

Sustaining Disruption? The Transition from Statistical to Neural Machine Translation

Dorothy Kenny



Dorothy Kenny
Dublin City University
dorothy.kenny@dcu.ie;
ORCID:
[0000-0002-4793-9256](https://orcid.org/0000-0002-4793-9256)



Abstract

If statistical machine translation (SMT) was a disruptive technology, then neural machine translation (NMT) is probably a sustaining technology, continuing on a trajectory already established by SMT, and initially evaluated in much the same way as its predecessor. Seeing NMT in this light may be a useful corrective to the hype that has surrounded its introduction.

Keywords: disruptive innovation; machine translation; statistical MT; neural MT; quality metrics; mobility.

Resum

Si la traducció automàtica estadística (TAE) va ser una tecnologia disruptiva, la traducció automàtica neuronal (TAN) probablement és una innovació incremental, que continua una trajectòria establerta per la TAE i que inicialment s'ha evaluat en gran part igual que la seva predecessora. Mirar la TAN des d'aquest punt de vista pot ser útil per matisar el bombo que envolta el seu sorgiment..

Paraules clau: Innovació disruptiva; traducció automàtica; TA estadística; TA neuronal; mètriques per a la qualitat; mobilitat.

Resumen

Si la traducción automática estadística (TAE) fue una tecnología disruptiva, la traducción automática neuronal (TAN) probablemente es una innovación sostenida, que sigue una trayectoria establecida por la TAE y que inicialmente se ha evaluado en gran parte igual que su predecessora. Mirar la TAN desde este punto de vista puede ser útil para matizar el bombo que rodea su nacimiento.

Palabras clave: Innovación disruptiva; traducción automática; TA estadística; TA neuronal; métricas para la calidad; movilidad.



1. Introduction

In 1997 Clayton M. Christensen published what was to become one of the most influential business books of the era: *The Innovator's Dilemma* is included in most major media lists of 'best business books' including those published by *Time Magazine* and *The Economist*, for example, and statistics provided by Kilkki et al. (2018) show how the terms used by Christensen gained widespread currency after the book's publication. In *The Innovator's Dilemma* Christensen set out to explain why "great" companies sometimes fail. Well-managed companies sometimes falter, he argued, not because other companies come along who can offer better products, but because they are displaced by new entrants who, in the short-term at least, offer products that result in worse performance than those of the incumbent. The explanation was originally based on the idea of 'disruptive technologies.' Disruptive technologies, according to Christensen (1997: xv) are ones that "generally underperform established products in mainstream markets. But they have other features that a few fringe (and generally new) customers value. Products based on disruptive technologies are typically cheaper, simpler, smaller, and, frequently, more convenient to use." In *The Innovator's Solution*, the sequel to his ground-breaking volume, Christensen and his co-author Michael Raynor replaced the term 'disruptive technology' with 'disruptive innovation' in a bid to stamp out incipient confusion with existing concepts (Christensen and Raynor, 2003: 66). The shift to the term 'disruptive innovation' is also an acknowledgement of the fact that technologies by themselves are not inherently disruptive. Rather it is the combination of business model and product, among other things, that can disrupt. The approach is thus generally consistent with that adopted in science and technology studies, which is beginning to make its mark in translation circles (Olohan, 2017), and which rejects the idea of autonomous technologies that are capable of independently determining social outcomes.

But even 'disruptive innovation' quickly became prone to "loose" usage, and disruption theory risks, according to its author, becoming a victim of its own success (Christensen, Raynor and McDonald, 2015: 46) with some commentators arguing that "the word "disruption" is now bandied about so much that it is losing all meaning" (*The Economist*, 2015). Worse still, a notion of 'disruption' that does not necessarily have much to do with Christensen's original conception has become fetishized in much commentary on the networked economy, leading to what Andrew Keen calls "a cult of disruption" (2015: 6). Industry definitions of 'digital disruption'—if proffered at all by users of the term—can also depart radically from Christensen's initial idea, not requiring, for example, that a disruptive innovation involve poorer product performance in the short term, or not allowing for some disruption to happen slowly. McQuivey (2014), for example, sees digital disruptors as building '**better** product experiences, that create **stronger** customer relationships, bringing it all to market **faster**' (emphasis in the original).

Notwithstanding the fact that contemporary understandings of 'disruption' have become diffuse, and even if Christensen's own approach is not without its detractors (see Kilkki et al., *ibid*: 276), there is much food for thought in his original conception

of disruption. Understanding the ‘forces’ of disruption, and being able to differentiate between disruptive and other kinds of innovation, can, according to Christensen, help managers to manage innovation or defend their companies in the face of certain challenges. Likewise, looking at changes in the translation market, and especially changes linked to the increasing use of machine translation (MT), through a Christensen-inspired disruption lens, might help providers of human and machine translation services alike to understand better the nature of the changes that accompany new types of MT and to find solutions that are appropriate to their own contexts. Of particular value are: the distinction between sustaining and disruptive innovations; the distinction between low-end and new-market disruption; the recognition that disruptive innovations are often accompanied by a shift in the main metric used to assess a product or service; and the classification of strategies to deal with disruption. These elements are dealt with in turn below, after a brief overview of how disruption proceeds in general.

Before moving on, however, it is worth mentioning some of the criticisms of Christensen’s work. Jill Lepore, one of his most strident critics, has accused Christensen of “hand-picking” examples, and claims that his sources are often “dubious” and his logic “questionable” (Lepore, 2014). Lepore also decries the whole rhetoric of disruption as “a language of panic, fear, asymmetry, and disorder” (*ibid.*), claiming that disruption is a theory borne of anxiety. It is this claim that is of particular interest in the current context. There is no doubt that the frenzied discourse of disruption subsequently encountered in some quarters could induce anxiety, especially among those who risk being ‘disrupted’, but I would argue that Christensen’s original approach to disruption is anything but frenzied. By comparison to the kinds of sources to which Lepore (*ibid.*) alludes, Christensen’s writing is positively staid, and provides a welcome antidote to anxiety-inducing hype.

In what follows, Christensen’s major concepts will be applied to the rise of rule-based, statistical and neural machine translation. Statistical machine translation (SMT) was state of the art in machine translation from the mid-2000s through to 2015, when it was displaced by neural machine translation (NMT) (see Bentivogli et al., 2016a, 2016b). The hype surrounding NMT in particular has since reached fever pitch, with industry sources claiming that NMT has already attained ‘parity’ with human translation (Hassan et al., 2018) and that it now has the potential to replace human translators (Shoshan, 2018a), and both industry and academic commentators writing about ‘disruption’ on a scale not seen before in translation circles (see Shoshan, 2018b), or that at least risks being underestimated (Way, 2018). A re-engagement with Christensen’s concept of disruption thus seems timely.

2. Disruptive Innovation

Christensen's (1997) model distinguishes between sustaining and disruptive technologies. The former "improve the performance of established products, along the dimensions of performance that mainstream customers in major markets have historically valued" (ibid. xv). In one of his best-known case studies, that of the disk drive industry, Christensen argues that technological innovations that resulted in improvements in total capacity and recording density (the latter measured in megabits per square inch) initially sustained the position of the leading manufacturers in the industry. When the upheaval came that toppled those leaders, it came in the form of disk drives that performed worse than the incumbents on these established metrics. The 5.25-inch drive, for example, was inferior to the 8-inch drive it eventually displaced, from the point of view of capacity, cost per megabyte and access time (ibid.,15), but it was small and lightweight and thus appealed to the emerging market for personal desktop computers.

But, as already noted, disruption is not just about new technologies, it's also about business models and how a company's customers respond to changing environments. In their 2015 reprise of disruption theory, Christensen, Raynor and McDonald summarize the process as follows:

"Disruption" describes a process whereby a smaller company with fewer resources is able to successfully challenge established incumbent businesses. Specifically, as incumbents focus on improving their products and services for their most demanding (and usually most profitable) customers, they exceed the needs of some segments and ignore the needs of others. Entrants that prove disruptive begin by successfully targeting those overlooked segments, gaining a foothold by delivering more-suitable functionality—frequently at a lower price. Incumbents, chasing higher profitability in more-demanding segments, tend not to respond vigorously. Entrants then move upmarket, delivering the performance that incumbents' mainstream customers require, while preserving the advantages that drove their early success. When mainstream customers start adopting the entrants' offerings in volume, disruption has occurred. (Christensen, Raynor and McDonald, 2015: 46)

Christensen's model also attempts to capture the different strategies that incumbents can use to respond to disruption. In short, they can choose (among other options) to migrate upward (and risk becoming uncompetitive with their existing customers) or downward, although he acknowledges that building a cogent case "for entering small, poorly defined low-end markets that offer only lower profitability" does not come easily to "[r]ational managers" (1997: 77).

3. Disruption and Machine Translation

An admittedly loose analogy can be made with the translation industry. For decades MT was not seen as a competitive threat by human translators because its linguistic quality could not compare with that of human translation. But when criteria such as 'ubiquity, mobility, connectivity, and immediacy' (Enríquez Raido, 2013) came to be increasingly valued in the networked economy, and MT could effectively be delivered for free online, it became clear that a particular group of consumers, who were perhaps not likely to have paid for a translation in the first place, could live with

sometimes unreliable linguistic quality. Performance metrics began to change. A key moment came when the Altavista search engine first made rule-based machine translation (RBMT) available to its users in 1997. Ten years later, another shift happened when Google shifted from RBMT to SMT. It is not self-evident that RBMT ‘disrupted’ human translation in Christensen’s use of the term: there was no great blood-letting in human translation because of it, and it can be argued that translation memory had a much greater impact than RBMT on human translation in the 1990s. But in both the introduction of online RBMT and the subsequent rise of SMT, there were elements of Christensen-style disruptions. Free online RBMT competed against ‘non-consumption.’ In other words, it did not displace incumbents (e.g. human translators with their translation memory tools), but rather helped create a new market. The ‘low-end’ market thus created was subsequently taken over by SMT. In the latter case, the new entrant technology was initially considerably inferior to the incumbent (RBMT) along well-established performance dimensions. When SMT was first presented as an alternative to RBMT by pioneers at IBM, for example, they were not even able to provide “actual results of French/English translation” (Brown et al., 1988: 1), and early models were extremely naïve (see Koehn, 2010). But the upstart technology eventually disrupted the incumbent along all of the dimensions recognised by Kilkki et al. (2018). In particular when ‘mainstream’ customers started adopting SMT, disruption can be said to have occurred within Christensen’s framework. (A full account of the recent history of machine translation is beyond the scope of this article. Interested readers are referred to brief treatments in Kenny (2018) and Poibeau (2017).)

In building his arguments, Christensen plots mainstream performance metrics against time to show the trajectories of sustaining and disruptive technologies. He also inserts dotted lines to represent the average performance of incumbent technologies in different markets or market segments. This average performance then serves as a proxy measure for the performance demanded in that market. As we have already seen, one of Christensen’s main arguments is that leading companies are often so focused on serving their high-end customers’ needs, that over time they actually exceed those needs, while simultaneously ignoring the needs of less demanding customers.

If Christensen’s model was transferred to the translation industry, and the ‘technologies’ in question were classified as human translation in combination with computer-aided translation (HT/CAT), RBMT, SMT and NMT, and dotted lines were further added to depict performance demanded in the different ‘segments’ of the market, the graph might look something like Figure 1.

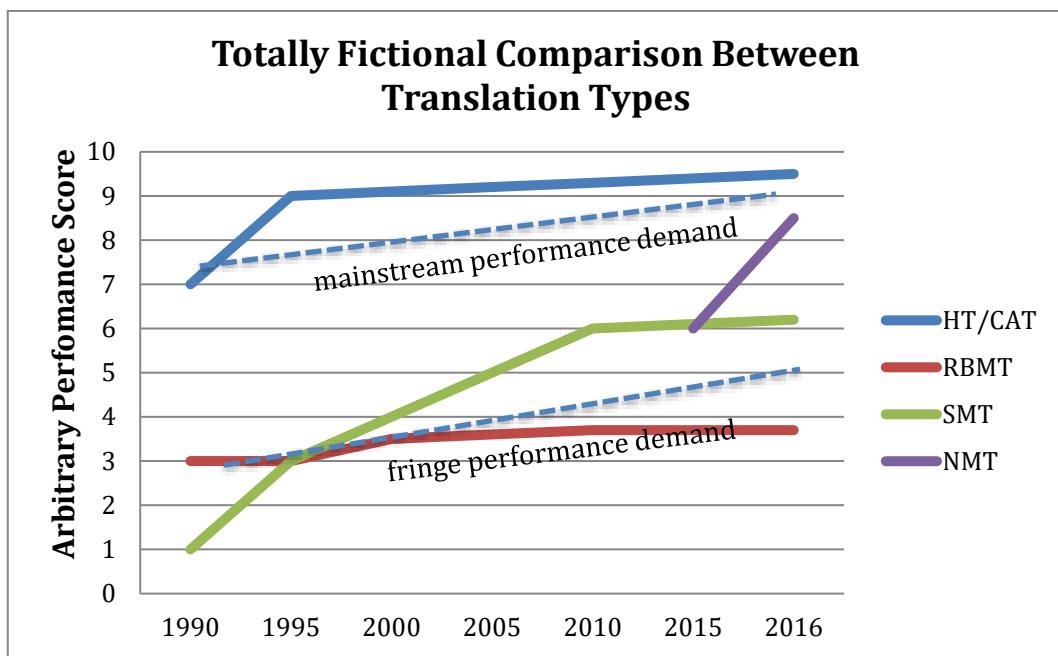


Figure 1: Fictional depiction of trajectories of performance demanded vs performance supplied by different translation technologies (inspired by Christensen 1997)

That Figure 1 is not to be taken too seriously is, it is hoped, indicated by the use of descriptors such as ‘fictional’ in its title and ‘arbitrary’ on the y-axis. It should be considered more as a thought experiment than as a record of what has actually happened in translation markets since the 1990s. For the graph to be in any way realistic we would need, first, to have full information on every translation market and market segment, and comparable performance metrics for different approaches to translation, classified according to the use of discrete technologies. It is doubtful that anyone has such full information, or that we can divide translation workflows into such neat non-overlapping categories in the first place. And even the existence of different market segments is not something that commentators can agree on, as we shall see below. (The translation industry is much harder to graph than the hard disk industry.) We would also need to agree on the precise metric to be depicted on the y-axis. This is easier said than done. Should we prioritize speed, cost or quality, for example? Or do we need a graph for each of these criteria? If we are most interested in quality of target texts produced using different technologies, then which metric would best capture this?

4 Quality Metrics

Translation Quality Assessment (TQA) is a vast area that has been studied for decades. There are multiple ways to evaluate both human and machine translation, discussion of which goes beyond the scope of this article (for an overview of approaches, see Castilho et al., 2018), and metrics tend to differ between academia, commercial production environments, and MT research and development laboratories. In production environments, for example, flexible frameworks like MQM-DQF (Lommel,

Uszkoreit and Burchardt, 2014; qt21, 2015) are often preferred and are used in the evaluation of both human and machine translation, while in MT labs, reference-based automatic metrics—of which there is a multitude—have tended to prevail in the evaluation of machine output. (González, Giménez and Girona Salgado (2014), for example, list some 60 automatic evaluation metrics, or variations thereof, based on similarity with a reference translation).

Having said that—and focusing solely on the MT research and development community—if one metric stands out as having risen to predominance in the first part of the current millennium, it has to be bleu (Papineni et al., 2002). Bleu was introduced as SMT gained ground, and computer scientists needed quick, inexpensive, language-independent methods to evaluate the effects of iterative, sometimes daily, changes to their own MT systems, and to compare multiple systems against each other. The requirement was for an automatic metric that would correlate highly with human evaluation and incur little marginal cost per run, and bleu, an n-gram precision-based metric, was born. Despite known problems with bleu (see, e.g., Way, 2018: 168), it remains “by some distance the most reported metric in papers involving MT experiments” (*ibid.*), and has been described as the “de facto standard for most research purposes” (Castilho et al., 2018: 26) and the “primary” metric used to rank competing MT systems in shared task evaluations (Bentivogli et al., 2016a: 16). Most notably, when NMT broke through in shared task evaluations in 2015 after a period in which it was “too computationally costly and resource demanding to compete with state-of-the-art Phrase-Based [Statistical] MT” (Bentivogli et al., 2016b: 1), its success was heralded in terms of a bleu score, one that was better than that of the incumbent technology by “a large margin”—+5.3 bleu points, to be precise—thus anticipating, according to Bentivogli et al. (*ibid.*) “what, most likely, will be the new NMT era.” (The bleu score in question, incidentally, was reported for English-German, which was known to be “a difficult language pair” (*ibid.*)).

So NMT took on the mantle of SMT without a shift in metrics. The breakthrough might indeed be era-defining, but from Christensen’s point of view, NMT is probably a sustaining rather than a disruptive technology. Its entry in 2015 on the fictional graph in Figure 1 is also thus depicted using a sustaining rather than a disruptive trajectory. Offering an incremental improvement on SMT—which had already begun to plateau—from the outset, it was initially judged on the same terms as SMT, as already noted. It has already begun to take over from SMT, in a neat demonstration of classic intersecting technology S-curves (Foster 1986; Christensen 1997), and in the way of all sustaining technologies.

This is not to say that the rise of NMT will not be followed eventually by a change in the preferred metric for ranking systems within the academic MT R&D community, or that other metrics, including human evaluation metrics, have not also been applied to NMT. MT specialists are already clamouring for bleu to be abandoned (Way, 2018), partly because it cannot do justice to NMT, and several papers have already been published that apply other, especially human, evaluation metrics to NMT (see, for example, Bentivogli et al. 2016b; Castilho et al., 2017). The point being made is merely

that NMT was first crowned the new state-of-the-art in MT using bleu scores as evidence. The attendant hype was such that industry source slator.com saw fit to explain bleu to its readers, who were no doubt more used to evaluating translation using metrics more commonly used in production environments (Pan, 2016).

5 The Translator's Dilemma

As in other areas, if there has, indeed, been disruption in the translation market (or markets), and whether or not it has been disruption in the Christensen mould, then the incumbents are faced with a decision: should they hold steady, or attempt to migrate upwards or downwards? And there is no shortage of advice to hand: speaking against a background of widespread automation anxiety, Google's Hal Varian (quoted in Brynjolfsson and McAfee, 2014: 200) counsels human workers in general to "seek to be an indispensable complement to something that's getting cheap and plentiful". In the context of machine translation, it is easiest to interpret such advice as meaning that translators should become post-editors, and Anthony Pym (2013: 488) has confidently predicted that "statistical-based MT, along with its many hybrids, is destined to turn most translators into posteditors one day, perhaps soon". Indeed, post-editing MT was recognised as one of the fastest growing segments of language industry even before NMT had gained widespread use (see, for example, Common Sense Advisory, 2016). But given concerns among post-editors about remuneration and boredom in particular (see Moorkens and O'Brien 2017), such a move risks being seen as a downward migration, and unappealing to many 'rational' incumbents. Meanwhile a strategy of upward migration, or holding steady if one is already serving top-end clients, is promoted by translators' organisations such as the Institute of Translation and Interpreting (ITI) in the United Kingdom and authors such as Moorkens (2017). This position has come under fire from some quarters however, including from Pym (2016), who appears to denigrate the ambitions of those who seek to serve high-end customers, and some translators have found themselves obliged to attest to the very existence of 'high-end' or 'premium' markets (see, for example, Sardon, 2017). A third way that combines upward migration with the use of machine translation is perhaps discernible in ideas about 'augmented translation' as envisaged, for example, by Lommel (2017). Here the human translator is presented not at the end of the translation chain, fixing errors in MT output, but rather at the centre of translation activity, drawing on key technologies—including adaptive NMT—that amplify his/her abilities and speed up the process of translation.

6 Conclusions

This article has argued that while it is possible to see SMT as a disruptive technology, as described by Christensen (1997), the more recently popular NMT, although greeted as a revolutionary achievement of 'electronic brains' by vendors, bloggers and the press alike, might be better viewed as a sustaining technology. Offering an incremental improvement on SMT from the outset, it was initially judged on the same terms as

SMT, and maintained a trajectory already established by SMT. This view is offered as a corrective to the hype that has surrounded the recent widespread adoption of NMT, and that surrounds 'digital disruption' more generally. Rather than expect the MT revolution that many have heralded, we might prepare ourselves to deal with many of the same challenges as attended to SMT (see Kenny and Doherty 2014). These include the need to source and profile (even greater quantities of) training data, to find suitable points at which translators can intervene in or control (à la Lommel, 2017) the translation process, and to educate new generations of translators/post-editors who are capable of working with NMT in a sustainable way. A Christensen-like lens might also help us to see the adoption of high-end or upwardly mobile strategies as representing a rational response to disruption, rather than deriding them as "quaint boutique" models (Pym, 2016). Of course we can expect more change: MT quality should improve as more systems shift to NMT; bleu might be displaced as the top automatic evaluation metric used in research (Shoshan, 2018a; Way, 2018); and augmented translation may catch on. But perhaps considering what kind of change is happening will help us face that change calmly and strategically, in a way that Christensen would no doubt approve of.

Bibliography

Bentivogli, L.; Bisazza, A.; Cettolo, M. & Federico, M. (2016b). "Neural versus Phrase-Based Machine Translation Quality: a Case Study". In: *EMNLP 2016*. <arXiv:1608.04631v1 [cs.CL]>. [Last accessed on April 30, 2018].

Brown, P. [et al.]. (1988). "A statistical approach to French/English translation". In: *Proceedings of the Second International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages (TMI 1988), June 12-14. Pittsburgh, PA*. <<http://www.mt-archive.info/TMI-1988-Brown.pdf>>. [Last accessed on April 30, 2018].

Brynjolfsson, E. & McAfee, A. (2014). *The Second Machine Age: work, progress and prosperity in a time of brilliant technologies*. 1st ed. New York: W. W. Norton & Company.

Castilho, S. [et al.]. (2017). "Is Neural Machine Translation the New State of the Art?". *The Prague Bulletin of Mathematical Linguistics*, v. 108, pp. 109-120. <<https://doi.org/10.1515/pralin-2017-0013>>. [Last accessed on April 30, 2018].

Castilho, S.; Doherty, S.; Gaspari, F. and Moorkens, J. (2018). "Approaches to Human and Machine Translation Quality Assessment". In: Moorkens, J.; Castilho, S.; Gaspari, F. & Doherty, S. (eds). *Translation Quality Assessment: From Principles to Practice*. Cham: Springer International Publishing. (Machine Translation: Technologies and applications; 1), pp. 9-38. <https://doi.org/10.1007/978-3-319-91241-7_2>. [Last accessed on August 30, 2018].

Christensen, C. M. (1997). *The innovator's dilemma: when new technologies cause great firms to fail*. Boston, MA: Harvard Business School.

Christensen, C. M. & Raynor, M. (2003). *The innovator's solution*. Boston, MA: Harvard Business School.

Christensen, C. M.; Raynor, M. & McDonald, R. (2015). "What is Disruptive Innovation?". *Harvard Business Review*. (01 December 2015), pp. 44-53.
<<https://hbr.org/2015/12/what-is-disruptive-innovation>>. [Last accessed on August 30, 2018].

Common Sense Advisory. (2016). *Global Market Research Firm Common Sense Advisory Finds Post-edited Machine Translation (PEMT) Among Fastest-growing Segments of the Language Industry*. Last updated: 08.02.2016
<<https://www.commonsenseadvisory.com/default.aspx?Contenttype=ArticleDet&tabID=64&moduleID=392&Aid=36546&PR=PR>>. [Last accessed on April 30, 2018].

The Economist. (2011). Business Books: Aiming High.
<<https://www.economist.com/node/18894875>>. Last updated: 30.06.11. Page consulted on: 30.04.18.

The Economist. (2015). Disrupting Mr Disrupter. <<https://www-economist-com.dcu.idm.oclc.org/news/business/21679179-clay-christensen-should-not-be-given-last-word-disruptive-innovation-disrupting-mr>>. Last updated: 25.11.15. Page consulted on: 30.04.18.

Enríquez Raído, V. (2013). "Teaching Translation Technologies "Everyware": Towards a Self-Discovery and Lifelong Learning Approach". *Revista Tradumàtica: technologies de la traducció*, n. 11, pp. 275-285. <<https://doi.org/10.5565/rev/tradumatica.52>>. [Last accessed on August 30, 2018].

Foster, R. N. (1986). *Innovation: The Attacker's Advantage*. New York: Summit Books.

González, M.; Giménez, J. & Girona Salgado, J. (2014). *Asiya: An Open Toolkit for Automatic Machine Translation (Meta-) Evaluation: Technical Manual Version 3.0*. Last updated February 2014. <http://asiya.lsi.upc.edu/Asiya_technical_manual_v3.0.pdf>. [Last accessed on July 12, 2018].

Hassan, H., [et al.]. (2018). *Achieving Human Parity on Automatic Chinese to English News Translation*. Cornell University Library. Last updated: 15.03.18.
<[arXiv:1803.05567v1 \[cs.CL\]](https://arxiv.org/abs/1803.05567v1)>. [Last accessed on May 30, 2018].

Keen, A. (2015). *The Internet is not the Answer*. 1st ed. New York: Atlantic Monthly Press.

Kenny, D. & Doherty, S. (2014). "Statistical Machine Translation in the Translation Curriculum: overcoming obstacles and empowering translators". *The Interpreter and Translator Trainer*, v. 8, n. 2, pp. 276-294.
<<https://doi.org/10.1080/1750399X.2014.936112>>. [Last accessed on July 12, 2018].

Kenny, D. (2018). "Machine Translation". In: Rawling, P. & and Wilson, P. (eds.). *The Routledge Handbook of Translation and Philosophy*. Milton Park, Abingdon, Oxon:

Routledge. (Routledge Handbooks in Translation and Interpreting Series), pp. 428-445.

Kilkki, K.; Mäntylä, M.; Karhu, K.; Hämmäinen, H. & Ailisto, H. (2018). "A disruption framework". *Technological Forecasting and Social Change*, v. 129 (April), pp. 275-284. <<https://doi.org/10.1016/j.techfore.2017.09.034>>. [Last accessed on July 12, 2018].

Koehn, P. (2010). *Statistical Machine Translation*. Cambridge, New York: Cambridge University Press.

Lepore, J. (2014). "The Disruption Machine: What the gospel of innovation gets wrong". *The New Yorker*, (June 23). <<https://www.newyorker.com/magazine/2014/06/23/the-disruption-machine>>. [Last accessed on May 30, 2018].

Lommel, A. (2017). "Augmented Translation Puts Translators Back in the Center" [Post], 19 May 2017. *The IILOG. How to Create and Deliver Intelligent Information*. <<https://intelligent-information.blog/en/blog/2017/05/19/augmented-translation-puts-translators-back-in-the-center/>>. [Last accessed on May 30, 2018].

Lommel A., Uszkoreit, H. & Burchardt, A. (2014). "Multidimensional Quality Metrics (MQM): a framework for declaring and describing translation quality metrics". *Revista Tradumàtica: technologies de la traducció*, n. 12, pp. 455-463. <<https://doi.org/10.5565/rev/tradumatica.77>>. [Last accessed on May 30, 2018].

McQuivey, J. L. (2014). *Digital Disruption. Unleashing the Next Wave of Innovation*. <<http://www.softtek.com/webdocs/others/application-innovation-2014/Digital-Disruption-McQuivey.pdf>>. [Last accessed on May 30, 2018].

Moorkens, J. (2017). "Under pressure: translation in times of austerity". *Perspectives: Studies in Translation Theory and Practice*, v. 25, n. 7, pp. 464-477. <<https://doi.org/10.1080/0907676X.2017.1285331>>. [Last accessed on May 30, 2018].

Moorkens, J. & O'Brien, S. (2017). "Assessing user interface needs of post-editors of machine translation". In: Kenny, D. (ed.) *Human Issues in Translation Technology*. London; New York: Routledge, pp. 109-130.

Moorkens, J., S. Castilho, F. Gaspari and S. Doherty (eds). (2018). *Translation Quality Assessment: From Principles to Practice*. Berlin: Springer. <<https://doi.org/10.1007/978-3-319-91241-7>>. [Last accessed on May 30, 2018].

Olohan, M. (2017). "Technology, translation and society: A constructivist, critical theory approach". *Target online*, v. 29, n. 2, pp. 264-283. <<https://doi.org/10.1075/target.29.2.04olo>>. [Last accessed on May 30, 2018].

Pan, H. M. (2016). *How BLEU Measures Translation and Why It Matters*. Slator, last updated: 26.11.16. <<https://slator.com/technology/how-bleu-measures-translation-and-why-it-matters/>>. [Last accessed on July 12, 2018].

Papineni, K.; Roukos, S.; Ward, T. & Zhu, W. J. (2002). "BLEU: A Method for Automatic Evaluation of Machine Translation". *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL): Philadelphia, Pennsylvania: July 07-12, 2002*. Stroudsburg, PA: ACL, pp. 311-318.
<<https://dl.acm.org/citation.cfm?id=1073083&picked=prox>>. [Last accessed on July 12, 2018].

Poibeau, T. (2017). *Machine Translation*. Cambridge, MA; London: The MIT Press. (The MIT Press essential knowledge series).

Pym, A. (2013). "Translation Skill-Sets in a Machine-Translation Age". *Meta: Journal des traducteurs = Translators' Journal*, v. 58, n. 3 (December), pp. 487-503.
<<https://doi.org/10.7202/1025047ar>>. [Last accessed on July 12, 2018].

Pym, A. (2016). *Getting it right, forever? Deconstructing a professional discourse on the role of translators*. <http://usuaris.tinet.cat/apym/online/training/2016_getting_it_right_2.0.pdf>. [Last accessed on April 30, 2018].

QT-21. (2015). *Multidimensional Quality Metrics (MQM): Definition*. QT-21, last updated: 30.12.15. <<http://www.qt21.eu/mqm-definition/definition-2015-12-30.html#dqf-mapping>>. Last accessed on July 12, 2018].

Sardon, V. (2017). *Redefining the premium market*. SIB Translations. Last updated: 14.06.17<<http://www.sjbtranslations.com/redefining-the-premium-market/>>. [Last accessed on April 30, 2018].

Shoshan, O. (2018a). *Make Neural Machine Translation Better, Faster – A New Way to Measure NMT Quality*. Slator, last updated: 19.04.18. <<https://slator.com/sponsored-content/make-neural-machine-translation-better-faster-a-new-way-to-measure-nmt-quality/>>. [Last accessed on May 30, 2018].

Shoshan, O. (2018b). *Neural MT and Blockchain Are About to Radically Transform the Translation Market*. Slator, last updated: 21.05.18. <<https://slator.com/sponsored-content/neural-mt-and-blockchain-are-about-to-radically-transform-the-translation-market/>>. [Last accessed on May 30, 2018].

Way, A. (2018). "Quality expectations of machine translation". In: Moorkens, J.; Castilho; S.; Gaspari, F. & Doherty, S. (eds.) *Translation Quality Assessment: From Principles to Practice*. Berlin: Springer. (MATRA; 1), pp. 159-178.
<https://doi.org/10.1007/978-3-319-91241-7_8>. [Last accessed on May 30, 2018].