Machine Translation: Technologies and Applications
Series Editor: Andy Way

Bernard Scott

Translation, Brains and the Computer

A Neurolinguistic Solution to Ambiguity and Complexity in Machine Translation

Springer

Machine Translation: Technologies and Applications

Volume 2

This book series tackles prominent issues in MT at a depth which will allow these books to reflect the current state-of-the-art, while simultaneously standing the test of time. Each book topic will be introduced so it can be understood by a wide audience, yet will also include the cutting-edge research and development being conducted today.

Machine Translation (MT) is being deployed for a range of use-cases by millions of people on a daily basis. Google Translate and FaceBook provide billions of translations daily across many language. Almost 1 billion users see these translations each month. With MT being embedded in platforms like this which are available to everyone with an internet connection, one no longer has to explain what MT is on a general level. However, the number of people who really understand its inner workings is much smaller.

The series includes investigations of different MT paradigms (Syntax-based Statistical MT, Neural MT, Rule-Based MT), Quality issues (MT evaluation, Quality Estimation), modalities (Spoken Language MT) and MT in Practice (Post-Editing and Controlled Language in MT), topics which cut right across the spectrum of MT as it is used today in real translation workflows.

More information about this series at http://www.springer.com/series/15798

Bernard Scott

Translation, Brains and the Computer

A Neurolinguistic Solution to Ambiguity and Complexity in Machine Translation

Springer

Bernard Scott
Tarpon Springs, FL, USA

ISSN 2522-8021 ISSN 2522-803X (electronic)
Machine Translation: Technologies and Applications
ISBN 978-3-319-76628-7 ISBN 978-3-319-76629-4 (eBook)
https://doi.org/10.1007/978-3-319-76629-4

Library of Congress Control Number: 2018935124

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Gratefully Dedicated
A. M. D. G.

Any intelligent fool can make things bigger and more complex. It takes a touch of genius – and a lot of courage – to move in the opposite direction.

– Albert Einstein

Preface

This book is about Machine Translation (MT) and the classic problems associated with this language technology. It is intended for anyone who wonders what if anything might be done to relieve these difficulties. For linguistic, rule-based systems, we attribute the cause of these difficulties to language's ambiguity and complexity and to their interplay in logic-driven processes. For non-linguistic, data-driven systems, we attribute translation shortcomings to the very lack of linguistics. We then propose a demonstrable way to relieve these drawbacks in both instances.

Throughout the book, we present a variety of translations by several of the most prominent linguistic, statistical and neural net MT systems in use today. Our object in doing this is to illustrate both the relative strengths and weaknesses of the various technologies these systems embody. The book's principal intent, however, is not to promote one particular translation system over against others (not even the one the author worked on for thirty years, described herein as Logos Model), but rather to examine the deeper and more critical question of the mechanisms that underlie the translation act itself, and to illustrate what can be done to optimize these mechanisms in a translation machine. We hold this to be the more fundamental issue that needs to be addressed if the classic problems associated with MT are to be solved, and consistent, high-quality machine output is ever to be realized.

Because the linguistic processes of the brain are singularly free of the classic difficulties that beset the machine, we have looked to the brain for possible guidance. We describe a working translation model (Logos Model) that has taken its inspiration from key assumptions about psycholinguistic and neurolinguistic function. We suggest that this brain-based mechanism is effective precisely because it bridges both linguistically-driven and data-driven methodologies. In particular, we show how simulation of this cerebral mechanism has freed this one MT effort, Logos Model, from the all-important, classic problem of complexity when coping with the ambiguities of language. Logos Model accomplishes this by a data-driven process that does not sacrifice linguistic knowledge, but that, like the brain, integrates linguistics within a data-driven process. As a consequence, we suggest that the brain-like mechanism simulated in this model has the potential to contribute to further advances in MT in all its technological instantiations.

These admittedly are controversial claims, and we recognize that the reader may be inclined to dismiss them out of hand, especially given the fact that the model being described, Logos Model, had its origins more than 45 years ago in the earliest days of MT. How, one will ask, can technology from so far back in time offer anything of interest to present-day MT? That is certainly a legitimate question, but it is one we trust this book will answer.

As readers work their way through this book, they will see that we are showing Logos Model at its best, seemingly at times at the expense of other translation systems. Our purpose in writing this book, however, has *not* been to prove that Logos is a better translation system. In terms of the general output quality of many MT systems nowadays, no such claim could be defended. Our purpose rather has been quite different, namely, to demonstrate that the technology underlying Logos Model offers a demonstrable solution to the problem that complexity poses for MT. As we argue throughout this book, complexity is the one issue that is most apt to limit the ultimate potential of any MT system, whether linguistic, statistical or neural. And we attempt to show that Logos Model, originally designed as it was to address the complexity problem, may offer a workable answer. Logos Model translations shown in this book are meant to demonstrate that the model must be doing something right in that regard, something that we trust would be of interest to MT developers generally. Allow me to repeat the point. Logos Model translations in this book are *not* intended to prove that Logos is a better system, only that underlying Logos Model technology may have something of genuine interest to offer. It is hoped that the MT community will understand this, and that the empirical data, arguments and personal testimony we have presented will be considered in constructive spirit intended.

One final matter. Translations shown in this book by Google Translate, Microsoft's Bing Translator, SYSTRANet, PROMT Translator and LISA Lab's neural MT system were carried out in the 2016–2017 timeframe. Readers should be aware that these translations do not necessarily represent output of these systems subsequent to this 2016–2017 timeframe. Readers will note that output from Google Translate and Bing Translator had to be marked as either statistical or neural, since both the Google and Microsoft systems transitioned to neural net technology in late 2016 as this book was being written.

Tarpon Springs, FL, USA Bernard Scott

Acknowledgements

The following individuals, among still others, are responsible for having turned the concepts of this book into a working system.

Linguists
Kutz Arrieta
Anabela Barreiro
Mary Lou (Christine) Faggione
Claudia Gdaniec
Brigitte Orliac
Karen (Liz) Purpura
Patty Schmid
Steve Schneiderman
Sheila Sullivan
Monica Thieme

Systems
Charles Byrne (Logos co-founder)
Daniel Byrne
Jonathan Lewis
Oliver Mellet
Barbara Tarquinio
Kevin Wiseman
Theresa Young

Corporate
Cecilia Mellet
Robert Markovits
Arnold Mende
James Linen
Charles Staples
Hans Georg Forstner

Special thanks to Anabela Barreiro and Fernando Batista (INESC-ID) without whose assistance this book would not have been possible. And much appreciation for helpful suggestions by two anonymous reviewers.

Permissions

All figures in this book were created by the author except for those listed below, used with permissions indicated:

Fig. 4.3 Permission granted by Dr. Francis Wong
Fig. 4.5 Permission granted by National Institute of Health under GNU Free Document License (GFDL)
Fig. 4.6 Permission granted by Oxford University Press
Fig. 5.1 Permission granted by Elsevier
Postscript Fig. 5G-1 Permission granted by World Scientific Publishing
Fig. 8.1 Permission granted by the Association of Computational Linguistics (ACL)
Fig. 8.2 Permission granted by Christoph Bugmer under GNU Free Document License (GFDL)
Fig. 8.4 Permission granted by Elsevier

Contents

Part II

Part I

Chapter 1
Introduction

Abstract This chapter illustrates some of the errors that translation machines may make in processing natural language. That these translation errors would never have been made by human translators raises the question as to why this is so. What is the translator's brain doing in its handling of language that is different from the machine? Is the brain processing language in ways that have yet to be understood? Could it be that neuroscience has overlooked evidence of a cerebral language process that is different from what cognitive science and neurolinguistics have traditionally proposed for the brain? What would such a hitherto under-recognized process look like and might it be simulatable in a translation machine? These questions constitute the topic of this book.

Most readers areaware that Alan Turing, at the very outset of the computer age, had suggested these new number-crunching machines ought also be able to do things like play chess and translate languages. Turing was prescient about chess, but as we know his predictions about computers and language have never been fully realized. Turing inspired many to try, but as the record shows, the history of the computer's struggles with language is long and legendary. And while telling progress has been made since the inclusion of statistics and neural net technology into the effort, huge, unsolved problems remain, making the goal of consistent, high-quality MT seem as illusive as ever. After so many decades of trying and never quite getting there, one has to wonder whether there is not something about language itself that is ultimately intractable.

Michael Halliday (2003) noted that language is among the most complex phenomena in the universe. To a non-linguist his statement must seem greatly overblown. After all, language cannot be that complex if children acquire it so readily, virtually without even trying. The world is filled with natural phenomena that seem far more complex than language, and more difficult to grasp. Weather is a good example, things like avalanches and earthquakes are another. But there's a difference here. The complexities of nature are invariably regular and law-like, even if the factors are non-linear and almost impossible to pin down. The complexity of language seems of an entirely different kind.

B. Scott, *Translation, Brains and the Computer*, Machine Translation: Technologies and Applications 2, https://doi.org/10.1007/978-3-319-76629-4_1

It is tempting to think otherwise of course, to treat language as essentially regulated by rules. After all, isn't that what grammar is all about, the regularities of language? To be sure, that's true, but the language that grammarians typically deal with is different from the language that the translation machine (or the brain) has to deal with; it has usually been purged of the ambiguities and irregularities that make for the troublesome character of raw language. By contrast, the language fed to the computer is ambiguous, complex, and often as not, irksomely resistant to rules. Just like the language, we might add, that gets input to the brain. What's so interesting about this is that none of these aspects of language that so bother the machine seem to trouble the brain in the least. Do we know why this is so?

It is easy to understand why Chomsky reduced linguistics to the simpler, more manageable level of syntax. And why his influence led linguistics and computational linguists after him to deal with language in terms of abstract, univocal symbols (N, V, PREP, etc., with word class subcategories). Chomsky's reduction is what made literal language suddenly amenable to generalized treatment and the formulation of axiomatic rules. MT couldn't have gotten off the ground were it not for this reduction. Apart from syntax, how else was one to deal with language in a generalized way?

But for all the supposed gains of Chomsky's syntax-based approach to language, it must be remembered (as statistical MT reminds us) that the language the translation machine must cope with comprises strings of literal words, vast quantities of them, not just a handful of syntactic symbols. And these literal words, apart from some context, often have no univocal meaning or even fixed grammatical function. Formal grammars are generally not obliged to deal with words at this equivocal level, but translation machines most certainly are. One might observe here that so are the brains of human translators.

To illustrate what the translation machine (and the brain) is typically confronted with, let's imagine a string of words, *w1, w2, w3,* etc., that constitutes an input sentence in English. Imagine that one of these words happens to be the morpheme *sound*. Out of context, neither the machine (nor the brain) has any idea what the morpheme *sound* signifies. Is it a noun, a verb, an adjective? If the context tells us that *sound* is a noun, there is still much a translation agent needs to figure out. Does *sound* as a noun denote something you hear, something used in a medical practice, or maybe a body of water?

If *sound* turns out to be a noun that denotes audible phenomena, should it be transferred in German, for example, as *Ton* (for *sound* in general); *Geräusch* (as in a strange *sound*); or *Klang*, (as in the *sound* of music). Obviously, in all these cases, the context holds the key, but that only raises a more troublesome issue. How does a translation machine get at something as nebulous as context? How does the translator's brain? And, most troublesome of all, what happens when the context itself is equivocal in some way, as it very often is? It is not hard to understand why a machine, whether driven by rules, statistics, or some other mechanism, will have its difficulties. It is less clear why brains are normally free of these problems.

To illustrate our point about the difficulties for the machine, consider the string of short sentences in (1). Notice that the morpheme ***sound*** occurs variously in all of its parts of speech.

(1) *John's heart is **sound**. John's heart **sounds** healthy. The **sound** of the heart **sounds** normal.*

We had these short sentences translated into German by the most prominent of present-day phrase-based statistical systems, Google Translate,[1] and also by the oldest and most well known linguistic system, SYSTRANet.[2] Output is given in (1)(i) and (1)(ii) below. Errors are underlined. Oddly, errors are curiously rampant for such simple sentences.

Translation by Google SMT Translate[3]

(1)(i) *Johns Herz ist **gesund**. Johns Herz **töne** gesund. Der Klang Herz **töne** normal.*

Translation by SYSTRANet

(1)(ii) *Johns Herz ist **Ton**. Johns Herz **töne** gesund. Der Ton der Herz **töne** normal.*

Commentary

In the first sentence of (1)(i), the statistical system nicely renders the predicate adjective ***sound*** with the German adjective ***gesund*** (*healthy*). In the first sentence of (1)(ii), the linguistic system misresolves the predicate adjective ***sound*** to a noun (***Ton***). In the second sentence of (1)(i) and (1)(ii), both systems misresolve the string *John's heart **sounds*** to a noun phrase instead of correctly recognizing it as a subject-predicate construction. Similar misresolutions of ***sound*** occur in the third sentence of both systems.[4]

That two such well-established translation systems should mishandle simple sentences like these is both surprising and telling. Surprising because both systems are generally able to translate much more difficult sentences quite well; telling because stumbles like this in the simplest of sentences, after many years of development, suggest some inherent difficulty underlies the process. Are such deficiencies signs of underlying weakness, or are they merely indications that work still needs to be done? Doubtless the latter will always be the case, but such errors may well suggest a deeper problem.

[1] In November, 2016, Google released a new neural MT version (GNMT) of Google Translate. Except where otherwise noted, a number of the Google translations in this book predate this release. See Postscript 6-B of Chap. 6 for the GNMT versions of Google Translate's SMT translations shown in this book.

[2] SYSTRAN is now said to be a hybrid linguistic/statistical system. Historically, its foundation is linguistic and algorithmic in nature.

[3] Google Translate's new GNMT output for the third sentence in (1) is now correct: *Der **Klang** des Herzens **klingt** normal.*

[4] Logos Model translation of (1) is syntactically correct: *Johns Herz ist **solid**. Johns Herz **klingt** gesund. Das **Geräusch** des Herzens **klingt** normal.*

No one can deny that shortcomings are intrinsic to MT, in all its stages and all its approaches. The long history of MT corroborates this clearly enough. In MT, output errors are a way of life, whether one's approach is linguistic, statistical or neural. And certainly, the developers of these systems can fix this particular problem, and the countless other output errors lined up behind it waiting to be dealt with. There's little doubt about that. The real issue is not whether a particular issue can be fixed. The real question is what happens when new fixes begin at times to undo old fixes, requiring them to be re-fixed, and this keeps happening until the cumulative process becomes harder and harder to maintain? Are the above errors a sign this might already be happening? We cannot know of course, but whatever the case, complexities like this are always lurking in the wings, threatening to frustrate progress.

Here's another question. Why is it that human translators never make these sorts of errors? What keeps them from virtually ever misresolving a part of speech? The quick answer of course is that translators have brains and machines do not. Language after all is a product of the brain, part and parcel of its cognitive life. But that quick answer merely begs the question why this is so. What is the brain doing that frees it from such blunders? Are its linguistic mechanisms data-driven or rule-driven, as the different schools tend to argue? Or is the brain possibly doing something that the machine is inherently incapable of, something so complex as to defy our programming capabilities, or possibly even our understanding? Perhaps, but to my knowledge no cognitive scientist or neuroscientist has ever suggested as much.

Cognitive scientists like Stephen Pinker (1994) tend to suggest to the contrary that linguistic brain function is best understood by symbol manipulation in accordance with Universal Grammar, i.e., by processes typical of the computer. And neuropsychologists like Angela Friederici, among a host of others, conduct their language research in the syntax-oriented, Universal Grammar framework.[5] Many neuroscientists have published neuroimage-based studies supporting a syntax-first, grammar-driven view of language processing in the brain, more or less in keeping with Chomsky's perspectives on language.

But other takes on brain function are also possible. For example, neuroscientists also recognize that data-driven word associations, frequencies and patterns play an important role in cognitive operations, especially in the learning and use of language. This is basically the perspective of this book. And certainly, the exhaustive, data-driven explorations of raw language that underlie statistical MT (SMT) are what account for its remarkable success, and for the improvements in hybridized, rule-based MT (HMT). Data-driven processes are also behind the immense promise apparent in the emerging neural MT (NMT) technology.

From a strictly historical perspective, then, it seems that we are left with only two recognized ways of processing language for translation purposes: processes that are either rule-driven or data-driven. And these ways seem to apply whether this process-

[5] The work of Optiz and Friederici (e.g., 2003) is a good example of this, as may be seen in Chap. 5.

ing is postulated for brain or machine. On one side we have the (by now) classical theory that holds that language processing must be driven by computer-like manipulations of symbols, which means at the abstract, syntactic level. Its proponents in computer science, cognitive science and neuroscience don't see how language can be dealt with in any other generalized way. (See fuller discussion of this issue in Chap. 5).

On the other side, in contrast to this theory, we have the practical evidence of various data-driven systems, among them SMT, NMT, and Logos Model described in this book. The successes of these systems demonstrate that language can be translated fairly effectively by a process driven not by rules of grammar but by patterns and frequencies extracted from the data of raw language itself.

It is tempting to think that a data-driven, non-linguistic approach is effective precisely because it might be closer to actual brain function. But even if that should partly be the case, data-driven MT remains deficient, as their developers freely acknowledge, and that deficiency has to do with linguistic competence. Something is still lacking in the data-driven process that the brain is able to do.[6]

So we continue to be left with the question how the brain processes language. Why does the brain not exhibit any of MT's manifest shortcomings? Why, for instance, would the translator's brain never make the resolution errors seen in (1)(i) and (1)(ii)? It is clear the translator's brain is doing something not yet accounted for in theory or in practice, neither by data-driven SMT and NMT nor by the rule-driven approaches of linguistic systems. Does anyone at all have a clue what this cerebral mechanism amounts to? Are there perhaps neurological data that might suggest what this brain function might actually look like (e.g., evidence that may have been overlooked or wrongly interpreted by neuropsychologists in the past)?

No one of course can expect a definitive answer to our question; neuroscience will not be coming up anytime soon (if ever) with a settled understanding of linguistic function in the brain. But suppose, for a moment, that neurological data do exist that suggest the brain really could be handling language differently than has been surmised. Such data could be suggestive, even if not definitive. And suppose, furthermore, that this suggestive cerebral process could be modeled in a machine so that this under-recognized cerebral operation might be tested and persuasively demonstrated. Suppose, most importantly, that this brain-based computer model were to show itself capable of dealing with language's ambiguities in a way that, for the very first time in any MT system, does not incur the usual complexity costs. What would it mean if this were so, if we had an MT process that was free from complexity in its dealings with ambiguity, free from the one factor that (in our view) most limits the MT process?

If it should turn out that the brain actually has this insufficiently recognized, complexity-free mechanism, and if indeed a computer model could be made to demonstrably persuade that this mechanism is simulatable in a machine, might it not mean that a way was now in hand that could free MT endeavors of the one thing

[6] What remains to be answered is the question of how one is to deal with language in a generalized way without doing as Chomsky did, i.e., without reducing language to syntax. This is the question this book attempts to answer, as it has a critical bearing on the underlying linguistic competence of any MT system.

that most constrains its success, viz., the growing complexity of these systems as they mature and strive for greater and greater improvement? Wouldn't it mean that fixes and improvements without end might now be implemented without incurring these complexity costs? If all this were so, wouldn't it suggest that machine output of human quality might now be ultimately foreseeable? Alan Turing saw no reason why such a thing ought not be possible. Nor did we who spent decades in developing and testing Logos Model as described in this book, naïve as this may sound.[7]

Admittedly, the above narrative is apt to strike the reader as far-fetched, if not utterly fanciful. After all, MT has been pondered and worked on from virtually every conceivable angle for over 50 years now. Are we to believe that some telling, beneficial clue from the brain has been overlooked? Are we to believe that simulation of an unappreciated brain mechanism could conceivably open up new, unforeseen vistas for MT? That must hardly seem likely, and yet that is the reason we have undertaken to write this book. We present both theory and evidence for precisely these assertions. We of course must leave it to our readers to make what they will of its findings and claims.

To illustrate something of the potential strength of Logos Model in a very small matter, consider the sentence in (2) and the French and German translations that follow. Sentence (2) once again involves the morpheme ***sound***, and has been artificially contrived so as to allow ***sound*** to occur in all three of its parts of speech in a single sentence.

(2) *The* ***sounds*** *of his heart* ***sound sound***.

We translated sentence (2) into German and French by the statistical and linguistic MT systems identified below, this time including output from Logos Model. Errors are underlined.

<u>Translation by Google SMT Translate</u>[8]

(2)(i) *Die* ***Klänge*** <u>*des*</u> *Herzens* <u>***Ton Ton.***</u>

(2)(ii) *Les* ***sons*** *de son cœur* ***sonnent*** <u>***sonore***</u>.

<u>Commentary</u>

In the German translation in (2)(i), only the first instance of the morpheme ***sound*** is handled correctly, viz., its noun sense. In the second instance, the verb sense of ***sound*** is misresolved to a noun (***Ton***). In the third instance, the adjectival sense of ***sound*** is also misresolved to a noun (***Ton***). By contrast, in the French translation in (2)(ii), all syntactic functions of ***sound*** are resolved correctly. The adjective ***sonore*** however incorrectly renders ***sound*** as *sonerous*.

[7] In Sect. 8.4 The Hippocampus and Continual Learning of Chap. 8, we discuss a paper by Kumaran et al. (2016) that describes a previously unrecognized power for *semantic* generalization in the hippocampus, and the relevance that this unappreciated cerebral mechanism has for AI's deep learning (and, in our view, more specifically for MT).

[8] Google Translate's new neural MT (GNMT) translation of (2) into German shows some improvement: (2)(i) *Die* ***Klänge*** *seines Herzens* ***klingen***____. The French translation however is poorer:

Translation by Microsoft's Bing SMT Translator[9]

(2)(iii) *Die* ***Klänge*** *des Herzens* ***klingt sound***.

(2)(iv) *Les* ***sons*** *de son bruit sonore du coeur*.

Commentary

In the German translation in (2)(iii), ***sound*** both as noun and verb are resolved correctly. However, the verb ***klingt*** is singular and should be plural (***klingen***). For some reason, ***sound*** as adjective is left untranslated. In the French translation in (2)(iv), the verb and adjectival functions of ***sound*** are both misresolved, completely destroying the translation.[10]

Translation by SYSTRANet

(2)(v) *Die* ***Töne*** *seines Herz* ***tontones***.

(2)(vi) *Les* ***bruits*** *de son* ***bruit*** *de* ***bruit*** *cardiaque*.

Commentary

In the German translation in (2)(v), ***sound*** as both verb and adjective are misresolved, completely ruining the translation. In the French translation in (2)(vi), the verb and adjectival functions of ***sound*** are also both misresolved.

Translation by Logos Model

(2)(vii) *Die* ***Geräusche*** *seines Herzens* ***klingen solid***.

(2)(viii) *Les****sons*** *de son coeur* ***semblent solides***.

Commentary

In the German in (2)(vii), all parts of speech for ***sound*** are resolved correctly. In the French in (2)(viii), again, all parts of speech for ***sound*** are resolved correctly. Note the translation of the verb ***sound*** as ***semblent*** (*seem*).

All systems have their strengths and weaknesses. Some are better at one thing, others at another. SMT systems normally do quite well, but not always and not uniformly. NMT systems tend to be better, but again, not always. In (3), below, we focus on another instance where translation depends upon linguistic competence. At issue here is the syntactic function of the morpheme ***that***, whether it is a conjunction introducing a ***that*** clause, or a demonstrative pronoun. Errors are underlined.

---

(2)(ii) *Les* ***sons*** *de son* ***son*** *du* ***son*** *du cœur*. See Postscript 6-B in Chap. 6 for the GNMT versions of Google SMT output shown here and there throughout this book.

[9] In late 2016 Microsoft also incorporated neural net technology into its Bing Translator. Bing Translator translations shown in this book are either from the SMT or NMT version, and are marked accordingly.

[10] Bing Translator's new neural net system appears to have regressed in its translation of (2), both in the German and French: (2)(iii) *Der* ***Klang*** *seines Herzens* ***Klang klingt*** (2)(iv) *Les* ***sons*** *de son cœur* ***son***.

(3) *He said* ***that*** *was not the case at all.*

Translation by SYSTRANet and Logos Model

(3)(i) *Il a dit que___ n'était pas le cas du tout.*

Translation by Microsoft's Bing SMT translator

(3)(ii) *Il a dit que* ***ce*** *n'était pas du tout le cas.*

Commentary

Both of the linguistic systems that we tested (SYSTRANet and Logos Model) mistakenly treated the demonstrative pronoun ***that*** as a conjunction, as shown by the underlining in (3)(i). So too did Google Translate (SMT translation, not shown). The exception was the translation by Microsoft's Bing translator in (3(ii)). Both the SMT and NMT versions of Bing translator correctly interpreted the morpheme ***that*** as a demonstrative pronoun (***ce***), and properly inserted the conjunction (***que***) missing in the English sentence. (French, unlike English, does not allow the ***that*** conjunction to be elided.) The excellent translation in (3)(ii) by Microsoft's Bing translator clearly suggests a significant degree of linguistic competence underlying both the SMT and NMT versions of this system.

Language being what it is, all systems are bound to stumble over unanticipated constructions of one kind or other. The question is not whether systems make mistakes but whether and to what extent these mistakes can be addressed, one after the other, endlessly, without these fixes causing new problems. The ability to absorb fixes without end seems to be the only legitimate measure of a system's potential. In short, if a translation model can absorb corrections and improvements literally without limit, and can do so without the complication of new logic fighting older logic, i.e., without the sort of complications that eventually arrest growth beyond a certain point, then it is fair to say that there are very few limits to what such a system can ultimately be made to accomplish. And theoretically, one day such a system would have the prospect of passing the classic Turing test as it might be applied to MT: machine output consistently indistinguishable from human translation.

After half a century of effort, there still seems to be no prospect of ever realizing this original ideal. In fact, one barely speaks of it any more. But need that be the case? We already see signs of potential breakthroughs in the recent merging of statistical and neural net technologies. The newly emerging neural MT technology (NMT) has already overcome some of the linguistic shortcomings of SMT, as we shall show in later chapters. But as we have already seen, problems still remain to be solved.

The present book seeks to further these positive developments by suggesting an entirely different kind of neural model, one with more of a psycholinguistic and neurolinguistic basis, the effectiveness of which we hope to demonstrate as we proceed. We will suggest that by combining the best elements of all these technologies, linguistic, statistical, and neural, a final breakthrough in the quality of MT output may yet become possible.

References

Halliday MAK (2003) In: Webster J (ed) On language and linguistics: collected works of M.A.K. Halliday. Continuum, London/New York

Kumaran D, Hassabis D, McClelland JL (2016) What learning systems do intelligent agents need? Complementary learning systems theory updated. Trends Cogn Sci 20(7):512. https://doi.org/10.1016/j.tics.2016.05.004. Accessed 12 Jan. 2017

Optiz B, Friederici AD (2003) Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. NeuroImage 19:1730–1737. https://doi.org/10.1016/S1053-8119(03)00170-8. Accessed 25 Jan. 2016

Pinker S (1994) The language instinct. William Morrow and Co., New York