A Translational Poetic Practice

On the uncreative potential of statistical computer translation

Samuel Eberenz

Thanks to

Francesca Broggi-Wüthrich and Stefan Scheidegger

Second Version: 31 January 2016

The first Version (17 January 2015) was written for the seminar
Uncreativity. Poetic and Cultural Practices in Literature
Mrs. Dr. Francesca Broggi-Wüthrich
HS 2014, ETH Zürich

Samuel Eberenz
Zürich, Switzerland

Let's imagine, we chase these words through random languages from all populated continents:

We think that we follow the words of the language of random people throughout the country;

We believe that random people throughout the country following language words.

We, language, people across the country continue to believe that.

We, the language, who continue to believe in the country.

We still believe in this country.

This sequence of words was produced using the free online service Google Translate[1] (GT) in the following way: I typed the first sentence and then used GT to translate it through different languages (Persian, Maori, Yoruba, Haitian) and back into English. The output is the second line. This was used again in the same way to produce the third sentence – this time using a different set of intermediate languages – and so forth.

At first glance, the resulting phrases seem to exhibit rather random changes in syntax and meaning as compared to the original. But is it really by chance that *imagine* shifts through *think* to *believe* – and what about the colon in the end of the line which transforms into a semicolon until settling to be a full stop? Why is it that the fifth sentence reads like the opening line of a manifesto? Or that the sentences become shorter and simpler with every round of translational merry-go-round?

The aim of this essay is not to answer these questions nor to examine the technical details of the algorithms and corpus of GT, but rather to argue that there is an uncreative poetic potential in the use of statistical computer translation.

To support this argument, I will first try to describe the functionality of statistical computer translation as a digital prototype of structuralistic practice as described by Roland Barthes[2] half a century ago. Subsequently, I will examine what I call the 'translational poetic practice' of using online translators to create literature. Examining the choice of source texts for this practice, I will also put it into a Situationist context. Integrating these three steps, I will then elaborate the uncreative aspects of this practice, including a brief speculative reflexion from a posthuman perspective.

Implemented structuralism

Contemporary online translation tools combine different approaches like classical dictionaries, the statistical exploitation of texts with translations available online (like classics of literature, but also movie subtitles, protocols of the United Nations and the European Council), as well as suggestions for better translations offered by users[3]. Briefly speaking, the corpus is a mixture of fiction and non-fiction using very different jargons

This means, that any text fed to the machine will be translated based on an accumulated corpus of known relations between expressions in the source language and the target language[4] (and, heavily, on the selection of corpus material). Let's consider these relations between two languages as the objects of a structural

1 https://translate.google.com/ (last accessed 16 January 2015)
2 Barthes, 1966.
3 Bellos, 2011.
4 If there is not enough data available, an intermediary language (e.g. English) is used (cf. http://en.wikipedia.org/wiki/Google_Translate#Translation_methodology)

practice as set out by Roland Barthes:

> “Das Ziel jeder strukturalistischen Tätigkeit [...] besteht darin, ein „Objekt” derart zu rekonstituieren, daß in dieser Rekonstitution zutage tritt, nach welchen Regeln es funktioniert (welches seine „Funktionen” sind).” [5]

According to Barthes, the reconstitution of an object happens not by reasoning but as a ‘practice of mimesis’, a ‘struggle against contingency’. The meaning of the object's structure is taken into account by reconstituting it, not by naming it:

> “Der Künstler, der Analytiker legt den Weg der Bedeutung noch einmal zurück, er braucht ihn nicht zu bezeichnen: seine Funktion, um Hegels Beispiel aufzugreifen, ist eine *Manteia;* gleich dem antiken Seher *sagt* er den Ort der Bedeutung, aber nennt ihn nicht.”[6]

In the case of statistical computer translation, it is not the human artist or analyst, as Barthes suggested, but computer algorithms carrying out this mimesis of structure by purely numerical means.

In light of the structuralistic perspective, this implies that the effects of multiple translations using an online translator (e.g. the lines shown in the beginning of this text) are not random at all.

One example of how statistical computer translation ‘reconstitutes’ the structures behind the texts found in the web, was exposed and subsequently tackled by Londa Schiebinger: In her case study *Machine Translation: Analyzing Gender*[7], she shows a male bias in the GT translation of nouns and pronouns: Since the male form is more often found in the linguistic corpus, it statistically wins over the female form. Even more, Schiebinger also found gender-related stereotypes being reproduced by GT: While the English term ‘a defendant’ is rendered into the male form when translated to German (‘ein Angeklagter’). ‘A nurse’ is rendered into the female form ‘eine Krankenschwester’. As a reaction to her findings, Schiebinger helped GT to improve their algorithms in order to overcome this bias. In this process, one structural forcing is replaced by another one – hopefully a less sexist one. This example of computer translation being biased towards a quantitative norm structurally implied by the corpus illustrates the crucial point: What we get out of statistical computer translation depends both on the object, that is the relationships between different languages as materialized in the linguistic corpus available, and on the algorithms being the structuralistic methodology of the respective translation engine, the tool working on the object.

In short, it is fair to say that the output of this technology is not merely the product of an arbitrary and imperfect dictionary and lost syntax, but embedded in the linguistic haystack of the world wide web. The algorithms don't examine every blade of grass from this digital haystack equally. Rather, an opaque selection of corpus (which might be based on pure availablility) and statistical rules and weightening determine the output.

In the following, I will time and again return to this description of statistical computer translation as a structuralistic practice (or medium, if you wish) while trying to elaborate its uncreative poetic potential.

---

5 Barthes, 1966:
"The aim of every structuralist activity [...] is an" object "to reconstitute such that in this reconstitution comes to light, the rules by which it works (which his" functions ")." *(translated by GT)*

6 ibid.:
"The artist, the analyst specifies the path of meaning back again, he does not need him to call: its function is to take up Hegel's example, is a Manteia; equal to the ancient seers he says the place of importance, but do not call him. " *(translated by GT)*

7 Schiebinger, 2013.

A translational poetic practice

Just looking at the simple piece of multiple translation in the beginning of this essay, it appears that this translational practice does not only induce structurally biased shifts and seemingly arbitrary errors in grammar and syntax: While reconstituting empirically dominant structures, it also creates new semantics – simply by adding up relationships as it statistically searches the most probably adequate translation of an expression. Furthermore, due to the missing understanding of the intention, narratives and context of the original text, this process is inclined to produce discontinuity between the single sentences or chunks of words that are treated as one unit of meaning by the tool and translated in relation to each other.

The reader's mind tries to construct a narrative while being confronted with a somehow dadaistic aesthetic oscillation between familiar figures of speech and the complete loss of structure and semantics. The familiar figures are probably those that occur most frequently in contemporary texts, the unfamiliar ones might be the crippled remnants of common figures of speech in one of the languages passed in the process. The translation is at the same time a process of alienation and creation.

While deciding on the input, the single human gives up control of the wording and phrasing to the web and the algorithms. Taking the text through multiple translation, the differentiation between writer and reader blurs. At the same time, the technology exploits the work of thousands of single human translators and interpreters working for the European Council or translating subtitles of Ben Hur to Hindi.

Tools like GT or Bing Translator[8] are often ridiculed for their tendency to deconstruct grammar and meaning. But it is precisely this ‘failure’, the unexpected shifts and distorted semantics, that constitute their poetic capability. In the essay *Pure Language 2.0: Walter Benjamin’s Theory of Language and Translation Technology*[9], Mathelinda Nabugodi discusses the potential of so called rule-based machine translation to produce a translation that, by being blind for the lure of meaningfulness, is able to translate what is essential, making the syntax the object of translation:

> “To fulfill the task of translation, one must disregard what the foreign text is about, and instead decompose its sentences into their building blocks, words and grammatical relations, and then transfer these relations–these foreign manners of meaning–into one’s own language through a literal rendition of the syntax.” [10]

The same could be said for statistical computer translation – the only difference being the fact that it uses statistical relations between the two languages instead of an arbitrary set of rules:

> “GT deals with translation on the basis not that every sentence is different, but that anything submitted to it has probably been said before. Whatever a language may be in principle, in practice it is used most commonly to say the same things over and over again.” [11]

This means that statistical computer translation essentially doesn't create any new translations[12]. It just mashes up millions of expressions written and translated before. If there is actually a word in the source

8 http://www.bing.com/translator/ (last accessed 12 January 2015)

9 Nabugodi, 2014.

10 ibid.

11 Bellos, 2011.

12 Nor do rule-based translation tools or hybrids of both statistical and rule-based approaches like Bing Translator. Rule-based tools just operate on a much more limited corpus of relations from arbitrary dictionaries and grammatical rules. While rule-based translation tools pioneered in the first years of this century, contemporary services (including GT) look to statistical approaches as being most promising.

text that is unknown – new – to the digital translator, it is simply left untouched by the algorithms. The same happens to misspelled words and inclusions from languages outside the respective pair of languages in translation. The output is thus determined by the linguistic corpus, the alghorithms, as well as the choice of languages and the initial input. They all define which of all the expressions found in the web will be combined in which way. Statistical computer translation defines a new, dynamical, set of rules that is not based on syntax but, none the less, on tradition of language use and translation.

The output of multiple translation can in turn be used as an insipiration, as building blocks for new texts – both in terms of semantic and poetic diction. The task of the human is no better than to select, to rearrange, to seize the aesthetics offered, and to elaborate the tendencies of meaning indicated by the translations – or more precisely, the meaning which the human mind tries to construct while reading a text. This symbiosis results to something, neither the algorithms nor the human could have come up with by themselves.

All of the steps involved from selecting a source text (input) to the rearrangement and framing of the output is what I call a 'translational poetic practice'. For me, this usually involves multiple translations, the translation of a source text through different languages, ending up in the language I want to eventually work in. What exactly the human writer does with the output of multiple machine translation is left to her discretion and poetic decision. It could be left untouched, reframed, rearranged, put into a proper syntax or even just be used as a starting point for an anthropogenic text. As shown by the example in the beginning of this text, the combination of text blocks from different stages of the translational process often proves to be fruitful, too. The human writer also decides whether she wants to reestablish a continuous narrative thread in the text or leave the search for one entirely to the reader.

Source code & Détournement

> "Ideas improve. The meaning of words participates in the improvement. Plagiarism is necessary. Progress implies it."[13]

Just as thousands of existing texts and translations online feed the corpus of GT, existing texts from any source can freely be used as input. The copyright dissolves in a swirl of translations and rearrangements. As Guy Debord and Gil J. Wolman put it in an early Situationist publication:

> "It goes without saying that one is not limited to correcting a work or to integrating diverse fragments of out-of-date works into a new one; one can also alter the meaning of those fragments in any appropriate way, leaving the imbeciles to their slavish reference to 'citations'."[14]

In the case of our practice even few rounds of translation will add an inspiring dadaistic touch to dry prose or bureaucratic correspondence. This can be seen as an enrichement compared to the original and inspire the prejudiced human mind for the postprocessing. Analogously, the most painful kind of source is probably poetry and beautifully composed prose, since the automatized translations will shatter a masterpiece before the eyes of the helpless admirer. While yielding grand potential, this incision will make it much harder for the prepossessed human writer to find a new aesthetic in the shambles of beauty still rememebered. The use of classics – or in general 'old' texts – as a starting point is especially interesting, as the translations resituate the text in contemporary (and historical) linguistic structures.

[13] Debord, 1967: Ch. 8, sct. 207.
[14] Debord and Wolman, 1956.

Using Situationist terms, we could call this a ‘détournment’ of words and expressions in a space of language, drifting through different linguistic relationships:

> “Détournement, the reuse of preexisting artistic elements in a new ensemble, has been a constantly present tendency of the contemporary avant-garde, both before and since the formation of the SI [i.e., Situationist International] . The two fundamental laws of détournement are the loss of importance of each detourned autonomous element – which may go so far as to completely lose its original sense – and at the same time the organization of another meaningful ensemble that confers on each element its new scope and effect.”[15]

> “Any elements, no matter where they are taken from, can be used to make new combinations. The discoveries of modern poetry regarding the analogical structure of images demonstrate that when two objects are brought together, no matter how far apart their original contexts may be, a relationship is always formed. Restricting oneself to a personal arrangement of words is mere convention. The mutual interference of two worlds of feeling, or the juxtaposition of two independent expressions, supersedes the original elements and produces a synthetic organization of greater efficacy. Anything can be used.”[16]

The détournement of texts using multiple computer translation creates a ‘mutual interference’ of the corpus and structures of the translator with the source text. Regarding the latter, the output can be read as a faint and distorted echo (i.e., an allusion), as a response, or even as a literal interpretation of the source text[17]. An interpretation based on semantics rooted in a corpus of ‘pure language’[18] rather than human-intended semantics; on a globalized context rather than the limited one of the original and its author. Of course, this new context is not universal. It is the interference with the corpus of the respective online translator. It is therefore also limited, but directly embedded in the core of the information age: the Internet.

While this kind of interpretation is rather hard to catch hold of in an analytical way, I argue that analytical understanding is not the point of this practice. The point is to exploit its potential in a poetic practice:

> “In itself, the theory of détournement scarcely interests us. But we find it linked to almost all the constructive aspects of the presituationist period of transition. Thus its enrichment, through practice, seems necessary.”[19]

Applied Uncreativity

But how does all of this relate to ‘uncreativity’? Many aspects of the translational poetic practice set out above point towards an updated approach to the creation of literature as compared to the romantic ideal of the creative, the poet as an original genius, drawing something new out of his inner source while over-looking the world as the object of his art in a prophetic pretense.

---

15 Situationist International, 1959.
16 Debord and Wolman, 1956.
17 This is also true for ordinary human translations of literature. Just that they usually focus more on the translation of meaning. Their interpretations operate thus on another level than those of statistical computer translation: They interfere with the arbitrary background of the translator rather than ‘Big Data’ and a structuralistic approach.
18 c.f. Nabugodi, 2014.
19 Debord and Wolman, 1956.

Artists like Kenneth Goldsmith and theorists like Marjorie Perloff argue for a more “unoriginal genius”[20]. Understanding “uncreativity as a creative practice”[21,22] in the light of mechanical and digital reproducability and the availability of an immense load of information and texts online challenges the classical idea of originality. It is not possible to master the oversupply of information with a synoptic view from the outside.

Marjorie Perloff argues for an unoriginal genius being the writer who masters this challenge by selecting information and pushing language around in a playful and emotional involved manner. In his project *Printing Out the Internet*, Goldsmith performed this new position of the unoriginal genius lying inside the heap of print-outs instead of overseeing it; He is aware of being small but not overwhelmed by it.

This is exactly what happens in the course of using online statistical computer translation for poetic creation. The artist selects source texts and languages to be revisited based on her own poetics. The database of the online translator manages the corpus of relations and expression available, the algorithms rearrange the source text according to the emprically based structure and relations of the languages involved. Like Charles Baudelaire's flâneur, the poet is a “passionate spectator”[23] observing the words drifting through different languages – some well-known to her and some indecipherable. She decides which turns to take and when to return home to her mothertongue. This can be viewed as a digitalized linguistic act of Détournment, recontextualizing the source text by strolling through random languages, following a mood or temporary interest for the fate of an expression in a certain sequence of languages.

But it is more than just a stroll, it is a process of reframing the source text by coating it into new words, decorating it with remnants from foreign languages. It is a multiple process of language acting upon language, a détournement moderated by the writer. This also makes it a game: The writer is playing with the hard-wired structuralistic statistical workings of GT or Bing, the fine-tuned nuances and intentions engraved in the source text, and eventually the own constituted expectations based on the knowledge about the functionality of the statistical computer translator, the languages visited, as well as the source text and its context.

Then, the output needs to be viewed, managed, rearranged and reframed again. This is where everything undergone by text and observer in the previous steps is reflected upon and processed. This involves poetic association and judgement. At this point it becomes obvious, that *uncreativity* doesn't imply *noncreativity*: The subjective workings of the human mind are required to decide on how to configure the final version of the text, its purpose.

In other words: The text that has undergone détournement by interfering with statistical computer translation is a material that needs to be further manufactured. The notion of the materiality of language can also be seen in a context of uncreativity:

> “... The flip side of digital language is its malleability, language as putty, language to wrap your hands around, to caress, mold strangle.
>
> The result is that digital language forgrounds its material aspects in ways that were hidden before.”[24]

The image of Goldsmith lying in a heap of print-outs from the internet pictures this aspect very impressively. Applied to what I have set out on the uncreative workings of multiple translations so far, we

[20] Perloff, 2010.
[21] Goldsmith, ?.
[22] Goldsmith, 2011.
[23] Baudelaire, 1964.
[24] Goldsmith, 2011.

could say that this practice foregrounds many material aspects of language. As the human subject in the progress, we are urged to do something with that material. But what?

First of all, we can use it to produce economical efficient poetry, a digitalized remake of Raymond Queneau's *Cent mille milliard des poèmes*, or, maybe even more fitting, his *Exercises in Style*, where one depiction of a couse of events is rewritten – or translated – into many different styles, perspectives, and narratives. Just taking a few words or a simple sentence and translating it along different paths of languages creates a multitude of remakes that are 'same same but different'[25].

Next to simple reproduction, another possibility is of course to reuse the output in order to create a new postprocessed text, engraving selected interpretations in form of syntax and narration on it. But it would be a pity to hide from the reader the playful process leading to the final output. It is much more interesting to invite her to a game:

Why not make the reader try to reconstruct lost narratives and meaning herself, allow her to compare them to the source text, and to search for traces of it in the result? Being aware of the methodology of the translational poetic practice, the reader will automatically speculate on the remnants of the source text, the particular expressions and modifications brought in by the online translator, the influences and remnants of the different languages passed in the process, as well as the extend of human postprocessing.

Posthuman poetics?

This open-source[26] approach also raises awareness of the ambiguous relationship of the human, the machine, and the Internet displayed in the translational poetic practice. In his 2009 publication *Man in Age of Technology*, Umberto Galimberti incapacitates humanity with respect to technology:

> "We are all used to thinking of technology as a 'tool' at the disposition of man (since the first ape got hold of a stick to open fruits) but the fact is that this position has now been completely reversed. It is technology that has become the subject of history while man has become a mere functionary of his technological apparatuses. If technology is the subject of history and man merely an obedient functionary, then we can only admit that humanism is at an end and that the humanistic categories that we have always used to read history are no longer capable of interpreting the age of technology."[27]

From this perspective, the internet is not only a tool for communication and a resource of information, it also acts as a subject changing our lives, our understanding of and our access to the world and other human beings. Yet, it is not a force lying outside the human capacities as it might be received from a postmodern perspective, it is rather a synthesis of human activity and technology. It can neither exist nor function without information technology, nor without humanity. The internet can be considered as memory and filter, thus intensifying man's capabilities. As we saw earlier, the online translator (e.g. GT) acts as a subject in the process of the translational poetic practice, and it acts fully automatized. But at the same time, it draws on the work of Google's programmers and experts as well as the writtings and translations produced by thousands of human beings.

---

25 *Same Same but Different* is the title of a movie telling the love story of two human beings from Cambodia and Germany. The title is apparently an English dictum commonly used in Thailand. http://www.imdb.com/title/tt1368443/?ref_=ttrel_rel_tt (last accessed 7 January 2015)

26 Obviously and unfortunately, GT and Bing are not open-source. But that is another story. Here, I refer to the corpus being expolited as an open source of language.

27 Galimberti, 2009.

> "Technology radically changes the way we think as, even if machines have been invented by man, they now contain an objectification of human intelligence that is superior to the competence of single individuals. The memory of a computer is vastly superior to our memories, even if it is a 'stupid' memory. Using computers modifies our thinking, transforming it from 'problem-solving' thinking to 'binary' thinking which follows the scheme 1/0 and renders us able only to say yes, no or at best, 'I do not know'."[28]

The use of a tool like GT in poetic practice goes beyond the capabilities of the indiviudal human writer. But by working as an unoriginal genius, by managing, pushing, and being emotionally and aesthetically involved, the writer also goes beyond the 'binary' capabilities of computers, and beyond the structuralistic engine of statistical translators. Therefore, I suggest that this practice is not antihuman as it could be feared from Galimberti's rhetoric, but posthuman, a literal 'software-update' for human literature as suggested by Steve Nichols in *The Posthuman Manifesto*[29]. I do not want to go into detail concerning the theory of Posthumanism here. I just want to suggest that the activity of the unoriginal genius, playing with the capacities and limitations of an 'octopus information'[30] like GT and Bing Translator, fits very well into that perspective, as does the whole discourse of uncreativity. Accepting and appreciating that creativity is not drawn from pure human originality but from the management of information and the emotionally involved play with influences opens up a whole new space for poetic practices.

Conclusion

Statistical computer translation is based on a corpus of linguistic material, a huge archive of human linguistic activity. At the same time, it works strictly 'binary', statistically implementing a structuralistic approach based on the assumption that everything it is asked to translate has been said (and translated) before. This workings can be exploited by a writer to transform a source text based on the relations of the expressions in the source text to expressions in different languages. These structural relations foreground certain material aspects of language and ignore factors like meaning or narrative continuity.

Multiple translations using online tools like GT can be seen as a détournment of the source text. The interference with the corpus and structuralistics workings of the online translator offer new possible meanings and aesthetics to the text.

Thus, the material produced has a high potential as 'putty' for uncreative poetic practice. What is made of it is up to the poetics of the writer. At the same time, being transparent about the translational practice and the choice of source texts opens the possibility to play with the reader and raise awareness of this updated view on creativity.

Translational poetic practice is but one possibility among many others of exploiting the uncreative potential of the concentration of information in the world wide web and the use of computational power. The posthuman claim of such practices is based on the appreciation of the potential and power of these digital resources and the playful engagement with them as unoriginal genii among others. But still, is it really not simply by chance that *imagine* shifts through *think* to *believe* in the first lines of this essay? Let's pass this question on to the mighty engines of GT: *But again, this is a coincidence, and a believer in the concept with the opening of the line this quarter?*

---

28 ibid.

29 Nichols, 1988.

30 *Input to GT:* 'Daten-Krake' (German dictum). *Multiple translation:*
> 數據八達通 > Мэдээллийн Наймаалжны > معلومات الأخطبوط > Informaçõe octopus > octopus information

Bibliography

- Barthes, Roland. *Die strukturalistische Tätigkeit.* Kursbuch 5, S. 190-196, 1966. Print.
- Bellos, David. *How Google Translate Works.* The Independant, September 2011. Web. 17 January 2015.
  http://www.independent.co.uk/life-style/gadgets-and-tech/features/how-google-translate-works-2353594.html
- Baudelaire, Charles. *The Painter of Modern Life, and Other Essays.* London: Phaidon Press, Ltd., 1964. Print.
- Debord, Guy. *The Society of the Spectacle*. Trans. Donald Nicholson-Smith. New York: Zone Books, 1967. Print.
- - - - , and Gil J. Wolman. *A User's Guide to Détournement*. 1956. Trans. Ken Knabb. In: *Situationist International Anthology*. Berkeley CA: Bureau of Public Secrets, 2006. Print/Web.
  http://www.bopsecrets.org/SI/detourn.htm
- Galimberti, Umberto. *Man in the Age of Technology*. The Journal of Analytical Psychology, 54(1):3-17, 2009. Print.
- Goldsmith, Kenneth. *Uncreative Writting: Managing Language in the Digital Age*. Columbia University Press, 2011. Print.
- - - - . *Uncreativity as a creative practice*, ?, Web. 16 January 2015.
  http://epc.buffalo.edu/authors/goldsmith/uncreativity.html
- Nabugodi, Mathelinda. *Pure Language 2.0: Walter Benjamin's Theory of Language and Translation Technology.* Feedback, Open Humanities Press, May 2014. Web. 17 January 2015.
  http://openhumanitiespress.org/feedback/literature/pure-language-2-0-walter-benjamins-theory-of-language-and-translation-technology/
- Nichols, Steve. *The Posthuman Manifesto.* UK, 1988. Print.
- Perloff, Marjorie. *Unoriginal Genius: Poetry by Other Means in the New Century.* University of Chicago Press, 2010. Print.
- Schiebinger, Londa. *Machine Translation: Analyzing Gender.* Case study, Gendered Innovations, 2013, Web. 11 January 2015.
  http://genderedinnovations.stanford.edu/case-studies/nlp.html#tabs-2
- Situationist International. *Détournement as Negation and Prelude*. Trans. Ken Knabb. In: *Situationist International Anthology*. Berkeley CA: Bureau of Public Secrets, 2006. Print/Web.
  http://bopsecrets.org/SI/3.detourn.htm