Dr Ben Molineaux

Mapping Mapudungun: documenting the words and sounds of an indigenous American language

Isabella Melking: This is Forward Thinking. I'm Isabella Melking.
How can we understand how language changes over time? Historical linguists aim to map this language change though time by looking at both the structure and sounds of the language as well as its social environment. Minority non-European languages, like indigenous American ones, are barely represented in research on language change, and limits our understanding of language change as a whole.
Linguistics researcher Dr Ben Molineaux is creating our first digitally based account of change in Mapudungun, the endangered ancestral language of the Mapuche people of Chile and Argentina.
Hi Ben, welcome to the Forward Thinking podcast!

Ben Molineaux: Hello, thank you for having me, I'm glad to be here.

Isabella: So, I was wondering, what drew you into your specific research area of linguistics, can you tell me a little bit about that?

Ben: OK, so I do what we call historical phonology, but before I say a lot about what that is, I'll tell you a little bit about how I got into this. So before going into linguistics I was a high school English as a second language teacher, I taught first in Chile, which is my home country, and then in New York City. And then in both places I taught mostly Spanish speakers - their first language was Spanish speaker and I was teaching them how to speak English, and because Spanish has a slightly more straightforward relationship between spellings and sounds, a lot of my students struggled with learning how to spell English. One of the things that really surprised me is that specifically the students that I worked with in New York City would very often not spell the final Rs in words, so in words like "mister", they would spell it MISTA. And this wasn't sort of trying to speak cool, or to try to just be very colloquial, what they were trying to do was actually to represent the way that they were actually learning how to speak the language, so their proper pronunciation of the language. Now that was surprising to me because my dialect tends to pronounce final Rs, so I say "mister", but, er, they weren't learning my dialect, they were learning their - the local New York dialect, which tends not to pronounce final Rs, and indeed they would call me "Mista". Now that seems like a relatively simple mistake to make, but what it does is it tells us a little bit about the dialects that these students were acquiring. What we had was that they were trying to spell using the spelling system of Spanish, and trying to map that onto their dialect of English, and indeed their dialect didn't match up with standard spelling of English. From the students' perspective, they just had to arbitrarily learn that the sound R at the end of a syllable was written using the letter R, which is the same that is used at the beginning of a word like "room", which is something that native English speakers who write are used to doing, and they don't even think about it that R can be used in these two different ways, but for a Spanish speaker, where "rrr" and "R" are really very different sounds, this was actually really puzzling. And for my students, being teenagers, they would get into really sort of heated debates with me about why we would want to spell a language in that way, in a way that isn't so intuitive, and that doesn't correspond to the actual pronunciation of the language, and well, I'd say well, it does correspond to the pronunciation of some dialects, indeed to my own, where I do pronounce "mister" as "mister" with a final R. So time and again I got into discussions with my students about why English had this kind of bizarre way of spelling its - its sounds, and that - the fact that very often these spellings didn't correspond to the actual form of the language, and so very often the answer to why that is the way it is had to be that at a certain

stage in the development of the language, there was a correspondence between the spelling and the sound, but the spelling got stuck, and the sound moved on, and hence there wasn't a correspondence - so there was language change. And indeed trying to explain this, and often failing to explain this to my students, meant that that I wanted to go back to graduate school, and study linguistics, and to study in particular language change. And that's what I did, I went and I started studying the history of English, and did my doctorate, and then I got interested in the study of other languages, and in particular in the history of Mapudungun, which is the language I'm going to be talking to a little bit about today, and looking at that language's history, so that's how I got into it.

Isabella: Well that's really fascinating, that - a debate with your English pupils kind of led you to want to find out more.

Ben: Yeah, so you think that, as a native speaker, you understand how these things work, but as soon as you start looking a little bit more deeper, you don't really understand why - why is "one" spelled ONE and "won" is also spelled WON. So these things all have their own histories, and it's really interesting to look into it, and the more we look into it we can see that spelling is so distant from the actual sound system that it tries to represent - at least for English today.

Isabella: And so - why do you think it's important to study historical language change in minority languages?

Ben: So most of the research that has been done on language change tends to be on sort of more well-known languages, and by well-known I mean languages that - whose speakers tend to have sort of a better standing in the global power relationships of the world. And minority, marginalised languages tend also to belong to historically disenfranchised peoples, who also have less political capital on the global stage. So very often, when these later languages, the minority languages are studied, they're seen as kind of exceptions, and seen as something exotic, while the normal way languages work, or the better-known, well-established languages. And that obviously there - there's a bias there. So on the one hand looking at the history of those languages is important for the field of historical linguistics because it broadens our sample of languages that we're actually studying, and it allows us to observe some changes that simply aren't instantiated in better-studied languages, and hence it gives us a broader view of the possibilities of the human language faculty. In that sense it expands our understanding of our own human heritage as language being a major part of our human heritage. On the other hand, looking into minority languages also helps us understand the present-day varieties of those languages. So I mentioned that for example in English, we have varieties that pronounce syllable-final R and others that don't, and it's useful to know that difference between those dialects that do and those dialects that don't comes from a relatively common process of loss of R, and this happens across languages in the world. And it's important to see that lack of R is not a sign of some kind of degraded or degenerate form of the language, so as linguists we don't see change as some kind of loss, or frown and sort of shake our heads at language change, as very often native speakers tend to do and say 'oh, people don't do it the way they used to'. It's a natural process, and indeed we see that a process like R loss shows up in traditionally much more prestigious forms of the language, such as standard Southern British English doesn't pronounce final Rs, right, but it also shows up in languages that are less prestigious, such as African American Vernacular English, or even Cockney. Notice that when I say that a language is more or less prestigious, that's a - a social factor, it's not a linguistic factor, so linguistics doesn't say that one language should be more, or one variety of as language should be more or less prestigious, it's actually a judgement usually made by the speakers and because we have a certain prejudice about a group of people, then we tend to associate their language, and to make

judgements about their language in that way. So looking at the language change in minority languages also helps us assess the different varieties of those minority languages and say 'oh well, actually these are just different varieties of the same language', and we avoid the kind of stigmatisation that can happen with looking at different varieties, and their social distribution as well. So finally another reason to study the historical linguistics of a minority language is the fact that this allows us to have another window into the history of those people. So indeed by studying the historical linguistics of languages of Europe, we've come to understand that there are a number of sort of strong links between Europe and the Indian subcontinent, so there is a common cultural ties and linguistic ties, indeed we speak varieties of the same language family, the Indo-European language family. And for minority languages, a lot of that work to discover what the links are between one language and another hasn't been done enough, and that's important work in the sense that it tells us a lot about how these minority cultures would have interacted with each other, before the historical record, in many cases. So an example from Mapudungun is that it acquired the words for hundred and the word for thousand from Quechua, which is the language of the Inca Empire. So you can imagine sort of what the cultural conditions in which that kind of transfer would have happened, and indeed you would have had different notions of mathematics, different notions of accounting in both cultures, which the - the Mapuche would have drawn new things from the Inca culture at that time. And indeed we get other kinds of transfer, not just words, we also get transfer of pronunciation, so at some point, for example, in the history of English, we became very strongly influenced by French, following the Norman Conquest. And indeed the fact that our pronunciation of the English language changed in contact with French, actually tells us a lot about the intimacy of those two languages at a certain point, the fact that there was a lot of bilingualism, the fact that there was a lot more interaction between two cultures as well. So having information about the history and the development of a language actually tells us a lot about the history of a people and the heritage.

Isabella: Thank you. So it's about increasing representation and thereby discovering the links, and that kind of improves understanding of language change as a whole.

Ben: Yeah, so you're looking at sort of how languages in general work, and then what the history of a particular language is at a more local level, at the level of that particular language, and what that tells us about the culture that surrounds it, and also about the social judgements that are made about a particular way of speaking a language.

Isabella: And what made you choose to focus on the endangered indigenous language of Mapudungun in particular?

Ben: I grew up in Chile, and Mapudungun is spoken in Chile and in Argentina. It's spoken by the Mapuche people, who are the main ethnic minority in that region, and in Chile there's about one and a half million people, or about 10% of the population that identify as ethnically Mapuche, and there is about a quarter of a million speakers of Mapudungun. I think that's a rather generous estimate, and indeed they are losing speakers left and right, and transmission is rather poor at this stage to new generations, which is something that of course we want to try and revert. But growing up in Chile, in the 80s and 90s, it's surprising how invisible that 10% of the population actually was for the broader part of the society, especially for young, urban, middle-class, non-Mapuche people such as myself at the time. And indeed there is a lot of prejudice against indigenous people and culture, especially at that time there was a lot more. I think there's a bit more of a resurgence now, and a little bit more pride in indigenous issues right now, in Latin America in general and Chile in particular. But at that time, outside of the rural indigenous communities, Mapuche identity and

culture were really sort of downplayed at that time. So - many years later, after I left Chile and during my graduate studies, I - when I was focusing on historical linguistics and trying to learn about how languages change, I'd started wondering about how a lot of things that I'd learned about English and its history, how that would work in a language like Mapudungun, and I picked up a few grammars, and I realised that although there is about a 400-year written record for this language, there had been very little work done on the history of the language. And this is just another way in which the Mapuche and their language are made invisible, also by the academic world, so I felt that it was the kind of thing that I could contribute, in a small way, to making these people and their language a bit less invisible. So I would up writing my dissertation on language change, and though I started working mostly on Old and Middle English, I ended up doing most of my work on Mapudungun, and ended up going to spend a few weeks in Mapuche communities in Chile, learning from the elders there, and also following my dissertation I've continued with my interest in Mapudungun, and now I'm - I've been given a 3-year grant to work on developing resources for the study of language change in this minority language, that is amazing and has very interesting features, but it's also been made kind of invisible because of the power dynamics of the region and of the world.

Isabella: And how do you actually go about mapping this language change?

Ben: So - when I say mapping language change, what I mean is actually tracing the paths of the language change, so one of the first things to say about language change is that it's a natural process. That's something that we have to assume is a reality for all living language, and insofar as a language lives, it changes. And that's not a bad or a good thing - they change insofar as they're being used, and appropriated by a language. So it would be worrisome actually if a language didn't change - that's actually a sign that the language is actually dying. But how much we can know about a language and how it changes, and what its history is, depends on the kind of evidence that we have for those languages. So in some cases, like in European languages, we have quite a bit of data for those languages, and it's easier to trace the roots of different languages, and one of the things that we have, for a lot of European languages for example, is several different dialects of the same language being spoken today. And that means that we can compare those dialects, and try to trace them back to a common ancestor. An example for that, between two dialects for example, is I speak a dialect where WH and W are pronounced exactly the same, so I say "which" and "witch". However in Scotland, you will find that people will say "hwhich" versus "witch". And that distinction, what a historical linguist will do is they'll go and they'll say OK, so what makes more sense? Is it the case that the "hw" sound has split off from the "w" sound at some point, or is it the other way around, that we have the "w" and the "hw" sound being the original, and then they collapse into one category, which is "w", such as in my dialect. And indeed, it's that second option that shows up as the most plausible, especially because when we look back, why would people write them differently at a certain stage? It's probably because in the early modern period, which is when the spelling of English became more stabilised, we would have had people pronouncing "hw" and "w" as different, and then again the spelling system got stuck there and some of the dialects went on and merged those two sounds as one, as the sound "w". So that's one way of tracing sort of the history of a language. Beyond comparing just dialects we can also compare full languages, and that can be done, for example in European languages we know that some languages, if we compare them, they have certain features in common, so for example for English, German, Norwegian the "f" sound is used for words like "father" and "fish", right, but in Spanish, Romanian and French you get the "p" sound, so think of the word "padre" and "pescado", right, so that tells us a little bit about how closely related English, German and Norwegian are, as opposed to French, Romanian and Spanish. But the fact that we can even compare those languages also tells us that these must be related at a

deeper level as well. We can't compare, for example, Zulu or Chinese in the same way as we compare French and German. So they have a common core, and indeed most of the modern - for most of the modern period we've known that all those languages belong to the same broader Indo-European language family. And we can identify the subgroups of that family by comparing the changes that they have undergone, and tracing them back to a common ancestor. Unfortunately, when we're dealing with languages that don't have known relatives, like the case of Mapudungun, these are languages that we call linguistic isolates, we can't use this comparative method. So one of the - the final tools that is left to us as linguists, is going through the written record and looking at that carefully, so crucially when there is a written record of a language, and particularly when there's an alphabetic writing system in place, we can compare writing from different periods and try to infer what changes must have happened to get from one stage to another, and between one stage and the next we can then reconstruct what the earlier stages would have looked like. Those are some of the ways in which linguists try to reconstruct the past.

Isabella: OK. And is that what you call a set of texts, so the corpora?

Ben: So yes, a corpus is a collection of items that can be studied, and in linguistics a corpus is a set of texts that we can explore to find linguistic information, as used by actual speakers or writers of the language. So the important thing about linguistic corpora is that they tend to have additional information besides the text itself, and today also these tend to be in digital format. And the additional information that we tend to have are usually what we call tags, so we add different information about, for example the parts of speech of each one of the words that are contained in these texts, or the lemmas, so that is the basic form of the word as you would find it in a dictionary. So say you were looking at a corpus of non-standard writing, and you were looking for the word "mister", and in the collection of texts you might be interested in both forms spelt as MISTER and forms that are spelt MISTA. And those would be subsumed under one lemma, which is "mister". And that way you could compare where and when, so in what text and by what speakers, but what writers, who was using the form "mister" and who was using the form "mista". So that can help sort of investigate the distribution of particular features throughout sort of the area where the language is spoken. Corpus is just a more general word to say a group of things to be investigated.

Isabella: OK. So who will be able to use the Mapudungun corpus, and what information will it contain?

Ben: So the research that I'm currently conducting is to construct a corpus of historical Mapudungun. And that's going to be a collection of texts that were written somewhere between 1606 and 1930, and there are about 300 thousand words, and since Mapudungun was only written following the arrival of Europeans, they didn't have their own writing system before that, the earliest texts were mostly written down by missionaries, then following by military men, explorers, ethnographers, and in the beginning of the 20th century we start getting the first texts that are actually written by Mapuche people. And the result is a collection of digital texts that are gonna cover the first 324 years of Mapudungun for all major language areas where we have information. So in terms of who's going to use them, the idea is that obviously this is going to be something that can be used by linguists who are interested in the kinds of changes that language has gone through, as well as the sort of contemporary forms of the language, it will also be open to people in similar areas of academic research, so historians, anthropologists might be interested in this as well, but, er, I hope that this resource will also be used by native speakers, by teachers of the language, by learners of the language, by advocates of Mapudungun to provide them with a repository of words, and sentences, and texts in general, for use in education, to improve pronunciation, to enhance

word-building strategies, to revitalise words in the dialects that have been lost, and ultimately to just contribute in some way to the process of revitalising this language. That's not the kind of work that I do personally, but I want that to be out there for people, and obviously the resource will be open to anybody who wants to use it online.

Isabella: OK, thank you very much. So you're hoping for a kind of legacy, for people to learn about the culture through the language?

Ben: Absolutely, and it will be something that there are a number of really interesting texts just from the perspective of history, from the perspective of anthropology, but there's also just a lot of linguistic material that can be used in order for, for people to practice their language on, to enhance their current use of the language and to - to basically keep the language in use.

Isabella: Thank you Ben, it's been great having you here -

Ben: Thank you

Isabella: - for the podcast, and we'll add in a link to your website, so our listeners can find out more about your research.

Ben: Yeah, absolutely.

Isabella: Thank you very much.

Ben: Thank you.

Isabella: If you want to know more about the topics discussed in this podcast, follow the links on the Forward Thinking blog, at forwardthinking.ppls.ed.ac.uk. You can also subscribe to our podcast on iTunes for more research news and views from philosophy, psychology and language sciences here at the University of Edinburgh.