International Consortium for Communication in Health Care (IC4CH)

Talking Health Online: Why it matters and what linguistics can contribute

Hosted by the ESRC Centre for Corpus Approaches to Social Science (CASS) at Lancaster University

Chair Prof Tony McEnery | Lancaster University, UK

Speakers
Prof Joanna Zakrzewska | University College London and Royal National ENT and Eastman Dental Hospitals, UK
Dr Tara Coltman-Patel | Lancaster University, UK
Prof Elena Semino | Lancaster University, UK
Dr Gavin Brookes | Lancaster University, UK
Prof May Lwin | Nanyang Technological University, Singapore

Technical support
Dr Luke Collins | Lancaster University, UK
William Dance | Lancaster University, UK

Transcript [00:00:00]

Tony: Hello everybody. I know we've got an audience here for the event today, which is great and I'd like to thank you all very much for joining me because today you're really joining me on a journey where I'm finding out about what some of my colleagues have been doing in the area of health care communication recently. I knew they were working in this area as I happened to be at a meeting where Elena Semino was present and was talking to somebody about establishing something called the International Consortium for Communication in Health Care. I thought it sounded like an interesting idea, it wasn't work that I was doing. But I know that they've then gone off and done things and put together a consortium and all sorts of interesting work seems to be coming from it. So the consortium has within it the Australian National University, Nanyang, Queensland, Hong Kong, and University College, London. So it's a really interesting group of researchers working on a topic that's very interesting indeed, though what I'm going to do today is try to find out exactly how linguistics comes into healthcare. I generally believe that linguistics comes into most subjects, but I'm really interested to find out how the people gathered today have worked with linguistics and how it fits into the healthcare context. So welcome to this journey of exploration, I'm really looking forward to it, and I hope you have the time to listen to all of the conversations that I have with the people gathered here because I'll be trying to find out from them what they're doing and how it fits into completing this picture for me, of how linguistics and healthcare communication interact. So without further ado, I'm going to start my conversation with the first guest that I have today, Professor Joanna

Zakrzewska. She is at University College London. She's also, I think it's true to say, a health care practitioner as well as a researcher, so I'll just pass over to you initially, Joanna, and say can you sort of maybe tell me something briefly about your background?

[00:02:12]

Joanna: Well, I'm one of these rare breeds, I'm both medically and dentally qualified and gradually drifted into the field of pain, pain medicine, which is a relatively new specialty and particularly facial pain. And in fact I first came in contact with linguists through a PhD student I was supervising, or co-supervising I should say, 'cause it was in the Arts: Deborah Padfield was a person that we were doing a PhD with and Elena then joined us for the post doctoral work. And it was here that I really began to understand the value of the humanities and particularly the language that's so important in healthcare. And it's made me very much more aware of what I'm doing and what I'm saying to my patients as well.

Tony: Great, so that's when you first met Elena Semino, who is here with us today. It's wonderful. And I think maybe that also explains why you then went on to establish an online forum for something called the Trigeminal Neuralgia Association in the UK. Now, Trigeminal Neuralgia - that's face pain again, is it?

Joanna: Yes, so trigeminal neuralgia is a really excruciating, severe, intermittent pain that has been termed and again, a word that I don't like using now, as a suicidal disease because it's like an electric shock going through your face when you touch it. And it can be intermittent, you can only get two or three attacks a day, or you can get a hundred of them. And then it may disappear, so it's unpredictability. And its impact on quality of life is horrendous. And so that made me realize that we needed to do something to talk to our patients and provide a bit more care, because it is a rare disease and people don't know enough about it. And the Americans set up a trigeminal neurologist support group, and their chairperson, who was a sufferer herself, came over to the UK to visit me and said you've got to set something up. You're an expert in this condition, you've written a book about it, set it up. So she found me a patient and we set this association up. And one of the things that was important there was to actually get our patients to be able to talk through here. So we set up helplines. So as well as having information, leaflets and now more based on the Internet, we also set up helplines. So there were telephone helplines, email helplines and a forum for those patients who joined the association. And then we got some feedback from our patients about the forum which we set up here.

Tony: Great. So the forum itself is something which you thought would have real value for the patients in terms of them managing their condition, talking to one another as - It had therapeutic value in that respect, do you think?

Joanna: Yes, because one of the things that, even though I've seen hundreds of patients, patients will always say to me 'Well, you're not a buddy, you don't understand my pain'. So the idea here is that these are people who've got the condition or potentially carers who can talk to each other, without geographic - any geographic boundaries and also with anonymity because you don't need to declare who you are and so that improves I think the discussions that you can have, because you don't have to declare who you are. You can

admit to quality of life, to moods and depression, which perhaps you don't want to share with your health care professional or share even with a with your intimate - your significant others and friends. But we started to get worried about the - what was going on there and we were getting feedback that there were some things that were being said that were upsetting people. I mean, if you put on for a newly diagnosed patient 'this is a suicide disease', I mean. That sort of set up some worries, so.

Tony: Yes, it would!

[00:06:32]

Joanna: So I sort of thought, well, how do I sort of work out what's on this forum? I haven't got time as a clinician to start looking at it, and in conversation with Elena, she said, "Well I can help" and we said "what can you do?". "Well, if you give me access to this", because obviously it's all on the out there in the cloud, "we can help". And so she's downloaded everything that we'd had between 2008 and 2013. So that was over 2 million words. And started to help us analyze it and it was just mind boggling.

Tony: So this gave you insights as a sort of clinician into what the patients were experiencing, and maybe change your own views?

Joanna: Yes, tremendously, because suddenly I realized what was being said and how important it was. And we realized that the language had to be moderated, which made us realize that we needed to have somebody who would keep an eye on it. But what was so important was the empathy and compassion that we were finding, and Elena was brilliant at being able to show us incidents of it, and also show us the humor that was there. And that made us realize that. it was an important forum, but that we needed to adjust things.

Tony: Okay. Can I ask you a few methodological questions then? One methodological issue you must have had - but I suppose Elena's dealing with that later, or maybe Tara - is the scale of the data, but I'll put that to one side. I suppose two sort of very trenchant methodological issues, which sort of immediately present themselves to me, would be A) how do you deal with problems of online communication, 'cause I suppose in your forums you had in microcosm what people are talking about on Facebook and Twitter nowadays, sort of anti social actors if you like. A second thought that occurs to me is: I can see it's great to observe the patients talking to one another, it's helpful but I wonder why a forum was useful, rather than say just a Facebook group or people communicating on Twitter. And finally, third questioning - I said two of course, academics always add a third - and the third question would be: for some of the insights, could you have used other methods like focus groups, for example? Or could you use those with this as well?

[00:08:58]

Joanna: So one of the dangers of social media, and we're finding that more and more now, is that there's uncontrolled - people post something up there, there's nobody to police it, so to speak. Whereas, after Elena had looked at this, we realized that it needed policing. So we have one of the members who polices what's on there and keeps an eye on what's on the forum. Now, if she finds - or he finds - something that they're not sure about how to

cope with or there are questions, healthcare questions, we have a medical advisory board. So I'm the chair of this one and I have six or eight different health care professionals, so they're surgeons, neurosurgeons, neurologists, general GPs. So there's a range of us, and we then tackle those questions. So what we're making sure on the forum is that the data that's out there, what we say is covered by evidence-based medicines. So we are making sure that we are checking what is on there, and I think that has been the big thing. As you say, focus groups are also a good way forward and that made us do focus groups as a result of that. We've now - I've got a PhD student who's looking at the importance of outcome measures and outcome domains and what we've used is, we've used for Trigeminal Neuralgia Association to help us search for patients who are willing to come to a focus group and then really specifically talking about what important domains there are, whereas currently the forum is much more open for any conversation that happens. So I think we're using a mixture of both, which I think is proving to be of invaluable help for us.

Tony: That's really helpful and this expert moderation, if you like, that you're giving to the forum, I can see how that has real advantages relative to the Wild West atmosphere that you can get on social. media sometimes. Although towards the end of the conversations today, we'll pick back up on the question of, what might social media be able to show us? Though I think in the context of your example I'm very thankful that those people can talk to one another, you can get insights from them, but also that they get that expert moderation, which keeps the conversation factually accurate. And morally, there's some moral I'm gonna move on, Joanna, to Tara because I'm really going to try and task Tara with dealing with this other issue that we mentioned, the issue of scale. So if we could switch over to Tara now, that would be great.

[00:11:42]

Tara: Okay, hi.

Tony: How are you?

Tara: I'm good, thank you. How are you?

Tony: I'm fine. Do you want to introduce yourself briefly to everybody?

Tara: Sure, I'm Tara Coltman-Patel. I'm a linguist and a Senior Research Associate at Lancaster University and I'm part of a research team looking at vaccination discourse, so the language around vaccinations online in social media, in parliamentary discourse and news media as well.

Tony: Okay. Well we've just heard from Joanna about the importance of the forum that she has and I buy that, that was clearly a very good explanation of why language has a role to play within the treatment, or at least the management of that particular condition. But what role do linguists have to play? The data is there, people are happily talking to one another, the clinicians are getting insights from what they see happening in the forum. Where do you see the role for, specifically, linguists to enter into this and look at the data that is being generated?

Tara: We have a lot to offer when it comes to health research, because exploring how we communicate about illness, how we communicate about medication can really be fundamental in understanding how much people know, how they feel. And forums in particular are a really great space to do that because - to get a bit more specific - what we as linguists can do, is we can track exchanges between people, ordinary people, having conversations with each other, and we can follow these exchanges and see how they end, whether they end in a resolution or more understanding. So maybe an individual will start a thread where they expressed like indecision around getting a vaccination, but by the end they have decided to get a vaccine or they've decided to consider it more. Or conversely, maybe they've decided not to get a vaccine. What we can do, using our specific methods, is track these exchanges. We can compare whether there are specific linguistic patterns associated with the threads where there is a positive resolution versus a negative resolution. And really, we can offer these types of insights and evidence-based insights. And we can help gain a better understanding of how the general public feel and what they already know about health, which is a very complex topic.

Tony: Yeah, that's helpful and I can sort of maybe see in there what I might call a sort of 'what works' strategy emerging, you're sort of looking at it perhaps, and finding out - well, through the examples that you find, through the methods that we'll talk about in a moment - where there were successful outcomes from your perspective, where there were unsuccessful outcomes and those link specifically to linguistic strategies that you can identify within the conversations that you're observing. Is that accurate?

Tara: Yeah, I mean there are different types of findings that we can arrive at and they're quite broad, but to give a couple of examples, at a more micro level we can examine the change in meaning of specific words over a period of time, 'cause we can work with quite large datasets. So we can follow a specific word and in forums we can look at how they're being used by people, in what context they're being used and what they're used to reference and then we can build up an idea of how - if it's developed some more positive connotations over time, more negative connotations over time, over ten-, twenty-year periods. So that's more of a micro level thing we can do and then macro level, more broadly, we can look at strategies. So for example, if I use an example with vaccination discourse, maybe we can compare whether people using personal stories to persuade people rather than scientific facts and figures, maybe that seems to work in order to dissuade people's fears around vaccinations.

[00:15:28]

Tony: I'll now ask you for an example, of course, which is always terrifying in such contexts, but if you can't think of one, that's fine, but going back to the example you talked of words shifting and getting more positive and negative connotations, and how that might be important for your work, do you have an example that you could share with me of such a word, or do you have possibly an example instead of one of those strategies that you've just talked about?

Tara: Sure, well I mean a good word in the context of vaccinations would be the word 'inoculation' because in previous times that was one of the most commonly used words to describe the vaccination. But now it's kind of been usurped by 'vaccine' and 'jab'. So then we can look at that. We can see how often these different words are being used, and we can then take that specific word and say to medical practitioners, public health campaigners, whoever we're working with: don't use this word or use this word because more people seem to understand and relate to this word, it seems to be more effective.

Tony: So that sort of shifting patterns of usage, yeah. Okay, I can get that. Now I'm going to come to the sixty-four-million-dollar question, Joanna talked about 2 million words of data, I think. How much data are you working with at the moment?

Tara: At the moment we're working with about 33 million words of data.

Tony: Okay, so the problem is even bigger. How on earth can you do what you've just talked about with the scale of data that you're looking at?

Tara: Yes, there are methodological challenges with working with datasets that large, but not necessarily the ones that you might expect. So in terms of - that sounds like a lot. It sounds like a lot of data and in terms of being overwhelmed by it, it is a bit easier to manage than you think. And that's because we have specialist linguistic software which is designed just for this. So what we'll do is we'll input all of this data into this software and the software will sift through the data for us and it can produce information, reveal linguistic patterns which would be near impossible for us to discern manually. And, you know, given the size of the data. So if we were to read 33 million words' worth, it would take months. But the software does it in seconds. So the types of things, what I mean by sifting, the types of things that we will get are, it could give us a word frequency list. So a list of words that are most frequent to the least frequent. It can tell us which words frequently, at a rate higher than by just chance, co-occur together. It can tell us, if we compare it to a separate data set, it can show us the words or cluster of words that are unusually frequent in our data set. And then what we can do with that information is, essentially what it does is it highlights the potential analytical avenues that we can then go and explore more qualitatively.

Tony: Right, okay so this is a sort of mixed approach. You're not just pressing a button and this clever machine gives you the answers. It's sort of directing you to potentially fruitful avenues of inquiry, and you then use your expertise as a linguist in order to see whether this is fruitful or not. And if fruitful, how is it fruitful? Is that a fair thing to say?

Tara: Yeah, that is. So essentially, we're able to manage these really large datasets, we're able to analyze these really large datasets, but we don't have to sacrifice the rigor and nuance that a closer, more qualitative analysis provides, cause we do the combination.

Tony: OK, great, I think. I've got that then. Let me ask one more question then, which will really just express a bit of skepticism, 'cause it's always healthy for keeping a conversation going to do so. Do you really need the scale of data that you've got? Could you not just maybe, I don't know, just do the focus groups or maybe just look at a few examples and

draw your conclusions from those. What do you feel the scale of data gives you that using other approaches alone, which don't rely on scale, might afford you?

Tara: The benefit of having a very large data set is that you are getting a more representative sample of online discussions about health. So if you have very specific and small datasets, you can make comments about the linguistic patterns you're observing within that data set. But to make more broad comments about how people understand a specific health issue would be more difficult to do if you're working with a much smaller data set because you don't necessarily have the evidence to back up that claim. Whereas if you work with a very large data set, you can make more evidence-based insights into how the general public are feeling about, or how much they know about a certain medical issue or a medication - just because you have a huge amount of data. So the benefit of having a large amount of data is that you acquire a lot more evidence and you can pick up on things that you may not have even thought of looking for.

[00:20:16]

Tony: Okay, and they're not necessarily that common so that you wouldn't find it in one or two texts?

Tara: Yes, because - and another thing is that the corpus linguistic software, which is the software that we use, it's really good at showing us what's really, really frequent and what's common. But also it shows us what's not frequent and what's not common. So in addition to showing us which words co-occur together frequently, we can also see which ones don't. And we can then draw conclusions from that as well. So what's not there is also very important.

Tony: So words are a bit like people in that they have favorite friends that they hang around with, other people that they shun and that tells you a lot about their character. Is that the type of reasoning?

Tara: Yeah, well you can draw meaning based off the company a word keeps, that's something that a linguist says. So yeah, you can determine how a word is supposed to be meant, what are the implications behind that word based off of the other words it's used with.

Tony: Great, well before I move on do you want to give us the name of one or two of the software packages that you use? You don't have to describe them, but just in case any viewers want to go off and sort of type the name in and see what the software does.

Tara: Yes, there's a good one that's good for if you've not really got much experience with this type of thing called AntConc. That's free, so you can just download that. And another good one is Sketch Engine. That's a good one that you can get a free trial on that one to see if you like this type of thing. But I think for beginners I would recommend AntConc because it's the easiest to use and it's quite user friendly, and it's free.

Tony: That's great. In that case, Luke or Tara, you could possibly just pop the URL for those two things into the chat so that people can go and find them easily if they want to. Well

thank you so much, Tara. I think I'll now move on to Elena because I think I've reached the point in the conversation where I'm convinced that practitioners do find language useful within their everyday work, they do get insights from using forums, for example, in managing pain as we heard. I've heard from Tara and now I'm fairly sure why you'd want to look at large volumes of data, and I'm fairly sure that's fruitful and also it's possible. But what I'd like to do now is turn to Elena, because I think Elena can give me some examples, or possibly just one really detailed example of how this all works in practice. And also, I think this is really important, refocus on the question of what benefit that patients get from this. We know that the clinicians benefits, I just wonder whether there might be some examples out there where patients have actually directly benefited from this. So let me turn to Elena and, entirely artificially, ask her to introduce herself, even though we've been working together for years.

Elena: Yeah, let's not say how many years, Tony. Hello everyone, I'm Elena Semino. I'm a professor in the Department of Linguistics and English Language at Lancaster and I direct the Centre for Corpus Approaches to Social Science here, which is part of the International Consortium for Communication in Health Care.

Tony: And who funds your work in your centre?

Elena: It's funded by the Economic and Social Research Council, which is part of UK Research and Innovation. So it's basically the research funds made available by the UK Government.

Tony: OK, that all sounds very grand, congratulations. Let me turn to you then. Can you sort of work through an example for me? Maybe the one that you mentioned before - that was mentioned before by Joanna, where you've engaged with clinicians and some of this data, you've gone through the work, you've produced the findings and you've helped the clinicians and the patients. Is it possible to give me a concrete example like that?

Elena: Yeah. If I can, I'm going to give you two, but I'm gonna start –

Tony: Okay, let's start with one. Maybe go to two if we've got time.

Elena: I'm going to begin with the one that Joanna mentioned because that is something that really happened in conversation between Joanna, myself and the people from the Trigeminal Neuralgia Association who are running the forum. And so as Joanna said, the forum they knew was lively, but they were also concerned about a small number of people who perhaps even inadvertently, were creating levels of anxiety or potential misinformation that Joanna [and her colleagues] were concerned about. And so at the time, they were wondering what to do about the forum - whether to even keep it open or not. And that was when, as Joanna said, I basically offered to take a look at what happened on this forum. So we - I was provided with a download and I used the kinds of methods that Tara has just alluded to to see what was distinctive about the conversations going on. So for the corpus linguists amongst you, I did a keyword analysis and what was striking amongst the keywords - of course I got lots of references to medication, the things you would expect - but I found

lots of language about empathy, about mutual support, about emotions, about relationships.

Tony: Okay, can I just ask you though what are these keyword things? 'Cause it sounds as though they're important to you. It sounds as though they produced sort of lists of words that sound entirely plausible given the nature of the problem as outlined by Joanna. But why do you find them? How do you find them? What's so special about these things?

[00:25:44]

Elena: Yeah, so the kind of software that Tara has mentioned can be used to compare a data set you're interested in, such as the trigeminal neuralgia. forum, to another data set that might be a good comparator. So I basically compared the forum to a corpus of general English to see what kinds of words - and keywords are words that are much more frequent, relatively speaking and statistically, in one in your data as opposed to the more general data set and so those distinctive words.

Tony: Yeah, that's clear. 'Cause for a moment I was slightly confused, I thought maybe you were doing what Raymond Williams used to do with his keywords, where he's choosing on the basis of his intuition, words of particular cultural salience, etc. But yours is driven by the data and, if you like, driven by the numbers, yeah?

Elena: Absolutely, yeah. it is driven by the –

Tony: Great, carry on. I understand it now.

Elena: And that is how I discovered that - partly, I could see what kinds of things people were talking about. They were talking about medication, they were talking about doctors. We found many references to Joanna herself, all positive, all positive. But also there was a core, a group of keywords, distinctive words that were about emotions, empathy, hope, relationships, even humor, as Joanna said. And so what I said to Joanna and the people from the Trigeminal Neuralgia Association was, one really important thing that's happening on this forum is that people are supporting each other. And so if you close the forum, they would lose that. Partly because, as Joanna said, it's a rare disease, it's difficult to find people in your community who have it. And so the decision was made, as Joanna said, to moderate it, not to close it.

Tony: Okay, so that's a very clear example of a direct benefit for the patients, in the sense that you maybe helped Joanna understand the value of this and all the people who might have thought of closing it understand the value of it, and therefore provoked a reaction which kept it open and did this moderation, which we heard before is terrifically important. But I must say, I'm surprised you went straight for that example because I think I've read something from you in the press about cancer and some of your work there. So am I wrong? Do I misremember that? And if I don't misremember that, can you tell me about it? And I suspect that was your second example.

Elena: It was my second example and then, you're correct, Tony in knowing that my colleagues and I worked on metaphors in communication about cancer. There, the starting

point wasn't a specific group of people and a specific forum, as in the case of the trigeminal neuralgia forum, but it was an awareness of long-standing debates about the dominant metaphor, in many cultures, of cancer as an enemy to be fought. So the battle against cancer that people can potentially lose. So these debates had been ongoing in academia, in the medical profession, in the media. And for example, the NHS, quite a long time ago, started talking about the cancer journey instead of a battle against cancer. So we collected a large corpus of communication about cancer, mostly online forum data from patients, carers, family carers and healthcare professionals and analyzed - the question was: What metaphors do people use and how do they use them?

Tony: Okay, very good question but let me step back into the classroom and say please Miss, what's a metaphor?

Elena: Okay, a metaphor involves talking and potentially thinking about one thing in terms of another - so, cancer as an enemy, or having cancer as a battle - when the two things are different, but you can perceive some similarities between them.

Tony: Okay, and it's distinct from a simile, is it? Where you sort of make that more explicit, you say 'it is like' or 'it is' and something like that, yeah? So cancer is like a battle.

Elena: Yeah, the most general definition of metaphor is about seeing one thing in terms of another, and so in that context even a simile comes under the general umbrella of metaphor. When you're talking about exactly how that seeing something in terms of something else is realizing language, then you might distinguish between a metaphorical expression and a simile, and then it becomes an important distinction.

[00:30:19]

Tony: Okay, great. In that case, that all sounds absolutely fascinating for a linguist, I'll say. For a linguist. Now, why would a member of the general public be interested in your entirely understandable fascination with metaphor? Is there anything in it for them, especially those members of the public who may of course be suffering from cancer?

Elena: Yeah, so first of all we know from research that metaphors can have framing effects so that the metaphors we're exposed to might influence how we think about things and how we experience things. Then if you look at linguistic data, one of our key findings was that there was indeed evidence that, at least for some people, the metaphor of battling cancer could be harmful. And that was particularly when people didn't get better and felt like a failure because they were losing.

Tony: Right, so it's a sort of psychological - it's not that it brings them out in red spots, it's a sort of psychological harm. It makes them feel that they have failed if they don't recover from this terrible illness.

Elena: Yeah, that's right.

Tony: OK, I can get.

Elena: One person on the forum says "I feel like a failure, that I'm not winning this battle". So that was the first finding, which basically then confirms people's misgivings about the metaphor, and then says to health care professionals for example, don't impose it on anybody. Or the media, try not to impose. But there was another finding.

Tony: Okay but I see you using the word 'impose' and I think that I can understand that because channeling my inner skeptic again, I'd say maybe it's a metaphor that really works well for some people.

Elena: And that was our second finding.

Tony: I keep on jumping ahead!

Elena: Our second main finding was that while there was evidence that for some people this metaphor was what we call disempowering, for other people, at least some of the time, it was empowering. Some people felt proud to be fighting the disease. Some people wanted - felt that they found meaning and purpose in that language. And equally the metaphor that was adopted by the National Health Service in the UK, the journey, the cancer journey - so some people, it was using an empowering way that they were, for example, looking at new scenery while they were traveling and for others, it was used to express frustration that they were on a journey not of their choosing. So the main finding of the project was that the crucial thing about metaphors and cancer is not necessarily which metaphor you choose, but whether they're empowering or disempowering and that different metaphors work for different people.

Tony: And is that message getting through to the patients?

Elena: Well, we have done as much as we could to communicate, but the most concrete thing we've done for patients, based on the finding that different metaphors work differently for different people, was to create a resource for patients called The Metaphor Menu for People Living with Cancer which is a collection of a range of metaphors, not just battles and journeys, there's music and nature and fairgrounds and unwanted visitors. So as a resource for patients to find either validation - Ah, I'm not the only one who sees it that way - or an additional tool - I hadn't thought of it that way, but it's it really works for me. So the Metaphor Menu is a resource based on our data.

Tony: So I suppose the idea is if you found that some metaphors can be positive or negative for some people, make them aware of as wide a range of possible metaphors as you can, and then guide them, if you like, through this advice to picking the metaphor that might work for them and produce the most positive framing. Is it something like that?

Elena: Yes, that's exactly it. It's like in a menu, so the metaphor menu is metaphorical, different people will like different things, but also people might be inspired to come up with their own recipes. So it's also inspiration for coming up with the best ways that work best for them, and equally for health care professionals too. So, for example, the Metaphor Menu's recommended by Cancer Research UK, and it has been used in, for example, hospices and in various contexts where cancer patients are being supported. So it is also an inspiration to come up with something that works for you.

Tony: Yeah. Well, I'm so glad you used the word inspiration because I find those examples inspirational. I think it's really nice to see linguistics having real world effects like that, especially in such sad circumstances and they can help people at this time in their life when they need help, that's wonderful. So thank you for those examples, Elena. I'm going to move on now to talk to Gavin. And I'm gonna shift gears slightly and think about a different type of data. Let's get Gavin on first. Hello, Gavin. I think you're muted.

[00:35:22]

Gavin: There we go. Hello Tony, sorry about that.

Tony: Why don't you tell us about yourself?

Gavin: Okay, I'm a Research Fellow in the ESRC Centre for Corpus Approaches to Social Science, so CASS, at Lancaster. I'm a linguist, so I'm interested principally in language and I have a particular interest in how language is used in relation to health and illness and in all different types of health care contexts, including online contexts.

Tony: Did I see some social media tweets recently about you winning some type of big prize?

Gavin: I had a feeling that might be mentioned, yes I recently won some funding and a UKRI Future Leader Fellowship, which will allow me to do precisely the kind of work I'm interested in so, to examine how dementia is represented across all different types of health care contexts, and public communication contexts. Yeah, that's the kind of work I'm interested in, and I'll be starting that in the new year.

Tony: Great, congratulations. Unfortunately, we're not talking about that today, I want to switch us to talk to another issue which I can think of. I think I've just heard three very persuasive presentations or had three very persuasive conversations with people who have convinced me that language has a role to play in healthcare. Linguists have a role in terms of being able to analyze that data and provide insights to clinicians, and also some really nice examples of having benefits directly for patients themselves. But of course, patients themselves often get the opportunity to give quite directed feedback on their experiences. So the forum is patients talking to patients with some moderation, I get that. But sometimes of course, I've certainly done this, people ask you to fill in feedback. You fill in the feedback and say, 'I like that doctor', 'I didn't like that doctor', or whatever, and give some explanations. Surely that's a useful source of evidence as well?

Gavin: Yeah, absolutely. So some work that I've previously done and some ongoing research that I've been involved in has involved looking at precisely that issue, so how members of the public use online platforms to talk about and evaluate their experiences of healthcare services. One example of that, some work that I've undertaken with Paul Baker, a colleague at Lancaster, is how patients in England use the website of the NHS, the National Health Service, to provide patient feedback. And it's just like these types of consumer reviews, like something you'd find on Trip Adviser. So patients give a score between zero and ten, rating the care they've received and then they also leave a written comment, a more qualitative comment where they can explain why they gave the score they did.

Tony: Is that the NHS Choices survey?

Gavin: Yes, that's right, yeah.

Tony: I think I may have filled that in once.

Gavin: Yes, I think that's likely, yes.

Tony: Okay, well let me be sceptical once more then. Okay, in the forum I can see that linguists might be helpful because the interaction is purely linguistic, I assume, and patients in that context aren't giving explicit numeric ratings. Whereas with Trip Adviser or in your case, NHS Choices, I guess they're clicking on a number to say, question: I'm happy about that. Question: I'm not happy about that, etc. And yes, there's a free text box, but could you just ignore the free text and go with the numbers because the numbers are telling us whether they're happy about the specific question. asked? So a little skepticism for you, can you counter it?

[00:39:10]

Gavin: Yeah, so there's certainly value in looking at the numbers. Like you point out, it is one of those quantitative ratings, so patients give scores between zero and ten and this can give us a broad overview of rates of satisfaction, but it can really only take us so far. So by analyzing the language that patients use in conjunction with those ratings, so the language they use in their free text comments, we could better understand why patients gave the scores they did and also generally better understand their experiences and not - On the basis of that kind of insight actually the NHS could learn a lot more about what it was that they were doing well, but also precisely which areas they could improve on. So I think that combination of quantitative and qualitative feedback together is really quite powerful.

Tony: Okay, well that's great. You've made it sound worth doing. Are there methodological issues that you have to face in using data like that? You've got a number which is very nice, so that number says what follows in terms of the text is about something I'm happy about, something I'm sad about. I can see that. But are there any challenges related to the data?

Gavin: Yes, certainly, and lots of them. But I think from the perspective of a linguist and a corpus linguist I think one of the major challenges that we've had to overcome was a lack of metadata. So by that I mean information about the data, about the people who were contributing the comments.

Tony: So you may get comments but you don't - there isn't - there aren't boxes for people to fill in to say I am a man. I am 57. I am born in Liverpool, whatever. You don't have that type of information?

Gavin: Precisely, so the NHS were interested in exactly how those types of issues, those variables so things like gender and age and sexuality, for example, how they affected the kinds of feedback patients gave but rather unhelpfully, they didn't ask for that information in the survey.

Tony: I can see that's a problem!

Gavin: Yeah, so basically we had to use - I suppose be methodologically creative and find ways of getting at that kind of information in the comments themselves. Though that requires –

Tony: Could you do it?

Gavin: Yes, certainly. I think with an asterisk. I think there were - we could certainly do it, it was a bit of a trial and error, but we had to find different approaches and test out different approaches to find cases where patients disclose those aspects of their identity. And we found that way into the data, which was quite effective actually, and that's what Paul and I have continued to do now with some more work we're doing with the NHS, at present.

Tony: I'll ask about that in a moment, but I can see the value of this now. So as well as being able to answer the 'why', related to the numeric score, you can also find out things about the respondents that people who filled in the questionnaire, that it'd be jolly useful to know but it wasn't asked in the first place, so you don't really know it. Can you be sure or is there any way you can check the accuracy of some of these insights that you're gaining? Say for example about age, I think was one of the things you mentioned. They don't tell you what their age is. You try to infer it from the text using all these clever techniques that you have. Have you ever done any types of experiments to find out whether that's a reliable thing to do?

Gavin: Yeah, absolutely. So age is, like you mentioned, one of the variables that we looked at and we've, in some subsequent work we've carried out some comparisons looking at cases where we know a patient's age from the metadata, from the boxes they ticked, but also cases where we've had to rely - Yeah, in a separate data set. In further work we've done focusing on cancer care feedback. So we found that generally the approach we took where patients disclose their age is effective, it is useful but also it has some limitations in that when people mention things like their age, they tend to do it in a specific context so in a context where they deem it to be relevant or rhetorically effective for making a certain point or a certain argument in their comments.

Tony: So I might say something like "I'm an old man with a back complaint and there was twenty stairs to climb" or something like that?

Gavin: Yeah, precisely so there, being old your age is effective for making that particular complaint or amplifying. the issue.

[00:43:52]

Tony: Okay. Well you've persuaded me once again that the data is interesting and that you can look at it. And in fact that you can provide some insights, for example, about the respondents themselves. Are there any other insights that you produced for the NHS, for example, based on the data that they found interesting, and if so, what were they?

Gavin: Yeah, absolutely. So there were lots of - I'll give you a couple of examples if I have time. So one thing that - our analysis in general could tell them some things they knew and we could provide stronger evidence for but also some new things, so some things they

didn't know. And one thing that they didn't know was about the feedback tool itself. So we found for example that there was an imbalance in the scale of the scores that patients gave. So we found that a score of 1 could be given if a patient had a single bad experience, whereas to get a score of 10, a provider would have to provide long term positive service, preferably not just to the patient, but their wider social networks so their friends and their relatives. So we found that this really was something of an imbalance in the score - in the scale, rather. So it's quite easy to get a score of 1, a low score because you could just give one bad episode of care, but to get the highest score of 10 you had to be consistently excellent for a very long time. And that was something that they didn't really know.

Tony: Okay, so it must be the case then that wasn't something which was captured in what I think you're calling the metadata, the information about the respondent. So they didn't say - there wasn't a question saying "how long have you been using this service?" Or "was this your first time", so you inferred that from the text and then got an insight into how the rating score works. Is that accurate?

Gavin: Yeah, precisely. And to sort of loop back to your first question, really I guess that is the value in going beyond the numbers, looking beyond the scores and looking at the language that patients used because that's simply insight that you couldn't get at without looking at those comments in linguistic detail.

Tony: OK, that's really helpful. You've convinced me once again that there's another interesting source of linguistic data that we can use to get insights into healthcare that's helpful for health care providers and ultimately therefore for patients. So thank you very much, Gavin. I'm now going to talk to May. So I'll just wait for May to pop up on the screen.

May: Hello, hi Tony. Can you see me?

Tony: Hi May, how are you?

May: I'm good, thank you.

Tony: Good. Would you start maybe by introducing yourself?

May: Yes, hi. I'm May Lwin, I'm a professor at the Wee Kim Wee School of Communication and Information at the Nanyang Technological University in Singapore. So it's night here and hello to all the audience joining us from different parts of the world.

Tony: Wonderful, isn't it? It's wonderful. And I've so enjoyed the conversation so far and I know I'm going to enjoy this too because I think here, May, we can take the opportunity to shift to yet another source of data, which we've mentioned already. But actually ask ourselves the question is there value to be gained there because we heard from Joanna how there may be issues with looking at certain things in terms of social media. But of course, there might still be value in terms of looking at social media and considering health care issues. So we've looked at forums, concluded that's helpful. We've looked at patient feedback, concluded that might be helpful. But I know that you've done some work on healthcare around covid I think, or work around covid on social media, so I'd really like to hear about that. Could you tell me maybe just briefly about that work?

[00:47:47]

May: Yeah. For about the past decade I have been looking at how digital communication, specifically social media and different types of messaging, can nudge populations' behavior and so - and one of the key areas has been in infectious diseases. So we've been looking at how people discuss dengue, which is a mosquito-borne disease, as well as other types of infectious diseases like influenza and zika and H1N1 and so on. And we just didn't expect the shape and the form of the communication to change so much because all those studies that we had done over the past decade, showed pretty similar trajectories in terms of how communication moves within the social media from government sources and so on. And really in January 2020, right it's now almost 20 months, we started noticing, and I'm sure every single one of us experienced this, that your [hand phone] and the different conversations that were coming out on your mobile platforms, particularly - and even your emails and Facebook and so on were being dominated by this particular disease. And it was a lot more than you know we had ever encountered before. So that's where I started doing work on looking at social media and COVID-19.

Tony: Great. And you mentioned before that you're interested in nudge, and you know when I hear 'nudge' I think of people like David Halpern, but maybe just for those people who are listening, could we maybe talk for a moment just about nudge. As I remember that's something where governments or other influencers try not to just legislate to control behavior, they try to do things which influence behavior gently in a certain direction. Is that a good way of putting it? Or would you prefer –

May: That's right. That's right. In areas like marketing communication you really have kind of the end behavior which is the purchasing or the consumption that's going on. But in health it's really difficult to have that one endpoint, because it could be a series of actions that occur. And so in health communication, particularly, we like to take baby steps to try to get the different segments to undertake certain types of behavior. And health communication, health education can play a role, and I mean if we are thinking, for instance, children and getting them to eat broccoli, for instance. Yeah, it's really kind of little baby steps. Maybe getting them to try a little piece or try piece that might be green but it's not broccoli and so on. We like to think of it in that pattern, so I guess that's where the terminology comes in and you know, so yeah.

Tony: Okay, that's really helpful. So that's nudge theory. I can see how it's helpful in the context of health care and in terms of healthcare communication, I suppose we may have heard some examples similar to it already today, where we know that language can influence behavior in positive and negative ways, etc. So we can choose language or advise people to use certain types of language to militate in favor of more positive outcomes. Yeah, I can get that. But let's talk - go back to your example and talk about covid so January 2020, Covid is everywhere in the human population. Covid is everywhere on Twitter. What did you do to look at it and what did you find out about it?

May: We realized that the type of social media conversations that were coming out where much larger in volume than anything that we had ever experienced. So when we look at the

numbers, even in January like a typical person in a city could be receiving 5 to 10 messages a day and this could be like repetitive, it could also have a lot of emotional elements. When I say emotionally, it's something that's fearful and so on. So these were the types of things that was unprecedented. And so we started gathering all the information that we could gather and I think Tara talked about big data, so we turned to big data. We partnered with one of the entities, scientific entities in Singapore, Institute of High Performance Computing, and we started off first with Twitter. So over the course of the year we collected public messages on Twitter, 55 million messages that were put out on Twitter in English on COVID-19 and then we - I mean over this year, we since expanded it to Facebook and Reddit and so on. But using Twitter as an example, we then tried to understand what was going on through something called sentiment. analysis. So in that particular space, what we do is that for every single message, we tried to identify which buckets or which categories the words would fall into. So some of the keywords.

[00:53:36]

Tony: So this must be related to emotions somehow, is it?

May: Yeah. So we use kind of the psychological facets of emotions. So we have something called fear, which is so common during a disease outbreak. So not just referring to fear itself, but feelings of being scared and feelings of being - unknown, unease and so on. So we've got the bucket of fear, we have the bucket of anger, which is when you feel very frustrated and when you're terrorized and so on. We have another bucket, which we - I think my linguist colleagues might tell me that maybe it's not the best word, but we call it 'joy', but it encompasses many of the components - and listening just now to Elena, she talked about hope and gratitude and inspiration and so on. So these are the types of words that we put - So we had like five different buckets and they were kind of streamed into these.

Tony: Okay. I love this image, May, that you have these sort of buckets and you're doing this sorting task. You've got all of this data and you're working with your colleagues in Singapore and you've worked out a way of putting the data into the buckets. So tweet by tweet, you're doing it. Okay, what did you find? This sounds fascinating. What did the English speaking world think of Covid?

May: I mean, it's been a fascinating journey. What we discovered in the early part of COVID-19 was no - we couldn't have predicted what was going to happen. So when we first started and we looked at this in January, February, March we were finding what we would have expected for myself, as a health communication scholar and that is that the highest, the most number of tweets were in the fear bucket, across the world. So we looked at the world in total and later on we were able to split that into different countries as well. But across the world, fear was the most dominant emotion, maybe something like 70 percent, 80 percent. Every single tweet was on fear, like what's happening, you know, what's going to happen when we're gonna have lockdowns and things like that.

Tony: Please, May, tell me that's changed.

May: That changed after about 2-3 months, when the government started to put in different types of measures. And of course we know the term lockdown and all that now. And then people were inconvenienced, really many, many people suffered, businesses suffered and then you see the - We use actually use a red line so you can kind of visualize the red line, which is anger started to rise. And by June we were predicting, so we publish a paper in June, but we were predicting that anger will overtake fear across the world. And indeed, by August, September anger was the dominant emotion. And you probably saw that, especially in the United States. So like a lot of news there it is and so on and anger like kind of dominated the emotions but what was kind of nice - and it was good from the research perspective, it's not all like doom and gloom - that we also started seeing these - we use yellow for joy because we wanted to, you know - it was also on the rise, that people started thanking the frontline workers, people were putting out messages of hope, gratitude, helping one another, inspiring words, and so on. I think there was also that kind of coming up across the world.

Tony: Oh that's great, that really is. It's so nice to hear that story. It's so nice to hear that sort of spreading of joy, towards the end of it. I suppose reflecting on what you're saying, I can see now how this type of work could be useful for policymakers. So if people are creating public health policy and they're wondering about the impact of it on peoples lives, or people's perceptions of it, the type of work you've described could be very interesting. But it also occurred to me, but I think you've mentioned it, that you'd probably want to break it down for those audiences into individual jurisdictions, in these different countries. So have you done that? I think you said you did. And maybe briefly at the end, can you tell us any differences between these countries?

May: So that work has just begun because we've been looking at a large number of countries and what was interesting is that - and now in hindsight, we know that every country had a unique set of circumstances and unique sets of policies and so on - so what we were able to do was to see that these spikes in emotions and the different types of volatility that the public face emotionally it was driven very much by the types of actions. It wasn't the only driver but it correlated quite a bit to the government activity, so you would see that anger for instance, rises when you have major lockdowns. If you have some types of positive news, for instance - I think Australia they just had Freedom Day - you could see the spike in positive sentiments and so on. And again now, when we look back we realize the whole world has been in such a volatile emotional state. We talk about mental health and many of the issues that so many of the populations are facing and the fact that we could then link it back to certain types of actions by the governments, made us realize that the governments really need to think about how they implement the types of policies or the messages could be conveyed to the public in a more sensitive way. I think that my colleagues have all talked about the power of words, and even how certain words like 'metal' and 'war' and so on can have negative connotations. So certainly there was a lot of learning experiences from there.

[01:00:04]

Tony: That's marvelous, and I can see perfectly how it links back into the work on nudge that you were talking about before as well in terms of developing communication strategies etc. That's so so interesting. Thank you so much for that example, May. I think now we can move over to the audience and start to take some questions. Here, I will be looking to Luke to guide us as to what we should be doing. So are there any questions from the audience, Luke, that you think you'd like us to consider?

Luke: Hi Tony, yes, we've had some very interesting responses. Lots of interest in the procedures as well as the outcomes of this type of research. So I think we've heard a great variety of examples of different approaches and different kind of linguistic methods for what can be discovered here. But we've had a couple of questions about the practicalities and the considerations that researchers have to kind of think about in terms of securing this kind of data. So whether there's any experiences of access issues or how you go about going and sourcing this kind of forum data.

Tony: Okay, could I maybe turn to Elena first of all and ask her if she wants to comment on that, as she's worked with this type of data for longest?

Elena: Yes, so with this kind of data, there is no single answer to whether you can access it and how to access it. But there are a number of things that you need to take into account and I should say that if you look, for example, at the - there is an Association for Internet Researchers, they have some guidelines that will take you through how you think about the data you might want to look at. So you clearly need some kind of consent, communication from the people who run the forum. So here it can vary from - in the case of the Trigeminal Neuralgia Association, I was in a situation where Joanne and her colleagues asked me to take a look because they were concerned. At the other end, you may become interested in a particular online forum, you have no contact yet with anybody, and you might want to make an approach. But to keep things short, and please consult the guidelines of the Association for Internet Researchers, you need to consider the status of the forum, how open is it? Is it possible for anybody to read the posts? Or do you need to be a member to read the posts? How vulnerable are the people communicating on the forum? What are the terms of reference in the privacy statements of the forum and so then you can decide how to make an approach and in what way. How would you use the data? Would you quote lengthy extracts? Or would you just present aggregate information, such as collocational patterns? But minimally, I would say you always have to contact the forum owners in order to make sure that it is okay for you to do the research. In terms of downloading the data, I can say some things, but I'm lucky to be in the position that I never do it myself. I always have a colleague who downloads the data for me, so I'm not sure that this is the best place to go into the technical side, but if somebody wants to answer the more technical question, I'd be happy to hand over to them.

Tony: Okay, that's great. I think we'll also pass over, just very quickly to Gavin, because Gavin, you were dealing with a different type of data, and I suspect that access for that is even more restrictive than the type that Elena has just been talking about. So, Gavin, could you comment on that?

Gavin: Yeah, I mean, I suppose it depends on the context. So the patient feedback work that I've discussed today, that comprises really two separate projects that are quite different in nature. So one, the NHS Choices website is a public domain, so people - it's a public context all of those comments are posted online for other patients or the members of the public to read. And in fact a lot of the comments are directed at other patients themselves, of the readers of the website. So there's an acknowledgement that these are public texts. That said, we still had to take care when we were reproducing those texts in our analysis, for example by not giving details of names or locations or other information by which specific individuals might be identified. On the other hand, in subsequent work that Paul Baker and I have done, again with the NHS but this time looking at cancer care feedback those comments, which are a similar type of comments or similar genre really, those comments are not public, but they were provided in private by patients to the NHS. So there, there were I guess more pressing ethical concerns and we had to get the approval of the NHS as the holders of that data and work closely with them on the project, really.

[01:05:20]

Tony: May, I don't know if you want to add anything on difficulties or issues with working with social media?

May: Yeah. For my team we have been working with public polls and when I talked about the world sentiments and so on, we like to think that we are able to assess what the macro level sentiments across the world are, right. But I also recognize that a lot of what's happening on social media are on platforms or in situations that is very difficult for us to gather the data. And I am concerned that in closed groups, so for instance, I think maybe with the audience we might be familiar with WhatsApp, Viber, WeChat and so on, right? So those are closed social media groups, chat groups and we were not typically able to gather any information from those types of groups. And in a little bit of my work on misinformation and so on, we're finding that those are the places where misinformation tends to be a lot more prolific and shared.

Tony: Right, that's really helpful, May. Interesting observation. Okay Luke, I went around the houses with that because data, of course, has varied across studies, but I think we've covered all the main types. Are there any other questions?

Luke: Yes, so there's obviously an interest here in health issues and a few of our guest speakers have talked about their interactions with practitioners. So there's questions here about, on forums, whether you have that medical expert voice, who's doing a kind of moderator role, but also then when you've kind of discovered your findings, how you get involved with practitioners in trying to implement what you've found and try and affect practice in that way. So like, I'm sure we have some expertise in the group who can talk to that kind of experience.

Tony: Maybe I'd turn again initially to Elena and perhaps Joanna for that one, because it clearly links most closely to the discussion of the work on the forum.

Elena: I think Joanna should go first here, so let's hear Joanna and then you can come back to me.

Joanna: Yeah, I think it is very difficult because, unusually, medics and health care professionals don't get much exposure to really, as I call it, the medical humanities. And they don't - really haven't come to appreciate the importance of working with people like linguists, like historians and artists. But I think it is very crucial that practitioners are informed of this data and it is better and I would encourage linguists to publish their data in mainstream medical journals. It is very difficult and I know particularly we're struggling with trying to publish focus data which is qualitative rather than quantitative, because the articles do tend to be very long and the rejection process is much bigger, unfortunately. But I think we have to try and push that more practitioners are aware of the value of the humanities in healthcare, and that's why I'm delighted to know about this consortium and the efforts internationally that will be done towards improving medics' appreciation, that we need other people to help us. And it's really transformed the way I communicate with my patients and even just the very simple things like when we end our interview, we tend to say "Is there anything else you want to say?" That doesn't bring out much in the way of conversation. If you change that one word, "is there something else", that completely changes the conversation and I think being aware of the language I use has been really transformed by working with linguists, and I hope more medics and other allied health professionals will be attuned to that.

Tony: That's a lovely example, Joanna of the type of nudge that May was talking about before, the 'something' nudge I'll think of it is now. Elena, do you want to add anything?

Elena: Yes, I think I want to add something from the perspective of linguists. Of course it is difficult sometimes to find collaborators or people willing to listen. However, there are people like that. And when - Joanna knows this - when you manage to find somebody like Joanna, and I'm lucky to know a few more people like that in different fields, who are interested in communication and who can see the value of studying language and the contribution that we can make, then that opens many doors and so - Joanna invited me to many events that - where I could reach people who were not already converted. So it is important to persevere, because there are people there who are happy to talk to us. It is really important also to do our part in shedding any technical vocabulary that is not needed, in trying to understand what is of interest and of value to the practitioners, as opposed to what might be necessarily our most pressing research questions and finding a common ground. So it is very much a two-way process, whereby we have to be just as open and flexible as the practitioners need to be.

[01:11:15]

Tony: Okay. Joanna there, talked about the importance of reaching out to the medical practitioners through mainstream medical journals. I wonder whether Gavin or Tara or May has any experience of that and might want to share any insights or advice on how to do it and what to expect?

Gavin: I don't mind starting on this one, perhaps. Actually based on the word that I've spoken about today, Paul Baker and I have published an article in the British Medical Journal, the BMJ Open and I think that was an interesting experience, it was quite different to publishing in linguistics journals. The format is much shorter, it was a four thousand word limit. The process was quicker as well, in terms of the rejections are a lot quicker so. We were rejected at first and then moved to a specific sort of BMJ journal. Originally, we applied, we submitted to the BMJ and then we were moved to the BMJ Open, which is still a great journal. But I think some of the issues that Joanna mentioned earlier are really pertinent here, that there is a bit of a reluctance to publish qualitative research in medical journals, particularly the BMJ has something of a policy really on this issue, which can be a bit of a barrier for sort of linguistic analysis. But I think corpus linguistics is. something of a bridge, a bit of a Trojan horse maybe, to get into these types of journals and do linguistic research and publish there. The key challenge there, as Elena alluded, is really about presenting the method in a way that will be accessible to that audience, to a non-linguist audience because - actually, when we're working on that paper, Elena advised me that I should write the article as if it was for a GP on their lunch break, eating a sandwich. And of course who wants to learn about corpus linguistics when you're eating a sandwich on a 10 minute lunch break? Certainly not me, so we had to find a way really of doing corpus linguistics, but not actually using terms like corpus linguistics, but presenting it in a more accessible way.

Tony: Great, thank you Gavin. That seemed to cover off the point that Joanna raised quite nicely so if there are no other comments there, I think I'll pass back to Luke. We have time if there are any other questions, Luke, we could take them.

[01:13:53]

Luke: Yes, absolutely. I think there's some interest in the nature of the data itself. So I think the types of methods that we've talked about today certainly favor kind of text-based analyses and we talked about forum data, which can come in various kind of modes. Of course, as corpus linguists, the other methods that we use are available for kind of multimodal analysis, so if there's any feedback on combining different datasets, whether that's using forum data in combination with the types of metadata that you have or social media data, whether there are - Or what are the advantages, really, of diversifying that text type, maybe in opposition to or in combination with building larger data sets, of the nature that Tara's talked about, those massive datasets. Yeah, so how do you make that decision between bringing in more text or perhaps thinking about combining text-based forum data with other data types?

Tony: Okay, maybe actually I'll pass over to Tara because I think that you're working on these lines. You're looking at different types of data which are all talking about the same issue I think you were saying. So for example in the vaccination work, I think you're looking at different types of data so could you maybe reflect, Tara, on the advantage of looking at the same issue in different types of data? And also the potential - Maybe the question of knowing when to stop: when have you found out enough and you don't need to collect anymore data? Is it fair to ask those questions of you?

Tara: Yeah, sure. So I guess the advantage of looking at different types of data sources are that - so for example, we're looking at social media data and online forums. Those can be quite informal, so you're looking at discussions between individuals and so they're more informal, and it's more like lay knowledge. Whereas we're also looking at parliamentary discourse, and that's going to be a lot more formal. And so looking at different types of interactions and different genres as well, that can really give you a good insight into how these health issues are discussed in different contexts. Is there any crossover? Could there be more crossover? If we are seeing in social media data say, for example, or the forum data that we're seeing - if we are seeing that there are certain language patterns that people are really responding to and it helps them to comprehend a really complex issue like health and like vaccinations but we don't see that in parliamentary discourse or we don't see that in the news media that then tells us that, well maybe these areas need to be employing this type of language, because that's how the general public is talking about it, but that's not being translated into the news media where a lot of people get their education about this and also parliamentary discourse where we look to our leaders to help us understand these types of things, particularly during a pandemic. So the benefit of looking at different data sets is that you can gain different levels of insight, and that can then help you get a more rounded view.

Tony: Okay, and when do you stop gathering? When do you think you've got enough?

Tara: That yeah, that is the everlasting question. I mean for us, with vaccination discourse, that is really tough because every week something new happens. So you do have to kind of think, you have to accept the fact that you might not get everything and that's fine. Because otherwise you're just going to be collecting data forever and you're not ever going to do any analysis. So I think as long as you think about what you want to find, what. you what you want to explore first, and then think about how what kind of data will I need to explore that? What time period do I need to cover to realistically explore this specific thing? Or what website specifically covers that? So think about what you're wanting to explore in the research questions that you want to try and answer, that will help you determine what data and how much you should get.

Tony: Okay, so I guess I don't know - if you're looking at something like yellow fever you might need to look through a hell of a lot more data to find and have examples of mentions of it, than say for example, like May was talking about, you're looking at Covid on social media, you're going to find lots and lots of examples very very quickly. Is that the type of thing you're talking about?

Tara: Yeah. Yeah, exactly so if you're interested in vaccinations for example, but a very specific vaccination like the yellow fever vaccination, thinking about search terms that you use in order to find data as well as really important. Because if you just did the word vaccination, you would get a lot of noise in your data, so thinking about what you're looking for and the search terms you use to find the sources of data is actually a really good point as well.

[01:18:40]

Tony: Okay, hopefully that was a useful ramble around that question, but can I just ask one last question of the colleagues gathered with me here: are any of you working on spoken interaction in healthcare contacts? We have talked about a lot of textual material today. Tara of course talked then about Hansard, which is sort of a transcription of the speech, but is anybody looking at any data relating to spoken interaction in healthcare settings?

Elena: I have done some work on spoken interaction, yes. We have a project in collaboration with colleagues at Durham, where we're looking at interviews with people who are in psychosis - being treated for psychosis, who hear voices. But I think the example I want to give is, relates to work that we've done with Joanna and others some time ago, which is we - a project that Joanna and Deborah Padfield, an artist, set up involved creating - working with people with chronic pain to create visual images that represented their pain. They were called pain cards, which were exhibited etc. And then this resulted in a collection of pain cards that were printed as postcards which were then made available to other patients in the waiting room of their pain clinician, with encouragement to look through them and take some into the consultation, if they were - thought they would be helpful. And so Joanna and others organized the video recording of these consultations with these pain cards and then an interdisciplinary group of people looked at - tried to answer the question: what difference do the cards make to these interactions? And my bit, as a linguist on that project, was to use corpus linguistic methods to try to answer that question: what difference do the cards make? And the headline finding is people are able to talk about their pain in a more fluent manner, but crucially, they engage in more emotional disclosure, which in the context of chronic pain is really important. So we can use corpus methods to look at conversational data as well, depending on the research question.

Tony: Great. And that I suppose is a perfect example of what, Joanna, you were talking about before when you talked about bringing in art as well as say, linguistics and you also mentioned history. I think into the study of healthcare communication and I think that's a really compelling example. So you found these pain cards very helpful in your own practice.

Joanna: Yes, and I continue to use them. But what I found fascinating as Elena said, when all of us looked at the same little section that we chose, how differently we approached it. So all of us found common ground, but then I, as a clinician, looked at one side of things. Elena mentioned something else, the psychologist looked at something else and it was that body of us putting it together that made us realize also how we look at or a conversation and we all read different things into it and therefore the need for us to learn to collaborate and work with each other. And potentially, this is what we do when we do appraisals for our staff and look at video consultations, particularly GPs in the UK have to have their video. consultations looked at and feedback given. And I think having somebody from linguistics looking at these as well could be incredibly helpful.

Tony: That's wonderful, thank you so much. And yeah, I find that a very exciting picture, intellectually, that you've outlined there. Those different disciplinary perspectives bringing different things makes a lot of sense to me and resonates with experiences I've had as well. Okay Luke, I think we've got time maybe for just one more question before we let people go on their way. So is there another burning question out there? So far the questions have

been really interesting and helped to bring out elements of the discussion that are really important to foreground. Is there anything else?

[01:23:11]

Luke: There's just one other thing I think we've got a good group of people to answer - in the UK in the last couple of days, there's been talks from a minister about removing anonymization. I think we're all aware of the problems of things like trolling and the dangerous side, the dark side of social media, for example. But it seems to me that we've had some comments here about the potential value of anonymization in this context. So with there being calls for the kind of insistence that people kind of disclose their identities in order to participate in these kinds of online spaces, what are the thoughts of the group here on how we navigate anonymization and the value of anonymization for the types of forums that we've been talking about?

Tony: Okay, that is a big question. While people think about it of course I give my own immediate reaction, which had been in the UK there's been a long-term problem about establishing your identity online because of course we don't have identity cards. There's no requirement for them. So in fact, if somebody was to do something like this, I think you'd have to bring in something like that and then you'd also have to establish ways of ensuring that these credentials weren't borrowed, and misused. So I think technically, it would be a very difficult thing to do in any way that had probity or was indeed credible. But let's imagine for a moment that it is possible, and let's turn over to the question of how it may impact on work like this. Does anybody want to comment on that?

Elena: We have - We have William here. William Dance. And I think he's perfect for this question, so here's William Dance, who is also working on the vaccination discourse project, like Tara. But I think he may be the best person to tell us few words about this.

Tony: Absolutely. I should have remembered he was there. William - fire away.

William: So this has come up a lot. So recently MPs have brought it up but in the past, such as with when the footballers are receiving abuse, lots of kind of sports people and social media influencers brought up this issue of verification. I think it's a very sensitive topic because, especially if you are suffering with the disease that means handing over lots of personal information which could identify you to these large companies when you may want to stay anonymous. And there's also the issue of while this may work in, for example, in Western contexts and in Western democracies - so we have strong data and information laws - it wouldn't necessarily work as well in countries where government seek to gain information about the people who use Internet, especially if they express opinions that disagree with them. So I think kind of verification, providing information - it needs some thought put into it, but we're not quite there yet in terms of the practicalities.

Tony: Well I can see those points, Will and it made me reflect on May's presentation. How at the moment you might assume that the information that we can see out there on social media say, in terms of people expressing anger might be an accurate reflection of that population. But if we know that they're being observed, those people may be less willing to

show certain emotions and other emotions may come to the fore, so we mainly in essence get a false reading in such circumstances. Anyway, I think we're done, Luke. We could try and take another question, but we'd be squeezing it in. So what I'd say is we should finish the journey at this point. I've thoroughly enjoyed the last hour and a half in an entirely selfish way, in that I've found out lots and lots of things about all of these topics that I was interested in finding out about. I hope that people listening in and watching also found this of interest, there are lots of really interesting projects here. I'm sure some of you will go off and read about them. Also, we've told you about some interesting tools that you might look at if you're unfamiliar with this type of work. Thank you very much for joining us, do join again. I know that the consortium does have events from time to time so keep an eye out on social media. I'm sure they'll be announced there. But for now it's farewell from myself, farewell from Joanna, Tara, Gavin, Elena, May and of course, thank you to Luke for being such a marvelous host for this, and also thank you to William for coming on at the end and also helping us with technical matters at the end. Thank you and goodbye.

[End]