

Elizabeth Stokoe: “Conversation analysis has many applications, from Silicon Valley to medicine”

*Anyone who has ever interacted with chatbots knows how frustrating they can be. The way to improve bot performance may be coming from an academic field called conversation analysis. This is the area of research undertaken by **Elizabeth Stokoe**, professor in LSE’s Department of Psychological and Behavioural Science. Her research focuses on understanding how social interaction works in settings as disparate as first dates, medicine, mediation, emergency service calls, sales and interactions involving conversational user interfaces. In this interview about her work, Elizabeth spoke with **Helen Flood** (LSE Consulting) and **Helena Vieira** (LSE Business Review).*

Academic Speakers Bureau Taster Talks Series ([find out more](#))

Elizabeth Stokoe will be joining Nick Robins, Professor of Practice in Sustainable Finance, and Alexander Evans, Professor of Practice in Public Policy in the ASB Taster Talks series. Wednesday 31 January, 5.30 to 7.30pm, LSE's Marshall Building.

Can you give us a brief introduction to the field of conversation analysis?

Conversation analysis is a six-decade-old field of research. It’s sometimes misunderstood as ‘just’ a research method for understanding social interaction, but it is much more than this: it is also a theory of human sociality. Given that we all spend a good deal of time interacting with one another, we might investigate whether we can really gain anything from a scientific analysis of something we “just do.” But obviously I think we can gain a lot from understanding how highly systematic and organised conversation is, when researched rigorously, and it tells us an incredible amount about the power of language to shape our daily lives.

Unlike much of what takes place in other social science fields, conversation analysts explore social interaction ‘in the wild’ rather than in the lab, through simulation or role-play, or via reports about conversation collected in interviews or surveys. We gather

audio and video recordings, from single cases to hundreds or even thousands, and unpack each encounter starting with a detailed technical transcription. Our job is to understand all the composite activities (the ‘actions’ and how they are built in and across turns and sequences) that comprise complete encounters from the moment they start to the moment they end.

I think it’s useful for people to know that “conversation” analysis is a bit of a misnomer since conversation analysts are not only interested in “the words”. This is because we interact using multiple resources: our bodies, gaze, intonation, material objects in the environment, all of these and more, to create, maintain, and repair intersubjectivity and to progress smoothly through an encounter.

What field of knowledge does conversation analysis belong to?

The origins of conversation analysis are in sociology in the late 1960s in the United States and, more specifically, in a field of sociology called ethnomethodology. The pioneer was Harvey Sacks, along with Emanuel Schegloff and Gail Jefferson. Sacks worked with Harold Garfinkel and Erving Goffman, whom a lot of people I’m sure will be familiar with, even if they aren’t sociologists. I’m a psychologist by background and my route into conversation analysis was via discursive psychology. You’ll find conversation analysts working in linguistics, communication, business schools, medical schools, pretty much anywhere there is work to be done on social interaction.

It must be important in the medical field.

Yes, conversation analysts conduct a great deal of research in healthcare and medical communication settings, with a great deal of impact. A now-classic example is research conducted (using conversation analysis within a randomised controlled trial) by conversation analysts [John Heritage and Jeffrey Robinson](#), who examined the impact when doctors changed just one word in a question: “any” to “some.” The issue: patients were frequently leaving appointments without voicing all of their concerns, resulting in dissatisfaction and inefficiencies. One reason might be that doctors’ opening questions, such as “What can I do for you today?”, typically elicit only one concern.

Recognising this problem, medical school training recommends that, after discussing the initial problem, doctors then ask, “is there anything else we need to take care of today?”

However, conversation analysts have shown, across many settings, that questions containing the word “any” typically receive negative or ‘no problem’ responses. In Heritage and Robinson’s trial, one group of doctors asked the “any” question and another used the word “some”: “Is there *something* else we need to take care of today?” That small change showed a statistically significant uplift in reported concerns. The finding tells us some other interesting things. First, while not every “some” works, outcomes cannot simply be attributed to individual differences (for instance, in age and gender). Second, it suggests what language *should* appear in communication training.

Is conversation analysis being used for marketing purposes?

It’s an interesting question. One of the projects I worked on some years ago, which started me down this path of working with organisations to address communication challenges, was an investigation of initial inquiry calls into community mediation organisations. When people have a neighbour dispute, what do they do? When they can’t resolve it for themselves, they try to get outside help. I have a few hundred recordings of people calling community mediation services. Hardly any of them start with, “Hello, I’d like to make an appointment with the mediator.” Instead, they more typically start by telling the mediator that they called somewhere else first. “I’ve just phoned the police and they gave me your number” or “I’ve just phoned the Council and they’ve put me onto you”. So, straight away, you can see that mediation services have a kind of marketing or PR issue because people with a neighbour dispute don’t always know to call them like they might know to go to a doctor if they have a persistent headache. Right from the get-go, mediators can be on the back foot and need to try to fit the person’s problem to what their organisation offers.

This means they must almost always explain what mediation is at some point in the call and they tend to do it in one of two ways. Either they explain the ideology or ethos of mediation (saying things like, “we’re neutral, we’re impartial, we don’t take sides, we don’t judge”) or they explain it as a process (“first this happens”, “then we do this”, “the final stage is this”). My research showed that process-based explanations were more likely to get people engaged than the ethos-based ones. And it’s in the call that people decide whether to become clients or not, in that very moment, responding to explanations of what the organisation does. This is very different to asking people in, say, a focus group, to imagine they had a neighbour problem and what type of

explanation of mediation services might engage them. This finding, and what works to engage potential clients, ended up being used not only by mediation centres themselves, including rewording their websites and leaflets, but also by the Ministry of Justice, who used the research to redevelop their promotional materials for a marketing campaign for family mediation services.

Is there a set of techniques that you stick to?

Conversation analysts only work with naturally occurring interactions and we always start with recordings. Sometimes this involves collecting recordings, but, in my own research, I've often been lucky to work (ethically, of course) with organisations where recordings are already made (when they say, for instance, "this call may be recorded for training and evaluation purposes"). The police, for instance, makes recordings as part of particular workplace practices. Audio recordings of telephone calls are quite easy to work with, because they're often digitally recorded.

But conversation analysts work with quite complex environments involving a high level of technical skill to collect data. For example, I'm currently a co-investigator on the ESRC-funded Centre for Early Mathematical Learning ([CEML](#)). One of the workstreams focuses on understanding the foundations of mathematical skills and learning before children are in school (when they are at home, walking to the park, shopping, etc). One of the postdoctoral researchers on the project has designed recording practices using multiple cameras and camera types, to capture, as best as we can, high quality recordings in a very dynamic environment! But this is always the first step, recordings that are either provided or that we collect.

The next step is to [transcribe](#) the recordings, and for that we use the Jefferson System, named after its inventor, Gail Jefferson. It has become a universal standard system with many symbols and marks that represent the detail and 'mess' of real talk, its pace, the pitch movement in words, where overlap starts and stops, gaps and silences, and so on. The aim is to represent and preserve not just the words uttered, but exactly how, when, with what pace, with what intonation and with what gestures, all those things that people use as resources to interact.

I often compare the transcription system to music notation. If you can read music, you can imagine what the music sounds like. It's somewhere between a tool for doing the

analysis and the first step in analysis itself. The recordings are the primary data, but transcripts enable other analysts to access the encounter too. We don't gloss or tidy them up, and we transcribe long stretches to see how something unfolded across a sequence. For example, going back to the mediation calls I mentioned earlier, you can identify the point at which the mediator begins to explain their service, and track what happens next in terms of a caller's engagement and disengagement, which can be subtle, but systematic. Once you have a recording and a transcript, you can ask lots of questions of the data.

What qualifications and training do people need to get into this field?

Most have a PhD. Conversation analysis is not often taught at undergraduate level. Sometimes you might find it within a postgraduate programme. However, the conversation analytic community, which is big and international, runs lots of conferences and training events that people attend from around the world. We're a community of lifelong learners, constantly honing our skills, like a craft – music is a good analogy again. Analysts (or musicians) might really know and understand interaction in telephone service encounters (or in classical piano), but never have worked with interaction within, say, a surgical team (or a jazz trumpet). Conversation analysis is a cumulative science, where the tools evolve as well as everything we know about interaction across languages, settings, modalities, and so on.

How is conversation analysis aiding the development of chatbots?

In 2018, I published a book aimed at a general audience called [Talk: The Science of Conversation](#). It turned out that people began to read it in Silicon Valley and its tech counterparts around the world, where conversational products of various kinds were being developed. As a result, I was fortunate to work as an industry fellow for six months with [Typeform](#), a data collection software company, where what it meant to be “conversational” was a big question. My experience with Typeform spread to other start-ups, and for the last couple of years I've been working with Dr Saul Albert (Loughborough University) and Cathy Pearl (Google) to bring together our collective expertise in conversation analysis and conversation design. We've been running expert classes with the [Conversation Design Institute](#), addressing questions like should conversational user interfaces interrupt, or should they make [so-called speech 'errors'](#)?

For example, real human conversation is littered with speech perturbations like hesitations, ums and uhs, pauses, gaps, repairs and so on. So, should a ‘conversational’ product do the same? Some say “no”, because they regard those perturbations as errors that should not be part of human social interaction. But conversation analysts have shown that they are crucial resources for doing all sorts of things. This is why we don’t tidy them away in our transcripts. Speech perturbations are data and they’re telling us something important about the action being built. For instance, in our recent expert class, we discussed Boris Johnson’s interactional style, which often comprises a kind of designed messiness that is littered with “ums”, making his speeches seem more improvised and spontaneous than they probably were.

More generally, one thing speech perturbations do is display (un)certainly or hesitancy and perhaps this would be something that chatbots could incorporate. ChatGPT issues a disclaimer about the limits of a response but what if it instead expressed uncertainty by saying “um”? This would be ‘conversational’ if we were leveraging what humans do to produce products. We’re not necessarily recommending this, but it’s a useful thought experiment which gets to the heart of what ‘conversational’ means, and how authentically ‘human’ we might want our conversational technologies to be.

When you interact with a chatbot outside of work, do you find yourself thinking of its flaws?

I spoke at an event a couple of weeks ago for a fintech’s conversation design group. We talked about how many organisations that have a chatbot are trying to get their customers onto the chatbot, rather than picking up the phone. The challenge is to identify and describe what it is that people see when they look at a chatbot that makes them immediately think, “this is not going to work”. Is it the placement of the chatbot on the screen, the words it’s using, what it’s trying to say to you...? This might be one of our next research questions: What is it about a chatbot that makes you think it is or isn’t going to address your needs? Of course, this is just the same as any encounter with a human interlocutor, too (talking to a human doesn’t guarantee a quality conversation!). If you phone an organisation, you similarly, often tacitly, know pretty quickly that “this is going to be hard work”.

Where do you see the field going in the next five years?

Conversation analysis is an exciting and rapidly growing field. Early research focused largely on telephone calls and largely on North American conversations. Over the years, the datasets that we're analysing has become a lot more diverse in terms of settings, languages and modalities. But one trajectory has been to conduct much larger-scale work and identify some of the universalities of the machinery of social interaction and, as a [recent paper](#) in *Nature* showed, "shared cross-cultural principles underlie human prosocial behaviour at the smallest scale." Another trajectory is to use conversational analysis to address issues of social and [racial justice](#), and the micro-workings of politics, power, and "-isms".

The development of conversation technologies is also drawing a lot of attention in the field, as discussed above, but, more generally, what we think we know about 'conversation' is already leveraged into the development of many conversational products, from communication guidance and assessment tools to skills training. But some of these products and activities don't have an evidence base at all, at least from my point of view. Instead, they're built off commonsense about conversation in terms of what we think sounds right, but there's nothing scientific about one's own anecdotal. In my work, my current goal is to push as much evidence into conversational products as possible and give them the integrity that they deserve.

This Q&A was edited for clarity and conciseness. The editor apologises for doing away with "ums" and "uhs" in both the interviewers' and interviewee's discourse.

-
- *This interview represents the views of the interviewee, not the position of LSE Business Review or the London School of Economics and Political Science.*
 - *Featured image provided by © Elizabeth Stokoe. All rights reserved.*
 - *When you leave a comment, you're agreeing to our [Comment Policy](#).*
-