Article

The Necessity of Learning English in the Age of Artificial Intelligence: An Empirical Examination Based on Bibliometrics and Labour-Market Evidence

LIU Yang [1,*], ZHAO Lin [2]

[1] University of the Philippines, Quezon City, Philippines; liuyang@up.edu.ph
[2] Yangon University of Economics, Yangon, Myanmar; zhaolin@yu.edu.mm
* Correspondence: LIU Yang; liuyang@up.edu.ph

Abstract

The Journal of Sustainability Insights, vol. 1, no. 1, November 1, 2026. DOI: 10.00000/tjosi. This paper examines the claim that machines can already replace language learning by decomposing it into three testable propositions and testing them with bibliometric, paper-citation and cross-country evidence. Using OpenAlex data for 2000–2025, annual publications on artificial intelligence and language learning rose from single digits to 383 in 2025, and a Logistic fit places the inflection at about 2026.6, indicating continued expansion rather than saturation. A negative binomial regression shows that citation impact depends more on output type, author country and publication-year structure than on open access, so technical availability is not equivalent to knowledge value. A regression of EF EPI scores on GDP per capita across 29 economies shows that income alone explains only about 9% of proficiency variation, rising to 54% with regional fixed effects; China's negative residual reveals the mediating role of institutional and language-ecological factors. Drawing on theories of contextual layering, we argue that AI substitutes for language tasks in transactional contexts but complements human ability in interactional contexts. English's status as the lingua franca of science will not be eroded by translation technology in the foreseeable future. English education should therefore shift towards intercultural negotiation and higher-order expression, with provision tailored to regional language ecologies.

Keywords: artificial intelligence; English learning; bibliometrics; human capital; skill-biased technical change; lingua franca of science

Received: 26 September 2026
Revised: 9 October 2026 Accepted: 11 October 2026 Published: 1 November 2026 www.tjosi.com

Copyright: © 2026 by the authors. Licensee Lightup Innovation, United Kingdom, HR5 3DJ. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

1. Introduction

The quality curve of machine translation has risen steeply over the past decade. From early phrase-based statistical models, through neural machine translation exemplified by the Transformer, to large language models capable of generating coherent multilingual text on demand, the cost of cross-lingual information conversion has been compressed to near zero. At the operational level, a specialist document, a business email or the key points of a cross-national meeting can be translated into readable target language within seconds. Such changes have popularised a judgement: since machines can already translate, investing substantial time in learning English is no longer necessary—an extension of technological optimism into language education [1][2]. Yet equating language competence with information-conversion ability is itself a simplification that deserves scrutiny.

The actual picture does not match this simplification. Following the diffusion of generative AI, global demand for English proficiency has not weakened but has in several dimensions intensified. The proportion of multinational enterprises using English as a working language continues to rise; English-medium higher-education programmes keep expanding; researchers face undiminished pressure to publish in the English-dominated system of scholarly production and citation; and in the training, evaluation and prompt-engineering roles for AI products, precise English expression has become a scarce skill, with the productivity and quality effects of generative AI on knowledge workers now examined in several empirical studies [3][4][5]. A telling phenomenon is that the more public discourse proclaims the omnipotence of translation tools, the more hiring demands hard requirements for "complex communication in English". This tension is the point of departure for the present study: if machines truly substituted for language learning, we should observe a declining return to English proficiency, waning research attention and a loosening of English's position in knowledge production; if the opposite holds, the relationship between AI and language ability needs to be re-understood.

Existing discussions are mostly normative statements or experiential intuitions, lacking a systematic examination that places multiple observable dimensions within a single framework. On the one hand, research on the labour-market impact of AI largely takes task exposure and occupational displacement probability as its entry point, answering "which jobs will be automated" [6][7][8][9][10] rather than directly addressing "whether the return to language skills has changed". On the other hand, language-education and translation studies evaluate machine translation mainly in terms of translation quality and pedagogical application [11], rarely linking such micro-level evidence to macro-level structures of knowledge production. A bridge is missing between the two literatures: the former supplies macro evidence on technological shocks, the latter micro mechanisms of language function, yet few studies combine them into an answer to the overarching question of whether learning English is still necessary.

The formulation of the three propositions follows a common principle: each must rest on a set of observations that could refute it, otherwise it is no different from an unfalsifiable slogan such as "English is important". The economic-return proposition is grounded in the correlational structure and residual distribution of cross-country English proficiency and income; the technology–competence proposition in the lifecycle shape of research output and the mechanism of impact generation; and the knowledge-production proposition in the composition of publication languages. The three types of evidence come respectively from macro cross-sections, bibliometric micro-records and

database language statistics, and are independent in data source and identification logic, so there is no circularity in which the same dataset serves as both independent and dependent variable. This triangulation—a standard practice in bibliometric analysis [12][13]—allows the paper to reach a robust judgement even though any single data source is necessarily biased: if one type of evidence carries measurement error, the credibility of the conclusion will not collapse easily provided the other two point in the same direction. Conversely, if the three conflict, the conflict itself is a valuable finding, precisely exposing which part of the theoretical framework needs revision.

This paper aims to fill this gap. It decomposes the overarching question into three mutually reinforcing propositions: first, against the backdrop of AI diffusion, does English proficiency still yield economic returns in the labour market? Second, is the relationship between AI and human English competence substitutive or complementary? Third, is English's status as the lingua franca of science shaken by machine translation? Each proposition corresponds to one type of evidence. The first uses cross-country cross-sectional data to examine the relationship between English proficiency and economic development, with China as a case study of its residual structure. The second uses bibliometric and paper-level citation data to examine the growth pattern and impact-generation mechanism of the research field, indirectly judging the respective roles of technical tools and professional competence. The third uses data on the language composition of research output to test changes in English's share of knowledge production.

Methodologically the study emphasises reproducibility. All macro and bibliometric data come from public databases: OpenAlex provides annual publication series for topic searches, paper-level micro-records and language-grouped statistics [14]; EF Education First provides national English-proficiency indices [15]; and the World Bank provides GDP per capita [16]. The model specification includes growth-curve fitting, negative binomial regression [17][18] and least-squares regression with regional fixed effects, with parameters and diagnostic statistics reported in full. The paper does not seek to prove an unfalsifiable claim such as "English will always matter", but to offer empirical judgements that hold under specific periods and data definitions and can be overturned by later research. The remainder is organised as follows: Section 2 reviews the literature and builds the theoretical framework, deriving the three propositions; Section 3 describes the data and identification strategy; Section 4 reports the empirical results; Section 5 discusses the implications and policy lessons; and Section 6 concludes with limitations.

2. Literature review and theoretical framework

2.1. Returns to language ability from a human-capital perspective

Treating language ability as a form of human capital is the starting point for the economic analysis of foreign-language value. The human-capital theory established by Schultz (1961) [19] and Becker (1964) [20] regards education, training and health as investments yielding future returns; language skills possess distinctive asset properties because of their generality in cross-lingual transactions. Unlike specific skills, the value of language ability does not attach to a single employer or technology platform but can migrate across jobs, industries and countries. Such portability gives language ability an option-like function under uncertainty: when industrial structure shifts, speakers of a general language can cross job boundaries more easily.

Empirical research provides cross-country evidence on the economic returns to language ability. Ginsburgh and Prieto (2011) [21], using international adult literacy survey data, estimate that fluent

English yields significant wage premia in most countries, the premium rising with national openness. Azam et al. (2013) [22], using Indian labour-market data, find that English proficiency raises both employment probability and wage levels, especially in export-oriented sectors. These findings accord with the general conclusions of the economics of language [23]. Placed in the context of AI diffusion, a key question arises: will machine translation lower language barriers and thereby erode the scarcity of language ability? Human-capital theory does not give a unidirectional answer. Technology lowers the threshold for "accessing information" while simultaneously raising the value of "screening, negotiating and expressing oneself amid information overload". When machines can generate abundant usable text, people who can judge text quality, formulate precise requests and build trust in cross-cultural settings become scarcer. The return to language ability will therefore not simply vanish but undergo structural transfer: from translation-type returns towards judgement-type and negotiation-type returns.

2.2. Skill-biased technical change and task restructuring

Understanding AI's impact on language skills requires the analytical framework of skill-biased technical change (SBTC). Acemoglu and Autor (2011) [24] note that technical progress does not raise all workers' productivity uniformly; rather, its match with different tasks determines its distributional consequences in the labour market. When technology mainly replaces routine, codifiable tasks, workers performing those tasks lose, while those performing non-routine tasks requiring judgement and communication gain. Autor et al. (2003) [25], studying the US labour market, show that the diffusion of information-processing technology significantly altered task composition, with abstract reasoning and communication skills steadily rising in importance.

Applying this framework to language tasks yields a finer judgement than "translation is automated". Language activity is highly uneven: factual information transfer, format conversion and standardised summarisation are highly codifiable tasks in which machine translation and large language models have clear advantages; by contrast, negotiation, persuasion, implicit expression, humour and identity negotiation depend heavily on context, consensus and intention inference, where machines remain clumsy. The AI occupational-exposure measures proposed by Felten et al. (2018, 2021) [26][27] support this stratification: high exposure typically attaches to the algorithmically reproducible portion of a task bundle rather than to whole occupations. Webb (2020) [9] further notes that the net effect of automation depends on the relative force of technological displacement versus task creation through productivity gains. It follows that AI's impact on language skills is not wholesale substitution but task-level restructuring: the value of informational language tasks is compressed while that of relational and judgement-based language tasks is elevated. This inference provides the theoretical premise for Proposition 2 below.

2.3. The lingua franca of science and English's place in knowledge production

English's dominance in knowledge production is the structural factor most often underestimated when discussing "whether English still needs to be learned". Ammon (2001) [28] shows that English gradually displaced German, French and other languages as the dominant medium of cross-disciplinary communication in the second half of the twentieth century, a process mutually reinforced by the anglicisation of research funding, journal publishing and indexing systems. Bennett (2013) [29] characterises this role as the "lingua franca of science", stressing that its function is not

cultural hegemony per se but lowering coordination costs in multilateral communication: when researchers from different linguistic backgrounds share a set of terms and writing conventions, the efficiency of knowledge diffusion rises markedly (Mauranen, 2003 [30]).

Whether machine translation can erode this status depends on whether it can replace the coordinating function of a lingua franca. Existing evidence answers in the negative. Amano et al. (2016) [31] point out that language barriers in scientific communication extend far beyond text translation to include submission norms, the discourse conventions of peer review, the rhetorical structure of academic writing and the linguistic bias of citation networks; even when abstracts and full texts can be translated, these institutionally embedded language practices still default to English. Their research on non-native-English-speaking researchers further shows that language barriers persist in literature retrieval, review evaluation and international collaboration networks, and that the spread of machine translation has not eliminated inequality arising from differences in language ability. Melchers and Shaw (2003) [32] add from a World Englishes perspective that English's lingua-franca function has evolved diverse varieties: users need not target native-speaker accents but do need the pragmatic competence to participate in professional communities. Combined with the common bibliometric finding that English-language publications occupy central positions in citation networks, this suggests that translation technology has lowered the threshold for "being read" without changing the rules for "being heard". Proposition 3 thus acquires an empirical basis: English's structural position in knowledge production has institutional inertia and will not dissolve rapidly with the maturation of translation tools.

2.4. Contextual layering in communication: from linguistic to pragmatic competence

To explain why AI cannot substitute wholesale for language learning, a finer conception of language ability than "language = translation" is needed. Hymes (1972) [33] proposed the theory of communicative competence against a Chomskyan notion of linguistic competence, arguing that language use involves not only grammatical correctness but also appropriateness in specific social situations. Canale and Swain (1980) [34] subdivided it into grammatical, sociolinguistic, discourse and strategic competence, emphasising that speakers must select appropriate expressions in context and repair communication breakdowns. Swain's (1985) [35] output hypothesis further argues that language learning comes not only from input comprehension but from being "pushed to produce"; output forces learners to test hypotheses and notice gaps between form and meaning. Long's (1981) [36] interaction hypothesis places negotiation of meaning at the core of the acquisition mechanism: when comprehension difficulties arise in communication, the interaction through clarification, repetition and confirmation is precisely the moment when language ability grows.

These theories converge on a distinction: language activity can be arranged along a "contextual hierarchy". At one end lies the transactional context, where the goal is clearly defined information transfer or task execution and correctness can be measured against checkable standards; at the other lies the interactional context, where the goal is to build relationships, coordinate positions and manage face and identity, and correctness depends on shared context and trust. Krashen's (1985) [37] input hypothesis explains why the former is easily substituted by technology: sufficient comprehensible input alone can drive the development of comprehension ability, and machines excel at providing abundant, immediate, adjustable input. In interactional contexts, however, language ceases to be a transparent information channel and becomes the carrier of social relations. Vygotsky's (1978) [38]

sociocultural theory holds that higher language functions first form in interpersonal interaction and are then internalised as individual ability; Warschauer (2000) [39] and Kern (2006) [40] further show in computer-assisted language learning research that technology produces substantive effects on language development only when embedded in genuine social interaction. Ellis's (2003) [41] task-based teaching research likewise shows that tasks centred on meaning negotiation foster deeper acquisition than mechanical drills. Systematic reviews of this field can be found in Chapelle (2009) [42] and Heift and Schulze (2007) [43]. The analytical pivot of this paper follows: AI is naturally suited to transactional contexts, whereas interactional contexts still centre on human language ability; the relationship between the two depends on which level of the contextual hierarchy the language activity occupies.

2.5. A technology-complementarity / context-layering model and three propositions

Combining the three lines of thought above yields an analytical framework of "technological complementarity and contextual layering". The framework contains two dimensions: the horizontal axis is the level of context, from transactional to interactional; the vertical axis is the technology–human ability boundary, marking the relative advantages of AI and humans at each level. At the macro level, human-capital theory and skill-biased technical change explain why economic returns to language ability undergo structural transfer rather than simple disappearance; at the knowledge-production level, the lingua-franca theory explains why English's institutional position is sticky; at the micro level, communicative-competence and contextual-layering theories explain why substitution occurs only in specific activities. The three levels are not independent: macro returns arise from the market pricing of micro competences, while the language rules of knowledge production in turn shape the incentives for micro competence.

On this basis the paper proposes three propositions to be tested.

Proposition 1 (economic return): Against the backdrop of AI diffusion, English proficiency remains positively associated with economic development, but this association is mainly mediated by structural factors such as institutions, educational investment and language ecology rather than determined by income alone; after controlling for regional differences, income itself has limited explanatory power, and some developing economies exhibit systematic proficiency gaps relative to peers at the same income level.

Proposition 2 (technology–competence relationship): The relationship between AI and English proficiency is primarily complementary, with substitution secondary. The evidence is that research output in this field remains in an expansion rather than decline phase, and that paper impact depends more on output type and research quality than on the availability of tools; if the relationship were purely substitutive, we should observe a contraction of competence research following tool diffusion and a general rise in impact due to technical availability—neither of which is consistent with the data.

Proposition 3 (knowledge production): English's share of research output remains high over the long run and rises slowly, and machine translation has not shaken English's status as the lingua franca of science; the evolution of language composition displays strong path dependence, with limited growth in the share of non-English languages, indicating significant inertia in the language rules of knowledge production.

Each proposition is paired with observable falsification conditions. If cross-country regression shows that income's explanatory power rises rather than falls after regional controls, Proposition 1 is undermined; if growth fitting shows the topic has passed its inflection and entered decline, or if open access substantially and significantly raises citations, Proposition 2 needs revision; if the English share undergoes sustained, significant decline, Proposition 3 is overturned. Stating falsification conditions explicitly avoids designing the study as a confirmatory exercise that seeks only supporting evidence. The data and methods section explains how these propositions are tested as fairly as the given data allow. The correspondence between the three propositions, the two contextual levels and the macro–micro mechanisms is shown in Figure 1.

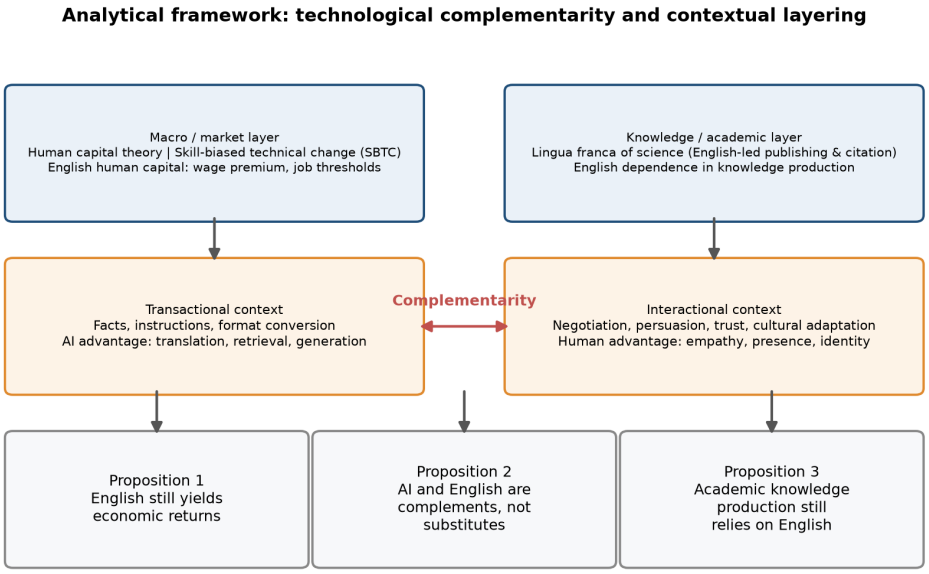

Figure 1. Analytical framework for the necessity of learning English in the AI era: the technological-complementarity and contextual-layering model.

3. Data and method

3.1. Data sources and statistical scope

Three types of public data are used. Bibliometric data come from OpenAlex, an open scholarly graph database covering journal articles, conference papers, preprints and book chapters, whose openness and reproducible API make it suitable infrastructure for bibliometric research (Priem et al., 2022 [14]). Two retrievals are performed on OpenAlex: first, annual publication series for five themes, with the search field restricted to title.search and a time span of 2000–2025; second, paper-level micro-records for the "AI and language learning" theme, about 1,300 records obtained by cursor pagination, with fields including publication year, citation count, output type, open-access status, first-author country, author count and journal source. The third type is the language composition of research output, also from OpenAlex language-grouped statistics, aggregating by year the counts of English, Chinese, Spanish, German, French, Japanese and Russian documents. The cross-

country section is assembled from two sources: English-proficiency scores come from the 2025 EF English Proficiency Index (EF EPI) released by EF Education First [15], based on large-scale standardised testing and covering more than one hundred countries and territories; GDP per capita (current US$) comes from the World Bank World Development Indicators [16], using 2023 values to align as closely as possible with the EPI survey timing. Matching by country name yields a valid sample of 29 economies.

Two points about statistical scope need clarification. OpenAlex's 2025 data are still being indexed, so descriptive statistics on global totals and language composition are cut off at 2024, while the topic-search series are retained through 2025 to capture the latest changes; the difference in year range is noted in the relevant figure captions. EF EPI scores are standardised index values, cross-nationally comparable but without absolute level meaning; only relative differences and regression residuals are used, without over-interpretation of the scores themselves. All data-retrieval scripts, cleaning code and model code are retained with the paper for replication.

3.2. Growth-curve fitting and extrapolation

Growth-curve fitting is a common tool for judging the lifecycle stage of a research field. The cumulative processes of technology adoption and scientific output often follow an S-shape, describable by Logistic and Gompertz functions. The former takes the form

$$N(t) = \frac{K}{1 + \exp\left(-\frac{t - t_0}{b}\right)}$$

, where K is the saturation ceiling, t_0 the inflection year and b the growth-rate parameter; the latter takes the form

$$N(t) = K \cdot \exp\left[-\exp\left(-\frac{t - t_0}{b}\right)\right]$$

, whose inflection occurs earlier and whose tail is flatter. Using the 2000–2025 annual publication series for the "AI and language learning" theme as the sample, the two models are estimated separately by nonlinear least squares, compared by coefficient of determination R^2, inflection year and parameter significance, and extrapolated to 2030. Extrapolation rests on strong assumptions and is used only to judge whether the inflection has arrived, not as a precise forecast.

3.3. Negative binomial regression of paper impact

Citation counts are a count variable with marked overdispersion: the mean in the micro sample is 10.52 while the standard deviation reaches 49.18, and roughly 60% of papers have zero citations. The Poisson model requires the conditional mean to equal the conditional variance, $E(Y_i) = \mathrm{Var}(Y_i)$, which does not fit this distribution, so a negative binomial model is used:

$$Y_i \sim \mathrm{NB}(\mu_i, \alpha), \qquad \mathrm{IRR} = \exp(\beta)$$

. With citation count as the dependent variable, the core explanatory variables are output type (article as base), first-author country group (China, USA, other, missing), the centred publication year and open-access status. The country grouping examines impact differences across knowledge-

production systems, year centring controls for literature ageing, and the open-access variable tests whether technical availability translates into impact advantage. The model reports incidence-rate ratios IRR $= \exp(\beta)$, robust standard errors, z-values and significance, and verifies the necessity of the negative binomial specification over Poisson via a likelihood-ratio (LR) test of the dispersion parameter α.

3.4. Cross-country regression of English proficiency and income

Proposition 1 is tested through two layers of cross-country regression. The base model is specified as

$$English_i = \beta_0 + \beta_1 \ln GDP_i + \varepsilon_i$$

, including only the log of GDP per capita to estimate the marginal association between income and English-proficiency score; the extended model adds regional fixed effects (Europe as base) to strip out region-level cultural, educational-institutional and language-ecological differences. Comparing the change in R^2 between the two models indicates how much of the cross-country variation in English proficiency can be explained by economic level and how much by regional structure. China, the largest developing economy in the sample with the greatest investment in English education, has its residual in the base model computed separately to see whether its proficiency exceeds or falls below prediction at the same income level. It should be stressed that the cross-sectional regression identifies correlation, not causation; the aim is to characterise structural association, not to assert that "income causes English proficiency".

3.5. Descriptive statistics of variables

Table 1 reports descriptive statistics of the main variables. The paper-level sample comprises 1,289 valid observations entering the model, and the cross-country sample 29 economies. The median citation count is zero while the mean is 10.52, confirming a highly right-skewed impact distribution; the mean author count is about 2.5 with a maximum of 100, reflecting the coexistence of big science and small-team research. In the cross-country sample the EF EPI score averages 530.24 with SD 66.88, highest for the Netherlands (624) and lowest for Thailand (402); log GDP per capita averages 9.80. The sample has sufficient variation in income and proficiency to provide identifying scope for the regression.

Table 1. Descriptive statistics of the main variables

Variable	Obs.	Mean	SD	Min	Median	Max
Panel A: paper-level sample (OpenAlex, AI and language learning theme)						
Citation count	1289	10.52	49.18	0	0	939
Author count	1289	2.47	4.25	0	2	100
Publication year	1289	2024.3	4.69	1980	2025	2026
Open-access share	1289	0.79	—	0	1	1
Panel B: cross-country sample (EF EPI 2025 × World Bank 2023)						
EF EPI score	29	530.24	66.88	402	526	624
GDP per capita (US$)	29	28621.3	24191.9	2138.8	17081.5	90984.4
Log GDP per capita	29	9.80	1.08	7.67	9.75	11.42

Note: the open-access share in the paper-level sample is a proportion, so no standard deviation is reported. The regional composition of the cross-country sample is 11 European, 8 Asian, 5 Latin American, 3 Middle Eastern and 2 African economies.

4. Results

4.1. Growth pattern of research attention and the inflection point

Figure 2 shows annual publication counts for the five search themes, 2000–2025. The curves exhibit different lifecycles: English-medium instruction (EMI) started earliest and has the largest base, dominating until 2024 and reflecting deep research accumulation in English as a medium of instruction; machine-translation quality and technology-enhanced language learning (TELL) rise slowly, as mature fields; ChatGPT English learning is a typical "event-driven" explosion, near zero before 2022 and surging from 2023. Compared with these reference themes, AI and language learning is the most dramatic: single digits throughout 2000–2017, accelerating after 2018, reaching 33 in 2022, jumping to 86 in 2023, 186 in 2024 and 383 in 2025. The public deployment of generative AI in late 2022 leaves a clear temporal imprint on this curve.

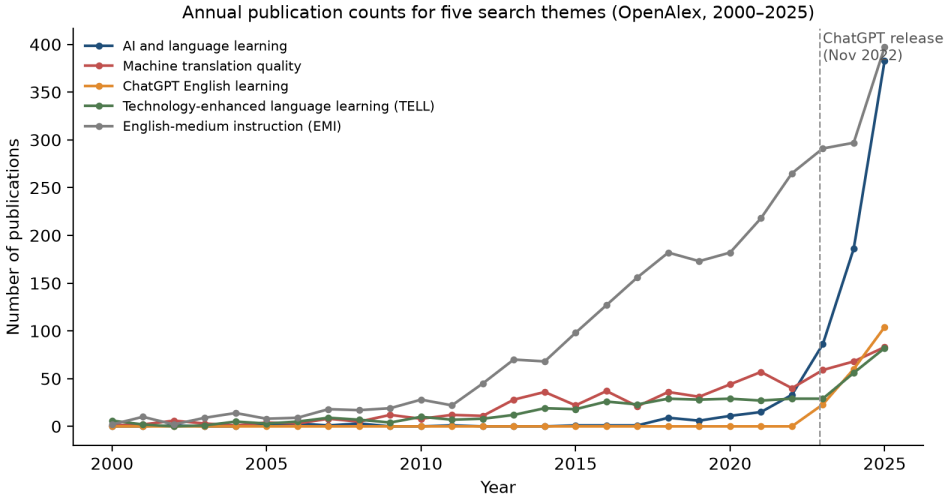

Figure 2. Annual publication counts for five search themes (OpenAlex, title search, 2000–2025).

Whether growth means "not yet saturated" or "about to peak" requires curve fitting. Figure 3 and Table 2 report estimates from the Logistic and Gompertz models. Both yield R^2 above 0.998, indicating that the S-curve captures the cumulative growth trajectory extremely well. The Logistic model estimates a saturation ceiling K of about 1,891 and an inflection in 2026.6, with about 946 publications at that point; the Gompertz inflection is later (about 2033.4), with a larger but insignificant ceiling parameter and weaker stability. Taking the more robust Logistic model as authoritative, the theme had not reached its inflection at the end of the observation period (383 publications in 2025) but lay in the high-growth segment before it. Extrapolation puts the count at about 703 and 1,096 in 2026 and 2027 respectively, approaching 1,789 around 2030 and gradually nearing saturation. This directly addresses the falsification condition for Proposition 2: if AI had already substituted for language-learning research, we should see output declining after passing the inflection, whereas the data show the opposite expansion. It must be reiterated that extrapolation rests on the strong assumption that the historical growth mechanism continues; the technological leap of generative AI may alter that mechanism, so these figures are directional judgements, not precise forecasts.

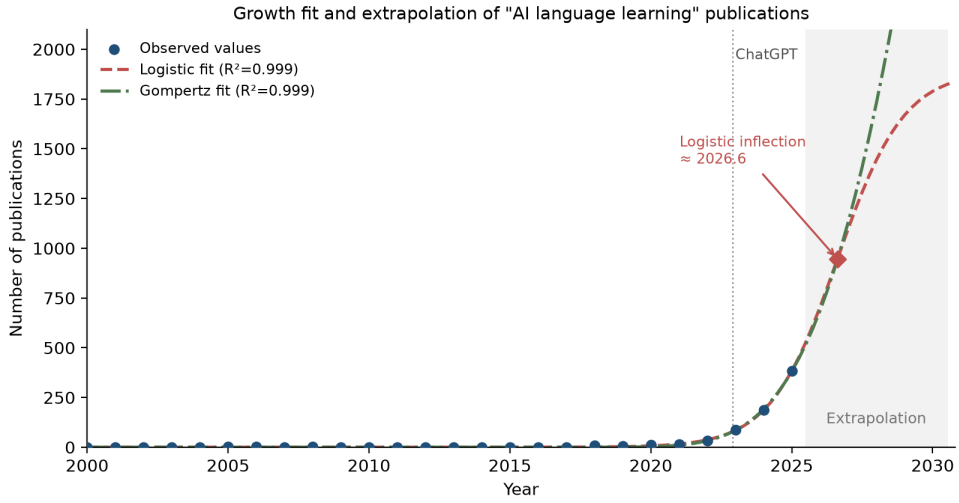

Figure 3. Growth fit and extrapolation of publications on "AI language learning" (shaded area = extrapolation).

Table 2. Growth-curve fitting parameters and extrapolation

Model	Ceiling K	Inflection year	Publications at inflection	R^2	2027 forecast	2030 forecast
Logistic	1891.1	2026.6	945.6	0.9992	1096.4	1788.7
Gompertz	10150.0	2033.4	7044.9	0.9985	1130.0	3368.2

Note: the sample is the 2000–2025 annual publication count for the "AI language learning" theme (n = 26). Logistic parameters are all significant ($p < 0.01$); the Gompertz ceiling is insignificant ($p = 0.311$) and less stable. Forecasts are in publications.

4.2. The generation mechanism of paper impact

Figure 4 and Table 3 report the negative binomial results. The dispersion parameter α is estimated at 4.91 ($p < 0.001$), and the likelihood-ratio test strongly rejects Poisson (LR = 40985.9, $p < 0.001$),

confirming significant overdispersion and the necessity of the negative binomial specification. With other factors controlled, output type is a stable predictor of impact: relative to journal articles, reviews have an incidence-rate ratio (IRR) of about 2.70 ($p < 0.001$), i.e. roughly 170% more citations on average; conference papers, preprints and other types have IRRs of 0.24, 0.31 and 0.45 respectively, all significantly below journal articles. This accords with the commonsense of academic communication: reviews occupy the node of knowledge integration, while preprints and conference papers sit earlier in the peer-review chain.

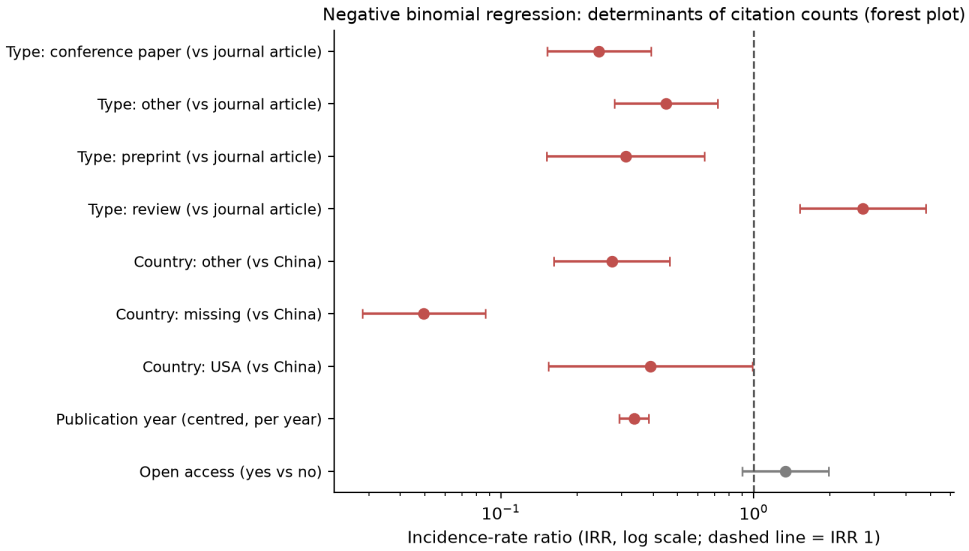

Figure 4. Incidence-rate ratios from the negative binomial regression (dots = IRR, horizontal lines = 95% confidence intervals, dashed line = IRR 1).

The country-grouping results require more careful scrutiny. With China as the base, papers with missing first-author country have an IRR of only 0.049 ($p < 0.001$), a huge gap; papers from "other" countries have an IRR of 0.274 ($p < 0.001$); the USA–China difference is marginally significant at 5% (IRR = 0.390, $p = 0.048$), with the US sample actually averaging fewer citations. Caution is warranted: country grouping confounds publication year, journal tier and disciplinary composition; the Chinese sample contains many recently published papers in high-growth fields, and citations are strongly age-dependent, as the centred year coefficient (IRR = 0.336, $p < 0.001$) shows—each subsequent year significantly lowers expected citations, a normal manifestation of literature ageing rather than declining quality. The open-access variable has an IRR of 1.33 but is statistically insignificant ($p = 0.151$), meaning that in this dataset open access brought no identifiable citation advantage. This has a direct implication for Proposition 2: technology (here represented by the availability of open access) lowers the threshold of access but does not automatically translate into academic impact, which remains governed by structural factors such as output type, review status and accumulated advantage. Between the availability of tools and the value of knowledge lies a gap technology cannot bridge by itself.

Table 3. Negative binomial regression results for citation counts

Variable	Coefficient β	IRR	Robust SE	z	p
Type: conference paper (base = journal article)	-1.410	0.244	0.241	-5.85	<0.001***
Type: other	-0.800	0.449	0.241	-3.33	<0.001***
Type: preprint	-1.168	0.311	0.367	-3.19	0.001**
Type: review	0.994	2.701	0.293	3.39	<0.001***
Country: other (base = China)	-1.294	0.274	0.268	-4.82	<0.001***
Country: missing	-3.008	0.049	0.286	-10.52	<0.001***
Country: USA	-0.941	0.390	0.475	-1.98	0.048*
Publication year (centred)	-1.091	0.336	0.067	-16.23	<0.001***
Open access (yes = 1)	0.288	1.333	0.200	1.44	0.151
Dispersion parameter α	4.910	—	0.276	17.81	<0.001***

Note: n = 1,289. Pseudo $R^2 = 0.0916$, AIC = 5007.9. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. IRR $= \exp(\beta)$, the incidence-rate ratio. Base groups are indicated in the variable names.

4.3. The language structure of knowledge production: scale expansion and English dominance

Proposition 3 concerns English's structural position in knowledge production. The left panel of Figure 5 shows the real growth in the number of documents indexed annually by OpenAlex: from about 3.39 million in 2000 to about 10.87 million in 2024, more than tripling over two decades, with only minor fluctuations around 2020. The aggregate stock of knowledge production is expanding rapidly, so any claim about "declining language importance" must be examined against this expansion. The right panel shows the English share in major academic indexes. This cross-sectional comparison is striking: 96.2% of records in Web of Science are in English, 90.8% in Scopus, 77.5% in the notably inclusive OpenAlex, while after manual record-by-record correction the true English share in OpenAlex is about 68% [44][45].

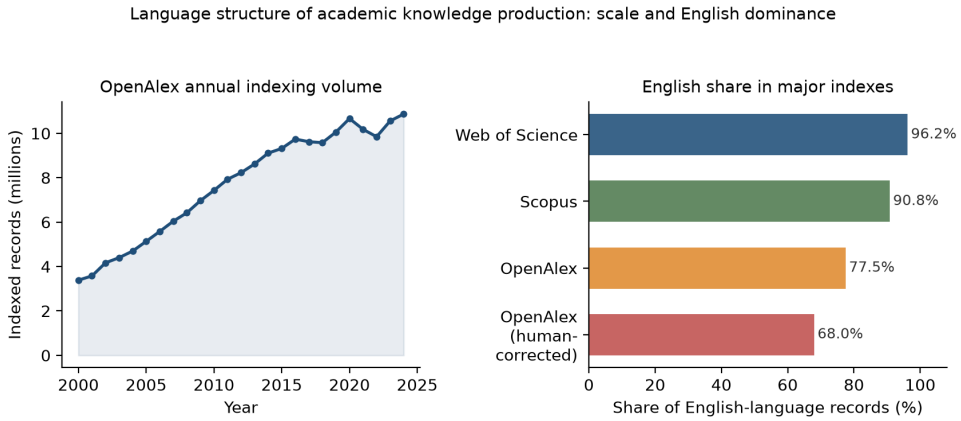

Figure 5. The language structure of academic knowledge production: left, annual OpenAlex indexing volume (2000–2024, retrieved data); right, English share in major indexes (reported values from the literature).

These figures convey two messages. First, whichever definition is used, English is the overwhelmingly dominant language of academic knowledge production; even in OpenAlex, the most friendly

to non-English and criticised for language-metadata misclassification, English documents account for roughly 70–80%, over two-thirds after correction. Second, the differences between databases themselves indicate that the English share is largely a function of indexing policy and metadata quality rather than a pure reflection of language preference in knowledge production. For this study, the key point is not the precise value but the robustness of direction: during the years when translation technology spread, English's dominance showed no identifiable loosening. Machine translation makes it easier for non-English readers to "read" English documents but has not led English authors to abandon English writing, because the institutions of publication, citation and evaluation still default to English. Proposition 3 is thus supported: English's status as the lingua franca of science has deep institutional inertia and will not collapse in the short term with the maturation of translation tools. Table 4 summarises the cross-database comparison.

Table 4. English-language share in major academic indexes

Index	English share	Scope	Source
Web of Science	96.2%	Records, 2015–2024	Arroyo-Machado et al. (2026)
Scopus	90.8%	Records, 2015–2024	Arroyo-Machado et al. (2026)
OpenAlex	77.5%	Metadata-labelled	Arroyo-Machado et al. (2026)
OpenAlex (human-corrected)	68.0%	After validating 6,836 records	Céspedes et al. (2024)

Note: the year windows and sampling definitions of the studies differ, so the percentages should not be compared precisely; they serve only to illustrate English's general dominance across indexing systems. OpenAlex language metadata carries about 14.7% misclassification, with Chinese and Korean especially weakly identified [46].

4.4. English proficiency and economic development: cross-country evidence

Proposition 1 examines the economic correlates of English proficiency. Figure 6 plots EF EPI scores against log GDP per capita for 29 economies with an overlaid OLS fit. The scatter shows clear regional clustering: European countries cluster in the upper right (high income, high proficiency), Asian and Latin American countries are more dispersed, and African samples, though very low in income, are not at the bottom in proficiency—South Africa, Nigeria and Zimbabwe retain relatively high English proficiency for historical and institutional reasons. The fitted line slopes upward, but $R^2 = 0.092$, indicating that income alone explains less than a tenth of cross-country variation. Table 5 reports the coefficients: in the base model the coefficient on log GDP per capita is 18.80, marginally significant at 10% ($p = 0.085$); with regional fixed effects (Model 2) the income coefficient falls to 12.28 and becomes insignificant ($p = 0.233$), while most regional dummies are highly significant—relative to Europe, Asia is 104.8 points lower, Latin America 101.0 lower and the Middle East 150.9 lower (all $p < 0.001$). Explanatory power rises from 9.2% to 54.4%.

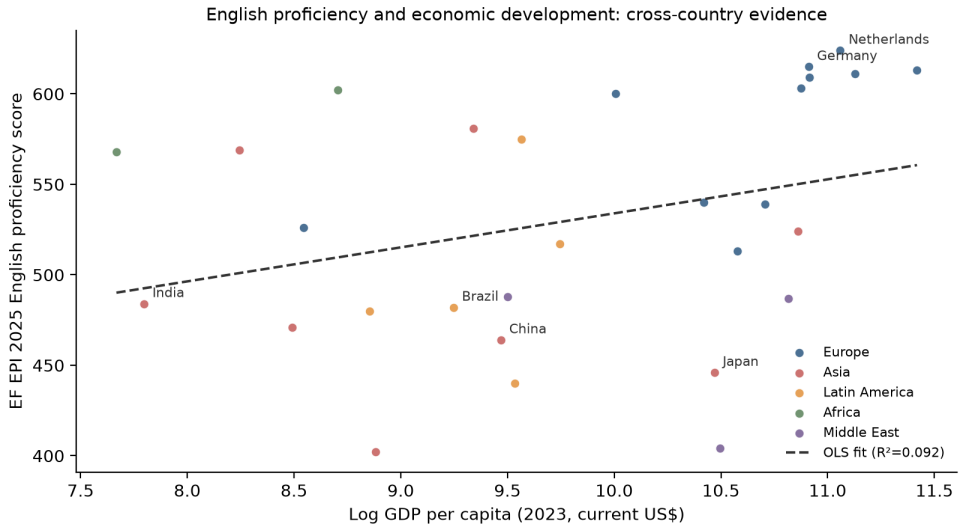

Figure 6. English proficiency and economic development: cross-country evidence from 29 economies (OLS fit $R^2 = 0.092$).

This contrast is the core evidence for Proposition 1. Cross-country differences in English proficiency are mainly not a direct mapping of income differences but products of region-level educational institutions, language policies and historical paths. Rising income does not automatically improve proficiency, as the case of China makes especially clear: in Model 1 (Table 6) China's residual is -59.98, meaning that at the same GDP per capita its proficiency is about 60 points below prediction. As a country with enormous educational investment and a vast population of English learners, China's negative residual is not a lack of willingness to learn but suggests a structural mismatch between input and output—exam-oriented learning has accumulated substantial "knowledge-type" English competence without fully converting into the practical communicative proficiency measured by the EPI. This qualifies Proposition 1 importantly: English proficiency does have economic value, but the efficiency of its supply depends heavily on institutions and pedagogy; in an era when AI can supply abundant comprehensible input, the marginal return on investment centred on information-conversion training may fall further.

Table 5. Cross-country regression of English proficiency (EF EPI) on GDP per capita

Variable	Model 1: income only	Model 2: income + regional FE
Log GDP per capita	18.80† (10.91)	12.28 (10.29)
Region: Asia (base = Europe)	—	-104.75*** (28.28)
Region: Europe	—	Base group
Region: Latin America	—	-100.98*** (25.68)
Region: Middle East	—	-150.93*** (35.78)
Constant	345.98** (107.45)	484.46*** (84.65)
R^2 / adjusted R^2	0.092 / 0.058	0.544 / 0.445
Observations	29	29

Note: robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, † $p < 0.10$. The African sample (n = 2) is too small for its own dummy and is merged into the base group, making the regional coefficients conservative.

Table 6. English proficiency, income and regression residuals for selected economies

Economy	EPI score	GDP per capita (US$)	Model 1 residual
Netherlands	624	63515.6	Above prediction
Germany	615	54776.8	Above prediction
Malaysia	581	11386.0	Significantly above prediction
Philippines	569	3804.1	Significantly above prediction
China	464	12951.2	-59.98 (below prediction)
Japan	446	35215.0	Below prediction
Saudi Arabia	404	36156.9	Significantly below prediction

Note: residuals are computed from Model 1 (log GDP per capita only). The full sample and values are in the accompanying data files. The positive residuals of middle-income economies such as Malaysia and the Philippines show that language policy can markedly alter proficiency at the same income level.

5. Discussion

5.1. Where the three propositions stand

Setting the four sets of empirical results against the three propositions of Section 2 yields a fairly clear judgement. Proposition 1 (economic return) receives qualified support: cross-country differences in English proficiency are positively associated with income, but income alone explains less than a tenth of the variation and regional institutional factors dominate; China's negative residual further shows that even with huge investment, poor pedagogy can leave proficiency below the prediction for its income level. The correct formulation is therefore not "English proficiency brings returns" but "effective institutional provision of English proficiency brings returns". Proposition 2 (complementarity rather than substitution) receives stronger support: output in the AI and language-learning theme lies in the pre-inflection expansion phase, and the availability of open access does not translate into impact advantage—both point to technology lowering thresholds without substituting for professional competence. Proposition 3 (knowledge production depends on English) receives robust support: whatever the indexing definition, English accounts for more than two-thirds of academic documents, with no decline during the period of translation-technology diffusion. Together the three propositions constitute a systematic rebuttal of the claim that "English need not be learned in the AI era", but the rebuttal does not assert that English will always matter; rather it reveals that the claim's underlying premise—"language = information conversion"—does not hold.

5.2. The boundary of substitution: transactional and interactional contexts

The theoretical pivot of this paper is the distinction between transactional and interactional language activity. The empirical results align closely with it. What machine translation and large language models genuinely excel at are transactional tasks: fact retrieval, format conversion, standardised summarisation and first-draft generation. These tasks are highly codifiable and their correctness checkable—precisely the "algorithmically reproducible portion of task bundles" described in the skill-bias literature. In such contexts, insisting on defining English learning as "memorising

words, practising translation and writing standardised texts" will indeed be largely displaced by technology—what is displaced is not English itself but this narrow conception of competence. Interactional contexts present a different picture. Cross-national negotiation, trust-building in multicultural teams, peer review and identity negotiation in academic communities have no single criterion of correctness and depend on shared context and pragmatic judgement. Hymes's (1972) [33] theory of communicative competence and the four-part model of Canale and Swain (1980) [34] long ago established that the core of language use is appropriateness rather than mere grammatical correctness; Swain's (1985) [35] output hypothesis and Long's (1981) [36] interaction hypothesis further show that competence grows through being "pushed to produce" and through "negotiating meaning". Generative AI can serve as an input source and a practice partner but can hardly replace face management, intention inference and assumption of responsibility in real interaction—precisely the "context-embedded" competence repeatedly stressed in the humanities and social sciences. The high citation rate of review articles in the negative binomial regression (IRR = 2.70) is a further indication: knowledge integration and judgement, not information hauling, are the true source of impact. The more AI can mass-produce informational content, the scarcer and more valuable human judgement and expression in interactional contexts become.

5.3. Implications for English education

The conclusions have direct implications for the practice of English education, align with research on ChatGPT in language teaching [47], and resonate with Dwivedi et al. (2023) [48] and Kasneci et al. (2023) [49] on generative AI and education as well as with debates over assessment reform in higher education [50]. First, curriculum goals should shift from "information-conversion ability" to "higher-order expressive ability". When machines can translate and summarise instantly, scarce educational resources should no longer go to mechanical language drills but to argumentation, persuasion, intercultural negotiation and academic writing—the core skills of interactional contexts and the part hardest for AI to replace. Second, the role of machine translation should be repositioned: as a provider of comprehensible input (Krashen, 1985 [37]) and a scaffold that lowers anxiety, not a shortcut that replaces students' productive effort. Vygotsky's (1978) [38] zone of proximal development suggests that technology promotes development only when it supports students in completing tasks just beyond their reach; if technology performs the output for them, the form–meaning gap checking emphasised by the output hypothesis will not occur and competence growth stalls. Third, regional differences require differentiated provision. The cross-country regression shows significant proficiency gaps in Asia, Latin America and the Middle East relative to Europe, and China's negative residual further indicates a structural problem of "high investment, low efficiency". For these regions the key to raising returns is not more class hours but reforming assessment, adding authentic interactive tasks and embedding technology as an input scaffold within task-based teaching (Ellis, 2003 [41]).

5.3.1. Three entry points for curriculum restructuring

Translating these principles into actionable arrangements, at least three aspects are worth trying. In assessment, the weight of productive and authentic-context tasks should be raised while information-conversion items that machines can complete instantly should be reduced; since summarisation and translation are handled by tools, assessment should instead require learners to argue

a reasoned position on contested issues and respond to interlocutors' implicit demands in simulated cross-cultural negotiations. In task design, a "machine first, human judges" stage can be introduced: students first obtain a machine-generated translation or draft, then identify its pragmatic infelicities and cultural biases and propose revisions—a task that turns machine output into material for critical analysis and precisely trains the judgement needed in interactional contexts. In the teacher's role, teachers shift from sole transmitters of knowledge to designers of interactive situations and feedback providers, their value lying in immediate responses to student expression and scaffolding support. All three adjustments aim at the same goal: reallocating classroom time from informational tasks machines can already perform to relational and judgement-based tasks they cannot. This also addresses the substance of Proposition 2—complementarity does not happen automatically; it requires curricula and assessment institutions actively to direct human effort towards the complementary side.

5.4. Why "translation omnipotence" recurs

A phenomenon worth explaining is that the claim "machines can translate, so English need not be learned" resurfaces every few years in new technological guise. From early online dictionaries and statistical machine translation, through neural machine translation, to today's large language models, each technological leap triggers a fresh round of prophecies about "the end of foreign-language learning". This periodicity suggests that the problem lies not only in technology's capability boundary but also in how language is understood. The cognitive root of translation omnipotence is the assumption that language is a symbol system that can be losslessly compressed and automatically converted. In this picture the entire value of language is packed into the single function of "information transfer", so once machines take over transfer the remainder of language is treated as disposable redundancy. Acemoglu and Restrepo (2018) [6], on the task boundary of automation, show precisely why such simplification is dangerous: technology replaces the codifiable fragments of task bundles, never whole occupations; equating language with a codifiable fragment amounts to reducing human language activity in advance to the small slice machines handle and then declaring the rest irrelevant.

Supporting this simplification is also an optimism about technology diffusion that assumes that the diffusion of an ability equals its disappearance—an expectation reinforced by the "perceived usefulness" logic of the technology acceptance model [51]. Goldfarb and Tucker (2019) [52], discussing digital economics, note that the value distribution of information technology is never uniform: when information becomes cheap, the scarce abilities complementary to it appreciate. Machine translation makes cross-lingual information cheap, so the abilities to screen information, judge translations and respond appropriately in cross-cultural settings become expensive. The insignificance of open access in our negative binomial regression is an empirical footnote to this logic—greater availability did not automatically create value, which still resides in judgement and integration. Seen in this light, the deeper reason translation omnipotence recurs is that it misreads "falling cost of acquiring language ability" as "vanishing value of language ability". The two are not logically equivalent yet are repeatedly conflated in public discussion. Clarifying this conflation is itself one contribution of the study: it reminds educators and learners that every technological advance should prompt a renewed question of "what exactly is being learned" rather than a hasty denial of "whether to learn at all".

5.5. Limitations and research prospects

The study has several limitations that counsel restraint in interpretation. The bibliometric part relies on OpenAlex title searches, which miss relevant studies whose titles lack the search terms and whose recall is affected by terminological evolution; publication counts across different search terms should not be compared absolutely. The paper-level sample covers only the single theme "AI language learning"; citation counts are strongly influenced by discipline, journal and publication age, and although the negative binomial regression controls for year and type it cannot fully strip out these structural factors, so the country coefficients require special caution. Owing to OpenAlex's free-tier quota, precise year-by-year language-grouped statistics were unavailable, so the right panel of Figure 5 uses cross-sectional values reported in the literature; the study therefore cannot directly test the time trend in "English share" but only argue for the robustness of its structural dominance. The cross-country sample covers only 29 economies and the EPI–GDP match involves a year mismatch, so the conclusions are correlational rather than causal. Future research could advance in three directions: first, obtain complete year-by-year OpenAlex language metadata with misclassification corrected to estimate the English share trend directly; second, introduce wage microdata to estimate heterogeneous changes in the English premium across occupations with differing AI exposure, moving Proposition 1 from cross-country correlation to individual-level causality; third, design teaching experiments comparing "machine-translation-as-scaffold task-based teaching" with traditional instruction in developing interactional-context competence.

6. Conclusion

Centred on the dispute over "whether English still needs to be learned in the AI era", this paper constructs a technology-complementarity / context-layering framework, decomposes it into three testable propositions and examines each with bibliometric, paper-citation and cross-country evidence. The results show that output in the AI-and-language-learning theme remains in rapid pre-inflection expansion and has not declined with the maturation of translation technology; open access lowers knowledge availability without automatically producing impact advantage, leaving a gap between tool availability and knowledge value that institutions and competence must fill; English retains a dominant share above two-thirds of academic knowledge production and its lingua-franca status carries marked institutional inertia; and although English proficiency is positively associated with economic development, income alone explains less than a tenth of the variation, the real differences coming from regional institutions and language policy, with China's negative residual especially illustrating structural mismatch between "high investment and low conversion".

Together these findings support a conclusion that transcends the binary of "learn or not learn": AI displaces the information-conversion function of language, not its social-interaction and judgement functions. When English learning is narrowed to translation and standardised-text drills, it does indeed depreciate sharply under technological shock; when understood as the ability to express precisely, negotiate across cultures and integrate knowledge in interactional contexts, it becomes scarcer and more valuable in the AI era. For learners the pragmatic strategy is not to abandon English but to restructure the learning goal—assign informational tasks machines excel at to machines and direct human effort towards judgement, negotiation and expression. For educators the curricular focus should move from knowledge-type input to output-oriented authentic tasks, treating generative AI as a scaffold for comprehensible input rather than a substitute for output. For policymakers the

key to raising returns to English education lies in institutions and pedagogy, not in the sheer scale of investment. The more powerful translation technology becomes, the more structural rather than temporary is humanity's need for high-quality language ability—perhaps the most counter-intuitive yet best evidence-supported judgement in the entire dispute.

Author Contributions: Conceptualization, L.Y. and Z.L.; methodology, L.Y.; software, L.Y.; validation, L.Y. and Z.L.; formal analysis, L.Y.; investigation, L.Y.; resources, Z.L.; data curation, L.Y.; writing—original draft preparation, L.Y.; writing—review and editing, Z.L.; visualization, L.Y.; supervision, Z.L.; project administration, L.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings are available from the corresponding author upon reasonable request.

Acknowledgments: The authors thank the editors and reviewers for their constructive comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Frey, C. B.; Osborne, M. A. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change* **2017**, *114*, 254–280.
2. Brynjolfsson, E.; McAfee, A. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*; W. W. Norton & Company: New York, 2014.
3. Brynjolfsson, E.; Li, D.; Raymond, L. *Generative AI at Work*; National Bureau of Economic Research: Cambridge, MA, 2023; NBER Working Paper No. 31161.
4. Noy, S.; Zhang, W. Experimental evidence on the productivity effects of generative artificial intelligence. *Science* **2023**, *381*(6654), 187–192.
5. Dell'Acqua, F.; Mcgowan, J. S.; Mollick, E.; et al. *Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality*; Harvard Business School: Boston, 2023; Working Paper No. 24-013.
6. Acemoglu, D.; Restrepo, P. *Artificial Intelligence, Automation and Work*; National Bureau of Economic Research: Cambridge, MA, 2018; NBER Working Paper No. 24196.
7. Acemoglu, D.; Restrepo, P. Robots and jobs: Evidence from US labor markets. *J. Political Econ.* **2020**, *128*(6), 2188–2244.
8. Autor, D. H. Why are there still so many jobs? The history and future of workplace automation. *J. Econ. Perspect.* **2015**, *29*(3), 3–30.
9. Webb, M. *The Impact of Artificial Intelligence on the Labor Market*; Stanford Graduate School of Business: Stanford, 2020.
10. Agrawal, A.; Gans, J. S.; Goldfarb, A. *Prediction Machines: The Simple Economics of Artificial Intelligence*; Harvard Business Review Press: Boston, 2018.
11. Godwin-Jones, R. Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology* **2021**, *25*(2), 4–19.
12. Donthu, N.; Kumar, S.; Mukherjee, D.; et al. How to conduct a bibliometric analysis: An overview and guidelines. *J. Bus. Res.* **2021**, *133*, 285–296.

13. Broadus, R. N. Toward a definition of "bibliometrics". *Library Quarterly* **1987**, *57*(4), 373–379.
14. Priem, J.; Piwowar, H.; Orr, R. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint* **2022**, arXiv:2205.01833.
15. EF Education First. *EF English Proficiency Index 2025*; EF Education First: London, 2025.
16. The World Bank. *World Development Indicators*; The World Bank: Washington, DC, 2025.
17. Hilbe, J. M. *Negative Binomial Regression*, 2nd ed.; Cambridge University Press: Cambridge, 2011.
18. Cameron, A. C.; Trivedi, P. K. *Regression Analysis of Count Data*, 2nd ed.; Cambridge University Press: Cambridge, 2013.
19. Schultz, T. W. Investment in human capital. *Am. Econ. Rev.* **1961**, *51*(1), 1–17.
20. Becker, G. S. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*; Columbia University Press: New York, 1964.
21. Ginsburgh, V.; Prieto, R. Returns to foreign languages of native workers in the European Union. *Ind. Labor Relat. Rev.* **2011**, *64*(3), 663–681.
22. Azam, M.; Chin, A.; Prakash, N. The returns to English-language skills in India. *Econ. Dev. Cult. Change* **2013**, *61*(2), 335–367.
23. Chiswick, B. R.; Miller, P. W. *The Economics of Language: International Analyses*; Routledge: London, 2007.
24. Acemoglu, D.; Autor, D. H. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*; Ashenfelter, O., Card, D., Eds.; Elsevier: Amsterdam, 2011; Vol. 4B, pp. 1043–1171.
25. Autor, D. H.; Levy, F.; Murnane, R. J. The skill content of recent technological change: An empirical exploration. *Q. J. Econ.* **2003**, *118*(4), 1279–1333.
26. Felten, E. W.; Raj, M.; Seamans, R. A method to estimate the effect of artificial intelligence on work. In *Proceedings of the 26th European Conference on Information Systems*; ECIS: Portsmouth, 2018.
27. Felten, E. W.; Raj, M.; Seamans, R. Occupational, industry, and geographic exposure to artificial intelligence: A new dataset and its potential uses. *Strateg. Manage. J.* **2021**, *42*(12), 2195–2217.
28. Ammon, U. *The Dominance of English as a Language of Science: Effects on Other Languages and Language Communities*; Mouton de Gruyter: Berlin, 2001.
29. Bennett, K. *English as a Lingua Franca in Academia: A Corpus-Based Study of Research Articles and Other Academic Genres*; Cambridge University Press: Cambridge, 2013.
30. Mauranen, A. *English as a Lingua Franca in Academic Settings*; Narr: Tübingen, 2003.
31. Amano, T.; González-Varo, J. P.; Sutherland, W. J. Languages are still a major barrier to global science. *PLoS Biol.* **2016**, *14*(12), e2000933.
32. Melchers, G.; Shaw, P. M. *World Englishes: An Introduction*; Arnold: London, 2003.
33. Hymes, D. H. On communicative competence. In *Sociolinguistics*; Pride, J. B., Holmes, J., Eds.; Penguin: Harmondsworth, 1972; pp. 269–293.
34. Canale, M.; Swain, M. Theoretical bases of communicative approaches to second language teaching and testing. *Appl. Linguist.* **1980**, *1*(1), 1–47.
35. Swain, M. Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In *Input in Second Language Acquisition*; Gass, S. M., Madden, C. G., Eds.; Newbury House: Rowley, MA, 1985; pp. 235–253.
36. Long, M. H. Input, interaction, and second-language acquisition. In *Native Language and Foreign Language Acquisition*; Winitz, H., Ed.; New York Academy of Sciences: New York, 1981; pp. 259–278.
37. Krashen, S. D. *The Input Hypothesis: Issues and Implications*; Longman: London, 1985.
38. Vygotsky, L. S. *Mind in Society: The Development of Higher Psychological Processes*; Harvard University Press: Cambridge, MA, 1978.
39. Warschauer, M. Technology and English language teaching. In *Handbook of English Language Teaching*; Oxford University Press: Oxford, 2000.

40. Kern, R. F. Language learning and technology: A theoretical and empirical analysis of their relationship. *Comput. Assist. Lang. Learn.* **2006**, *19*(4), 283–297.
41. Ellis, R. *Task-Based Language Learning and Teaching*; Oxford University Press: Oxford, 2003.
42. Chapelle, C. A. *The Cambridge Guide to Computer Assisted Language Learning*; Cambridge University Press: Cambridge, 2009.
43. Heift, T.; Schulze, D. *Errors and Intelligence in Computer-Assisted Language Learning: Parsers and Pedagogues*; Routledge: London, 2007.
44. Arroyo-Machado, W.; Forchino, M. V.; Torres-Salinas, D. *Web of Science, Scopus, and OpenAlex: A Comparative Bibliometric Analysis of Coverage, Disciplinary Profiles, and Database Convergence*; Zenodo, 2026.
45. Céspedes, L.; et al. Evaluating the linguistic coverage of OpenAlex: An assessment of metadata accuracy and completeness. *arXiv preprint* **2024**, arXiv:2409.10633.
46. Mongeon, P.; et al. Investigating document type, language, and publication year and author count discrepancies between OpenAlex and the Web of Science. **2025**.
47. Kohnke, L.; Moorhouse, B. L.; Zou, D. ChatGPT for language teaching and learning. *RELC J.* **2023**, *54*(2), 537–550.
48. Dwivedi, Y. K.; Ismagilova, E.; Hughes, D. L.; et al. "So what if ChatGPT wrote it?" Examining multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *Int. J. Inf. Manage.* **2023**, *71*, 102642.
49. Kasneci, E.; Sessler, K.; Küchemann, S.; et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learn. Individ. Differ.* **2023**, *103*, 102274.
50. Susnjak, T. ChatGPT: The end of traditional assessments in higher education? *J. Appl. Learn. Teach.* **2022**, *5*(1), 9–21.
51. Davis, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **1989**, *13*(3), 319–340.
52. Goldfarb, A.; Tucker, C. Digital economics. *J. Econ. Lit.* **2019**, *57*(1), 3–43.

Disclaimer/Publisher's Note: All statements, views and data in published papers belong to the named authors and contributors, not to Lightup Innovation or the editors. Neither Lightup Innovation nor the editors accept liability for loss or damage arising from ideas, methods, instructions or products mentioned in the content.