Supplemental Digital Content

Table of Contents

Appendix 1. Protocol for preparing transcripts

Convention	Description
Encounter ID	One unique numerical ID that can be used to link other data sources. Use name of file.
Speaker Identification	Each line begins with an identifier for the particular speaker. I have shown the main identifiers below. This is not an exhaustive list and other participants could be encountered. The main thing is to give each speaker a distinctive (within the conversation) identifier, and to give each speaker category a consistent (across conversations) identifier. D1: = first doctor to speak in the consultation D2: = second doctor P1: = patient N1: = first nurse N2: = second nurse R1: = first resident A1: = first person who is accompanying the patient (e.g., family member, friend) Keep log book if new codes added.
Paragraph breaks	If one speaker says multiple sentences in a row, or has long pauses, do NOT skip onto the next paragraph. Continue within the same paragraph.
Signal start of conversation	Immediately before the transcribed conversation (i.e., after any descriptive metadata written by the research team at the start of the transcript) insert the following text on a separate line: START_CONVERSATION
Signal end of conversation	Immediately after the transcribed conversation (IF there is postscript) insert the following text on a separate line: END_CONVERSATION
Preamble in documents	All text before conversation is considered 'preamble.' Before running the linguistic analysis program, delete all text before the phrase: START_CONVERSATION
Overlapping speech	If multiple people are speaking at the same time, keep the text on separate lines as usual. However, add < and > to indicate the overlapping section.

	e.g., P1: So I have to choose today, <is that right?> D1: <Yes, let's get> a move on.
Personal health information (PHI)	PHI is any name or location of other data that could serve to identify the patient, date of meeting, etc. If required, a system is needed to handle personal information of patients. This system will enable us to remove all personal info automatically, or to remove certain categories automatically, and enables us to automatically count the number of occurrences of each type. The list below is probably NOT exhaustive and this should be extended as new categories are encountered. Email me anytime if you have a query. If someone's name is mentioned: immediately before the name write PERS_NAME_START immediately after the name write PERS_NAME_END If someone's address is mentioned: immediately before the name write PERS_ADDR_START immediately after the name write PERS_ADDR_END If someone's employment is mentioned: immediately before the name write PERS_EMPL_START immediately after the name write PERS_EMPL_END If someone's date of birth is mentioned: immediately before the name write PERS_DOB_START immediately after the name write PERS_DOB_END
Avoid non-alphanumeric characters	Keep a record in a log book of any special characters (non-alphanumeric) that are inserted into the transcripts.
Financial numbers	Insert financial amounts in a consistent manner. Keep a record of this notation in a log book. e.g., €100, $100
Spelling	Be consistent in spelling (e.g., American spelling).
Speech fillers	Consistency is the key: e.g., please use a consistent way if you are transcribing - One form for “Hmm, one form for “Um”, one form for “OK”, etc.
Long Pause	If there is a long pause, indicate by inserting <*> within the paragraph, where * = number of seconds

Laughter	Include laughter by typing *laughter*, *crying*, other?
Unintelligible	Indicate a section of the transcript as unintelligible with <unintelligible>
Preparation of file for analysis	- save the transcripts as a .txt file (preferable) or as a MS Word document.

Appendix 2. Final codebook

Key:
Italics = code created via inductive coding
Light Blue = positive framed code
Red = negative framed code

Short Description	Federal Plain Language Description[38]	Code Description	Code
Words --> Verbs			
Active voice - negative code	"Active voice makes it clear who is supposed to do what. It eliminates ambiguity about responsibilities. Not "It must be done., but "You must do it." Passive voice obscures who is responsible for what and is one of the biggest problems with government documents."	Does the surgeon use the passive voice? Passive statements often do not identify who is performing the action.	neg_words_verbs_does not use active voice
Active voice - positive code	n/a	Does the surgeon use the active voice?	pos_words_verbs_uses active voice
Present tense	"The simplest and strongest form of a verb is present tense. A document written in the present tense is more immediate and less complicated."	Does the surgeon use the present tense?	neg_words_verbs_does not use present tense
Hidden verbs	"A hidden verb is a verb converted into a noun. It often needs an extra verb to make sense. So we write, "Please make an application for a personal loan" rather than "Please apply for a personal loan."	Does the surgeon use a hidden verb?	neg_words_verbs_uses hidden verb
Words --> Nouns and pronouns			
Clear noun placement	*"Readability suffers when three words that are ordinarily separate nouns follow in succession. Once you get past three,*	*Don't use strings of nouns. Look for examples of surgeon using strings of nouns.*	*neg_words_nouns_noun placement is not clear*

Short Description	Federal Plain Language Description[38]	Code Description	Code
	the string becomes unbearable. Technically, clustering nouns turns all but the last noun into adjectives. However, many users will think they've found the noun when they're still reading adjectives, and will become confused."		
Clear nouns	*n/a*	*When nouns such as "it" "that" and "this" are used, it is clear what is being referred to.*	*neg_words_nouns_does not use clear noun (it/that)*
Pronouns - negative code	"Pronouns help the audience picture themselves in the text and relate better to your documents. More than any other single technique, using "you" pulls users into your document and makes it relevant to them. When you use "you" to address users, they are more likely to understand what their responsibility is. Using "we" to refer to your agency makes your agency more approachable. It also makes your sentences shorter and your document easier to read."	The we/you in the sentence is not clear.	neg_words_nouns_does not use clear (we/you) pronoun
Pronouns - positive code	See pronouns negative code description.	The we/you in the sentence is clear.	pos_words_nouns_uses clear (we/you) pronoun

Short Description	Federal Plain Language Description[38]	Code Description	Code
Abbreviations/acronyms (positive code)	"Minimize abbreviations" "The best solution is to find a simplified name for the entity you want to abbreviate."	Minimize use of uncommon (or common) abbreviations. Nicknames may also be given to abbreviations so they do not need to be remembered. For instance, a drug named FHQ-8 could be referred to as "the drug."	pos_words_nouns_uses abbreviation w/ explanation
Abbreviations/acronyms (negative code)	"Minimize abbreviations" "The best solution is to find a simplified name for the entity you want to abbreviate."	Does the physician use an acronym or abbreviation without explaining or making sure the patient understands what the abbreviation means?	neg_words_nouns_uses abbreviation w/out explanation
First name	*n/a*	*Does the physician call the patient by their first name?*	*pos_words_nouns_uses first name*
Words --> Other word issues			
Use short, simple words - positive code	"Use short, simple words: Prefer the familiar word to the far-fetched. Prefer the concrete word to the abstraction. Prefer the single word to the circumlocution. Prefer the short word to the long. Prefer the Saxon word to the Romance word. Phrases should avoid using unnecessary words and should be concise and clear. Omit information that the audience doesn't need to know. This can be difficult for a subject matter	Does the surgeon use clear simple language when explaining concepts?	pos_phrases_uses simple clear language

Short Description	Federal Plain Language Description[38]	Code Description	Code
	expert, so it is important to have someone look at the information from the audience's perspective."		
Use short, simple words - negative code	See short, simple words positive code description	Does the surgeon not use clear simple language when explaining concepts?	neg_phrases_does not use simple clear language
Putting in common terms	n/a	Putting ideas and thoughts into common, familiar terms can be helpful for understanding. Examples include metaphor, analogy, colloquialism, etc.	pos_phrases_put in common terms
Excess modifiers - positive code	"Similarly, we often use excess modifiers such as absolutely, actually, completely, really, quite, totally, and very. But if you look closely, you'll find that they often aren't necessary and may even be nonsensical."	Does the surgeon avoid excess modifiers where they might have otherwise or are often used?	pos_words_other_does not use excess modifiers
Excess modifiers - negative code	See excess modifiers positive code description	Does the surgeon use excess modifiers?	neg_words_other_uses excess modifiers
Doublets and triplets	"Avoid doublets and triplets. English writers love to repeat the same concept by using different words that say the same thing."	Does the surgeon use doublets or triplets?	neg_words_other_uses repetitive concepts

Short Description	Federal Plain Language Description[38]	Code Description	Code
Uncommon words	"Give common words their common meanings and don't define them. Never define a word to mean something other than its commonly accepted meaning."	Physician uses uncommon word	neg_words_nouns_uses uncommon word
Same terms for the same thing - negative code	"You will confuse your audience if you use different terms for the same concept. For example, if you use the term “senior citizens” to refer to a group, continue to use this term throughout your document. Don't substitute another term, such as “the elderly” or “the aged.” Using a different term may cause the reader to wonder if you are referring to the same group."	Does the surgeon use the different terms for the same concept? For example, do they use lumpectomy plus other concepts such as breast conservation surgery or partial mastectomy. Acceptable to do this in the context of explaining that these words all mean the same thing and that the patient might see varying terminology.	neg_words_other_does not use the same words for the same thing
Same terms for the same thing - positive code	See same terms negative code description	Does the surgeon use the same term for the same concept?	pos_words_other_use the same words for the same thing

Short Description	Federal Plain Language Description[38]	Code Description	Code
Technical jargon - negative code	"Avoid legal, foreign, and technical jargon"	Does the surgeon use a medical/technical word without explaining it or ensuring that the patient understands what they are saying?	neg_words_other_uses technical jargon w/out check or explanation
Technical jargon - positive code	See technical jargon positive code description	Does the surgeon explain or check understanding when using a technical or medical word?	pos_words_other_uses technical jargon w/ check or explanation
Phrases ("Sentences")			
Short phrases - positive code	"Express only one idea in each sentence. Long, complicated sentences often mean	Phrases should be short and only express one idea. They should be short	pos_phrases_uses short phrase

Short Description	Federal Plain Language Description[38]	Code Description	Code
	that you aren't sure about what you want to say. Shorter sentences are also better for conveying complex information; they break the information up into smaller, easier-to-process units."	and concise. Short is defined as 15 words max. Does the surgeon use a short phrase (sentence)?	
Short phrases - negative code	See short phrases positive code description	Does the surgeon use a long phrase (sentence)?	neg_phrases_does not use short phrase
Double negatives	"Avoid double negatives and exceptions to exceptions" and ideas should be expressed using positive language."	Does the surgeon use double negatives or exceptions to exceptions? See examples from FPL guidelines.	neg_phrases_uses double neg or exception to exception
Does not begin phrase with the main idea	"When you start a sentence with an introductory phrase or clause beginning with “except,” you almost certainly force the reader to re-read your sentence. You are stating an exception to a rule before you have stated the underlying rule. The audience must absorb the exception, then the rule, and then usually has to go back to grasp the relationship between the two. Material is much easier to follow if you start with the main idea and then cover exceptions and conditions."	Begin phrases with the main idea, rather than after exceptions and conditions. If the exception is only a few words then it may be ok to put it in the beginning to avoid misleading the patient.	neg_phrases_does not begin phrase with main idea

Short Description	Federal Plain Language Description[38]	Code Description	Code
Changing direction in the middle	*n/a*	*Ideas are most clear and easy to follow when they are carried out and followed to the end. Phrases should not change direction in the middle of the phrase.*	*neg_phrases_changes direction mid-phrase*
Chunks ("Paragraphs")			
Topic phrase - negative code	"If you tell your readers what they're going to read about, they're less likely to have to read your paragraph again. Headings help, but they're not enough. Establish a context for your audience before you provide them with the details. If you flood readers with details first, they become impatient and may resist hearing your message. A good topic sentence draws the audience into your paragraph."	Chunks should begin with a topic sentence or introductory form of some sort. Establish a context for what is about to be spoken about.	neg_chunks_does not begin chunk with topic phrase
Topic phrase - positive code	See negative topic phrase description	Does the surgeon begin the chunk with a topic phrase	pos_chunks_begins chunk with topic phrase
Transition words - positive code	"...a transition word or phrase (usually in the topic sentence) clearly tells the audience whether the paragraph expands on the paragraph before, contrasts with it, or takes a completely different direction."	Different ideas should be connected by transition words. Most sentences in fact should contain transition words that help guide the listener. 3 types of transition words. 1) Pointing words - this, that, these, those, and the. These refer directly to something already spoken about. 2) Echo links - often work together with pointing words. If you talked about strip mining removing the top surface of the earth, the next spoken clause as an echo link could be, "this scarring of the earth." 3) Explicit connectives - words whose chief	pos_chunks_uses transition word

Short Description	Federal Plain Language Description[38]	Code Description	Code
		purpose is as a transition. Further, also, therefore, it follows that, etc. Does the surgeon use a transition word?	
Transition words - negative code	See transition word positive code	Does the surgeon not use a transition word?	neg_chunks_does not use transition word
Short chunk - positive code	"Long paragraphs discourage your audience from even trying to understand your material. Short paragraphs are easier to read and understand. Writing experts recommend paragraphs of no more than 150 words in three to eight sentences. Paragraphs should never be longer than 250 words. Vary the lengths of your paragraphs to make them more interesting. As with sentence length, if all paragraphs are the same size your writing will be choppy."	Chunks should be short, to the point, and easy to follow. Chunks are defined as 5 phrases max. Does the surgeon use a short chunk?	pos_chunks_uses short chunk
Short chunk - negative code	n/a	Does the surgeon use a long chunk?	neg_chunks_does not use short chunk
One topic per chunk - positive code	"Limit each paragraph or section to one topic to make it easier for your audience to understand your information. Each paragraph should start with a topic sentence that captures the essence of	Chunks should only cover one topic.	pos_chunks_uses one topic within chunk

Short Description	Federal Plain Language Description[38]	Code Description	Code
	everything in the paragraph."		
One topic per chunk - negative code	n/a	Does the surgeon use more than one topic in the chunk?	neg_chunks_uses more than one topic within chunk
Miscellaneous			
Examples - positive code	"Examples help clarify complex concepts. Good examples can substitute for long explanations and can help listeners without the relevant background information to understand easier. An example is commonly given in spoken English when asked for clarification of something."	Does the surgeon use an example to help the patient understand complex concepts?	pos_other_uses example to clarify
Validates feelings - positive code	*n/a*	*Validation - Physician explicitly validates patient feelings and concerns by using phrases such as, "I understand that this is difficult" or "I wish things were different" or "I would have many questions if I were you. What questions do you have?"*	*pos_other_validates patient feelings*
Validates feelings - positive code	*n/a*		*neg_other_does not validate patient feelings*

Short Description	Federal Plain Language Description[38]	Code Description	Code
Absolute risk - positive code	*n/a*	*Risk communication - Risks can be discussed in absolute vs relative terms. Physician uses absolute terms to discuss risks, such as saying that out of 100 patients with the surgery, 1 will get a complication vs 2 complications in the other treatment option. This gives more context compared to saying the surgery is two times as likely to result in a complication.* *This will include discussion of survival.*	*pos_other_uses absolute risk*
Absolute risk - negative code	*n/a*	*Could include non concrete language such as "lower risk" "higher risk"*	*neg_other_does not use absolute risk*
Asks for questions - positive code	*n/a*	*Questions are elicited - Physician asks for questions from the patient.*	*pos_other_asks for questions*
Asks for questions - negative code	*n/a*		*neg_other_does not ask for questions*
Asks a question	*n/a*	*The physician asks the patient a question, not medical related.*	*pos_other_asks question*
Asks a medical question	*n/a*	*The physician asks the patient a medical-related question.*	*pos_other_asks medical question*
Vague question	*n/a*	*When questions are asked they should be clear and unambiguous.*	*neg_phrases_asks vague question*
Checks understanding - positive code	*n/a*	*The physician in some way confirms that the patient understands what they are saying.*	*pos_other_physician checks understanding*
Checks understanding - negative code	*n/a*	*The surgeon does not check to make sure the patient understands what they are saying.*	*neg_other_physician does not check understanding*

Short Description	Federal Plain Language Description[38]	Code Description	Code
Provides explanation - positive code	*n/a*	*Explanations should be provided by the patient in response to questions or when the physician is discussing material that requires background.*	*pos_other_provides explanation*
Provides explanation - negative code	*n/a*	*The surgeon does not provide an explanation.*	*neg_other_does not provide explanation*

Appendix 3. Major Theme 2 Full Quote

Full quote for major theme 2 demonstrating a long turn with multiple topics.

"Well, I'm just going to go through the two options side-by-side. Again, I'm not used to using this little tool, but it's very similar to what I used to do on my own. Basically, the two options are doing a lumpectomy, where we're removing the area of the cancer leaving most of the breast tissue alone versus a mastectomy, rather, which is removal of all the breast tissue. There are pros and cons to each. The first thing to know is that your chance of surviving this breast cancer is really, really good. It's no better with a bigger surgery. Mastectomy isn't better than lumpectomy, double mastectomy isn't better than single mastectomy or lumpectomy. They're all the same in terms of the outcomes. So, we really determine if a woman is a candidate for lumpectomy really just based on how much breast tissue we need to remove versus what we leave behind. In your case, we should be able to leave you with most of your breast tissue. We would expect you to have a pretty good cosmetic result with the lumpectomy. That's really how we determine when we offer a lumpectomy. If we do a lumpectomy, even at the operating room, I don't trust that I can feel these areas and I certainly want to make sure that we're getting what we need to. If we do a lumpectomy, what the radiologists do is something called, needle localization or a wire localization. On the day of surgery, they repeat your mammogram to find the markers. Of course, in your case, you've got two little markers in there. They find those markers. Like your biopsy, they numb up the breast. They place flexible wires and basically showing me where to start and stop based on your mammogram. We then take a mammogram of the tissue we removed to make sure that it looks like we got what we needed to. Now, anybody with a lumpectomy, this is how we make sure we're in the right area, but microscopically, we never know where the cancer starts and stops. About 20% of women, the cancer cells are ultimately found right at the edge. Those 20% of women have to have a second lumpectomy, which is a second surgery. That's one of the downsides of a lumpectomy. The other thing about lumpectomy is that radiation is always recommended. Radiation doesn't make you sick, doesn't make you lose your hair. Just treats the breast itself to reduce the risk of this coming back. It's a really big nuisance. It's Monday through Friday for up to six weeks. It can be a little shorter, but given that you have two separate areas, I would expect probably standard radiation for you and again, six weeks. It's really, really annoying. It takes ten minutes, but it's Monday through Friday. Again, it's inconvenient. As far as recurrence, recurrence, just like survival, is exactly the same. Even though we're leaving you breast tissue, the radiation is what sterilizes it and keeps your recurrence risk just about as low as it is with a mastectomy. Again, I'm sorry. I'm just not used to this side. I can't read upside down. This is where we are, so more than one surgery. Again, 20% of women with a lumpectomy need a second surgery to take more tissue. Women, who have a mastectomy, if they're having reconstruction, they'll pretty much always get at least one additional surgery as well. By law, insurance has to cover everything for reconstruction done in the setting of cancer, a mastectomy, removal of all the breast tissue. Usually, for women who are doing reconstruction, we save as much of the skin of the breast as possible. What we're usually doing – I'll draw this. What we typically do is make a small incision over the center of the breast to remove – we're typically removing the nipple and areola, but we save as much of that skin as possible. A plastic surgeon comes in while you're still asleep and they place a temporary saline implant. That's the most common out of the gate reconstruction. Typically, women will then have a second surgery down the road when all the cancer treatment is done to exchange that temporary implant for permanent implant, or from moving around your own tissue. Again, reconstruction is almost always at least one more surgery. We're covering

a little different between these two as well. With a lumpectomy, it's outpatient surgery. You go home the same day with very little down time. Most women don't have a huge amount of pain. Mastectomy is obviously a bigger surgery. Women with or without reconstruction typically stay in the hospital overnight and go home with drains in. Those drains are the worst thing because they're in for a couple of weeks. Typically, the down time with a mastectomy with or without reconstruction is close to about four weeks, so definitely more down time. One potential benefit of a mastectomy is that you may not need radiation. We're removing all this breast tissue, so there shouldn't be anything left to sterilize, but I never make that guarantee because there's always a chance of a surprise like tumor size, tumor and lymph nodes, that sort of thing. I tell women, "Don't choose a mastectomy just to get out of radiation because I can't make that promise." Now, I did just bring up the fact like lymph nodes. There are two lymph nodes that can change recommendations for radiation, so I'm going to talk about the lymph node part of things. Whatever we do with the breast, lumpectomy or mastectomy, we're going to check the lymph nodes to make sure that they are as okay as we think they are at this point in time. That is something called a sentinel lymph node biopsy. If cancer goes to lymph nodes, it goes to certain lymph nodes first. We actually figure out which lymph node we need to check by injecting radioactive dye into the breast on the day of the surgery. The dye travels to the underarm lymph nodes. Any of the lymph nodes that pick up the dye, we surgically remove. With a lumpectomy, it's usually a separate small incision under the arm. With a mastectomy, it's usually all through the same incision, but we remove those. If those are okay, we've confirmed that your lymph nodes are negative, that helps finalize your stage and just, again, hopefully confirms what we're expecting now, which is that your lymph nodes are okay. That will be done at the time of the surgery. That's, overall, the nuts and bolts of mastectomy, lumpectomy. What sort of questions do you have so far before I move on?" - 7007