low health literacy of millions of adults in the USA has been referred to as a ‘silent killer’. A popular approach to improving health communication and health promotion to low health literate consumers has been to simplify the language of health information. The expected result has been that individuals and groups will better understand information and will then make informed decisions about their health and behaviors. This expectation has grown to include the belief that the public will be better prepared to take appropriate action in complex natural and man-made emergencies. Demonstrating the efficacy of this approach remains, in large part, uninvestigated. And it is becoming more evident that health literacy itself is complex and multifaceted. This article applies linguistic and sociolinguistic models in order to better articulate the role of simplification in health communication and health promotion. Focusing on two models from sociolinguistics—pragmatics and text theory—the article discusses their usefulness in rethinking message simplification. The discussion proposes that a richer, more theory-based understanding of text structures and functions, along with other powerful constructs, including cultural appropriateness, relevancy and context, are needed to close the gaps between health messages, health messengers and patients/the public. The article concludes by making recommendations for future study to empirically test the strengths and limitations of these models and constructs.

Key words: health literacy; simplicity; plain language; text theory

INTRODUCTION
Who would argue against clear communication, plain and simple language? From the nighttime cry ‘The red coats are coming!’ to the ubiquitous H1N1 message, ‘Sneeze into your sleeve’, few would argue against the virtues of simplicity. Surely it works. Take Australia’s skin cancer campaign, ‘Slip, Slop, Slap’ or New York Metropolitan Transportation Authority’s anti-terrorism slogan, ‘If you See Something, Say Something®’. Simple wins the day.

OR, WOULD THEY?
For almost 20 years there has been growing effort to simplify the language of health information, often referred to as the ‘plain language’ or the ‘clear and simple’ language movement. Other models used to improve health communication include, cultural competence, eHealth and social networking, and social marketing (Andreasen, 1995; U.S. Department of Health and Human Services Office of Minority Health, 2001; Chou et al., 2009; Jones and Fox, 2009).
The simplicity movement has gained much traction. A popular assumption is that simplification will lead to greater patient/public comprehension, improved health literacy, and thus improved health behaviors and outcomes [Doak et al., 1996; McGee, 1999; National Cancer Institute (NCI), 1994; Partnership for Clear Health Communication, n.d.; Plain Language Action and Information Network (PLAIN), n.d.; Root and Stableford, 1998; Schwartzberg et al., 2005]. While there is no lack of complex and hard to understand health information disseminated, in the USA most federal agencies have ‘clear and simple’ health communications initiatives or guidelines, which have been slowly working their way into practice. These include the National Cancer Institute (NCI), the U.S. Food and Drug Administration (FDA), the National Library of Medicine (NLM), the Environmental Protection Agency (EPA), and countless major philanthropies and societies, including the American Lung Association (ALA), American Cancer Society (ACS) and the American Diabetes Association (ADA).

There is absolutely no doubt that in some situations, plain and simple language is vital and woefully rare. We can think of examples such as storm evacuation instructions, directions for taking medication, informed consent explanations and health insurance forms, just to name a few. However, surface level language simplification as it is generally practiced in health promotion and communication, is not ideal for communicating complex information effectively. And much contemporary medical and disease is complex.

Increasingly, health promotion and communication experts acknowledge the need for more evidence based and theoretically substantiated approaches to developing and evaluating health communications across modalities (Thompson et al., 2005). The goal of this article is to present two constructs from sociolinguistics, pragmatics and text theory, and discuss their usefulness in articulating the role of surface-level simplification of health information.

The central question—Is it possible that the effort to create clear and simple health messages has over-promised? What if the complexity of health information in the first half of the 21st century requires more than simple representations of science, medicine and decision-making? What if the intention to simplify is good, but many of the executions are done poorly and in consistently correctable ways? [(Zarcadoolas, 2006), p. 33].

This article proposes that a richer, more theory-based approach to text, its structures and functions, along with other powerful constructs, including cultural appropriateness, relevancy and context, will yield health communications needed to close the gaps between messages, messengers and audiences. The article concludes by suggesting the need to empirically study the strengths and limitations of applying these models and text constructs.

HEALTH LITERACY

A major challenge to individual and public health at the start of the 21st century is to effectively communicate complex health and science information to a diverse public (U.S. Department of Health and Human Services, 2000; Nielsen-Bohman et al., 2004). The term ‘health literacy’ refers to the wide range of skills and competencies people use in order to seek out, comprehend, evaluate, and use health information and concepts. The public uses health literacy to make informed choices, reduce health risks, and improve health outcomes and overall quality of life (Rudd, 2002; Zarcadoolas et al., 2006). The low health literacy of millions of adults in the USA has been referred to as a ‘silent killer’ (Zarcadoolas et al., 2006). Improving health literacy is seen as a necessary goal in improving health and reducing health inequity. Therefore, the question of how effective broad approaches to information simplification are in communicating to the public is a timely one.

THE NEED FOR GREATER HEALTH LITERACY IN CONSUMER CULTURE

Patients and health consumers face challenges and opportunities in a society where science is often changing and contradictory, health care delivery systems are in flux, and consumer preferences, understanding and decision making are often driven by powerful forces such as direct to consumer marketing, online information and social media (Salmon and Atkin, 2003; Thompson et al., 2005; Duffy and
In general, the ‘plain language’ or ‘clear and simple’ imperative focuses on adjusting health messages at the word and sentence level. Guidelines for clear communication generally include recommendations such as the example in Figure 1 from plainlanguage.gov.

**READABILITY: A VERY BRIEF HISTORY**

In the field of health literacy, accessibility and appropriateness of health communication have generally been discussed in terms of ‘readability’ and grade level [(Nielsen-Bohlman et al., 2004), pp. 46–51]. Reading scores of consumer materials are calculated by syllable and sentence length counts (Flesch, 1951, 1974; Fry, 1968). Likewise, an individual’s health literacy level is often calculated by using reading tests, such as the Test of Functional Health Literacy in Adults (TOFHLA) (Parker et al., 1995) or the Newest Vital Signs (NSV) (Weiss et al., 2005).

Readability formulas, popular in the 1950s and 1960s, were originally developed as gross screening tools. While more sophisticated scales such as the Golub Syntactic Density Scale (Golub and Kidder, 1974) rely on Natural Language Processing Systems, the most popular tools used in the health field rely on measuring vocabulary usage based on the number of syllables in a word and sentence length. These include the Rapid Assessment of Health Literacy (REALM) (Davis et al., 1993); the Simplified Measure of Gobbledygook (SMOG) (McLaughlin, 1969), the FRY Readability Scale (Fry, 1968) and the Flesch–Kincaid Reading Grade Level (Flesch, 1951, 1974). Flesch, a developer of the Flesch–Kincaid, stated that he hoped users ‘won’t take the formula too seriously and won’t expect from it more than a rough estimate’ (Flesch, 1974). However, during the 1990s, in the field of health literacy, these formulas became the gold standard for assessing and critiquing health communications. Published analyses of health communication messages from disease information to consent forms are often based on readability scores [Allen and Deming, 1995; Rudd et al., 2000; (Nielsen-Bohlman et al., 2004), p. 47].

More recently, some are acknowledging that we need to better understand and articulate the complexity of health literacy and develop new methods and strategies for advancing people’s health literacy (Nutbeam, 2000; Nielsen-Bohlman et al., 2004; Zarcadoolas et al., 2006; Pleasant and Kuruvilla, 2008; Freedman et al., 2009). Evolving ecological models of health literacy are broadening the lens through which we connect health literacy with a range of social, cultural and environmental determinants of health (Nutbeam, 1999; Kickbusch, 2001; Kerka, 2003; Zarcadoolas et al., 2005; Freedman et al., 2009).

---

**Fig. 1:** Recommendations for writing clear language from Plain English at a Glance (Smith and Wallace, n.d.).

### Use These Tools to Write Clearly

- Active voice with strong verbs
- Short sentences
- Personal pronouns
- Concrete, familiar words
- No surplus words
- No legal jargon
- Tabular presentation of complex information
- Use a design and layout that increases comprehension
THE PROBLEM WITH FOCUSING ON WORDS AND SENTENCE LENGTH

Most would agree that the primary goal of clear and simple language is to use familiar and accessible language to link new ideas to existing knowledge. Thus, we could argue that the overly complex language of the U.S. Food and Drug Administration’s (FDA) explanation of the January 2009 peanut butter salmonella contamination problem should be simplified and clarified for the general reading public (Figures 2 and 3).

In Rewrite A (Figure 3), familiar words and shorter sentences simplify the text. Rewrite A also uses new-old information logic, where one sentence refers to information in the previous sentence by starting the next sentence with the main noun of the previous sentence.

However, when we apply text theory analysis methods to the rewrite (Grice, 1957; Grice, 1967; Austin, 1962; van Dijk, 2001) gaps and deletions in the simplified information become apparent (Figure 3). These gaps and deletions can actually make the text harder to read and comprehend because the reader must make more inferences. Here is a list of the categories of information missing (deleted):

(i) Who did the testing?
(ii) What authorities endorsed the findings?
(iii) What was the infectious agent?
(iv) How did the science knowledge evolve and what role did any given study play?

Figure 4 presents two different ways to simplify the original text, while still including these four categories of information.

In Rewrite A (Figure 3), a core strength of language, its ability to cohere, has been hobbled through surface-level simplification. Reading the text becomes a process of making big inferences and filling in gaps. Conversely, Rewrites B and C (Figure 4) use basic and cohesive text rules that actually make the text more readable (Grice, 1957; Austin, 1962; Grice, 1967; Searle,
(Rewrite B)

Where did the problem start?

The FDA (government Food and Drug Administration) is now sure that the dangerous peanut butter and peanut paste came from a company in the state of Georgia (Peanut Corporation of America in Blakely, Georgia). The FDA tested the peanut butter and paste from this company and it contained Salmonella. Experts at the FDA and the CDC (government Centers for Disease Control and Prevention) worked with experts in Minnesota and Connecticut to do the lab tests and studies to find out where the problem started.

(Rewrite C)

The Food and Drug Administration (FDA) and the Centers for Disease Control and Prevention (CDC) did tests. The tests helped find out where the Salmonella Typhimurium (Sal mo NELL A / Tif for ee um) in peanut butter and peanut paste came from. Salmonella Typhimurium is a dangerous type of bacteria that can make people sick. Sometimes it can cause death. The FDA had help from state officials in Minnesota and Connecticut. The FDA now states that the Salmonella Typhimurium contamination came from the Peanut Corporation of America (PCA) in Blakely, Georgia.

Fig. 4: Two different ways to rewrite.

1969; van Dijk, 2001). Rewrites B and C (Figure 4):

(i) use rhetorical devices to engage the reader/listener and allow them to better predict what the text is going to be about;
(ii) make the most important message obvious, salient and in simple syntax and vocabulary;
(iii) put less important information later;
(iv) (I would add a fourth to the list) prevent the legal department from having the final say in language intended for the general public.

Researchers have amply demonstrated that 'simplified' texts are often not any more understandable, and in fact are often more difficult to understand (Charrow and Charrow, 1979; Davison et al., 1980; Duffy and Kabance, 1982; Kintsch, 1994; McNamara et al., 1996).

A prime culprit in this simplicity paradox is the almost exclusive focus on word and sentence length. These two elements have been the central definers and predictors of text readability. The sentence is seen as the maximum unit of description and the text is seen as a sequence of sentences. However, in the fields of linguistics and reading research, reading is understood as an active process of constructing meaning that involves different levels of interaction between reader and text elements (Mandler, 1984; Paris et al., 1991; Farstrup and Samuels, 2002; du Pre, 2005). In fact, most meaning is contained in beyond-the-sentence chunks in the text and the interaction between the text and the reader/listener (pragmatics).

As language users we comprehend written and spoken language, in large measure, as a result of things we do across sentence/utterance boundaries (Olson, 1977, 1994; van Dijk, 1977; van Dijk, 2001). For example, a person can read or hear, ‘If you see something, say something’ in entirely different ways: (1) as a birdwatcher on a morning’s outing; (2) as a subway...
A commuter reading a poster en route to work in post 9/11 New York City (Figure 5).

We use a wide range of skills to decode and make meaning, including expectations for the message and the topic, motivation, context, past experience, language comprehension and trust in the messenger, just to name a few. Recycling an example used by Zarcadoolas, Pleasant and Greer (Zarcadoolas et al., 2006), we ‘read’ the following message with surprising fluency:

Aoccdrnig to a rscheearch at Cmabrigde Unervtisy, it deosn’t mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a total mses and you can stil raed it wouthit a porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey ltteer by istlef, but the wrod as a wlohe. Amzanig, huh?

When a text is hard to read, we will generally persist, rearranging the text to make some meaning, even if it is not the author’s intended meaning. [This is not true of extremely low literate readers (third grade and lower) who are overwhelmed by any amount of print.]

Often the mandate to write or revise text to meet formal readability criteria leaves writers and materials developers in a Catch-22, and can result in actually trying to game the system by artificially dividing sentences and using sentence fragments [(Redish and Seizer, 1985), p. 4; Ancker, 2004]. Adding the very words or sentence types that would make the text more comprehensible unhappily increases the readability score of the material, and thus is judged inappropriate. The very things that give language its power have been deleted, making it more difficult to represent complex and nuanced concepts and information in the carefully clipped and truncated, staccato-like code of simplified language. The idea that shorter and simpler is better is inadequate as a representation of what the reader needs in order to negotiate the text (Kintsch, 1994; McNamara et al., 1996). Ancker (Ancker, 2004) demonstrates this with the clever example ‘Be prepared to die next month’. This utterance (text) scores lower (easier-to-read) than ‘Call for an appointment next month’ because the words in the latter are shorter.

**PRAGMATICS: LANGUAGE AS WE ACTUALLY USE IT**

Early models of communication (Figure 6) depict communication as a one-way flow of information from sender to receiver (Shannon, 1948). Such models are reflected in more traditional public health ‘dissemination’ approaches, which often privilege the message and do not actively take into account the receiver of the message.

Sociolinguists study how people use language to make meaning in different kinds of settings (Gumperz and Hymes, 1964; Labov, 1972; Hymes, 1974). They have amply demonstrated that communication is, in fact, not a direct transmission line between the speaker–writer and the listener–reader. A successful message gets co-developed by the Sender and the Receiver, as depicted in Roman Jakobson’s (Jakobson, 1960) elaborated model of communication as seen in Figure 7.

Both reading and listening are very active human behaviors. People are always working to make some meaning from what they read, hear and see. While the least fluent reader may tackle a text one word at a time with little if any comprehension, most readers use elements, clues and feedback and feed forward activities. We regularly skip over words we do not know.

An important limitation in the simple sentence approach is that there is very little feed-forward, anticipatory mechanism in the simplified text. There is not much to propel the reader/listener forward and back. Without this engagement, people are just reading words and lists of sentences.
Linguists and reading researchers have long known that it is inadequate to describe text solely through surface structure (words and sentence length). Gordon (Gordon, 1980), in his readability analysis of Plato’s Parmenides, used two popular readability formulas (Dale and Chall, 1948; Fry, 1968) to show that passages from Plato scored between fourth and sixth grade reading levels. The sophistication required to understand Plato, however, is clearly beyond the average fourth grader’s conceptual and life experience. Short words are not always easy to understand nor are long sentences always hard to understand. The word ‘waive’ as in ‘We will waive your premium’ counts exactly the same on a Flesch test as ‘we’, ‘will’ and ‘your’ [(Redish and Seizer, 1985), p. 4].

Readability formulas deal with the text as a static object not acted upon by the reader (Davison et al., 1980; Holland and Campbell, 1982). By focusing on discrete surface-level elements of a text, and backgrounding the reader’s active meaning making, readability formulas ignore the reader (Redish and Seizer, 1985; McNamara et al., 1996; Ancker, 2004). Sociolinguistic models offer richer perspectives for talking about texts in terms of their ‘context of use’—what listeners/readers bring to the text and what they intend to get out of it (Hymes, 1974; Lakoff, 2004).

**TEXT CREATES CONTEXT AND CONTEXT HELPS UNDERSTANDING**

In linguistic theory and research, simplicity and complexity are often more defined by context than by sentence structure. For example, a passive verb (the patient was denied care by the hospital, versus the hospital denied the patient care) taken at face value, is a complex verb form. Yet, experiments have shown that children comprehended passive verbs in context even when they could not comprehend them in isolation (Coleman, 1964). Words, ambiguous alone, once embedded in context, are surprisingly decodable. A primary reason—decoding is an active process that engages the reader.

**REFOCUSING EFFORTS TO SIMPLIFY AND CLARIFY**

In linguistic theory, a text is any passage or chunk of spoken, written, or visual language that belongs together and forms a unified whole (van Dijk, 1977). The text is also a unit of ‘language in use’. Linguists describe it as a ‘semantic unit’, rather than a grammatical unit. The important glue that holds a text together is its cohesion, the ‘ties’ between sentence elements and the meaning the reader/listener
makes. A text becomes more understandable or ‘decodable’ when it has elements that allow the reader to interact with it. Cohesion and context are central to comprehension (Grimes, 1975; van Dijk, 1977; De Beaugrande, 1981). Cohesion signals to readers that they can look elsewhere in the text for the referent. Applying poorly understood language principles and overusing readability formulas often results in content that is simplified on the word and sentence level but has poor logic or cohesion and thus is actually harder to read. Figure 8 is an example of a text simplified poorly, without much cohesion. For the poorer reader, the sentences on the left leave a lot of gaps. What does the test do? It is not clear that the test checks for blood sugar levels or that too much blood sugar can cause a big baby. And, is the delivery harder because of the blood sugar problem or because the baby is big?

Conversely, a text with surface-level cohesion and no context is also hard to read. The following ‘cat’ text has cohesion, but is very hard to read:

The cat saw the mouse. The mouse was in the field. The field was covered with many dry leaves. They had fallen from the trees. The trees grew from seedlings. They had been planted by the landlord. The landlord was not at home. He was away in Idaho. Idaho is a state located...

Figure 9 presents an example of easy-to-read consumer health information (Scores 5.1 on Fry readability scale). But, is it really easy to read?

The Hepatitis C text in Figure 9 is constructed on a new/old information pattern, and has syntactic cohesion, but it is not very coherent and, therefore, is hard to read. It is written with artfully clipped rows of sentences, no ditches and no places the reader could fall in and get tripped up. Each sentence is like the one before it and after it. However, there are numerous new ideas or concepts in each short message, an element shown to decrease reader comprehension (Kintsch and Vipond, 1977; Charrow and Holland, 1980). Is this text very readable? On the surface, it looks and sounds readable and often consumers say that it appears more readable in focus groups. However, text theory would judge this as a text that is hard to read with good comprehension. In rewrites such as this less fluent readers must rely on the very things that they do most poorly—decode vocabulary and surface structure of sentences, without the assistance from other text elements, such as context and cohesion. A similar lack of cohesion undermines the comprehensibility of the H1N1 swine flu information in Figure 10.

If we accept that readability and usability of information reside more in the text as a whole rather than in individual words and sentences, it becomes essential for us to interrogate our simplified texts and develop new ways to ‘simplify’ and communicate health information and advance public health literacy. One method of text analysis that holds promise has been developed by the author and is referred to as a ‘health literacy load analysis’ (Zarcadoolas and Pleasant, 2008). A health literacy load analysis entails conducting a structural and functional analysis of any given text. The goal of the load analysis is to use linguistic text models along with an ecological model of health literacy to identify the likely demands the message/material places on the patient/consumer.
health literacy load analysis ‘unpacks’ the text and requires one to identify the surface-level, textual and pragmatic elements likely to impact the person’s comprehension and use of the information. The method is modeled on ‘task analysis’ in human factors engineering, in which the tasks a user performs with a system are deconstructed into steps in order to specify user requirements for accomplishing the task objectives (Hackos and Redish, 1998). When combined with usability testing, load analysis can offer new avenues to assessing text understanding and the roles of simplicity and complexity.

**EXAMPLE OF HEALTH LITERACY LOAD ANALYSIS**

Figure 11 is an example of the health literacy load of a text about Human Papilloma Virus (HPV) from the US Centers for Disease Control and Prevention (CDC) website. (This is a partial load analysis and is used here only for demonstration purposes.) The website content presents facts about the benefits of vaccination and guidelines (Figure 11).

Using a health literacy load analysis, the text in Figure 11 uses and assumes the reader possesses some degree of science, civic and cultural literacy, in addition to fundamental (reading) literacy.

**Science literacy**

(i) There are types of viruses (the difference between a virus and a bacterium).
(ii) Medical recommendations and what they are based upon? How trustworthy they are.
(iii) Risk/benefits of vaccines generally, and of the HPV vaccine in particular.
(iv) Medical effectiveness and how it is measured.

**Civic literacy**

(i) The value of professional association recommendations and guidelines.
(ii) Safety as determined by government organizations such as the FDA.
(iii) Understanding the health system and coverage of the cost of the vaccine.
(iv) Vaccine requirements for schools and daycare facilities.
Cultural literacy

(i) Different attitudes and perceptions regarding sexually activity.
(ii) Beliefs about young girls getting vaccine before they become sexually active.
(iii) Attitudes towards sexually transmitted disease.
(iv) Other sexual behaviors—abstinence, mutually faithful relationship.

Simplifying the language of the HPV health messages will not address the underlying health literacy demands of the content. Unless the message is framed appropriately, with an understanding of the science, civic and cultural load as well as the related literacies of the reader, the text is not as effective as it can be. Using a health literacy load analysis of existing content and then developing consumer messages and materials with an ecological model of health literacy moves us away from singular focus on surface-level language, and can result in messaging that is more linguistically and culturally appropriate, engaging and effective. We are finding this is true when constructing messaging and materials about complex health issues such as:

**Fig. 10:** H1N1 Flu (Swine Flu) text from MedlinePlus (MedlinePlus, 2009b).

**Fig. 11:** Excerpts from Vaccines and Preventable Diseases: HPV Vaccine - Questions & Answers (CDC, 2009).
as adult and childhood obesity (Zarcadoolas et al., under review). We believe the approach is promising and will hold up as it is studied more rigorously.

DISCUSSION

The goal of this article has been to apply linguistic and sociolinguistic models and constructs to better understand the roles and limitations of current approaches to simplifying health communications. In most approaches to simplification the sentence is seen as the maximum unit of description and most simplification occurs at the word and sentence level. When you apply richer models of text and users it becomes clear that simplification often does not capitalize on the overall structure and function of the text and the activities of the reader/decoder. The emphasis on simplifying surface-level language backgrounds important linguistic text elements and reading/listening behaviors and minimizes the intricate, exploratory and self-confirming strategies people use to comprehend and interpret information.

While research remains to be done, there is ample evidence from theory and practice that questions surface level language simplification approaches as currently practiced. We need further empirical study to investigate how theoretical principles can be applied to the development and implementation of communication strategies. Such research could integrate knowledge from fields such as health communication, health literacy, linguistics, social psychology, adult education, new media and human factors engineering and would contribute to the ongoing definitional dialogue about health literacy and its practical application for communicators.

After closer examination, we may conclude that popular simplification methods have yielded a public health communication by subtraction. By adopting a universal and un-nuanced strategy of simplifying information, we may be unwittingly perpetuating limited access to the information commons necessary for people to make informed decisions about their health.

ACKNOWLEDGEMENTS

I would like to acknowledge and thank Wendy Vaughan for her diligence in preparing this manuscript, and Joslyn Levy for her comments and guidance.

REFERENCES


