# Decoding Human Interaction: Transcription Conventions for Conversation

## Transcript

Video: <https://youtu.be/oHgyVgc_0Qs>

Full resource: https://www.ncrm.ac.uk/resources/online/all/?id=20831

Welcome to this tutorial on transcribing for social interaction. I hope to show you that transcribing, although sometimes seen as a less glamorous task for qualitative researchers, can actually be a fascinating journey into the intricate dynamics of how we get things done and talk. So for me, rather than a mundane chore, it’s a magical engagement with the unfolding moments of everyday life. And I kind of see a transcript as a kind of living record, capturing the essence of human communication.

So just to give you an overview of what’s to come, I’m going to first begin with telling you why I think interaction matters, before outlining some of the key insights of conversational analysis, and then I will go into the many different transcription conventions that are relevant. And we’ll finish up with just a short description of some of the tools that I use for transcriptions.

So social interaction is at the heart of everyday life, it’s how we share ideas, express emotions, build relationships, and get our daily jobs done. But beneath the surface of everyday conversation, lies a world of orderliness that often goes unnoticed, even by people who actually study human behaviour quite closely. So it’s this orderliness that the tools of conversation analysis allow us to uncover.

So conversation analysis is like the key that unlocks this hidden orderliness, revealing how we structure our conversations, present ourselves in particular ways, and make sense of the others that we’re interacting with. So over the next half hour or so, we’re going to explore the role of the transcription conventions specifically, and see how they are actually crucial in achieving the goals of conversation analysis.

So we all know that it’s not what you say, but it’s the way that you say it that counts, and CA researchers have shown over the past six decades that how something is said can completely change the action that you’re doing in talk. And this is why Gail Jefferson, pictured here, developed these conventions to capture the sounds that are interactionally relevant. And these Jeffersonian transcriptions, as they’re now called, have provided a set of agreed upon conventions that we all adopt. So how do we even get started?

So one problem is that talk goes by really fast and it’s very hard for us as researchers to recreate post-hoc the details of how something was delivered. So it follows that recreating actions that are being done in talk can be almost impossible, and yet, this is what most research actually relies on, our memories of what happened, whether it be in a self-report questionnaire or philosophical theorising about language and the mind and our place in the world and so on.

The solution is to work with some actual data and transcribe it using the Jeffersonian transcriptions. And that will give us a set of conventions that will capture these aspects of interaction, that are often hard to notice and have been shown to be interactionally relevant. So let me just give you a short example.

So the following clip shows Brett Kavanaugh, at the time a Supreme Court nominee, being questioned by Rachel Mitchell, Head of the Special Victims Division. And Kavanaugh was providing testimony regarding allegations that he had assaulted Christine Blazey-Ford while the two were teenagers. Mitchell was hired by the Republicans to question Kavanaugh, and so this was a little piece that I did with Sun Hee Park, so I’m giving you the reference here where you can read more about it. But we compared the basic transcript that appeared on various journalistic sites at the time, with the Jeffersonian transcript. So let’s take a look at the clip.

Mitchell: “Did you ever wake up with your clothes in a different condition or fewer clothes on than you remembered when you went to sleep or passed out?”

Kavanaugh: “No, no.”

The Jefferson transcript is not only double the length, but also appears a lot more complicated. It includes numbered lines and various symbols that give us specific information about the timing and delivery of everything, and that’s without adding in other relevant details like facial expressions, gaze, and so on, which in this tutorial, we don’t have time to go into, but would typically be added into the transcript.

For now we’re just focusing on the sound, but why, why, why do we go to all this trouble? There are many ways of answering this question, but let’s just note one or two simple observations. We have a question here and conversation analysts have spent a lot of time looking at questions, and one thing that they’ve shown is that negative polarity words such as ‘ever’, “Did you ever wake up,” with this little bit of emphasis on ‘ever’, within this kind of a yes/no question, tilt the recipient towards a ’no’ response. You can read more about the references behind this in the paper that we’ve done.

Secondly, researchers have shown that silences of more than 0.2 of a second between one initiating turn and one responding turn can suggest some kind of trouble. Here we have a second and a half. We can then consider the response itself, which after the preferred, “No,” response, includes what sounds like the start of a, “Yes,” answer.

Mitchell: “No, y…no.”

And more silence before the final no, followed by more delay and then some breathy laughter.

So when we rely solely on this kind of orthographic transcript, which just gives us the words, you can see how there might be a bit of a misleading impression given, where, for example, if we look at Kavanaugh’s response, using only the basic transcript, it seems like he handled the question perfectly effortlessly, but as we’ve seen, this overlooks the interactional specifics, the fact that Kavanah’s response was delayed, delivered in a hesitant, halting manner, and obviously included some equivocation.

So now we’re going to delve into the transcription conventions. And in a recent book with Galina Bolden, we divided these conventions into five categories. And the first of these is the transcript layout, so how we set out a transcript. The second focuses on the timing and sequencing of speech. The third focuses in on aspects of speech delivery, so pitch changes, volume changes, tempo, different types of emphasis and so on. The fourth relates to how we represent things like transcriber comments on certain hearings, that kind of thing. And related to this is how we go about representing things like laughter, sighing and crying, that are not verbal.

And of course, over the decades, as more and more people have collected video instead of just audio, it’s become really important to represent the visual elements, so gestures and facial expressions and so on, and Lorenza Mondada has developed a really neat set of conventions for doing this, which we don’t have time to go into, but these are the multimodal conventions that she’s set out to accompany the Jeffersonian transcripts. And you can read about her fantastic work in the link that I’m giving you here.

So it’s important to include a header that allows you and others to locate the specific clip that’s being referred to, so here we have the data source, CSpan, which is a publicly available set of recordings of events like this, identifying features, so ‘Kavanaugh facing forward’, and the timing of the clip, which is kind of crucial for locating it because it was a very, very long recording.

You may have also noticed that the lines are numbered, and speakers are clearly identified, with a three letter name abbreviation across each turn, and things are carefully placed with non-proportional font to make sure everything lines up properly.

And this allows us to mark the beginning and ending of the onset and offset of simultaneous talk or overlapping talk, as we call it, and we’ll cover more about that as we go through.

And finally, talk is represented exactly as it was produced, not as the transcriber thinks it ought to have been produced, so for example, Kavanaugh’s answer to the question. So as you can see, the layout of transcripts is not arbitrary and it helps us understand various interactional dynamics, which brings us now to the second of our categories, which is temporal and sequential relationships, and the first element of these is overlapping talk, where two or more people end up speaking at the same time. So we show this like in lines eight and nine with square brackets, and the start of the overlapping talk is shown by the left square bracket, and the end of it by the right square bracket. And this allows us to see some really neat things, like how speakers can predict in a split nanosecond of timing when the other person might stop, and aim to start just at that right moment. And when they misjudge this, they might have various ways of dealing with that, such as, you know, shouting to make themselves heard if they feel like what they’re saying is more important, or dropping out of overlap quite quickly. So careful transcription of overlapping talk is quite revealing of what speakers are actually getting up to, and the little tricks that they can use to get around the kind of normative rules of turn taking.

I’m going to show you the clip with this bit of transcript, just so that you can hear how it sounds.

Susan Page: “Kamala Harris, Senator Harris, I mean, I’m sorry.”

Kamala Harris: “It’s fine, I’m Kamala.”

Susan Page: “No, no, you’re Senator Harris to me. Um…”

So usefully, this clip also has a lot of latching in it. So latching refers to a kind of normative sense that we’ve discovered in CA, that between one turn, unit of turn and another, or between one turn and another, there will be a very slight amount of silence, usually between 0.1 and 0.2 of a second. And this in the past has been referred to the kind of beat of silence between units of talk. And this can be between one person and another in the interaction, or it can be within the same speaker’s turn, where they’re rushing through to get extra units into their talk. So latching means that this very small amount of silence is missing, and we represent this by equals signs. So let’s have a look again in the following example. So here we can see, like on line eight, that there’s, “Senator Harris, I mean, I’m sorry,” two units of talk here that are latched together. And Harris says, “It’s fine, I’m Kamala,” again, no hearable silence in there. And then between the Kamala Harris on line nine and Susan Page on line 10, again, we have equals signs, so indicating that there’s no discernible silence between one turn and another. Let’s have a listen to how this sounds.

Kamala Harris: “It’s fine, I’m Kamala.”

Susan Page: “No, no, you’re Senator Harris to me. Um…”

The final broad category within temporal and sequential relationships is silence, and as you’ve probably seen, we time these and put them to the nearest tenth of a second and put that number in parenthesis within the transcript. But we make a distinction between two different kinds of silence: one is called a pause, and one is called a gap. So what do we mean by this?

So if we look at this clip that we played earlier, we can see this difference emerging. So line one, Mitchell hasn’t completed a turn when she pauses. So, “Did you ever wake up with…” – no discernible question here, so this is 0.7 of a second of a pause. She continues, and then having finished her question on line three, there is a silence. Now, this silence is placed on its own line because it’s interactionally significant that it took a second and a half for Kavanaugh to respond. And then within his own turn, there is a pause on line five, and another gap on line six, having finished his response. So you can see the difference here playing out between pauses and gaps, and these are interactionally significant in various ways.

So in this next section, I’m going to delve into various aspects of speech delivery, and so this will be about how we represent at the end of turns, how to represent sounds that occur at the end of turns for units of talk, and also how we represent words that are emphasised, stretched, cut off, reduced in volume, reduced in pitch, speeded up, slowed down, or having various kind of aspiration/breathiness within them.

So turn endings are quite an important part of how we represent speech delivery and we do this with punctuation. So we don’t use the punctuation grammatically, but rather we use it to indicate intonation. So I’m going to show you some examples of this, but I’ll just explain the different elements here. So the period or full stop indicates falling intonation, so not necessarily at the end of a whole turn, but it could be. So for example. “It’s raining,” would have a, if it comes down at the end, it would have a period. Similarly, a question mark indicates rising intonation, not necessarily a question. So, “It’s raining,” would be a question mark. And you can see what we mean by the fact that how you say something, there are two completely different actions here, so, “It’s raining,” might be a complaint or informing, whereas, “It’s raining,” is, like, possibly a question.

So a comment indicating continuing intonation means slightly rising, so not, again, not necessarily a question. So it could be something, like, “It’s raining,” but I’m going to go out anyway, so it kind of can sound as if you’re continuing on with whatever it is you’re saying. But you sometimes see questions with not much uprising intonation.

So you might also see inverted question mark, and this is somewhere in between a comma, but not quite up towards a question mark. And an underscore at the end of a turn means that it’s delivered with the same intonation all the way through, so it’s, like, “It’s raining, it’s raining,” you can sound bored or annoyed or something with that kind of intonation.

So I’m going to show you a clip with various turn endings here, and I’ll just read through it and try and replicate it, just so that you can tune in when you hear the actual extract. So it starts with Brianna saying, “Okay,” we’ve got that final (inaudible 00:21:38) there “Remember Val’s first boyfriend?” Okay, we’ve got this upside down question mark here, which suggests not quite, not quite, “Boyfriend,” it’s more, “Boyfriend,” and Lydia says, “The cute one with the brown hair?” Hair, sorry, I came up too high there. “The cute one with the brown hair.” And then Ashley says, “The blonde guy?” And Brianna says, “Yeah, Jeff,” okay, two periods there. Okay, I think I did that justice, just about. But let’s hear it as it actually sounds.

Brianna: “Okay, remember Val’s first boyfriend?”

Lydia: “The cute one with the brown hair?”

Ashley: “The blonde guy?”

Brianna: “Yeah, Jeff.”

So when we’re representing a bit of talk that has continuing intonation, as in it doesn’t go up or down, we use an underscore, as we can see at the end of the word Harris here. So these are quite hard to find and it doesn’t really count as being a turn ending, but it will give you a sense of how this hears. So let’s pretend ‘Kamala Harris’ is the end of a turn and listen to how this sounds.

Susan Page: “Kamala Harris. Senator Harris, I mean.”

So that’s given us a sense of turn endings, and I’m sure you noticed there was a lot of other symbols in that transcript, so we’re going to get on with some of those, we’re going to start with emphasis. So now we’re going to have a look at emphasis, and this is shown somewhat intuitively by underlining And obviously the more underlining we’ve got, the more emphasis we’ve got. So if we think about emphasis, it can be like a little bit of volume, sometimes a little bit of pitch as well. So emphasis, where it’s, like, I just put a bit of emphasis on the start of ‘emphasis’, so it’s a bit louder, maybe slightly higher in pitch.

Let’s have a look at this extract. So we’ve got Brianna, she’s got a bit of emphasis on, “Okay,” at the very start, “Okay,” then it comes down with a full stop. And so we’ve got, “Remember. Val’s. First. Boyfriend?” So we’ve got a little bit of emphasis on the colons of, “Val’s. First,” which I’ll talk to you about further along. And we’ve got a bit of emphasis on, “Blonde,” Ashley says, “The blonde guy?” And Brianna says, “Yeah, Jeff.”

Okay, so let’s have a listen to the actual extract.

Brianna: “Okay, remember Val’s first boyfriend?”

Lydia: “The cute one with the brown hair?”

Ashley: “The blonde guy?”

Brianna: “Yeah, Jeff.”

So hopefully you could hear the bits of emphasis there, and this kind of represents the sense in which we’ve got a sliding scale. So one end of the scale we’ve got very quiet and that’s indicated by degree signs, so just about audible, so might be whispering or something sotto voce is two degree signs, something just a bit quieter than the surrounding speech is one degree sign, and then we go up the scale with some underlining for emphasis, and then extreme differences in volume, we can give capital letters, and also underlining if it warrants it. And always with the symbols, we are orienting to the speaker’s natural volume. So for example, if we looked at the Kavanaugh interaction, in a normal interaction, the speakers might be considered to be shouting, but in the context that they’re in, not so much, because that’s the normal delivery for that kind of context. So we’re always going with what the kind of natural flow is, and we’re adding in details that are markedly different.

So these symbols aren’t to be overused with every little thing, so another issue with things like emphasis is that you can get words that have a natural emphasis, like information. So you might have heard me going up a little bit on ‘mation’, information, that’s a normal way of saying the word. So we wouldn’t typically mark it unless it was a bit more pumped up than just ‘information’.

So the default is always, “What’s natural for this speaker?” And in terms of what we can hear, one way of imagining it is to think of it as, like, a robot or something. Which doesn’t sound very natural, but it’s, like, “Okay, remember Val’s first boyfriend?” So what you can hear added into that is things that are remarkable.

So I’m going to give you a couple of examples, one of elevated volume and another of decreased volume, so we can see on lines 28/29/30 and 31 that not only is there overlapping talk from three speakers, all kind of trying to speak at once, but there’s also elevated volume. So let’s have a listen to this.

Brianna: “Because he’s also a rapper.”

Lydia: “Oh, I know, I know.”

Ashley: “I didn’t know.”

Brianna: “Okay.”

And you may not hear the next one, because it’s, again, overlap, but it’s the speaker is saying, “I meant to say,” which is just about hearable over the headphones, in overlap with Kamala Harris saying, “It’s fine, I’m Kamala.” Okay, so let’s just have a listen.

Susan Page: “Sorry, I meant to say…”

Kamala Harris: “It’s fine, I’m Kamala.”

Susan Page: “No, you’re Senator Harris to me. Um…”

So as with volume, and the sliding scale with emphasis and volume, we also have a bit of a sliding scale with pitch. So we conserve the arrows for markedly upward pitch or markedly downward pitch, again, relative to that speaker’s normal way of speaking. So if you’ve got a very high voiced person then you don’t just mark everything with upward arrows, and similarly, fi you have a very deep voiced person. And going down the scale a little bit, we can talk about intonation counters, where it’s not like a big change in pitch, but there’s some interesting stuff going on in and out of the words, so it's, like, “Hello,” so you can hear, you can start high and then gradually go down and slide up again, “Hello.” So it’s not a big change in pitch, but there’s a counter, and people do a lot of interesting things with contours. So we distinguish between a down-to-up word, and an up-to-down word. So when it’s down-to-up, we underline the colon, because generally when you’re doing contours, there’s a little bit of stretch, which we’ll talk about in a minute, and then when it’s up-to-down, you’ve got a little bit of underlining, but no underlining on the colon, so it’s, “Word.”

So that’s the contour difference, so one issue that a lot of people find difficult is this distinguishing between a sharper pitch rise or fall, and a contour. So let’s have a listen to some of these. So this is just to show you a couple of things about the use of arrows, so we’ve got on line 25 a word that doesn’t start particularly upwards in pitch, but the last syllable does, so ‘Al, it’s underline, so, “Also a rapper,” “Also, a rapper.” And then on line 28, we’ve got, “Oh, I,” surrounded by the arrows, so it’s, “Oh, I”, said, “Oh, I,” with the same pitch level, elevated pitch level. And then we’ve got, “Know,” much higher in pitch than what I just did, “Know,” “Oh, I know.” So we’ve got two arrows at the start of, “Know,” showing that it’s a lot higher than just normal elevated pitch. And then we’ve got on line 30, “I know,” a similar kind of thing, and again on line 31.

And another use of example is the Kavanaugh example, where we’ve got two different version of the same word, one with higher pitch. So on line five, we’ve got Kavanaugh saying, “No,” so it’s higher in pitch, “No.” And then he latches on a, “Ye…” and then he just says a normal, “No,” with a bit of emphasis. Okay, so let’s listen to that one.

Bruce Kavanaugh: “No, ye…no.”

And in our next example, we’ve got a lowered pitch, so we’ve got, Susan Page, I think it is, doing a repair on, “Kamala Harris, uh, Senator Harris,” and then she drops her pitch throughout, “I mean, I’m sorry.” So we’ve got the downward arrows surrounding, “I mean,” and, “I’m sorry.” So let’s have listen to this.

Susan Page: “Senator Harris, I mean, I’m sorry.”

Okay, so now we’re going to take a look at the contours, which is a little more subtle than pitch changes. So I’m going to start with an up-to-down contour, that you’ve probably heard a few times now, where on line 10, the speaker goes, “You’re Senator Harris to me,” and so how we would represent that sound is, so if we look at the word, “Me,” we have the vowel underlined, followed by a non-underlined colon. So that shows us that there’s a contour here across the vowel, that goes from up-to-down, “Me.” But also because we’re at turn ending, we’ve got a very slight rise after the ‘me’, and she indeed does continue speaking after this. So it’s, so the whole sound is what you’ll hear is like a slight, it starts on a high and goes down, and then comes up with a slight rise, “Me.” So let’s have a listen to this.

Susan Page: “…Harris to me, um…”

Okay, so now we come to the down-to-up contours, where basically we’re underlining a colon. So let’s have a look at how this pans out. Okay, so we’ve got this speaker going, “Val’s. First. Boyfriend,” so we can hear, “Val’s – First -,” so down-to-up, down-to-up. And then we’ve got a similar one with the next line, with the word ‘hair’, but it’s a kind of a sharp in breath, “Oh, the cute one wit the brown hair?” So it’s kind of going, “Hair?” sliding up, from down-to-up. Okay, let’s have a listen to these.

Brianna: “Okay, remember Val’s first boyfriend?”

Lydia: “The cute one with the brown hair?”

Ashley: “The blonde guy?”

Brianna: “Yeah, Jeff.”

Alright, so that’s pitch and contours, and now we come to another element of speech delivery, where we’re looking at tempo. And you may have noticed in the immediately prior example, these greater and less than symbols. And so I’m going to explain these to you, if you haven’t figured it out already.

Okay, so here’s the bit of slowed down phrase here in line six, “Val’s first,” is said much slower than the surrounding talk. So as usual with making your judgments about this, it’s, like, “Well, how slower a speaker are they normally?” and if their speech suddenly becomes slower, then you’re justified in using the greater than/less than symbols around it. So with the stretching out of the symbols, we’re slowing it down, so, “Val’s. First.” And we could just represent this with stretch sounds, but there’s something about the formation of the word that’s said more slowly and deliberately as well. So it’s a judgement call whether you just use the sound stretch, which I’ll come onto in a minute, or show that it's also a surrounding, a phrase or a sentence or something that is somehow stretched. If it’s just a word on its own with the slow down symbols, then you can just use the colons to show a sound stretch. So we’ll have a look at that in a minute.

Okay, so here’s the slowed down and of course, you probably notice the speeding up one. So here’s the speaker saying, “The cute one with the brown,” so she says, “The cute one with the brown…” And from other speech that she does in this video, we know that she doesn’t normally speak quite so quickly. So there’s interesting things that speakers can do with these speeding up/slowing down things. They can sound excited, if they’re speeding up, or they can sound very deliberative if they’re slowing down. And there’s a bunch of kind of important interactional reasons why we might do this. So again, it’s something that conversation analysts would pay attention to.

So because this is more of a ‘how to’ kind of video, I don’t really have the space to tell you all about the why of the transcribing, but let’s have a listen to both of these examples of slowing down and speed up.

Brianna: “Okay, remember Val’s first boyfriend?”

Lydia: “The cute one with the brown hair?”

Ashley: “The blonde guy?”

Brianna: “Yeah, Jeff.”

Okay, so now we’re continuing with speech delivery and looking at some examples of sound stretched through a word, and the cut-off sound. Okay, so let’s have a look at some of these. Okay, so here we can see in the Kavanaugh video, that we’ve got on line two, the word ‘on’. So you’re making ga judgement here usually about how much stretch to give something. So can you hear a stretch on the vowel sound and the consonant sound, or is it just the vowel sound? And also, what is the speaker’s normal rate? What is the normal rate to say the word ‘on’ and what is the speaker’s normal rate of speaking? So if you’ve got a really slow speaker, you’re not going to do a transcript that’s got loads and loads of sound stretches all the way through. That’s just their natural rate of speaking, so you can put a note maybe on your transcript that the speaker is a very slow speaker. So what you’ve got here is a markedly different way of saying ‘on’ than this speaker would normally have. And you can see the sound stretch here is, she’s maybe searching for words, so she’s doing, “Or fewer clothes on,” ‘on’, there appears to be, maybe this is a transcription error, but there appears to be no contour through the word, it’s just, “On,” “Or fewer clothes on.” So she’s holding the turn here, with an in breath, “Than you remembered when you went to sleep or passed out.” So she’s showing by the stretch that she’s mid-turn, that she needs to take a breath, and that she’s going to continue and perhaps that she’s maybe doing a bit of a word-search as well. So these are the kinds of things that you will see when you have a stretch through a word.

Okay, so let’s have a listen.

Brett Kavanaugh: “No.”

Mitchell: “Did you ever wake up with your clothes in a different condition or fewer clothes on than you remembered when you went to sleep or passed out?”

Kavanaugh: “No….ye, no.”

Okay, so the cut-off sound is where we can hear a speaker kind of stopping themselves, so it’s not just like the word ‘with’ in line one, “Did you ever wake up with…” silence, “Your clothes in a different condition?” It’s more like line five, which we’ve talked about before, which is like cutting off, cutting yourself off when you’re about to say something maybe that you might regret, who knows. But line five is, “No….ye, no.” “No….ye, no.” Except there was a second pause as well in there. So you can track it for, like, does it sound cut off, and there’s a bunch of phonetic work as well that talks about these sort of abrupt stops and…I think they call them glottal stops. But for us it’s, like, “What is the interactional relevance of that here?” and that’s the thing that we would be focusing in on.

Okay, so I’m going to use this clip another couple of times, so let’s just take a look now at the transcriber comment issue. Okay, so if you want to distinguish as the transcriber something that you’re doing as a description, versus something that one of your participants said, then you put that comment in double parentheses. So you noticed probably in the Kavanaugh transcript that I said, I had in double brackets, “((Breathy laughter)),” to show that there was a bunch of Hs there, “Hh-hh-hh,” that didn’t, you know, it’s not voice laughter, “He he he,” it was more, “Hh-hh-hh.” I hope that picks up on the recording. But, so you can disambiguate this with a description.

So with guesses, or, “It sounds a bit like this,” you can just enclose it in single parentheses, so in one of the transcripts, we have the speaker in overlap, saying what sounds like, “I meant to say,” but I wouldn’t bet my life on it, so I put it in single parentheses. And you can also do, like, if it sounds like, “I meant to say,” or it also sounds a bit like, “I should have said,” or something, you can put, “Try one/try two” in single parentheses as well, just to show that it could have been one of these things and you’re not sure.

Okay, so that’s transcribers comments. And there’s a couple of things that I want to say about void quality, so I’ll get onto that.

So here’s some other, maybe slightly less common elements of voice quality that we can notice. So the first of these is smiley voice, where we can hear through the sound – and this is often useful in telephone calls or calls where people aren’t co-present. Smiley voice is where you say something without really laughing, but you’re smiling, so there’s a certain change in the quality of your delivery that we call smiley voice. So I think we’ve seen an example of this, and I’ll just play it to you in a minute.

We also have creaky delivery. I think these days it might be called vocal fry in everyday context, but we also hear it in pain cries and in upset, maybe even post-laughter when you’re trying to recover yourself, or in trail-offs at the end of a sentence. So sounding creaky. Tremulous delivery is the thing that is typically related to some kind of upset, so when your voice goes a bit tremulous, we can show that with the tilde sign.

Laughter can have various voiced vowels in it, as we’ve seen with Kamala Harris laughing was mostly, “He he he,” but we can have other vowels in there as well. And laughter can also be interpolated into words, so it can be a kind of a plosive sound, where we put Hs in parentheses, so, like, “Lau(h)ght(huh)er(huh),” versus a more breathy sound, where it’s more a kind of, “Lhaughtehr,” so a more gentle, aspiration thing going on.

And then at various points in the clips I’ve shown you, you might have noticed the Hs, some of which are preceded by a period, and that marks an in breath, like, “.hhh,” and without any period sign, we’ve got just Hs, and that’s usually just a, “hhh,” like a sigh or something like that. Okay, so I have a couple of examples of these that I’m just going to play for you. Okay, so we have a couple of examples here, one of smiley voice on line 10, where the speaker says, “You’re Senator Harris to me,” and at the same time, Harris is laughing. And so let’s have a listen to this.

Susan Page: “Kamala Harris, Senator Harris, I mean, I’m sorry.”

Kamala Harris: “That’s fine, I’m Kamala.”

Susan Page: “No, no, you’re Senator Harris to me. Um…”

Okay, and now here we have a creaky turn ending on line three, where the speaker says, “When you went to sleep or passed out?” I think there might be a teeny bit of creak on passed as well, but listen out for it, it’s more pronounced on ‘out’.

Brett Kavanaugh: “No.”

Mitchell: “Did you ever wake up with your clothes in a different condition or fewer clothes on than you remembered when you went to sleep or passed out?”

Brett Kavanaugh: “No….ye, no.”

Also I noted earlier with the very sharp in breath on line seven from this little fragment, where the speaker is going, “Hhh.” So we can put in a little exclamation mark there as well, just to show that it’s kind of exclamatory, it’s not just a big loud in breath, it’s more of a kind of a, “Hhh! The cute one with the brown hair?” So it’s a kind of excited sounding, I guess you could say.

Okay, so that’s just some examples of different elements of voice quality. And I’m going to wrap things up with some concluding comments. But before I wrap up, I just want to say something about the tools that I’ve been using for transcribing. So early on, I used Audacity, which has been free open source software for a while, which gives you a visual kind of picture of your soundwave, so it’s really useful if, even if you have video, it’s a good idea to start by making a Jeffersonian transcript, because then you can add in the relevant and line up the relevant visual details to add in.

So this is the Audacity wave showing, you know, you can zoom in to see, like, just time your pauses and the gaps, or you can listen over little fragments and see when overlapping talk starts, when it finishes, you can look up things like the pulses of laughter and see how many of those you’ve got. So it’s a useful kind of visual guide for helping you transcribe. And it also has the volume represented as well, so the larger the size of the wave, the more the volume, so it’s kind of reassuring. And then the Adobe Audition and Elan are obviously, if you can get them through a university, that’s great. So that’s, they’re good because they co-ordinate the visual elements with the audio elements.

Okay, so missing from this tutorial has been the kind of consideration of the range of different ways that emotion comes into the transcript, and the specific types of delivery that relate to different kinds of emotion. So also missing from this tutorial is a second language transcription, which is increasingly important as more and more people around the world start doing conversation analysis and working with these kinds of interactional transcripts. Also, as I mentioned earlier, the multimodal transcription and I’ve given you another link to that to follow up.

You can read more about all of these things in our book that I mentioned earlier, the book that I did with Galina Bolden, where we have a chapter on emotion, on second language transcription and the different types of multimodal transcription. And increasingly, the Mondada transcription system that I told you about is the one that has become standard, so I think that’s the one to go for if you do go for a multimodal transcript.

So hopefully I’ve managed to convey the fascination that I certainly feel when I’m transcribing, and the way that it gives you a kind of fascinating glimpse into the intricacies of social interaction and how people get things done in talk. And this provides us with a standardised framework for representing interaction and sharing our work through things like data sessions or journal papers and so on. And the idea here is that we can look at the transcript that has followed these conventions, and in the absence of having the actual data, we can get a really good idea of what the evidence is for the claims being made.

Over 50 years of working with Jeffersonian transcripts has facilitated findings about the underlying structures and principles that govern conversation, and reveal the intricacies that bind us as communicative beings. These findings, which we’ve only touched on, are things like turn taking, sequence organisation, preference organisation, repair, affiliation, action formation. I mean, I haven’t really mentioned these but many of the transcription conventions that I’ve outlined here have been instrumental in showing the relevance of these things. And these discoveries stand as a testament, challenging and reshaping traditional notions of what language and social behaviour are, and serving as a launching pad for the continued evolution of a comprehensive science of social interaction.

National Centre for Research Methods (NCRM)
Social Sciences
Murray Building (Bldg 58)
University of Southampton
Southampton SO17 1BJ
United Kingdom

**Web** www.ncrm.ac.uk
**Email** info@ncrm.ac.uk
**Tel** +44 23 8059 4539
**Twitter** @NCRMUK