

Raising the Bar Sydney 2018

Nick Enfield – And the word was 'Um'.

Welcome to the podcast series **Raising the Bar Sydney**. **Raising the Bar in 2018** saw 20 University of Sydney academics take their research out of the lecture theatre and into bars across Sydney, all on one night.

In this podcast, you'll hear Nick Enfield's talk "And the Word was Um."

Enjoy the talk.

[Applause]

Thanks very much. Hopefully you can hear me. So it's my pleasure to be here.

What I'm going to talk to you about tonight is some aspects of language which you all know but may not all know you know. And we'll look at some things which should sound familiar to you, but I'll talk about the research that I've been involved in that kind of delves into the very everyday aspects of you know, what it takes to string sentences together, and in particular what it takes to have a conversation.

So we think of language often as being this kind of static form. We think about the written word, we think about dictionaries and grammar books and these kinds of things.

So when we think about the science of language, those are the things that we often gravitate towards. And what I've been doing is trying to get out of that sort of basic idea of language as being the written form and the grammar and so forth, and looking at language in the wild.

What that means is going out into the field with video recorders and sound recorders and recording language in its natural setting. And the natural setting of language is human interaction in the form of conversation. Now obviously we use language in other forms like this. It's just me talking at the moment, although you will have an opportunity to ask some questions when we're done. But this is obviously exceptional, so the real setting in which language is learnt and in which we use language is face to face interaction, and that's what I've been studying. And it's not easy to study it.

So if you're studying the written word, if you're studying grammar, the history of words and stuff, you can use books, you can go to the library. But if you want to see what happens in real time in language, you've got to go out and record it in real time, and you've got to transcribe it and translate it if you're looking at other languages and really pick it apart in great detail.

So I'm going to be talking about some of the findings of that work and talk a bit about how we've done that work. So the basic idea behind what I've been doing in my work with many colleagues is to say language is not primarily a form of information transfer. So we think of it sometimes in psychology or cognitive science as a kind of tool for encoding information. Instead -- I mean, it is that. But in addition to that, it's also a kind of social glue. It's something we use to attach ourselves to others in all sorts of ways. And in many ways, when we're using language in a conversation, we're kind of sticking ourselves onto other people and getting sort of connected within a bit of machinery, the conversational machinery that I study, which keeps us together as if we're sort of sitting in a single vehicle, each playing our part. So an idea that we could start out with in order to kind of get at this is the idea of interpersonal commitment.

Okay? So when we think about language, we don't immediately think about interpersonal commitment. Interpersonal commitment usually has to do with social relationships, these kinds of things. Well, just think about what happens when you tell a joke for example, okay?

So I might tell a joke about you know, a horse walked into a bar and ordered a beer and the bartender said, "Why the long face?" So you've all heard that joke. Now if I announce to you that I'm going to tell a joke, what happens is that you by looking at me and by taking up my invitation to listen to my joke, you're essentially making a social commitment to me. You're essentially agreeing to pay attention to me, to listen as I go through the joke and then when I get to the end of the joke or the story, the narrative, whatever it is, you are signing up to my sort of project of telling this and you're agreeing to monitor what I'm saying for the funny part or the amazing part if it's a story. And then to react accordingly, okay? So if it's a good joke, you'd better laugh.

But the point here is that I'm inviting you to commit to this sort of project that I'm starting. Well, jokes and stories are something we use all the time, but obviously we're not doing that kind of thing with language all the time. So that commitment we see in something even more basic -- so think of something simple like just asking a question. Okay? So I might not have any relationship with you. I might never have met you and I won't meet you again. But if I come up to you and look at you and ask you a question like, "Excuse me, what time is it?" Or, "Does this bus go to central?" or something like that, I'm actually drawing you in to a commitment to me, that is to pay attention to my question and to give me a response to that question. And this is partly why we feel so weird and uncomfortable if somebody asks us a question and we just walk off, okay? A number of us have experienced doing this to people and it feels bad. And just imagine if you did that to people who you did have relations with. Someone asks you a question, if you just walk off, don't pay attention to them, it has a very strong meaning, right? Because we're going against that basic commitment to respond to a person. So you see there that with something as simple as a question, it's not a grammatical structure that I could describe, but it creates a momentary social relationship which you have to respond to.

So there's a similar kind of commitment in something even more basic in conversation and that is what we refer to as the turn-taking system in social interaction. So when I am having a conversation with you, whether I'm telling you things or asking you things, just imagine two friends having a good conversation. We have this thing called turn-taking which means basically we have a system of one person talking at a time. There might be a little bit of overlap. There might be some small pauses, but usually if we're on the phone and we're talking, it's going to be either you or me talking and it's not going to be silence.

So there's a kind of puzzle for scientists of language to how that system works, how is it that we are able to switch back and forth between speakers. Well, we've discovered in looking at this turn-taking behaviour across languages of the world, that there's a very well-defined and very sort of tight system for doing this. So here's what we did. We went to countries all around the world, places like including Laos and Ghana and Japan and Korea and the UK and Italy and Russia and Southern Africa, very different languages. And the different researchers that were involved in this project and other projects had to go and spend time in the field recording people's interactions and then annotating them using computer programmes that allow us to precisely measure the timespan between my turn finishing in a conversation and the next person's turn starting up. So you can just imagine sometimes there will be a little gap. Other times there would be a little bit of overlap. Sometimes there's variations on those things.

So if you go and you look in people's sort of descriptions of different cultures and their travels around the world, sometimes people will comment on the different conversational style that

different cultures have. So they might say, you might read in descriptions of life in Scandinavia that people are extremely slow at responding to questions. They take a couple of minutes to respond to an invitation to have a cup of tea. You know, or they might answer you in the afternoon if you ask them a question in the morning. I mean, these are the things that you read about in the literature. Well, we were interested in testing these questions. I mean, on the other hand you have people saying, "New York City, you can't get a word in edge-wise. People are talking on top of each other." So we thought, "Well, let's go measure exactly what it is that people do."

And what we found was something quite different from what you would be led to expect in our sort of literature. So we found that if you take all these languages and you put them all together with all the measures we had, many, many of these transitions between one speaker and the next in conversation.

And we found that the average gap that is left between the end of one person's turn and the beginning of the next person's turn in a conversation is 200 milliseconds. Okay? So 200 milliseconds is a fifth of a second and it's the amount of time it takes for your eye to close and then open again in a natural eye blink. Okay? So it's a very short period of time. And in fact, we don't even hear that time as silence, okay? So you measure it using instruments. You can tell it's a little bit of silence, but it doesn't sound like that. So what's interesting is that across the world, the grand average is in this very fine overlapping in speech from one speaker to the other.

Now you're asking, "Well, what about across languages? What if you divide those results up among the different languages?" We found that languages do in fact differ, but they differ very slightly. So English happens to be around the average. So if you just collect conversations in English and time that turn transition, you find that it's about 200 milliseconds, maybe 230, something like that. If you go to Japan, you'll find it's much tighter.

We found that it was almost zero time delay. The people were very, very quick to respond. And at the other end of the scale, you had Danish, right? So the prediction in a sense was correct that Scandinavian language would be at the slower end. But it really isn't very slow at all. So the average response to somebody's question or to somebody's remark in a Danish conversation is around 400 milliseconds. So it's still a less than half a second gap. Now what's interesting here is that on the one hand, the languages of the world are showing a very sort of strong tendency to act in a similar way, that people do time their responses around an average sort of turn-taking target. But there are small differences across languages. For example, the difference between the English average for turn taking and the Danish average. That's only a difference of 200 milliseconds, but it gets blown up in our minds, very exquisitely sensitive to this small timing difference. So we go on a trip to Denmark and we come back reporting that people take an age to respond to our questions. Okay? But actually the differences are just the blink of an eye, but we're so sensitive to it that we can feel it.

So what we also found when we dug a little bit deeper in that particular study was a principle of turn-taking that again sort of related back to this commitment in social interaction. Everybody is in a way committed to trying to hit this target of no gap and no overlap in this sort of to and fro of interaction. So we asked, "Well, what happens within our language when you get differences around that average?" What's going on in English for example when someone takes a bit longer to respond to your question versus when they respond more quickly than average? Well, you probably got a hunch about why this might be. So if I say something to you like, we're on the phone and we're going to go out to see a really great talk at Raising the Bar and you know, I say, "Well, you know, what about swinging by my place on the way?" Because you're the one with a car. And then there's a gap. Sort of the 200-millisecond point passed and the 400-millisecond point's passed. And so we actually find that

when we've got examples in our data that by the time the gap in that context goes up to about one full second, okay, which is not really long, but it's an age in the sort of cognition of language. You get people redoing the question in a kind of negative way. "Or maybe that wouldn't give you enough time to get there." The other person says, "No, sorry, I won't be able to do that." So people resist giving the response that is sort of not the preferred one as it were. So it's actually quite a powerful generalisation.

So we found across all of the languages we looked at in that study I mentioned, that if you divide all of the responses to questions into the answers on the one hand -- so if there's a yes/no question, if you say yes or no, that would be an answer. That would be directly addressing my question. And you separate them from all the non-answer responses, right? So we get plenty of those. You might say, "I don't know." Or "I'm sorry, go ask John," or what have you. We find that all of those non-answers are slower to arrive. People delay them. And whether it's because they're processing whether they can answer the question or not, or whether it's because they're sort of trying to put a buffer between your question and my kind of not really perfect response to your question that's an open issue. But we find that people are quite consistent about that not only within languages, but across very different languages. So it doesn't matter if it's Japanese or Danish or English; if you answer the question, your response will come more quickly than if you do something else, right? There's this great sort of sensitivity to it. So what that means is that if you're delayed for some reason in responding, you're vulnerable to being understood in a certain way, right? Like with the example of I'm asking you to pick me up on the way, if I delay because I'm distracted for example or I'm reaching for a word that I can't think of, then there's a danger you're going to think I'm about to refuse.

So what do I do? Right? What I do is I say, "Um, okay." Now this um or uh, these are examples of hesitation markers that all languages have. And we think of these as being these sort of rubbish language, right? Or we think of this as something in language that we need to eradicate from our speech. But actually in this context you see that words like um and uh in conversation play an important role. What they say is, "I'm attending to what you've just said." Or "I'm continuing what I was just saying." And I'm just processing it. It's not that I've you know, phased out. It's not that I'm not kind of living up to my commitment in this interaction. But for one reason or another, I'm reaching for a word. I'm formulating my thoughts. There's something going on in here and the thing is you can't hear or see what's happening in my brain so I've got to externalise that in some way. So all languages will give you that kind of a signal like um, which makes that private issue public. And it makes it public in a situation where I actually need to let you know that something is coming. I am going to respond but you just need to give me a second. And again, the timing of those things is really very, very fine. So researchers in the US did a big study on um and uh in English, and they got thousands of examples of these from recordings of English. And they found that if you look at the average time of resumption of fluent speech after um and compare that to the average time of resumption of fluent speech after uh, you find that the ums take longer to get back on track.

So they argued that um and uh on that basis, um and uh actually have a distinct function in English and that they should be treated as words in the English language because of course other languages don't have these same words. These words don't just arise. They're not like cries or shouts. They're actually a word that children have to learn when they're learning the English language. And I have this very subtle, very fine function in social interaction. So we're all taught of course not to say um in public speaking. We're taught not to say it in job interviews and things like that. And there's a reason for that, right? Because it reveals that I'm struggling in a way to get on track. But when you're in an informal interaction with people who you talk to all the time, it's actually an extremely useful device to have, and there's a very good reason why um and uh exist.

You're listening to Raising the Bar Sydney 2018.

So what if in the sort of to and fro of interaction you miss what somebody else said? You didn't hear it for one reason or another or you didn't recognise a name that somebody used or you didn't understand an expression. If that's the case, then you've got a problem, right?

Because interaction is such that it moves forward, inexorably forward, okay? So when someone says the next thing, you know they're going to say something next after that and keep going in another direction if you've missed a crucial reference, for example who the person is talking about, you'd better fix it right there, right then. So what do we say? We say things like, "Huh? What? Pardon me. Excuse me."

So we have this whole array of words that we use to get people to back up and repeat what it is that they were saying. So huh is one of the words that again we get told, "Don't say huh. Say pardon me or excuse me, or could you please repeat that?" More polite sort of versions of that. But the real sort of fundamental basic word for that function is indeed huh. And that's a word that we studied again in an international team of people that looked at languages across the world. And we got data from 30-plus languages from very different parts of the world and found that not only do they all have a word that has this function, but the word actually sounds like huh everywhere you go.

So it has very subtle differences. In Spanish, sometimes you find it's more like heh? In some languages it's more like hah. But it's always sort of around that area of the possible vowels that a language can make. You never get any language that says hee or who or something else for that particular function. So what we found there -- it's really a remarkable finding from our point of view because it's a universal word in human language as far as we know. And as far as we know, there are no other ones. Right?

So everywhere you go, whether you're talking about words for objects, words for actions, words for qualities, these will obviously sound very different in different languages. That's the nature of language, right? We can't understand other languages when we hear them. But if you go amongst the people in any language in the world, the findings of our work suggest that you will always know when it is, while you might not know what's being spoken about, you'll know when someone is getting the other person to repeat what they said by saying huh, heh? Okay? So going back to this commitment in conversation idea, when you say huh, what you're doing is you're imposing upon the speaker for one reason or another to go back and repeat what they said. And sometimes it's my fault, okay, or I wasn't paying attention or I didn't recognise what you said. Sometimes it's your fault. You referred to a person by name and I didn't know their name. But regardless of that, what I'm doing is taking you up on this commitment to stay on track in the interaction. And that's exactly what we find people do. If you say huh to someone once -- if you say it like three times in a row, it's a bit of a problem. But if you say it once, you know, then you will find people are very good about just going back and repeating what it is that they said. So there's a sense in some of the examples that I've been talking about of accountability to each other in social interactions. So we're accountable for trying to hit that point in interaction where we're sort of on the average quick to infer the turn-taking. If we can't do that for one reason or another, then we make a signal that says, "Okay, sorry. I'm not going to quite come in on time." And that's oriented to this kind of accountability that we have to each other.

If I say huh to you and you repeat what you said, you're showing that you're accountable to making sure I know what you're saying before you go on to the next step. So if I tell you a joke and you didn't hear, I say, "A horse walked into a bar," and you didn't hear what animal it was or what person it was, then that would be a problem, right? Because you wouldn't get the punchline. So I am committed to making sure you understand what's going on as we go forward through to the end of whatever the interaction is, whether it's a narrative, a story, a

joke or just a conversation about our friends or about something that happened. So I want to talk about another little kind of traffic signal. So all these kind of words like um and ah and so forth that are like traffic signals that help us to direct the traffic of the social interaction. And one of the really interesting ones that brings us back to this idea of things like stories and narratives and jokes is uh-huh, mm-hmm and these kinds of words. So we're very familiar with those.

We've all probably had the experience of watching somebody on the telephone. We can't hear what the other person is saying, but this person is saying mm-hmm, uh-huh. Mm-mm. You know, all these things. Prompting the other person. So we can imagine what's being said. Obviously we don't know what the other person is saying, but what we know is that that other person is on some kind of a journey towards some form of a punchline, okay? They're either telling a narrative and they're going to get to the bit which is evidence of why you know John is a real jerk. Or they're getting to the funny bit, or whatever it is that they're getting to. They're not there yet. We know that by hearing that someone is prompting them to continue with the uh-huh and the mm-hmm.

So there's something really important about those types of what we call feedback words or backchannel words or continuer words that orient to the mechanism of moving forward in the social interaction. Okay? So as a listener, when I'm listening to your story, no matter whether I'm excited, I can't wait to hear what you're talking about, or whether I'm absolutely bored and I just would love nothing more than hanging up the phone, I'm still committed. As soon as I started that conversation, I committed to seeing it through. I've got to find a way to get out of that conversation in sort of a natural way. I can't just hang up. Or if I do hang up, it's a serious problem, right? It would be a serious problem for the relationship between the two people on the phone. Just try it.

[Laughter]

Just try hanging up in the middle of the conversation. It's funny because it's so true, right? You really couldn't even do it as a joke or as an experiment on your friends. Or if you did, you'd have to have a major kind of debriefing session and definitely clear it with the ethics department at the university.

So I want to talk about a study that was done on these kinds of words, again down in the states by a psychologist by the name of Janet Bavelas. And she was interested in how when people are telling these narratives and somebody is saying things like, "Uh-huh, mm-hmm" to them and paying attention to them, she had this hunch that somehow the responses of the listener were affecting the fluency, were part of a system that the speaker was also sort of part of.

So she got this experiment going where she had pairs of people come into her lab and she recorded them talking to each other. And she said, "Okay, you know, you are going to tell a story about a near-miss experience. Like you almost got killed or you almost had a bad accident or something like that. So just think about that for a second and we're just going to record it and you're just going to listen to their story."

Okay, so this is a very natural thing to do, right? If I have a story about a near-miss experience, you know what to do. You listen, you pay attention, you nod, you say uh-huh and these kinds of things. And it's exactly what people did. So she recorded those and then she had a second condition. And in this second condition, she secretly told the listener of the story an instruction that the speaker didn't know about. So they said to them, listen, "Okay, you're going to listen to this person's story, but there's this button here underneath the desk and you have to press this button every time the person says a word starting with the letter T."

Okay, so the, table, tree, whatever the word is that is written with the letter T, they've got to press this button one time. And they said, "I'm going to test you later on how well you did." So what that does is it causes the listener to stop paying attention at all to the content of what the speaker is saying, right? They just completely tune out of what's actually being said because they're too busy trying to think of, was that a letter T? And press the button, press the button, press the button. And the result was that you know, so what that team did is that they then studied the stories that these people told in those two situations.

So the only difference in the two situations was the behaviour of the listener, okay? But what they found was that in the normal condition where I'm actually paying attention to you, the speaker is quite fluent. They're quite able to get through this story. They're quite able to reach their conclusion in a fluent way, okay? And these are obviously untrained people. They're just regular people like you and me telling a story. In the other condition, when the person was not paying attention, what they found was breakdown in the fluency of this person's story. So people would sort of go on a different route and they would repeat themselves. They would become less fluent. They'd get to the punchline and then they'd keep going. They'd circle back to the punchline because they'd be waiting for like the person to go, "Oh wow," "Gee," "You're lucky," or whatever the appropriate response is. And people were missing that.

And so the conclusion to that was you know, people when in interaction we form a kind of a single mechanism, a single unit and the use of language is not something that is just situated in the individual. Language is something that actually operates in a diad, that is to say in a two-person or a larger group kind of mechanism. And that also really sort of comes back to this idea of social commitment because when I am listening to your story, I'm committed to paying attention. I'm committed to giving you what you need in order to get through this task and get to the end of it in a way that is fluent and appropriate and comes off well. The interesting thing about this is that all of us here are capable of doing that to probably greater or lesser degrees, but with a bit of individual variation, all of us are capable of doing that without really even thinking about it. So the conclusion to the things that I've been saying to you this evening is really that the little words of language, the words like the ums and uhs, the huhs, the mm-hmms – these little bit of language, they're the last bits of language we would think of sort of leading us to the core of what language really is all about. But I actually think, and the people who I've been working with over the years agree that in fact these are the kinds of words that really direct us to what's essential about human language.

The human language is not just this system for packaging information to tell people things. It's actually a system for connecting with people. I don't just mean in the general sense of being their friends, but actually hooking up with them in a kind of mechanism where you are playing a role as listener, I'm playing a role as speaker and that role is constantly reversing in such a way that we're committed to each other in the course of that. So in these ways, language really is in its home in conversation and in conversation there are all of these commitments. So things like showing that you're paying attention is really – deep down it's a kind of moral commitment. Because if you give it up, if you walk off, if you fail that, then it actually becomes a question of who you are to me and what the consequences will be to our relationship. So I think what's really telling and what's really interesting about this is if we then reflect on why it is that we click with certain people, why it is that we feel awkward with certain people, that we would predict you can examine those interactions and look at these seemingly quite technical aspects of the interaction, but locate those problems in the fine timing and the fine sensitivity to the collaborative task that is cooperation. So thank you. I'd be very happy to take some questions.

[Applause]

Thank you for listening to the podcast series of Raising the Bar Sydney. If you want to hear more podcasts from Raising the Bar, head to raisingthebarsydney.com.au.