**INTERVIEW JONAS**

*Kim*: First of all, thank you very much for participating. I really appreciate it since I recognize that your time is also valuable. Do you maybe have some questions beforehand? Before we begin?

*Jonas*: No, not really.

*Kim*: Just as a disclaimer, of course since I have my questions here, I will be looking down sometimes but I'm still listening. Yeah, just to begin, could you maybe explain how you got into the European Parliament, because I saw in your questionnaire that you first worked at another unit?

*Jonas*: Well not really at another unit. Just for your information, I'm 61 years old. I have a degree in interpretation. So actually, I'm a conference interpreter. I studied Dutch, which was compulsory, English and modern Greek and then I worked for 10 or 11 years as a freelance translator and interpreter. And in 1999, I joined, I passed a competition and in January ‘99, 25 years ago exactly, I joined the Dutch translation unit of the European Parliament. For one year I was accorded to the Budget Committee as an administrator. This had nothing to do with translation. I came back and I stayed here, and I will stay here until my retirement.

*Kim*: That's nice.

*Jonas*: And apart from, yes that's actually so I am one of the senior translators in the Dutch translation unit.

*Kim*: So, you still like it, I suppose, because otherwise ...

*Jonas*: I still like it. I've always liked it, I still like it very much, but for you as a young person I must admit that the translation profession, the translation market has evolved enormously. I date from pre-history times. So, in my time, when I was a student, when I started working, there were no computers, there was no internet, there was nothing. Everything was on paper, encyclopaedia, dictionaries, newspapers, magazines, etc. That's it. So, my first translations really were with pencil then it had to be tagged, that gave the first word docs. And then the first computers, the first word processing systems, CAT-tools, etcetera, etcetera. And now we are in the era of machine translation, artificial intelligence, so it has evolved enormously. And very often I think, well, we've seen everything, but I'm afraid we haven't seen everything yet.

*Kim*: No, probably not.

*Jonas*: But the future will tell us.

*Kim*: Yeah, that's very true. And in those years, all those experiences, did your view on machine translation change? Because you also mentioned that you've worked on machine translation for a couple of years already.

*Jonas*: Yeah, yeah it definitely has changed, because as I told, well, I have always been interested in technology, in technological tools, in tools and such that can help us in doing our work. And so, I started with translators work bench, with Studio. I've been a trainer for those tools for many years. And I've always been interested also in the evolution of machine translation. As you know, at the very, very beginning everybody was laughing like, well, it's ridiculous. But over the years, it has evolved enormously, quality has improved enormously and now we even reached the point that one starts considering, starts asking, as a translation, do I still have a future?. Will machine translation or will AI and/or a combination of both take over my job? But I think it’s not only valid for the translation form, you could also ask the same question for journalist, for instance, or for other types of professions. But my view has always been that it's a very useful tool, but it's one of the many tools we have at our disposal. It's not perfect. I don't know whether one day it will ever be perfect for legislation or for specific document types. But in my view, it's a very useful tool. But of course, you have to look at it and I also consider it as a, from time to time, as a useful source of information, because sometimes machine translation pops up with suggestions, with solutions, that I would not have thought of myself. But then of course, I start realizing, okay, we worked with, so our systems, the eTranslation that we and the European Commission uses has been trained with certain data. The more data you have, and actually, the data we have is everything that has been translated by humans. Not only at the Parliament, but also the Commission admins use it etcetera. So, sometimes you get a very good translation, that I say, wow, I wouldn't have thought about it myself. But then I realize, okay, this has been invented by one of our many colleagues, in one of the institutions.

*Kim*: Oh yeah, I think that's a nice way to look at it.

*Jonas*: And apart from that, sometimes there are, you have to be very, very attentive. Because, well, we are not allowed to use DeepL for security reasons, but of course DeepL for Dutch is very good. And the risk is that when you have a lot of work, a lot of time pressure, your attention is decreasing, because you have to speed up, you have to continue, you have to go on and certainly in the beginning, DeepL, it read, you read something as if it were written by a human. But when you start analysing it, it says just the opposite than that what is meant in the source text etcetera. So, you have to be, I think you have to be more attentive than you translate yourself everything from scratch. But of course, humans translating from scratch can make mistakes as well. So, you have to be very concentrated. And I, when we translate legislation, and there is a reference to a regulation or a directive, it's never translated correctly. I don't understand why, because we work with texts from the European institutions. So, I don't understand why they have not succeeded in coupling, let say, the database of EUR-Lex, that contains all the titles of directives, regulations, etc. That's, so there you can be sure, that you have to check it you have to correct it etc. And what I also mostly do when I translate a document in Studio, I add a specific QA checker that you probably know, that you might also use already, which is very useful for identifying more technical mistakes like dates or references or double spaces or that kind of things. And I also always add a term base, the IATE term base, that is available. Because let's say, the combination of having machine translation on the one hand, and then other integrated tools in Studio like the QA checker and the IATE term base, also saves a lot of time, because we have a term, a validated term that is recognized in IATE, you have it immediately. You can check where the machine translation has taken the right terminology, has used the right terminology or not, so that's the way I work.

*Kim*: Alright, alright. And you also mentioned that you use a couple of tools within the Parliament, with which of those do you apply machine translation?

*Jonas*: Well, let's say the two main translation tools we have are Studio, SDL Studio and CAT4trad, an inhouse-designed tool. I use it for all my texts, I use machine translation for all my texts, for all my documents.

*Kim*: Alright.

*Jonas*: And for some types of documents the results are better for others they're not. But for me, I'm so used to it, because even when you have a text, you have information coming from different translation memories, for the different sources, sometimes you have a match, maybe 65% or 85%, sometimes you don't have a match. So actually, I only use machine translation, it pops up when there are no matches of at least 65% in the translation memories, but if there is nothing, I use machine translation that is offered automatically.

*Kim*: Alright, alright. And you mentioned that the outcomes can vary on the type of texts that you translate. Would you say that there are certain types of texts where you would, where you would recommend using machine translation more than other types of texts?

*Jonas*: Well, let's say for more typical texts for Parliament or for legislation, I would recommend it. For more creative texts, especially our eTranslation is not that good. But even then, even then, it might always come up with some suggestions. Let's say, you can use it as a basis, but it will take you more time to have a good translation than for technical translations or other documents. But I've noticed over the years that quality has improved enormously, yeah.

*Kim*: So, in general, you would say that you are content with the machine translation?

*Jonas*: I'm satisfied.

*Kim*: Satisfied, that's the word I was looking for.

*Jonas*: I'm satisfied, yeah definitely, definitely. But I don't rely, let's say, only on machine translation. For me. It's a very important tool, that is integrated in our own translation tools, that I use in combination with other integrated tools, like the term base, like the translation memories, etc.

*Kim*: Alright.

*Jonas*: So, I don't use it as a stand-alone solution.

*Kim*: No, exactly.

*Jonas*: I don't use it as a stand-alone solution, but when I have something that has to be translated from scratch, so this means that nothing comes from the memories or has been translated completely or partly by someone else, that is contained in one of the memories, I rely on machine translation. And sometimes you don't have to change anything. And sometimes you have to delete it and start from scratch, yeah.

*Kim*: Alright. According to you, what are the main difficulties when translating with machine translation?

*Jonas*: The main difficulties, well I wouldn't call it difficulties, but as I told already, you have to be very, very attentive, because the risk is when it sounds like good or normal Dutch, and that's mainly, I would say, the pro-, not the problem, but when you use DeepL, very often it sounds as if it were written by a native speaker. So, when you don't have a lot of time and you read, okay it sounds good and then with control enter you confirm, etcetera, etcetera. So, you have to be very, very attentive, and you always have to check, let's say, references to directives and regulations, etc.

*Kim*: And then on the other side, what do you think are the main advantages of machine translation?

*Jonas*: The time you gain. Well, not only the time. So, you gain a lot of time. Even, let's say, if you have to post-edit something or to change something, you don't have to type or to translate the sentence yourself completely. So, when you're quite keyboard literate, when you know some shortcuts etcetera, you can very easily change the position of words or characters or delete a word or part of a sentence in one go. So, when you know some shortcuts, you can gain a lot of time. And very often also a very useful source of inspiration I would say.

*Kim*: Alright, alright. Oh yeah, that was another question that I was curious about. Does it also matter from which language you translate from, how good the machine translation output is?

*Jonas*: Yes, because English has the best results, because it's the biggest database. And, let's say, I do Polish, I do Greek, I don't know whether you know that, but when I have to translate something from Greek into Dutch, and you get machine translation, it is first translated from Greek into English, and then from English into Dutch. So, it's like in the interpretation booth where they don't cover all languages directly, they use a so-called relay language. So anyway, when you have to translate something, or you have to process something more than one time or two times or three times, then there's always something lost. So definitely, English-Dutch, for me has the best results, because it's based on the biggest set of data that we have at our disposal and for all, for many other or so-called exotic languages, or less used languages, it's always first a translation into English and then from English into Dutch.

*Kim*: I didn't know that that was the way it works.

*Jonas*: Yeah, that's the way eTranslation works.

*Kim*: Okay. And if there was anything that you could change about the machine translation used here, what would you like it to have?

*Jonas*: Well, as I mentioned already, I don't understand why references to directives and regulations are not taking over literally from, let's say, from CELEX or from EUR-Lex, from the official journal, or whatever, and that it has to be reinvented or retranslated. And what I also notice very often is that, I know how machine translation works, but sometimes you have to repetitive words, repetitive sentences in a big document and well at the beginning it is translated this way, a few pages further down it's translated, the same thing is, or more or less the same thing, is translated using other words. So, it would be nice if machine translation could be more context based.

*Kim*: Yeah, that makes sense, yeah.

*Jonas*: I know, I know it always looks at a few sentences or a few words before that and a few words after that. But it would be nice and maybe with artificial intelligence it would be possible in the future, in the near future, I don't know, that machine translation not only looks at the sentence level or part of the sentence level, but also at a bigger level, at a higher level, let's say at the document level. And when, I think that can be done with AI, when it sees that it's the same term, a specific term comes back five times that it's, if possible, that's also maybe one of the improvements, that it's coupled to IATE databases. That that specifical term is given based on the IATE terminology and that in those five cases you are sure that you get five times the same exact translation. But on the other hand I would say, when you don't use machine translation, and in our context and projects we work in you have your translation memories with input from everywhere, from in house from outside of the Parliament, the Commission, the Council, the committees etcetera and even then you are not 100% sure that all the input you get used the same terminology, so that's one of the main tasks of the human translator, of the human revisor, post-editor, call it the way you want it, to see to it that when you opt for, let's say, term A or term B, I don't care, but at least see to it that within the same document, within the same context, you always opt for either A or B.

*Kim*: Alright, yeah, I can imagine that that would help a lot.

*Jonas*: But that's not only for machine translation, that's also for human translations. You get so many inputs from so many different sources, and it's not always streamlined. That's human of course, it's not always wrong. It also has to do with personal style or preferences or whatever. It's not necessarily wrong, but sometimes it's useful to have more streamlined terminology within one document.

*Kim*: Yeah. And then I have some questions about clear language guidelines. Because in your survey you said that you basically apply them always.

*Jonas*: I try to apply them, yes.

*Kim*: Could you explain a little bit about how you apply them and what the clear language guidelines are for you?

*Jonas*: Well, for me, clear language, first of all, or for me a good translation, is, let's say, a text that reads and sounds as if it were written by a native speaker in a clear, understandable way. So not too many... It's not always easy with the kind of texts we have to translate. But there, for instance when you have a very long, Latin style, half page sentence. Well, I will cut it in pieces, to make it clearer, more understandable and that of course, machine translation doesn't do it. When you have a half page sentence, well you will get a half page sentence translated as well. Then of course you start changing it, and well that's one aspect of clear language to me. Another one is also, although I'm one of the oldest colleagues, I try to use vocabulary that is modern. So, I won’t use, I hate archaic words or archaic style or whatever. Although I also try to use, let's say, the *je*-form instead of the *u*-form in Dutch. So not really the polite official, but it depends on the kind of text of course, so that kind of things. I always try to imagine, but that's very general, clear language specifically for Dutch, language that is comprehensible for both Dutch people and Flemish people.

*Kim*: Yeah.

*Jonas*: Sometimes my Dutch colleagues, I'm Flemish, my Dutch colleagues use words or expressions that even I don't know. So, if they use that, let's say, the average Flemish or Dutch reader will not understand it either.

*Kim*: Yeah.

*Jonas*: So that's also an aspect of clear language specific to Dutch. Then, of course, let's say, something that is comprehensible, let's say, for the average intelligence quotient of readers.

*Kim*: Yeah.

*Jonas*: Of course, sometimes when you have technical legislation, of course, we have to use that vocabulary, but if it is more a press release or something for the bigger, the wider public, then you can be more free and you can apply more, let's say, the clear language rules than in a really technical or legal texts.

*Kim*: Yeah, I can imagine. And since you mentioned differences between Flemish and Dutch, do you also experience some difficulties, I guess, with machine translation when it comes to those types of things?

*Jonas*: Yes, yes. For instance, let's say, I won't be the first one to mention that example. SME's, small and medium enterprises, we say *KMO's*, *kleine en middelgrote bedrijven*, in Holland it is *het MKB*. And as I told already, when you have a text about SME's, depending on the source, or the part of the dataset where the machine translation gets its information you sometimes might get *KMO* and sometimes MKB and whether you choose *KMO* or *MKB*, I don't care, although in Flanders most of the people don't know *MKB*, they don't know what it is. We know, but they don't. But at least within one, within the same text, see to it that you opt either the one or the other.

*Kim*: Yeah, alright.

*Jonas*: But sometimes, it depends of course on the text, *bijvoorbeeld ‘het College van burgemeester en schepenen*’, we say *‘schepenen’*, you say *'wethouders’*. Okay, in Belgium, constitutional law it's called *'schepenen’* in Holland it's called *'wethouder’*. What do you do with it, I don't know.

*Kim*: Yeah, but I think that’s also a general...

*Jonas*: This, in say, it says has nothing to do with machine translation.

*Kim*: Yeah.

*Jonas*: Also, for the human translations you have the same problem.

*Kim*: Yeah, I can imagine. And then I think my last question for now, do you experience more difficulties when, I guess, applying clear language guidelines when using machine translation?

*Jonas*: A bit more, yes. Because very often, it depends on the kind of text you have, but very often machine translation is still quite literate. It follows, let's say, more or less the structure of the original. And sometimes you have to change it.

*Kim*: Like in the longer sentence you were talking about earlier.

*Jonas*: Yeah, yeah. But especially long sentences. Then you make two or three sentences of it. But it's getting better and better. Because at the very, very beginning, well, very often the verb was missing, or parts were missing etcetera. Or let's say the subject was in singular and the verb was in plural, or the other way around. But the last few years, we don't have very often that kind of errors anymore.

*Kim*: That's good. Alright, and do you have any other things you would like to add or questions?

*Jonas*: No, not really.

*Kim*: Then I would like to thank you very much again for the interview.

*Jonas*: You're welcome. I wish you good luck.

*Kim*: Thank you.