Neural machine translation and language teaching – possible implications for the CEFR

Alice DELORME BENITES & Caroline LEHR
Zürcher Hochschule für Angewandte Wissenschaften
Institut für Übersetzen und Dolmetschen
Theaterstrasse 15C, CH-8401 Winterthur
delr@zhaw.ch, lehc@zhaw.ch

L'arrivée de la traduction automatique neuronale (TAN), dont la qualité surpasse largement celle des systèmes précédents, a bouleversé le champ de la traduction professionnelle, notamment commerciale. Par ailleurs, l'accès gratuit et instantané à des systèmes de TAN performants (DeepL, Google Traduction) introduit une nouvelle inconnue pour l'enseignement-apprentissage des langues, allant pour certain·e·s jusqu'à remettre en question le besoin d'apprendre une langue étrangère. Les enseignant·e·s doivent s'adapter à ce changement et à ses conséquences. Or, les objectifs d'enseignement, surtout en Europe, reposent désormais majoritairement sur les standards établis par le Cadre européen commun de référence pour les langues (CECRL) publié il y a vingt ans par le Conseil de l'Europe et sur ses mises à jour publiées plus récemment, notamment le Volume complémentaire du CECRL (2018, 2020). Le présent article se penche sur l'impact de la TAN sur l'applicabilité des descripteurs du CECRL pour l'enseignement-apprentissage des langues aujourd'hui. En effet, les nouvelles technologies invitent à une relecture critique des compétences définies pour les différents niveaux du CECRL.

Mots-clés:
traduction automatique neuronale, enseignement des langues, apprentissage des langues, CECRL, niveaux de langue, évaluation.

Keywords:
neural machine translation, language teaching, language learning, CEFR, proficiency levels, language assessment.

1. Introduction

With the advent of neural machine translation (NMT), the quality of machine translation (MT) output has made astounding progress in recent years. Freely available online translation engines, such as Google Translate and DeepL, have become widely-used tools now able to facilitate reading, writing and mediation across languages in various contexts. In view of this progress, some people have even started to call into question the necessity of learning foreign languages at all. In the media, for example, technological developments in NMT are frequently discussed, including their impact with respect to language learning. Seemingly provocative and illusory headlines, such as "Don't bother learning a foreign language!" (O'Callaghan 2014) have multiplied in this context and sometimes reveal an undifferentiated view of language learning and ignorance of the complexity of multilingual communication. It therefore seems important to (re-)consider language learning and teaching in light of NMT.

New questions have to be addressed in language teaching in order to adapt to rapid technological developments because the challenges posed by this

progress are multifaceted. Language teachers have reported that the performance of machine translation has a negative impact on students' motivation to learn foreign languages (Zhu 2020). Moreover, students have been found to employ free neural machine translation solutions in language courses to produce texts (O'Neill 2019), making it difficult to evaluate their progress with conventional assessment methods. Conversely, even if MT tools are familiar to teachers, they rarely explicitly use them in language teaching, which leaves the didactic potential of these tools widely unexploited. Sometimes teachers are reluctant to adopt new technologies that they regard as disruptive or even a threat to language teaching. Furthermore, we may find that machine-translated output now fits the purpose for certain contexts of communication, which calls for a redefinition of the goals of language teaching and the competences we are aiming for in different contexts. These problems demonstrate the importance of rethinking and adapting didactic methods and tools when faced with recent technological developments.

In this article, we set out to contribute to this objective by focussing on the widely employed Common European Framework of Reference for Languages and its Companion volume (CEFR, Council of Europe 2001, 2018, 2020[1]). Providing a systematic approach to the definition of language proficiency levels, the framework serves as a reference for language teaching in Europe and is also referred to in professional contexts in which language competence needs to be assessed. Not only is the CEFR therefore a central tool, it may also guide further reflection on the development of language teaching in light of NMT and its capacities.

2. The relevance of the CEFR for language teaching and didactics

The CEFR is a widely-used framework with which countries, institutions and language teaching professionals can efficiently set goals, activities and milestones for foreign language learning. It was first published in 2001 by the Council of Europe in 40 European and non-European languages and has been updated on a regular basis since then. The framework describes various levels, ranging from A1 (which can be considered beginner level) to C2 (considered fluent). These levels apply to different categories of communicative competence (linguistic, pragmatic, sociolinguistic), to four language communicative activities (reception, production, mediation, interaction) and to four sub-competences (reading, writing, listening, speaking). The fine-grained differentiation allows a very precise yet simple implementation of these categories even in job advertisements and applications. As a result, many European countries have

---

[1]   In the following, we refer to the 2001 publication of the CEFR and the 2018/2020 versions of the Companion volume with new descriptors as "CEFR" and only distinguish between them when relevant.

been drawing their guidelines on language teaching from the CEFR (Pym et al. 2013), fostering its standard-setting status.

However, the CEFR "does not represent a revolution but is part of an evolution of practice" (Council of Europe website[2]). It was first thought of in 1971, when Europe was growing together and its high linguistic diversity threatened to slow down European integration. It was not until 1991, though, that the project was started – and it took another 10 years to officially launch the CEFR in 2001. This is why a brief overview of the evolution of modern language didactics in Europe should help contextualise the creation of the CEFR.

In Europe, modern language teaching has been subject to various streams and tendencies over the last decades. These trends have encompassed not only different understandings of the learning process and the learners themselves, but also varying conceptions of the relationship between language learning and translation. One of the oldest approaches, still applied by some, is the Grammar-Translation-Method. It was developed in Europe in the 19th century and mainly comprised applying methods of ancient language studies to modern language classes (Neuner & Hunfeld 1993). This approach was primarily chosen for students in secondary schools, who were considered "cognitively" strong and could be confronted with metalinguistic explanations directly from the start. As the name clearly states, the Grammar-Translation-Method focussed heavily on theoretical grammar teaching and on translation as a more practical activity in which grammatical rules could be applied. In this context, translation fulfilled several purposes:

- Translation of single sentences into the target language to train specific grammatical constructions in the target language (see examples in (Neuner & Hunfeld 1993: 23);

- Translations of continuous texts into the target language to train various grammatical constructions at once;

- Translations of literary texts from the target language to train text reception, vocabulary and style.

This approach has been strongly criticised since the end of the 19th century, amongst other reasons, because of its heavy focus on grammar as a fixed set of rules. Viëtor (1886: 6) argues that grammar is rather what society makes of it and that grammatical correctness changes over time. This statement introduces the notion of use and society in the considerations. Especially Viëtor's assertion that "what is true today will be false tomorrow" (ibid.) paves the way for diachronic variation and more communication-oriented teaching. As a result, the Direct Method (Berlitz 1887) and the Audiolingual Method were developed in the

---

[2]   Council of Europe. *Historical overview of the development of the CEFR.* https://www.coe.int/en/web/common-european-framework-reference-languages/history (13/07/2021).

USA. Around the same period as the Audiolingual Method, the Audiovisual Method was developed in Europe in the second half of the 20th century. All three methods focussed strongly on the spoken word and on a strict monolingual learning environment in which students are immersed in the target language. While the Grammar-Translation-Method was probably the approach that included the most translation tasks in the classroom, translation activities were abandoned with subsequent methods and imitation tasks, as well as so-called "drills", became central. Interestingly, the Audiolingual Method was actually born out of the dire need to quickly train North American interpreters from and into Asian languages for military purposes in the context of WWII (Neuner & Hunfeld 1993). This shows how complex and changing the relationship between language teaching and translation/interpreting is. While translating written texts had been considered a good way to learn a foreign language, later approaches recommended excluding the first language from the teaching situation, even to train future interpreters.

In the 1970s, foreign language didactics took a major turn with the introduction of the Communicative Approach (Savignon 2007). While communication purposes had already been central to previous methods, the notions of situativity (Roche 2009) and authenticity now became paramount. It was in this context that the four core competences – as they are still listed in the CEFR – were first introduced into language teaching: listening, speaking, reading and writing. The pragmatic orientation of this approach, however, does not include multilingual communication contexts – and does not foresee translation, either as a purpose or as a learning task.

In the late 1980s, the pragmatic-functional basis of the Communicative Approach finally gained an intercultural dimension (Rösler 2012: 81). To take the reality of the learner's cultural and linguistic background into account more, intercultural foreign language teaching should encourage students to compare their own world with the foreign reality. They should reflect on what makes a foreign culture "foreign" in all aspects: interesting, fascinating, attractive but also weird, unfamiliar and even threatening (Neuner & Hunfeld 1993: 109). This approach tends to focus more on the reception of texts and the perception of their cultural specificities, but translation is not explicitly part of the intercultural language teaching concept.

The above very brief historical overview shows how translation (including interpreting) has shifted from being a primary didactic activity to being almost systematically excluded from didactic considerations in language teaching. The act of switching back and forth between languages no longer seemed to be part of language competence or its acquisition for a long time. Correspondingly, the CEFR, which emerged from a communicative perspective, strongly focussed on the language activities to be carried out in a foreign language. Still, over the years, the four core competences (reading, writing, listening, speaking) have

been broadened and redistributed into four core activities (Council of Europe 2020: 231): reception, production, interaction and mediation. In the 2018 and 2020 updated versions of the CEFR, these activities are divided into two newly defined domains of competence: communicative language competence (linguistic, pragmatic, sociolinguistic) and plurilingual and pluricultural competence (pluricultural, plurilingual comprehension and repertoire). These changes are significant: since mediation activities include translation and interpreting tasks, the use of the learner's first language is now officially taken into consideration in the CEFR. One could argue that this reconciles some elements of the Grammar-Translation Method with newer aspects of communicative didactics, especially the intercultural approach. This is consistent with recently increasing calls to reinstate translation as a part of the language learning and teaching process (Cook 2010; Fountain & Fountain 2009). As the CEFR seems to closely follow the evolution of tendencies in language didactics, it comes as no surprise that the latest versions of the companion guide to the CEFR (Council of Europe 2020) now includes new descriptors.

12In the newer versions of the CEFR Companion volume (Council of Europe 2018, 2020), the intercultural dimension plays an important role. The new descriptors, however, separate mediating activities such as translation from reception and production activities – which are generally the main purpose of language teaching – thus indirectly excluding translation as a text-producing learning activity. Many justifications for and against including translation as a learning task or purpose could be presented, but that is not the point of this article. In fact, such considerations would risk limiting the discussion to human translation – which is currently being challenged by the disruptive progress of NMT. This article therefore sets out to tackle the new emerging relation between translation and language learning, putting the role of machine translation into focus. While most of the aforementioned didactic developments have been included in the ongoing reflection that fosters the creation and regular adaptation of the CEFR, the question of high-quality machine translation is still absent – although the routine of many experienced language teachers is already being jeopardised by NMT systems. In other words, which *reception*, *production* and even *mediation* activities described in the CEFR are likely to be changed or even made obsolete by the rise of NMT?

3. Recent developments in neural machine translation

Machine translation has existed for some time but it is only in the last few years that its quality has reached a level that has made it a widely-used tool. Recently, a surge in interest in NMT has been witnessed, displacing its predecessor statistical machine translation (SMT). Both SMT and NMT are corpus-based or corpus-driven forms of machine translation trained on huge corpora of

sentences which exist both in a source and a target language. Ideally, the translations they are built with were produced by professional translators, as the output can only be as good as the training data. Despite this similarity, the computational approach underlying NMT is a different one.

As its name implies, NMT is based on so-called neural networks which are composed of thousands of artificial units. These units are words or sub-word units, such as character sequences, and resemble neurons in that they represent a certain type of knowledge in a network. When a source text sentence is read by the MT engine, their activation depends on the stimulus the engine receives and the strength of connections within the network. When activated, these units build distributed representations of words and their contexts, so that outputs computed from them in the target language are as close as possible to the corresponding reference translation in the training corpus. Moreover, for each position of the target sentence being built by the MT engine, and for every possible word in the target vocabulary, the system provides the likelihood that the word is a continuation of what has already been produced. The system then builds the best translation by picking the most likely word at each position (Forcada 2017).

In fact, neural networks in artificial intelligence (AI) terms seem to have more in common with a text completion device than a human brain. This becomes obvious whenever we read machine-translated texts that failed to convey the right meaning for a polysemic word because, contrary to humans, NMT systems are unable to take semantic and contextual information into account. Nevertheless, recent developments in language-related AI have allowed for considerable improvements in the quality of machine-translated output. The remarkable increase in quality is illustrated by the growing significance that NMT-based tools now have for the work of language professionals (European Union Association of Translation Companies 2019). In addition, ready access to freely available online tools, such as Google Translate or DeepL, has led to new practices in multilingual communication among the wider public. Even though the use of voice translators is increasing, MT tools are much more widely employed for written texts. In this article, we therefore focus on the use of language-related AI for written texts and the written aspects of language reception, production and mediation. Although some language pairs produce better quality than others, NMT is used today in a wide variety of contexts and for a range of language combinations. As a consequence, it also affects language learning and teaching as well as language assessment.

4. The CEFR descriptors in light of NMT

Although such efforts still seem to be at a developmental stage, various attempts to include MT activities in language classrooms have been documented in recent years (Briggs 2018; Garcia & Pena 2011; Pym et al. 2013). Based on

Nino's (2008) and Garcia & Pena's (2011) observations, Yamada (2020: 185) claims that "MT can be useful as L2 writing support for university level language learners". Nevertheless, the goals pursued by language teachers, whether they include MT in their classes or not, generally follow the standards set by the CEFR, which currently ignores the role and use of MT. The addition of descriptors for translation and interpreting as mediation activities in its last version paved the way for the inclusion of language technologies, especially MT, in the framework. In their study on the role of translation in the teaching of languages in the European Union, Pym et al. (2013) defined translation as including "the reception and/or production and/or reworking of spoken or written bi-texts (paired discourses in two languages) within the classroom situation" (*ibid*.: 12), explicitly including the "identification of problems in machine-translation output, and their correction" (*ibid*.: 13). However, there is no mention of MT in the latest version of the CEFR. The tremendous progress made by NMT in terms of output quality in the last few years leads to many questions that reach beyond the scope of teaching methods and calls for a redefinition of language learning goals in an era in which multilingual communication is so frequently technologically-enhanced. In the following sections, we analyse key documents published within or alongside the CEFR to demonstrate how the current standards for language proficiency levels are called into question by the rise of MT. Many parts of the CEFR are not thought of as providing exhaustive criteria and goals, giving leeway to learners, teachers and assessors to include their respective priorities. Nevertheless, given its standard-setting status and influential guidelines, the content and formulation of core documents of the CEFR merit scrutiny when the possibilities for multilingual communication in society are undergoing potentially profound changes.

4.1 Global scale and self-assessment grid

The CEFR is a very large and detailed reference – but the daily work of many language teaching professionals relies on two condensed documents: the global scale (see Figure 1) and the self-assessment grid (Figure 2). Since these two tools are widely used for assessment purposes and also for curricular decision-making, we first examine their contents in light of the newest progress in NMT.

4.1.1 Global scale

PROFICIENT USER	C2	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
INDEPENDENT USER	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.
BASIC USER	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

Figure 1: Global scale of the CEFR (Council of Europe 2001: 24)

The description of every level, from A1 to C2, includes goals for two main communication activities: understanding (or reception) and interaction. In addition to oral interaction, writing activities are explicitly mentioned for some levels (B1 to C1), but there is no distinction made between reading and listening activities, both being summarised as "understanding". Today, in many situations, multilingual reading activities are performed with the help of MT. For example, Nurminen & Papula (2018) found that many users of MT systems do so for assimilation or dissemination purposes: to understand a document, to verify that they understand it correctly or to allow someone else to understand it. However, this implies that understanding a text in a foreign language is now facilitated by the use of MT tools and is possible for a wider range of people, as it no longer depends entirely on their personal language skills. The same is true for the production of written texts. Furthermore, voice translation systems, nowadays based on NMT technology, might play a significant role in oral communication activities – understanding and interacting – as well. For example, the system BabelDr (Hôpitaux Universitaires de Genève 2020[3]), now in use at the university hospital in Geneva, allows health professionals to communicate directly with foreign patients who cannot speak French. In such a context, the foreign language requirements for new and future staff could be

---

[3] Hôpitaux Universitaires de Genève. (06 2020). *BabelDr.* https://www.hug.ch/medecine-premier-recours/babel-dr (13.07.2021).

more difficult to map onto the CEFR global scale. As a result, it might be useful to resort to a more fine-grained evaluation tool, such as the self-assessment grid.

4.1.2 Self-assessment grid

The self-assessment grid is well-known and one of the most-frequently used documents produced within the context of the CEFR. It presents, in a very condensed form, which target activities a learner can carry out when he or she achieves level A1 to C2. Unlike the global scale, the activities are categorised in a way that encompasses the main sub-competences. This is why this document is often used as a basis for course design, assessment preparation as well as recruiting purposes when assessing the required language competences of applicants. The second and last rows of the grid are of special interest in our context, as they describe reading and writing competence – both activities where NMT can play a direct role. Nevertheless, as briefly shown in the example with BabelDr, oral interaction is very likely to be affected by NMT technology sooner or later as well.

		A1	A2	B1	B2	C1	C2
U N D E R S T A N D I N G	**Listening**	I can recognise familiar words and very basic phrases concerning myself, my family and immediate concrete surroundings when people speak slowly and clearly	I can understand phrases and the highest frequency vocabulary related to areas of most immediate personal relevance (e.g. very basic personal and family information, shopping, local area, employment). I can catch the main point in short, clear, simple messages and announcements.	I can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc. I can understand the main point of many radio or TV programmes on current affairs or topics of personal or professional interest when the delivery is relatively slow and clear.	I can understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar. I can understand most TV news and current affairs programmes. I can understand the majority of films in standard dialect.	I can understand extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly. I can understand television programmes and films without too much effort.	I have no difficulty in understanding any kind of spoken language, whether live or broadcast, even when delivered at fast native speed, provided I have some time to get familiar with the accent.
	Reading	I can understand familiar names, words and very simple sentences, for example on notices and posters or in catalogues.	I can read very short, simple texts. I can find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus and timetables and I can understand short simple personal letters.	I can understand texts that consist mainly of high frequency everyday or job-related language. I can understand the description of events, feelings and wishes in personal letters.	I can read articles and reports concerned with contemporary problems in which the writers adopt particular attitudes or viewpoints. I can understand contemporary literary prose.	I can understand long and complex factual and literary texts, appreciating distinctions of style. I can understand specialised articles and longer technical instructions, even when they do not relate to my field.	I can read with ease virtually all forms of the written language, including abstract, structurally or linguistically complex texts such as manuals, specialised articles and literary works.
S P E A K I N G	**Spoken Interaction**	I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.	I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.	I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).	I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.	I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers.	I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.
	Spoken Production	I can use simple phrases and sentences to describe where I live and people I know.	I can use a series of phrases and sentences to describe in simple terms my family and other people, living conditions, my educational background and my present or most recent job.	I can connect phrases in a simple way in order to describe experiences and events, my dreams, hopes and ambitions. I can briefly give reasons and explanations for opinions and plans. I can narrate a story or relate the plot of a book or film and describe my reactions.	I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.	I can present a clear, smoothly-flowing description or argument in a style appropriate to the context and with an effective logical structure which helps the recipient to notice and remember significant points.
W R I T I N G	**Writing**	I can write a short, simple postcard, for example sending holiday greetings. I can fill in forms with personal details, for example entering my name, nationality and address on a hotel registration form.	I can write short, simple notes and messages relating to matters in areas of immediate needs. I can write a very simple personal letter, for example thanking someone for something.	I can write simple connected text on topics which are familiar or of personal interest. I can write personal letters describing experiences and impressions.	I can write clear, detailed text on a wide range of subjects related to my interests. I can write an essay or report, passing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.	I can express myself in clear, well-structured text, expressing points of view at some length. I can write about complex subjects in a letter, an essay or a report, underlining what I consider to be the salient issues. I can select style appropriate to the reader in mind.	I can write clear, smoothly-flowing text in an appropriate style. I can write complex letters, reports or articles which present a case with an effective logical structure which helps the recipient to notice and remember significant points. I can write summaries and reviews of professional or literary works.

Figure 2: Self-assessment grid of the CEFR (Council of Europe 2001: 26-27)

The highest level in reading (C2) includes all kind of written texts and in particular manuals, specialised texts and literary works as the most salient instances of "abstract, structurally or linguistically complex" texts. However, these criteria no longer seem to pose a problem for NMT systems: manuals are nowadays very often machine-translated as are specialised texts (Cochrane

France 2021[4], WIPO 2021[5]). NMT has even been investigated regarding its potential for literary translation (Toral & Way 2015; Hansen 2020). Moreover, the purpose of the described reading activities seems to be to understand the content but not necessarily to grasp the aesthetic aspects of the author's style – or at least this is not stated in the assessment grid. For comprehension or gist purposes, machine-translating a text into the learner's first language can spare them almost all effort. A logical question that arises in this context is thus: Does acquiring C2 reading competence in a foreign language yield the same concrete results as a good command of machine translation solutions coupled with a C2 reading competence in one's first language? While this does not replace the value or actual motivation for developing reading skills in a foreign language, it might challenge the evaluation of such skills in concrete situations, not only for reading but also for writing purposes.

The texts suggested for writing purposes (last row of the grid in Figure 2) follow a logical learner progression from A1 (e.g. writing a simple postcard) to C2 (e.g. writing summaries and reviews of professional or literary works). However, the possibility to use NMT for writing purposes induces a shift of perspective. As shown in Figure , levels A2 up to C1 focus on the following text types:

- Short, simple notes and messages relating to matters in areas of immediate needs

- Simple connected texts on subjects that the learner is familiar with

- Clear detailed texts on a wide range of subjects

- Essays

- Reports

These kinds of texts can already be handled efficiently by the leading NMT systems, as they can rely on a large amount of data covering numerous subjects regardless of their complexity. However, all levels up to B2 also include various text types that are rather problematic for NMT systems:

- Postcards and holiday greetings (A1)

- Letter to thank someone (A2)

- Personal letters (B1 and B2)

These text genres presuppose not only lexical and grammatical knowledge but also a certain knowledge about conventions and politeness in the target culture. The most salient case is, paradoxically, the postcard, as it usually contains a very high proportion of formulaic speech. For example, the French expression

---

[4]   Cochrane France. https://france.cochrane.org/simpliquer/traductions/des-%C3%A9tudiants-du-master-2-ilts-participent-%C3%A0-leffort-de-traduction-des (24. 02. 2021)

[5]   World Intellectual Property Organization. *WIPO Translate*. https://www.wipo.int/wipo-translate/fr/ (25.02.2021)

"*Bon baisers de Saint-Malo*" is typical of postcards but can become "*Good kisses from Saint-Malo*" when machine-translated into English[6], which does not have the formulaic meaning intended in the French version.

The more advanced CEFR levels tend to describe writing activities that have to be guided by a human writer but could (and some already do) rely on human-machine interaction. Level C1 foresees the capacity to "select style appropriate to the reader in mind", implying several strategic writing decisions that can currently not be made by NMT systems without human intervention. For the last level (C2), the texts described seem to be mostly producible by an NMT system under the guidance and control of a human writer through so-called pre-editing or post-editing (Delorme Benites & Lehr 2021). Moreover, the last writing activity listed under C2 concerns the production of "summaries or reviews of professional or literary works". As automatic summarisation is currently one of the most investigated challenges and several major advances have been made in the field (Klymenko et al. 2020), it is very likely that this task will be carried out by machines rather than humans in the foreseeable future.

These primary observations show how a logical learner-centred progression scheme can be disrupted by the rise of AI. As a result, it might prove difficult to use this general self-assessment grid with learners who are used to working with NMT systems – and possibly further AI solutions such as automatic summarisation. However, the self-assessment grid is avery simplified overview based on a much broader and more finely granulated framework that includes numerous descriptors for four communicative language activities/strategies (reception, production, interaction, mediation). A closer look at the fine-grained descriptors therefore is necessary to better understand how such systems challenge the theoretical frameworks underlying language teaching strategies. Since the focus of the present article is on the impact of the currently most widely-used written automatic translation systems, we concentrate here on the written aspects of reception, production and mediation.[7]

---

[6]   Google Translate (12.07.2021)

[7]   The activity/strategy called "Interaction" mainly focusses on oral, spontaneous interaction and therefore corresponds to real-life situations where the learner is less likely to use written NMT systems (automatic speech translators are not part of the discussion in the present article).

4.2 Writing in L2

Overall written production		
	Overall written production	
C2	Can produce clear, smoothly flowing, complex texts in an appropriate and effective style and a logical structure which helps the reader identify significant points.	
C1	Can produce clear, well-structured texts of complex subjects, underlining the relevant salient issues, expanding and supporting points of view at some length with subsidiary points, reasons and relevant examples, and rounding off with an appropriate conclusion. Can employ the structure and conventions of a variety of genres, varying the tone, style and register according to addressee, text type and theme.	
B2	Can produce clear, detailed texts on a variety of subjects related to their field of interest, synthesising and evaluating information and arguments from a number of sources.	
B1	Can produce straightforward connected texts on a range of familiar subjects within their field of interest, by linking a series of shorter discrete elements into a linear sequence.	
A2	Can produce a series of simple phrases and sentences linked with simple connectors like "and", "but" and "because".	
A1	Can give information about matters of personal relevance (e.g. likes and dislikes, family, pets) using simple words/signs and basic expressions. Can produce simple isolated phrases and sentences.	
Pre-A1	Can give basic personal information (e.g. name, address, nationality), perhaps with the use of a dictionary.	

Figure 3: Descriptors for overall written production (Council of Europe 2020: 66)

In the CEFR, written production is not linked to the mediation category, so translation, human or otherwise, is not considered writing or a text-based activity. It might therefore not have been considered relevant to establish *how* learners successfully perform the writing activities described in Figure . Whether a learner first drafted a text in their first language or whether they directly wrote it in their foreign language is, according to the way the descriptors are formulated, not deemed essential to evaluate the resulting product, and hence the learner's level of competence. However, with the generalised use of automatic text generation, in our case NMT, the evaluation of the product is not enough to tell if learners can perform the described tasks on their own or with the aid of a machine. As the CEFR does not mention AI explicitly, one can assume that the guidelines do not give central consideration to recourse to any kind of AI and can be interpreted to be performed by humans only. However, by doing so, one ignores the reality: NMT might be used by the majority of language learners, even when it is prohibited (O'Neill 2019). This calls for a more balanced stance regarding the goals described in the CEFR, taking into account the possibilities that NMT offers and that learners all over the world probably already make use of. From such a perspective, we can make a clear distinction between activities that can be performed by NMT alone and activities for which NMT

becomes part of a human-guided writing process. The former correspond to the levels pre-A1 to B2, since learners could write the texts in their first language and obtain acceptable translations for many types of texts that would fulfil the text's communicative function in the L2. Human-guided writing processes would include activities listed for levels C1 and C2 that require language competence by the learner even if NMT is part of the writing process. Especially the ability to "employ the structure and conventions of a variety of written genres, varying the tone, style and register according to addressee, text type and theme" (C1) and to choose an "appropriate and effective style" (C2) imply knowledge about genre, register and style conventions. Such language-specific textual competence is not included in current NMT systems and is a prerequisite for the competent post-editing of machine translation outputs (Depraetere 2010).

This two-fold situation raises several interesting issues. First, if all levels up to B2 are achievable with the help of NMT – or rather by using NMT output as is without further human action in the target language, the linear progression described through the CEFR level scale might no longer be followed by many learners. If so, how can teachers provide learners with the necessary proficiency to perform the goal activities described in the upper levels? This issue was raised by Garcia & Pena (2011: 486) before MT went neural. Since then, NMT has developed into an unavoidable technology in our daily lives, and it seems utopic and counterproductive to completely exclude it from the early stages of language learning. Quite the contrary, we need to take it explicitly into account and make use of it didactically. Challenges that arise include how NMT can be included in levels A1 to A2 in order to achieve the objectives described in C1 and C2, and how assessment can be redesigned around the use of MT. Moreover, language students should learn about the chances and risks of producing texts with the help of NMT across proficiency levels as a component machine translation literacy (Bowker & Buitrago Ciro 2019) and informed use of these tools.

4.3 Reading in L2

	Overall reading comprehension
C2	Can understand virtually all types of texts including abstract, structurally complex, or highly colloquial literary and non-literary writings. Can understand a wide range of long and complex texts, appreciating subtle distinctions of style and implicit as well as explicit meaning.
C1	Can understand in detail lengthy, complex texts, whether or not these relate to their own area of speciality, provided they can reread difficult sections. Can understand a wide variety of texts including literary writings, newspaper or magazine articles, and specialised academic or professional publications, provided there are opportunities for rereading and they have access to reference tools.
B2	Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms.
B1	Can read straightforward factual texts on subjects related to their field of interest with a satisfactory level of comprehension.
A2	Can understand short, simple texts on familiar matters of a concrete type which consist of high frequency everyday or job-related language. Can understand short, simple texts containing the highest frequency vocabulary, including a proportion of shared international vocabulary items.
A1	Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required.
Pre-A1	Can recognise familiar words/signs accompanied by pictures, such as a fast-food restaurant menu illustrated with photos or a picture book using familiar vocabulary.

Figure 4: Descriptors for overall reading comprehension (Council of Europe 2020: 54)

The questions raised for writing in L2 apply even more for reading activities: the objectives described from Pre-A1 to C1 can today be achieved by machine-translating a foreign text into one's first language. The C2 requirement that the learner can appreciate "subtle distinctions of style and implicit as well as explicit meaning" (Figure 4) is, as of today, not achievable without reading the original text. However, the descriptors at all levels stress that the acquisition of vocabulary is related to reading activities (for example in B2: "*Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms*"). This suggests an implicit goal entailed in reading activities. If reading in L2 for comprehension purposes can indeed be substituted with reading machine translation output in L1 for almost all levels, what happens to the interdependency between vocabulary acquisition and reading competence? Further, as was the case for writing levels, how can learners achieve level C2 without linearly following the progression described in the previous levels?

Given the ubiquity of MT, reading in a foreign language will in the future have to be considered as a deliberate activity for didactic purposes, separate from reading just for comprehension.

Such questions are also relevant for evaluating language competence in work environments: If job applicants indicate a reading level C1 on their CV using the support of MT tools, this no longer means that they have enough competence to achieve level C2 soon. On the contrary, based on the current version of the self-assessment grid, their independent reading competences can no longer be properly assessed and remain rather unpredictable. In addition, recruiting companies might today be more interested in an applicant with a Pre-A1 level in the required language plus a good user understanding of MT tools than in an applicant with a B2 level without any skills regarding language technology. This could be the case whenever languages and intercultural communication are not central to the job profile but knowledge in foreign languages is used occasionally to access information. In the media and communication industry, for example, this would apply to social media and community managers. All of the above demonstrates that the current descriptors of various sub-competences in the self-assessment grid no longer give us a clear picture of language proficiency and its usability in the practice of communication in an era of NMT.

4.4 Mediation

For mediation activities (see Figure), the descriptors include both the translation of written text in writing and in speech. As these competences focus on the activity of translating itself, they are affected by the use of MT tools in several ways. Levels A1 to B2 foresee the translation of texts about everyday themes and factual texts with frequent vocabulary. Even if the accuracy of the translation of even simpler content is never guaranteed by MT developers, state-of-the-art NMT is able to perform these mediation tasks satisfactorily. At the lower levels of proficiency, MT tools may even outperform language learners in this mediation task. This comes with opportunities but also with some challenges. While the various ways in which MT could be employed as a didactic tool remain to be fully explored, its performance also has motivational implications for language learners. Moreover, it causes problems for some methods of assessment in language teaching which will have to be rethought.

At the higher levels of language proficiency (C1-C2), the requirements for mediation activities in writing can be performed with the support of MT, but only if post-editing of the machine-translated text is carried out. For the latter, a highly advanced command of language as outlined in the descriptors for levels C1-C2 is required, as the post-editor needs to check and correct the machine-produced output. In a similar vein, both MT tools and voice translation systems can support spoken language mediation when only lower levels of proficiency are required. Whenever more specialised topics, complex argumentative structures, emotions and nuances matter, these tools have to be monitored by someone

who is highly proficient in both languages. In view of these developments, the following questions arise: Are language mediation competences still relevant for the assessment of language proficiency as such, and should they be included in the descriptors? Given that human intervention in language mediation is above all required for specialised communication, should those contexts be reserved exclusively for language professionals? If these competences are to be reformulated in the light of NMT, should other aspects be included instead that allow language learners to make use of MT tools in an appropriate way? For example, would controlled language writing in language A, which allows MT tools to provide better results in language B, be another useful competence? In addition, does the concept of MT literacy, which includes knowledge about the use of MT tools and an awareness of their possibilities and risks, belong in the CEFR as a pre-requisite for mediation? And finally, if language mediation exercises are used for language learning purposes, should they be carried out specifically without NMT and focus on suitable didactic goals, such as the acquisition of intercultural competence?

TRANSLATING A WRITTEN TEXT IN SPEECH	
Note: As in any case in which mediation across languages is involved, users may wish to complete the descriptor by specifying the languages concerned.	
C2	Can provide fluent spoken translation into (Language B) of abstract texts written in (Language A) on a wide range of subjects of personal, academic and professional interest, successfully conveying evaluative aspects and arguments, including the nuances and implications associated with them.
C1	Can provide fluent spoken translation into (Language B) of complex written texts written in (Language A) on a wide range of general and specialised topics, capturing most nuances.
B2	Can provide spoken translation into (Language B) of complex texts written in (Language A) containing information and arguments on subjects within his/her fields of professional, academic and personal interest.
B1	Can provide spoken translation into (Language B) of texts written in (Language A) containing information and arguments on subjects within his/her fields of professional, academic and personal interest, provided that they are written in uncomplicated, standard language.
	Can provide an approximate spoken translation into (Language B) of clear, well-structured informational texts written in (Language A) on subjects that are familiar or of personal interest, although his/her lexical limitations cause difficulty with formulation at times.
A2	Can provide an approximate spoken translation into (Language B) of short, simple everyday texts (e.g. brochure entries, notices, instructions, letters or emails) written in (Language A).
	Can provide a simple, rough, spoken translation into (Language B) of short, simple texts (e.g. notices on familiar subjects) written in (Language A), capturing the most essential point.
	Can provide a simple, rough spoken translation into (Language B) of routine information on familiar everyday subjects that is written in simple sentences in (Language A) (e.g. personal news, short narratives, directions, notices or instructions).
A1	Can provide a simple, rough spoken translation into (Language B) of simple, everyday words and phrases written in (Language A) that are encountered on signs and notices, posters, programmes, leaflets etc.
Pre-A1	*No descriptors available*

TRANSLATING A WRITTEN TEXT IN WRITING	
C2	Can translate into (Language B) technical material outside his/her field of specialisation written in (Language A), provided subject matter accuracy is checked by a specialist in the field concerned.
C1	Can translate into (Language B) abstract texts on social, academic and professional subjects in his/her field written in (Language A), successfully conveying evaluative aspects and arguments, including many of the implications associated with them, though some expression may be over-influenced by the original.
B2	Can produce clearly organised translations from (Language A) into (Language B) that reflect normal language usage but may be over-influenced by the order, paragraphing, punctuation and particular formulations of the original.
	Can produce translations into (Language B, which closely follow the sentence and paragraph structure of the original text in (Language A), conveying the main points of the source text accurately, though the translation may read awkwardly.
B1	Can produce approximate translations from (Language A) into (Language B) of straightforward, factual texts that are written in uncomplicated, standard language, closely following the structure of the original; although linguistic errors may occur, the translation remains comprehensible.
	Can produce approximate translations from (Language A) into (Language B) of information contained in short, factual texts written in uncomplicated, standard language; despite errors, the translation remains comprehensible.
A2	Can use simple language to provide an approximate translation from (Language A) into (Language B) of very short texts on familiar and everyday themes that contain the highest frequency vocabulary; despite errors, the translation remains comprehensible.
A1	Can, with the help of a dictionary, translate simple words and phrases from (Language A) into (Language B), but may not always select the appropriate meaning.
Pre-A1	*No descriptors available*

Figure 5: Descriptors for mediation activity: Translating a written text (Council of Europe 2018: 114)[8]

---

[8]   The 2020 version of the descriptors is identical to this one but presented on several pages, so it could not be reproduced here.

5. Conclusions

In this article, we discussed the latest version of the widely-used CEFR in light of the recent rise of NMT. We showed how the usefulness and applicability of some of the CEFR descriptors are being called into question by technological developments. The analyses of the descriptors for the writing, reading and mediation competences showed that, while the more advanced levels encompass tasks that cannot be performed by a machine alone, many of the tasks in the early levels (up to B2) could be done today by or with the help of NMT. This shows the need for a new definition of these language proficiency levels, which should take the recent advances in language technology, especially NMT, into account. As of today, the adequacy of the level descriptors for in-class activities (reading, writing, translation for learning purposes) as well as real-life situations is no longer guaranteed, and this can lead to serious difficulties regarding assessment and curricular development.

One could argue that all activities described in the various levels of the CEFR are exemplary for in-class, didactic tasks and that they are to be performed by humans only. However, such an approach would lead to a divide between language teaching objectives and language use reality, much contrary to the overarching aim of the CEFR. In the digitalised world, the widespread use of NMT suggests that we need to integrate the role of technology into our conceptualisation of foreign language proficiency and also rethink the relevance of L1 skills. Foreign language proficiency can no longer be conceptualised without a component that includes human-machine interaction. The rethinking process has already started for language learning and teaching approaches (Briggs 2018; Garcia & Pena 2011; Zhu 2020; Yamada 2020; O'Neill 2019; Niño 2008). For example, the concept *literacy* has been redefined and expanded with the growing digitalisation of information sources. Bowker introduced the concept of *machine translation literacy*, that "is primarily about developing critical thinking skills in regard to technology use" (2020: 15). This concept could provide a substantial basis to include NMT in the description of foreign language proficiency levels. Technology-oriented components of language learning could be complemented by a focus on the uniquely human skills that are necessary for successful communication in different languages, such as knowledge about intercultural differences. The latter illustrates particularly well the importance of language learning as an indispensable asset that widens people's perspectives and opens up new opportunities for them. We hope that, in this article, we convincingly demonstrated the importance of integrating technology into our view of language learning without challenging the continued need for language teaching.

REFERENCES

Berlitz, M. D. (1887). *Methode Berlitz*. New York: Berlitz School.

Bowker, L. (2020). Machine translation literacy instruction for international business students and business English instructors. *Journal of Business & Finance Librarianship, 25(1)*, 25-43.

Bowker, L. & Buitrago Ciro, J. (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Bingley: Emerald Publishing. doi:https://doi.org/10.1108/9781787567214

Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *JALT CALL Journal, 14(1)*, 2-24.

Cook, G. (2010). *Translation in language teaching: An argument for reassessment*. Oxford: Oxford University Press.

Council of Europe. (2001). *Common European framework of reference for languages*. Cambridge: Cambridge University Press.

Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching, assessment - Companion volume with new descriptors*. Strasbourg: Council of Europe Publishing.

Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment – Companion volume*. Strasbourg: Council of Europe Publishing.

Delorme Benites, A. & Lehr, C. (2021). Mehrsprachigkeit und Technologie: Who's lost in translation? In ZHAW Angewandte Linguistik (ed.), *Angewandte Linguistik in Sprachberufen* (pp. 118-130). Berlin: DeGruyter Mouton.

Depraetere, I. (2010). What counts as useful advice in a university post-editing training context? Report on a case study. *Proceedings of the 14th annual conference of the European association for machine translation*. Saint-Raphaël. https://aclanthology.org/2010.eamt-1.11.pdf

European Union Association of Translation Companies. (2019). *Language industry survey. Expectations and concerns of the European language industry*. https://euatc.org/wp-content/uploads/2019/11/2019-Language-Industry-Survey-Report.pdf

Forcada, M. (2017). Making sense of neural machine translation. *Translation Spaces, 6(2)*, 291-309. doi:https://doi.org/10.1075/ts.6.2.06for

Fountain, A. & Fountain, C. (2009). A new look at translation: Teaching tools for language and literature. In A. Fountain & C. Fountain (eds.), *Empowerment through collaboration: Dimension 2009* (pp. 1-15). Roswell GA: Southern Conference on Language Teaching.

Garcia, I. & Pena, M. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning, 24(5)*, 471-487.

Hansen, D. (2020). Artificial intelligence and machine learning: How a better grasp of AI can frame and facilitate the dialogue about its role in the years to come. Paper given at *CIUTI Conference*. Paris.

Klymenko, O., Braun, D. & Matthes, F. (2020). Automatic text summarization: A state-of-the-art review. *Proceedings of the 22nd International Conference on Enterprise Information Systems, (1)*, 648-655.

Neuner, G. & Hunfeld, H. (1993). *Methoden des fremdsprachlichen Deutschunterrichts*. Kassel: Universität Gesamthochschule Kassel (GHK).

Niño, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning, 21(1)*, 29-49.

Nurminen, M. & Papula, N. (2018). Gist MT users: A snapshot of the use and users of one online MT tool. *Proceedings of the 21st Annual Conference of the European Association for Machine Translation,* 199-208. Alacant, Spain.

O'Callaghan, J. (2014). Don't bother learning a foreign language! Skype will soon translate spoken foreign words in real time. *Daily mail.* https://www.dailymail.co.uk/sciencetech/article-2641653/Dont-bother-learning-foreign-language-Skype-soon-let-translate-spoken-foreign-words-real-time.html

O'Neill, E. (2019). Online translator, dictionary, and search engine use among L2 students. *CALL-EJ: Computer-Assisted Language Learning-Electronic Journal, 20(1)*, 154-177.

Pym, A., Malmkjaer, K. & Gutiérrez-Colón Plana, M. (2013). *Translation and language: The role of translation in the teaching of languages in the European Union. A Study.* Directorate-General for Translation, European Commission.

Roche, J. (2009). Emergente Textualität in der Lernersprache - von Chunks und Situativität zum Text. In M. Foschi, M. Hepp, E. Neuland & M. Dalmas (eds.), *Text und Stil im Kulturvergleich* (pp. 47-65). München: Iudicium.

Rösler, D. (2012). *Deutsch als Fremdsprache. Eine Einführung.* Stuttgart: J.B. Metzler.

Savignon, S. J. (2007). Beyond communicative language teaching: What's ahead? *Journal of Pragmatics, 39*, 207-220.

Toral, A. & Way, A. (2015). Translating literary text between related languages using SMT. *Fourth Workshop on Computational Linguistics for Literature, NAACL.https://aclanthology.org/W15-0714.pdf*

Viëtor, W. (1886). *Der Sprachunterricht muss umkehren! Ein Beitrag zur Überbürdungsfrage.* Heilbronn: Verlag von Gebr. Henninger.

Yamada, M. (2020). Language learners and non-professional translators as users. In M. O'Hagan (ed.), *The Routledge Handbook of Translation Technology* (pp.183-199). London: Routledge.

Zhu, X. (2020). Machine translation in foreign language learning classroom - Learners' indiscriminate use or instructor's discriminate stance. *English Linguistics Research, 9(4).* https://doi.org/10.5430/elr.v9n4p1