

Training and Learning

Maria Laura Ferroglio (Università di Torino) & Aurora Trapella (Università di Torino)

marialaura.ferroglio(at)unito.it, aurora.trapella(at)unito.it

Abstract

This paper analyses the terms *training* and *learning*, examining their historical evolution and usage as recorded in English and Italian monolingual and bilingual dictionaries. The analysis then focuses on two specialised domains, namely machine translation (MT) (both through Neural MT and generative AI technology) and Second Language Acquisition (SLA). We examine how the meanings of *training* and *learning* are adapted in each field, highlighting overlaps between their human-related and machine-related uses. English and Italian have been selected as instances of a Germanic and a Romance language to compare their respective encoding of the concepts of *training* and *learning*, analysing when and how these terms began to be applied to machines rather than exclusively to humans. This contribution to the glossary of AI terms discusses how linguistic choices shape perceptions of machine capabilities. The paper underscores the role of language in mediating technological change across the evolving meanings and uses of *training* and *learning*, and invites reflection on the consequences of anthropomorphic discourse in AI-related communication.

Keywords

artificial intelligence (AI), Second Language Acquisition (SLA), Neural Machine Translation (NMT), critical AI discourse studies, lexicography, anthropomorphism

1 Introduction

When Artificial Intelligence (AI) became popular in 2022, specialised terminology increasingly entered public discourse. Among terms originally associated with human cognition are *training* and *learning*, which historically belong to human-centred domains like instruction, physical preparation, and other human-related activities. In contemporary AI discourse, however, these terms have acquired new meanings and have been frequently used to describe computational processes underlying machine learning (ML), machine translation (MT), and generative AI systems. This semantic shift raises questions about how language shapes conceptual understandings of AI and whether such terminology and metaphorical use of words encourage anthropomorphic interpretations of machine behaviour.

This paper analyses the meanings and uses of *training* (Section 2) and *learning* (Section 3) and their diachronic development through an English-Italian contrastive lexicographic analysis. The etymology and definitions of these words in major English dictionaries are examined to identify their primary meanings and the semantic domains in which they are traditionally attested. Particular attention is paid to whether general monolingual dictionaries of English and Italian include references to the technical meaning of *training* and *learning* in relation to AI or ML. Bilingual English-Italian dictionaries are also analysed to identify the main



Ferroglio, Maria Laura & Trapella, Aurora. 2026.

Training and Learning

AI-Linguistica, Vol. 6 No. 1

DOI: 10.62408/ai-ling.v6i1.83

AI-Linguistica. Linguistic Studies on AI-Generated Texts and Discourses

ISSN: 2943-0070

CC-BY-NC-SA 4.0

Italian equivalents and semantic fields associated with the English terms, considering all attested meanings and translation equivalents to determine which of these forms have been adopted in AI discourse.

This step is also preliminary to a final reflection on anthropomorphism resulting from the use of human-oriented terminology to describe AI processes. Combining the authors' research fields, definitions of the terms in publications on MT and AI-assisted Second Language Acquisition (SLA) are examined to trace the semantic extension of *training* and *learning* from human-centred practices to computational processes. This helped investigate whether and how human-centred terms are extended and potentially reinterpreted, and assess how such linguistic choices may contribute to anthropomorphic representations of AI systems.

2 Methodological framework

The lexicographic analysis is based on the Oxford English Dictionary (OED) as the principal English reference source because of its extensive historical coverage, including early attestations, semantic developments, and diachronic changes in usage. Definitions of the terms are drawn from general-purpose English dictionaries, namely Cambridge Dictionary (CDE), Collins English Dictionary (CED) and Merriam-Webster Dictionary (MW). The Longman Dictionary of Language Teaching and Applied Linguistics (LDLTAL) was consulted for specialised usage in SLA. Italian monolingual resources include Treccani (DT), Il Nuovo De Mauro (NDM) and Lo Zingarelli (ZING). Bilingual dictionaries include Il Ragazzini (DIR) and the English-Italian component of CDE. The comparison considers both the meanings associated with human-centred domains and the specialised meanings in relation to AI.

The discussion of the potential anthropomorphic implications of AI terminology is informed by critical literature addressing the discursive framing of AI. In order to capture current perceptions in AI discourse, the analysis draws on relevant scholarly discussions to contextualise the observed lexical and semantic developments, with particular attention to the use of human-centred terminology. However, the current analysis does not include empirical data, as such an approach would require an updated Italian corpus¹ or the construction of an ad hoc one, which would represent a valuable direction for future research.

¹ The largest Italian corpus currently available is the *itTenTen20* (<https://www.sketchengine.eu/ittenten-italian-corpus/>) and includes data compiled before the emergence and widespread diffusion of Generative AI.

3 Training

3.1 Etymology and diachronic development of *training*

According to the OED, *training* is a noun derived by suffixation from the verb to train (“to drag along”), expressing the action denoted by the verb. The earliest attested use of the noun *training* dates back to the Middle English period (1150-1500). According to the OED, the first recorded occurrence appears in 1440 in the *Promptorium Parvulorum*, the first English-Latin dictionary, with the now-obsolete sense of “protraction, delay”. From the sixteenth century, the meaning progressively shifted toward forms of instruction aimed at improving abilities or behaviour. For instance, by 1537, *training* relates to discipline and instruction (given or received) for the development of character, behaviour, or ability. By 1598 it indicated the “instruction and practice (given or received) in an art, profession, occupation, or procedure, with a view to proficiency in it”, and more generally, as a period of instruction or practice, a *training session* (OED). By the 1570s, *training* was also attested as “military drill” for preparing soldiers. It is attested by 1581 as “physical preparation of oneself or another for athletic or sporting competition”, and by 1677 as the “teaching of a particular behaviour, esp. obedience to orders, to an animal”. In 1794, the act of training is also described as “a thorough education in a subject, profession”.

Around this period, we observe that the initial meaning of protraction and movement towards a certain direction has shifted toward senses involving improvement or development. Indeed, the meaning of *training* in connection with movement, as used in its very first attestation, appears to be extremely rare. It is only recalled in one other instance, dating from 1812, where it describes the act of pointing, directing, or aiming a firearm, specifically through horizontal movement, or by indicating the direction in which the weapon is aimed. If we think of the way we intend *training* today, this sense can also be interpreted metaphorically: the “direction” intended within the sense of movement becomes the target, i.e. the ‘proficiency’ to which one aims when *training*. The movement is oriented toward a specific objective, and the direction can be seen as the “path” or “course” along which one trains. Indeed, modern dictionaries such as CDE define *training* as “the process of learning the skills you need to do a particular job or activity”, with examples including *a training course*, *a teacher-training college*, *have/receive training in*. MW and CED confirm this dominant meaning, situating the concept of *training* within human-centred domains, particularly professional education, sports, and behaviour.

From 1959, *training* has acquired a new meaning in relation to the development of ICT and computing technology to refer to “the technique or process of configuring a program or device so that it responds to inputs in the desired

manner” (OED). The noun derives from the transitive verb *to train* which, by 1958, was employed in the field of computing with the following meaning:

To configure (a program or device, esp. one based on neural networks) so that it will respond to inputs in the desired manner, by modifying the parameters controlling its outputs in the light of its responses to a series of test inputs.

Indeed, CDE also records this more recent specialised sense in computing, including a lemma related to ML, thus with the sense of ‘supplying data’, defining it as “the process of creating or improving a computer representation of a system or process by supplying it with data”. Here, expressions such as *training data* and *model training* are attested as established in AI discourse.

3.2 Equivalents of *training* in English-Italian bilingual dictionaries and Italian monolingual dictionaries

The semantic profile of *training* can be further examined through its equivalents in English-Italian bilingual dictionaries and Italian monolingual lexicographic sources. As for English-Italian bilingual dictionaries, CDE presents the following equivalents of *training*: *tirocinio*, *addestramento*, *preparazione* and *formazione*. The ICT meaning is presented in the fixed phrase *computer training*, translated as *addestramento informatico*. Conversely, in relation to the meaning associated with the preparation for a sport or competition, the Italian equivalent attested in this dictionary is *allenamento* in the direction English-Italian. DIR adds the words *educazione*, *esercizio* and *pratica* to the equivalents of *training* referring to the process of developing abilities through practice or instruction. These meanings are attested across military preparation, professional education, and sports preparation. Additional senses extend to animal conditioning, with the equivalent of *ammaestramento*, and *dressage* in equestrian practice. The Italian-English direction also includes the word *training* as an anglicism in the domains of management (meaning *apprenticeship*), sports, and psychology (*training autogeno*).

Modern Italian monolingual dictionaries such as NDM, DT and ZING further confirm the semantic scope of *addestramento*, which denotes the process of learning through practical instruction aimed at developing competence in relation to human and animal subjects. DT and ZING also include a lemma for the anglicism *training*, adding new information to its etymology: it derives from *to train* ‘*allenare, addestrare*’, which in turn comes from the French *trainer* ‘trascinare’.

This lexicographic overview of the meanings of *training* in major monolingual and bilingual dictionaries highlights that the primary meaning of the word remains that of instruction for the development of certain abilities, and it is closely connected to human- or animal-centred domains. Despite the increasing prominence of AI, the technological sense of *training* appears more established in specialised usage than in general-purpose usage.

As a result, the near absence of the technological meaning of *training* in general-purpose dictionaries also raises questions on the circulation of English terminology in Italian in the form of anglicisms, particularly in technological domains. From a historical perspective, the Italian equivalents of *training* emerged not as direct borrowings but as calques introduced during a period characterised by linguistic purism, which favoured the domestication of foreign terms (Pulcini 2023: 50). Thus, the Italian equivalent for *training*, *allenamento*, entered the Italian language as an indirect borrowing in contexts related to professional or sport preparation well before the rise of digital technologies (Pulcini 2023: 39). When *training* was later used in relation to computer science, Italian did not import the non-adapted anglicism because it drew on *allenamento* and *addestramento*, which were already established in human- and animal-related domains. More recently, *training* entered Italian as a non-adapted borrowing in the domain of AI technology to refer to the optimisation of a model's parameters. This meaning of *training* in AI discourse involves a semantic narrowing – from a general to a more specialised sense – with a conceptual metaphor (Lakoff and Johnson 1980): computational processes are framed in terms of human *learning*, where machines are described as being *trained* to improve their performance (“aim to proficiency”). When these concepts circulate beyond technical contexts, however, expressions such as machines being “trained” to improve their performance may encourage comparisons with human learning and abilities. Vieira's (2020) framing analysis of the newspaper coverage of MT shows that positive representations of MT often foreground technological progress and, at times, compare MT capabilities with those of human translators. This potentially reinforces a human-centred understanding of machine performance and may contribute to anthropomorphic interpretations.

3.3 *Training* in the domain of Generative AI (GenAI) and Neural Machine Translation (NMT)

The expansion of AI and ML has led to a specialised use of the term *training*, which has, however, retained the core meaning identified in general-language dictionaries, namely a progressive improvement of abilities through an instruction path and practice.

The conceptualisation of machine capabilities in terms of training can already be observed in early reflections on computing. In his discussion of intelligent machines, Alan Turing compared the preparation of a machine to the education of a human expert, noting that “a human mathematician has always undergone an extensive training” and suggesting that this process could be seen as analogous to “putting instruction tables into a machine” (Turing 1950: 394). From this perspective, the machine does not independently generate knowledge but acquires its capabilities through structured instruction and interaction with humans.

Moreover, Turing emphasised that machines must maintain contact with human users to adapt to their standards. Although neither language nor translation are the central topic of his 1950 paper, Turing envisages the possibility of teaching language to a machine in a way comparable to the education of a child, where linguistic competence can be developed. Such observations anticipate the terminology that is now central in AI-related discourse.

An initial indication of the consolidation of *training* within AI terminology is the emergence of derivational forms and fixed expressions. For instance, the OED records the word *pretraining*, defining it as

training of a neural network or other AI system which sets the initial values of its parameters, after which further training with a different dataset can be performed to give the parameters their final values.

The word acquired this meaning in 1965, when it was mainly used in the field of psychology to describe the training taking place before an experiment or a test. This affixation process also illustrates the productivity of *training* as a lexical base within AI terminology, where it also frequently appears as a modifier in a range of technical expressions. For instance, we may sometimes encounter the phrase *model training*, indicating the goal in ML to enhance a model by training it on algorithms (Taulli 2019: 51). Another expression is *training process*, namely the process that ML algorithms undergo using large datasets in order “to learn forms and disambiguate meanings” (Cheng, Nanjundan and George 2025: 10). Hence, the data used to create the associations in the AI algorithm are called *training data* (Taulli 2019: 67). Once the *training phase* is complete, the trained model is used on the test dataset to make predictions that satisfy the *training objective*, a mathematical function that defines the goal a ML model aims to achieve during training. Approaches such as *supervised*, *non-supervised* and *reinforcement learning* define different *training methods*.

Within the field of Natural Language Processing (NLP), such terms are used in relation to linguistic data, such as associations of words and sentences. In Neural Machine Translation (NMT), the term *training* belongs to a dense terminological network used to describe the optimisation of translation systems. Translation models are described as being *trained* on large collections of parallel (aligned) texts, commonly referred to as *training data* or *training corpora*. Through this process, the model *learns* statistical or neural associations between linguistic units across languages, allowing it to generate translations for the inputs (Koehn 2020: 72). Such tools are said to be approaching human-level quality and constitute a standard practice in many professional workflows, thanks to their cost-effectiveness and output quality (Sun, Liu and Moratto 2025: 1). Nowadays Large Language Models (LLMs), i.e. general-purpose tools which can perform a variety of tasks when given an instruction named *prompt*, are also being adopted for translation. They are trained on a vast amount of multilingual data and are said to “demonstrate better capability compared with [...] NMT engines” (Jiang and Zhang 2024: 7) and “excel

at capturing semantic relationships and generating coherent text” (Jiang and Zhang 2024: 9).

Not only do these developments have profound implications for the traditional notions of translation expertise and for the role of human experts, but they also influence the language used to describe computational processes. Since NMT has made increasing progress in output quality, claims have emerged that such tools have reached human-level performance (Popel et al. 2020: 2), sometimes accompanied by suggestions that the translator profession may eventually be replaced by machines “in the form of a narrative that opposes humans to machines” (Larsonneur 2025: 61). Consequently, the increased performance of these tools is perceived not only as technological progress but also as a potential threat to human expertise. Moorkens and Lewis (2019: 5) argue that the growing availability of NMT technologies, both free and commercial, is reshaping the role and status of human translators. At the same time, NMT technologies raise broader issues of transparency, accountability and standardisation (Larsonneur 2020: 265). Translation may come to be perceived less as a specialised, labour-intensive form of professional work and more as an immediately available automated service. This narrative is further reinforced by the opaque functioning of LLMs, whose internal mechanisms are difficult to interpret, as Bender and Koller (2020: 5186) state:

[...] the way we speak about what neural LMs are doing is misleading to the public. Part of the reason for this tendency to use imprecise language may well be that we do not yet fully understand what exactly it is about language that the large LMs come to implicitly represent.

The opacity in LLMs’ functioning can further intensify concerns about the future of the translation profession, particularly when AI tools are presented as matching or even surpassing humans in tasks traditionally associated with human expertise. Such comparisons can encourage anthropomorphic interpretations by encouraging machine capabilities to be interpreted through concepts traditionally applied to human cognition, especially when discourse around AI draws on vocabulary traditionally associated with human cognition, describing machines as capable of “understanding” or “comprehending” language and *learning* its “meaning”. As the terminology used to describe these machines remains human-centred, in the case of *training*, the metaphorical meaning acquired during its diachronic evolution is applied to the idea that machines can be trained as humans can. Nevertheless, interpreting anthropomorphism solely as a means of making complex notions more relatable and vivid may obscure the broader social and economic dynamics that shape the circulation of AI terminology.

As the lexicographic overview pointed out, *training* evolved from an earlier sense related to movement or direction toward meanings associated with the development of skills through practice and instruction. When applied to computational systems, the term contributes to presenting AI systems as capable of

forms of human cognition, which can reinforce perceptions of autonomy, intelligence and societal relevance. Such representations can play a role in legitimising and promoting technologies whose economic value increasingly depends on their ability to penetrate in the labour market and professional workflows. The perceived human-like competence of AI not only supports the commercial project of AI companies of developing and deploying systems designed to perform tasks traditionally performed by humans, but also reinforces concerns of human replacement by AI. This happens if technology companies have access to and the ability to process large-scale linguistic data. As Larssonneur (2025: 61) argues, many discussions surrounding AI, particularly those focused on the supposed cognitive abilities of machines, reinforce the economic structures underpinning these technologies, notably the corporate appropriation and monetisation of language exchanges. For large technology companies, the ability to process language accurately is central to business models based on the control and monitoring of global digital interactions. The terminology used to describe AI systems, therefore, does not merely reflect metaphorical extensions but also participates in broader narratives that shape how intelligence, language, and technological progress are conceptualised and marketed in contemporary society, but also how the replacement of human activities by automated systems is presented and potentially legitimised. As Bridle (2023: 9) states:

The definition of intelligence which is framed, endorsed and ultimately constructed in machines is a profit-seeking, extractive one. This framing is then repeated in our books and films, in the news media and the public imagination – in science-fiction tales of robot overlords and all-powerful, irresistible algorithms – until it comes to dominate our thinking and understanding.

Indeed, this overview of *training* reflects a complex interaction between diachronic semantic and metaphorical developments and contemporary technological discourse. The agency attributed to algorithms both mirrors societal structures and shapes our perception of the world. Anthropomorphic framings of AI should therefore be considered not only in relation to how machines are perceived, but also within the economic and social structures in which these technologies are developed and deployed. Presenting AI with anthropomorphic framings fuels a broader system of commodification of language solutions platforms. Thus, this raises further questions about the role of language in shaping our perception of the world, highlighting the need for greater critical awareness of how AI terminology is conceptualised, governed and developed.

4 Learning

4.1 Etymology and diachronic development of *learning*

According to the OED, the base term *learn* originates from the Old English verb *leornian*, which means to “acquire knowledge”, to “be cultivated, study, read, think about”; the noun is derived from the verb. The two entries are defined as follows:

To learn (verb): “To gain or acquire knowledge of or skill in (something) as a result of study, experience, or being taught”

Learning (noun): “The action or process of acquiring knowledge, understanding, or skill”

The verb form was first recorded in an Old English (OE) Northumbrian dialect before 1150, and the main semantic domain was religion, in documents such as the Lindisfarne Gospels. The noun is attested earlier, from Ælfred the Great’s translation of Pope Gregory’s Pastoral Care. This origin may be related to the role of knowledge and literacy in the religious sphere, in particular after the Benedictine reform. The noun reflects the use of existing terms to express new, domain-specific concepts, as with *lēran* (to teach) (Baugh and Cable 2002: 81). The first meaning of the noun form *learning* encompasses knowledge, understanding, and skills acquisition. Around 1530, the meaning of the verb form evolved, describing “skills or abilities” acquired following experience and training. Since 1800, both the verb and the noun have been used to describe the capacity to modify thinking or behaviour. The noun acquired its psychology-related meaning in 1901, to describe mental processes that result in modifications of behaviour or the acquisition of skills in addition to natural development by growth or maturation. Another meaning of *learning* is *to teach*, classified as “rare” in the OED.

MW defines the verb *to learn* as “to gain knowledge or understanding of, or skill in, by study, instruction, or experience”, while the noun definition includes the aspect of behavioural modification. Indeed, in MW *to learn* may also mean *to teach*, a meaning attested from the 13th century to early 19th century. What is also worth noticing in the definition by MW is the reference to *understanding*, defined in a secondary entry from the same dictionary as *knowing*, *intelligent*.

Thus far, it can be seen how the verb has remained polysemous. Human-centred connotations have survived across specialisations in distinct fields, and meanings have shifted from cognitive to experiential and behavioural domains. A further observation is that both the verb and the noun retained older knowledge-related meanings after entering the psychological domain.

According to the OED, both word forms entered the field of computing in 1950. When the term *Artificial Intelligence* appeared in John McCarthy and colleagues’ proposal for Dartmouth College on machine intelligence, it was explained as follows: “every aspect of learning or any other feature of intelligence

can in principle be so precisely described that a machine can be made to simulate it” (McCarthy et al. 1955: 1). Based on this citation, both learning and intelligence appear as phenomena which are replicable by machines. Both the verb *to learn* and the noun *learning* describe the process whereby a computer or a machine acquires and stores information until it can adapt their operations to new, unseen input without being explicitly instructed. While in the OED the verb entry attributes this meaning specifically to AI itself, in the noun entry this particular meaning is associated with *machine learning*, a subfield of AI.

4.2 Equivalents of *learning* in English-Italian bilingual dictionaries and Italian monolingual dictionaries

Acquiring knowledge and information through education or experience remains a core meaning of the verb *to learn* in bilingual dictionaries. Both CED and CDE translate *to learn* as *imparare* or *venire a sapere* (‘come to know’), with the latter source adding *apprendere* (‘learn’) as a synonym of *imparare*. DIR includes *imparare a proprie spese* (‘learn the hard way’), and ZING reports the idiomatic expression *così impari!*, which indicates the aftermath of punishment or a life lesson learnt. The meaning differs between English and Italian. The English equivalent is *that’ll teach you!*: while Italian construes the phrase with the learner as both the recipient of the lesson and the culpable subject, English construes the event or experience as responsible for the lesson learnt.

In Italian monolingual dictionaries, ZING defines the lemma *imparare* as “acquisire una serie di conoscenze mediante lo studio, l’esercizio, l’osservazione” [‘acquire knowledge through study, exercise, observation’, our transl.], as “apprendere qualcosa tramite l’esperienza” [‘learn something through experience’, our transl.] and as “venire a sapere” (‘come to know’), identified as colloquial. The meaning of *to teach* is listed as fourth, and is more common in central or southern areas of Italy. Interestingly, the noun *apprendimento* is defined as the process of acquiring new knowledge, and the entry includes the noun phrases *apprendimento automatico* and *apprendimento profondo* from the IT domain, translated as “machine learning” and “deep learning” respectively. In particular, *deep learning* is a “specialised anglicism” (Pulcini 2023: 177) in Italian: the term is used despite the existence of an Italian equivalent, *apprendimento profondo*, as is the case with *machine learning* and *apprendimento automatico* (ZING).

Similarly, NDM provides the same meanings for *to learn*, namely acquiring, retaining in memory through study or exercise, and coming to learn through experience. This entry also includes learning by paying the consequences of an action, as outlined above. The meanings of *come to know* and *teach* are identified as more typical of central and southern areas of Italy (“mbè te lo imparo, come no” (Pasolini) [‘Well, of course I will teach you!’, our transl.]. in ZING. Etymologically, the Italian equivalents of *to learn* (*imparare* and *apprendere*, from Latin *im- p̄rāre*, i.e. to find, to procure, emphasised by the suffix -im,

apprehendere, to catch) reveal both the concepts of *finding* and *gaining information* in the connotations of the term. In brief, meanings of the entries *learn* and its nominalised counterpart *learning* in English and Italian encompass the concepts of knowledge gain, acquisition, exercise or study, and experience in both languages.

4.3 Learning in Second Language Acquisition (SLA)

In addressing the construct of *learning* in SLA, Brown (2007: 7) summarises definitions of *learning* as “acquiring or getting knowledge of a subject or a skill by study, experience or instruction”. Besides the ideas of acquiring, obtaining, and retaining skills or information, Brown (2007) highlights that retention is strictly linked to storage and memory; as a consequence, its connotations evoke ideas of cognitive organisation, permanence or transience. According to Brown, practice is another key feature of learning, alongside self-directed focus and action “upon events outside or inside the organism” (Brown 2007: 7). Another notable way of defining learning is as a “change in behaviour” (Brown 2007: 8). Learning is therefore conceptualised as described in dictionary entries above, that is obtaining, possessing, and retaining information or skills, as well as promoting behavioural change.

Features in the pedagogical conceptualisation of learning are, instead, the cognitive engagement required for retention through organisation, the risk of forgetting, active focussing, and the need for practice. The definition of *learning* in LDTAL (2010: 328) reads: “the process by which change in behaviour, knowledge, skills, etc., comes about through practice, instruction or experience and the result of such process”. Learning is identified as both a process and an outcome; while knowledge and skills are mentioned, behavioural change is the core concept. The entry for *language learning* from the same source is also worth consulting; it directs the reader straight to the entry for *language acquisition*, where it is specified that the words *learning* and *acquisition* are interpreted differently in the field. The two terms may be either synonyms or have distinct meanings and belong to different domains. More specifically, *learning* involves conscious endeavour and is associated with typical classroom settings characterised by explicit study of rules and progress monitoring. Instead, *acquisition* refers to the process of internalising rules as the result of exposure to comprehensible input coming from a natural environment where the target language is spoken and where meaning is prioritised over form (LDTAL). *Acquisition* may also be used to refer to the learning of one’s first language (LDTAL 2010: 312). This distinction is aligned with that originally proposed by Krashen (1982: 10) between language acquisition and language learning, where the former refers to a subconscious process similar to that involved in first-language development, and the latter relates to conscious knowledge of rules.

However, as explained by Ortega (2009: 5), both *learning* and *acquisition* have been increasingly used interchangeably. This overlap can be linked to the conceptualisation of language learning as a form of cognition, that is, the acquisition and processing of information (Ortega 2009: 82). Research in the field of learning strategies contributes to this meaning: effective learners actively associate new information with existing knowledge, progressively storing it in their long-term memory, thus constructing increasingly complex mental structures (Oxford 2002: 125). These findings reinforce a conception of learning as a process whereby information acquired from instruction and experience is connected with prior knowledge, organized in memory, and used to construct increasingly nuanced understandings of reality.

When considering recent research on AI-assisted or chatbot-assisted language learning, AI tools are often framed as catalysts for increasing learning opportunities outside traditional classroom settings (Li et al. 2024: 9; Godwin Jones 2024: 11). Digital and AI tools may also be framed as both a means and a companion: for example, in a study by Jacob, Tate and Warschauer, a participant making competent use of ChatGPT for support during her writing process framed the chatbot as a “thinking partner” (Jacob, Tate and Warschauer 2025: 14). The participant had the impression of having a continuous exchange with “someone who is always there”, despite being well aware that the machine makes “a probabilistic prediction of the next word based on the previous word” (Jacob, Tate and Warschauer 2025: 11).

The above debate about the terms *learning* and *acquisition* focused on how instruction and experience shape acquisition and retention of new information, and research in GenAI in language learning may contribute to more nuanced meanings. Besides, AI agents are easily personified, even by users who are conscious that they are machines and understand how they work.

4.4 Learning in the domain of Generative AI

Turing (1950: 458) explained that “[a]n important feature of a learning machine is that its teacher will often be very largely ignorant of quite what is going on inside, although he may still be able to some extent to predict his pupil’s behaviour”. This explanation supports reflection on terminology which may favour anthropomorphisation of the term *learning*: “pupil” and “behaviour” are associated with a “learning” machine and suggest that the machine itself is capable of gaining new knowledge, just like a pupil. This comparison encourages a personification of the machine.

Russell and Norvig (2021: 22), too, explain learning in machines through vocabulary that foregrounds behaviour and adaptation, stating that “knowledge representation and reasoning enable agents to reach good decisions”, and that, as humans, we need “to be able to generate comprehensible sentences in natural language to get by in a complex society”. Capraro (2026), too, observes that humans

operationalise thoughts precisely through language, and notices that language and cognition may both be attributed to either human or synthetic agents.

In addition, as humans, “we need learning not only for erudition, but also because it improves our ability to generate effective behavior, especially in circumstances that are new” (Russell and Norvig 2021: 22). From these perspectives, what seems to characterise the meaning of learning in AI is the extension to machines of a notion originally associated with human behavioural change, now used to refer to adaptation in outputs or task performance. Two other key terms featuring the word *learning* in the AI domain are *machine learning* and *deep learning*. More specifically, ML is defined as a method for identifying patterns among vast amounts of data to develop a model (Kurni and Mohammed 2023: 3), or the process of finding relationships among labelled data (Taulli 2019: 42); on the other hand, deep learning is defined by Taulli (2019: 71) as “a subfield of machine learning” which involves the “processing of huge amounts of data to find relationships and patterns that humans are often unable to detect”.

Learning is also found in *supervised learning*, *unsupervised learning* and *reinforcement learning*. In short, supervised learning involves models learning from labeled data by identifying patterns and associations within it. By contrast, in unsupervised learning, data are not labelled, and the model identifies patterns or groupings autonomously. The method of reinforcement learning is described as “constantly enhancing the model based on feedback” (Kurni and Mohammed 2023: 3). In reinforcement learning, models are created from data and improved through trial and error. When a task (for example, preventing self-driving cars from crashing) is completed successfully, learning is rewarded (Kurni and Mohammed 2023: 3).

To sum up, the meaning of *learning* across traditional and AI-assisted SLA and AI domains describes a process of information acquisition and retention leading to pattern recognition, information (re)structuring and behavioural change. Signs of anthropomorphisation have appeared since early metaphors for “learning machines” (Turing 1950: 48). The research in AI-assisted SLA selected for the present work distinguishes between machine-driven learning and learning involving human intervention. Yet, the term in both domains is characterised by implicit anthropomorphism: computational “representation” and “decisions” (Russell and Norvig 2021: 22) could – for example – be compared to cognitive mental states (Oxford 2002: 125) and behavioral potential (Brown 2007: 8). Cheng and colleagues (Cheng et al. 2024: 810), developed the tool *AnthroScore* to examine how texts describe technologies using terms that describe human qualities, and showed that verbs like “understand” drive high-anthropomorphism scores. While “learning” may retain its general meaning, authors warn that its structured use in GenAI may promote a pervasive, implicit anthropomorphisation, and recommend continuous investigations.

As remarked in section 2.3, the education sector, too, is not immune to commodification. Monnet and Paquet (2026: 7-8) denounce the turn of education into “knowledge business” which favours forms of “technological autocracy”. To avoid the risk of indiscriminate AI adoption, they propose placing more emphasis on humanistic education, reclaiming educators' agency and rediscovering the value of time in performing tasks accurately as opposed to delegating them to AI tools. The present analysis has targeted the term *learning* and its evolution in light of GenAI and signs of potential anthropomorphisation have been detected, laying the ground for further investigation.

5 Conclusion, limitations and future work

In this paper, we analysed the terms *training* and *learning* in parallel, tracing their etymology and historical development, and examining their adoption in MT and SLA research. The lexicographic analysis of the term *training* reflects a complex interaction between its semantic developments over time and its metaphorical extensions affecting both our current understanding of the concept and its use in AI discourse. The metaphorical meaning of movement towards an objective (‘proficiency’) has been extended to describe computational optimisation processes, thus facilitating the framing of AI technologies through familiar terms related to human cognition. However, in the domain of MT, anthropomorphisation should be understood in a broader socio-economic context. The use of human-related terms referring to NMT tools and GenAI systems supports an economic model centred on the control and commodification of language and translation solutions: presenting these tools as capable of learning and understanding language through *training* contributes to their public acceptance and, consequently, to the profit and influence of language service providers.

The term *learning* shows a clear historical progression across domains: if the meaning of knowledge acquisition has remained predominant over time, the term also gained prominence in the field of psychology, extending its meaning to processes of behavioural change. In the field of SLA, the word retains its behavioral change connotations as well as its connection to the acquisition and application of rules, which is also central to the meaning of the term in AI. The behavioural change-oriented characterisation of the term is confirmed in the studies on learning strategies, where the creation of new cognitive patterns is also involved in learning processes. In AI-assisted language learning, synthetic agents may easily be personified – for instance, as learning companions or assistants. Within the scope of the present analysis, it would appear that the notion of *learning* in GenAI is consistent with the meaning of acquiring information in the form of data and applying rules for problem-solving purposes. In the case of *learning*, anthropomorphic framing has been evident since Turing’s (1950) early conceptualisation of “learning machines”. Today, users may perceive GenAI tools as companions even though they may adopt an impersonal mindset towards them.

Finally, the risk of commodification of GenAI tools as substitutes for human expertise is also evident in education (Monnet and Paquet 2026).

We hope that this analysis contributes to inviting reflection on the role of language in shaping our perception of the world, stimulating greater critical awareness of the ways AI terminology is conceptualised, communicated and marketed. Due to space constraints, however, this analysis can only be considered a starting point. First, the investigation is restricted to the English-Italian language combination for the terms *training* and *learning*, which limits the extent to which the findings can be generalised to other AI terms and linguistic contexts. Second, the lexicographic analysis is necessarily dependent on the dictionaries considered and cannot fully capture the circulation and development of these terms. Finally, the discussion of anthropomorphism and its possible socio-economic implications is based on a critical, interpretive reading of existing literature rather than empirical observations. To obtain empirical discourse data, an up-to-date corpus of scientific and news texts could be compiled to test whether the meanings discussed here recur on a larger scale. A corpus-based analysis of recent AI discourse could provide empirical evidence of how the terms are used across different contexts and how their meanings evolve as AI technologies become increasingly widespread.

Future research could include a corpus-based analysis of recent AI discourse to provide empirical evidence of how the terms are used across different contexts and how their meanings evolve as AI technologies become increasingly widespread. Future research could also extend the investigation to other languages to examine whether the anthropomorphic effects associated with AI terminology are specific to particular linguistic systems or reflect broader patterns in the conceptualisation of AI. Such corpus-based and cross-linguistic approaches could provide a stronger empirical basis for assessing how terminology contributes to the ways in which AI systems and their capabilities are understood.

Attributions

Maria Laura Ferroglio: Conceptualisation, methodology, lexicographic analysis. In particular, she wrote Sections 1, 4 and 5.

Aurora Trapella: Conceptualisation, methodology, lexicographic analysis. In particular, she wrote Sections 1, 2, 3 and 5.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this contribution.

AI disclosure statement

The authors acknowledge the use of OpenAI's GPT-5.3 in assisting with the editing and refinement of this paper, enhancing both its clarity and presentation in English. However, we ensured that the resulting text adheres strictly to the standards and ethical guidelines of academic writing.

References

- Bender, Emily M., & Koller, Alexander. 2020. Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. In Jurafsky, Dan & Chai, Joyce & Schluter, Natalie & Tetreault, Joel (eds), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.463>.
- Brown, Douglas. 2007. *Principles of Language Learning and Teaching*. (5th edition). New York: Longman.
- Baugh, Albert C. & Cable, Thomas. 2002. *A History of the English Language* (5th edition). London: Routledge.
- Capraro, Valerio. 2026. LLMorphism: When humans come to see themselves as language models. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2605.05419>
- CDE = 2026. *Cambridge Dictionary*. Cambridge: Cambridge University Press. <https://dictionary.cambridge.org/dictionary/english> (last accessed on 20/03/2026).
- CED = n.d. *Collins English Dictionary*. <https://www.collinsdictionary.com/> (last accessed on 20/03/2026).
- Cheng, Myra & Gligorić, Kristina & Piccardi, Tiziano & Jurafsky, Dan. 2024. AnthroScore: A computational linguistic measure of anthropomorphism. In Graham, Yvette & Purver, Matthew (eds), *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics* (Volume 1: Long Papers). St. Julian: Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.eacl-long.49>
- Cheng, Xiaochun & Nanjundan, Preethi & George, Jossy P. 2025. *Introduction to Natural Language Processing: Exploring Techniques, Applications, and Challenges* (1st edition). New York: River Publishers. <https://doi.org/10.1201/9781003660675>
- DIR = Ragazzini, Giuseppe. 2023. *Il Ragazzini 2023: Dizionario inglese-italiano, italiano-inglese* (4th edition). Bologna: Zanichelli. <https://u-ubidictionary-com.bibliopass.unito.it/viewer/#/dictionary/zanichelli.ilragazzini16> (last accessed on 20/03/2026).
- DT = 2022. *Il vocabolario Treccani*. Istituto della Enciclopedia italiana. <https://www.treccani.it/> (last accessed on 20/03/2026).

- Godwin-Jones, Robert. 2024. Distributed agency in second language learning and teaching through generative AI. *Language Learning and Technology* 28. 5-31. <https://doi.org/10.48550/arXiv.2403.20216>
- Jacob, Sharin. R. & Tate, Tamara & Warschauer, Mark. 2025. Emergent AI-assisted discourse: A case study of a second language writer authoring with ChatGPT. *Journal of China Computer-Assisted Language Learning* 5(1). 1-22. <https://doi.org/10.1515/jccall-2024-0011>
- Jiang, Zhaokun & Zhang, Ziyin. 2024. Convergences and Divergences between Automatic Assessment and Human Evaluation: Insights from Comparing ChatGPT-Generated Translation and Neural Machine Translation. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2401.05176>
- Koehn, Philipp. 2020. *Neural Machine Translation*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108608480>
- Krashen, Stephen D. 1982. *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.
- Kurni, Muralidhar & Mohammed, Mujeeb Shaik. 2023. *A Beginner's Guide to Introduce Artificial Intelligence in Teaching and Learning*. Cham: Springer International Publishing. <https://doi.org/10.3991/ijet.v17i13.28873>
- Lakoff, George, & Johnson, Mark. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
- Larsonneur, Claire. 2020. Neural machine translation: From commodity to commons? In Desjardins, Renée & Larsonneur, Claire & Lacour, Philippe (eds), *When translation goes digital: case studies and critical reflections*, 257–280. Palgrave Studies in Translating and Interpreting. Cham: Springer International Publishing. <https://hal.science/hal-03998644v1> (last accessed on 05/10/2026)
- Larsonneur, Claire. 2025. Representations of language and translation in artificial intelligence. In Baumgarten, Stefan, & Tieber, Michael (eds), *The Routledge Handbook of Translation Technology and Society*, 53–65 (1st edition). London: Routledge. <https://doi.org/10.4324/9781003271314-5>
- LDLTAL = Richards, Jack C., Schmidt, Richard. 2010. *Longman Dictionary of Language Teaching and Applied Linguistics* (4th edition). Harlow: Longman.
- Li, Belle & Lowell, Victoria Lynn & Wang, Chaoran & Li, Xiangnin. 2024. A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching. *Computers and Education: Artificial Intelligence* 7. 100266. <https://doi.org/10.1016/j.caeai.2024.100266>
- McCarthy, John & Minsky, Marvin Lee & Rochester, Nathaniel & Shannon, Claude Elwood. 1955, August 31. *A proposal for the Dartmouth summer research project on artificial intelligence*. Stanford University. <https://www-formal.stanford.edu/jmc/history/dartmouth.pdf> (last accessed on 05/10/2026).

- Monett, Dagmar, & Paquet, Gilbert 2025. Against the Commodification of Education—if harms then not AI. *Journal of Open, Distance, and Digital Education* 2(1). 1–24. <https://doi.org/10.25619/wazgw457>
- Moorkens, Joss. & Lewis, David. 2019. Research Questions and a Proposal for the Future Governance of Translation Data. *The Journal of Specialised Translation* 32. 2–25. <https://doi.org/10.26034/cm.jostrans.2019.154>
- MW = n.d. Merriam-Webster Dictionary. <https://www.merriam-webster.com> (last accessed on 20/03/2026).
- NDM = De Mauro, Tullio. 2026. *Il Nuovo De Mauro*. <https://dizionario.internazionale.it/> (last accessed on 20/03/2026).
- OED = 2000. *Oxford English Dictionary*. Oxford University Press. <http://dictionary.oed.com/entrance.dtl> (last accessed on 20/03/2026).
- Ortega, Lourdes. 2009. *Understanding Second Language Acquisition* (1st edition). London: Routledge. <https://doi.org/10.4324/9780203777282>
- Oxford, Rebecca L. 2002. Language Learning Strategies in a Nutshell: Update and ESL Suggestions. In Jack C., Richards & Willy A., Renandya (eds), *Methodology in Language Teaching: An Anthology of Current Practice*, 124–132. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511667190.018>
- Popel, Martin & Tomkova, Marketa & Tomek, Jakub & Kaiser, Lukasz & Uszkoreit, Jakob & Bojar, Ondřej & Žabokrtský, Zdeněk. 2020. Transforming Machine Translation: A Deep Learning System Reaches News Translation Quality Comparable to Human Professionals. *Nature Communications* 11(1). 4381. <https://doi.org/10.1038/s41467-020-18073-9>
- Pulcini, Virginia. 2023. *The Influence of English on Italian: Lexical and Cultural Features*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110755114>
- Russell, Stuart J. & Norvig, Peter. 2021. *Artificial Intelligence: A Modern Approach* (4th edition). Hoboken, NJ: Pearson. <https://ebookcentral.proquest.com/lib/unigeit/detail.action?docID=6563568>
- Sun, Sanjun & Liu, Kanglong & Moratto, Riccardo (eds). 2025. *Translation Studies in the Age of Artificial Intelligence* (1st edition). London: Routledge. <https://doi.org/10.4324/9781003482369>
- Taulli, Tom. 2019. *Artificial Intelligence Basics: A Non-Technical Introduction*. (1st edition). Berkeley: Apress. <https://doi.org/10.1007/978-1-4842-5028-0>
- Turing, Alan. 1950. Computing machinery and intelligence. *Mind* 49. 433–460.
- Vieira, Lucas Nunes. 2020. Machine translation in the news: A framing analysis of the written press. *Translation Spaces* 9(1). 1–21. <https://doi.org/10.1075/ts.00023.nun>
- ZING = Zingarelli, Nicola & Cannella, Mario & Lazzarini, Beata & Zaninello, Andrea & Canepari, Luciano. 2024. *Lo Zingarelli: vocabolario della lingua italiana* (12nd edition). Zanichelli. <https://u.ubidictionary.com/viewer/#/dictionary/zanichelli.lozingarelli16> (last accessed on 20/03/2026).