Vraila, Stavroula Paraskevi. 2026. Political Translation in the AI Era: Journalists as Dual Agents and Systemic Vulnerabilities in Cross-Linguistic Political Communication. In: L10N Journal. 1(5), pp. 39-61.

Political Translation in the AI Era: Journalists as Dual Agents and Systemic Vulnerabilities in Cross-Linguistic Political Communication

Stavroula Paraskevi Vraila

Ionian University

svraila@ionio.gr

Abstract

This study examines how German political statements change when rendered for Greek audiences and what AI systems return for the same material. A corpus of sixteen documented cases from 2010 to 2025, thirteen in which the Greek departs from the German and three in which it holds, yields a taxonomy of five patterns: context isolation, systematic tone modification, narrative-driven translation, weaponization under polarization, and semantic manipulation. Three of the cases were submitted to GPT-4o and GPT-5.3 Instant under five prompt conditions, three runs per cell, and once each to DeepL and Google Translate, giving 96 coded outputs. Under the two conditions that ask only for a translation, seventeen of eighteen language model outputs in each carry none of the five patterns in the author's coding, as do five of the six machine translation outputs. An instruction to translate for a Greek news website returns adaptations in fifteen of eighteen outputs, against one of eighteen under a bare instruction, and no headlines.

Keywords: political translation, journalistic translation, neural machine translation, large language models, prompt design, cross-cultural communication

1 Introduction

The translation of political discourse is an ideological practice, and translational decisions are shaped by institutional constraints and by the norms of the context in which the text is received (Schäffner 2012). In journalism, the translator is seldom a separate figure. Journalists who translate act as what Bielsa and Bassnett (2009) call double agents, and the texts they produce reframe political statements for a new readership.

AI systems now take part in that work. In a survey of Greek journalists, 51% regarded AI as useful for translation and 32% reported using it (Triantafyllou et al. 2025).

The article compares the patterns documented in human journalistic translation with what AI systems return when the task put to them changes, and asks three questions. Which recurrent patterns characterize problematic translation of political statements in journalism? Do AI systems reproduce these patterns, and under which instructions? Which institutional, temporal and technological conditions make political mistranslation systemic?

2 Political Translation, News Translation and the Platform Environment

Venuti (1995) argued that the ideal of fluent, transparent translation conceals the translator's interpretive work, a premise recently revisited by Freeth and Treviño (2024) to account for contemporary sociotechnical and institutional settings, while Baker (2006) describes translation as renarration, through which a text foregrounds particular actors and evaluative frames. Schäffner (2004, 2012) ties such choices in political texts to the constraints that govern what is translated and how political actors are framed. Lexical choices that look neutral carry evaluative weight that accumulates across a text (Munday 2012).

News translation research, a recognizable subfield for some twenty-five years (Valdeón 2025), sits between journalism studies and translation studies, which have not conceptualized translation in the same terms (Valdeón 2023); Kalantari (2022) locates journalistic translation where the two meet. This article draws on both, taking its linguistic categories from one and its account of editorial practice from the other.

Machine translation had been shown to reproduce gender bias before language models entered the newsroom (Prates et al. 2020). The closest published work is Conway et al. (2025), who argue that AI-based machine translation has changed the conditions under which journalists could be assumed to compose news texts and not translate them, and who measure lexical distance between parallel corpora of AI-generated news translation. The present study shares the premise and differs in method: a distance measure registers how far wording diverges, whereas a controlled design can tie a given shift to the instruction that produced it.

Automatic translation also operates after publication. On social media, the translation offered to a reader depends on a language-identification step that can misfire, and at least one platform gives the reader no way to contest the translation it displays (Paullada and Haroutunian 2026). The journalist is then one translating agent among several between the source and the reader's screen.

Newsrooms that test headlines select among variants by click-through rate (Hagar and Diakopoulos 2019), and a headline can depart from the source while the body beneath

it stays closer. This article uses quality in the narrower sense of fidelity to the source statement and treats engagement as a separate criterion that need not coincide with it.

3 Methodology

3.1 Research design, corpus and analytical framework

The study combines a corpus of documented cases in which German political statements were rendered for Greek audiences, analyzed qualitatively, with a set of AI translation outputs produced under controlled conditions and coded. Framing analysis (Entman 1993) supplies the account of how a rendering positions actors and assigns responsibility, and the comparison of source and rendering follows House's (2015) procedure for setting a translation against its original along lexical, syntactic and pragmatic dimensions.

The corpus holds sixteen cases from Greek coverage of German political speech between 2010 and 2025: thirteen in which the Greek departs from the German, dated between 2011 and 2025, and three in which it holds. The window opens in 2010, with the onset of the debt crisis. Cases were selected on four criteria: a documented diplomatic response or public controversy; an available German source; Greek press coverage carrying the disputed rendering; and a difference between source and rendering that can be described linguistically. Table A1 in the appendix gives each case with its German segment, its Greek rendering, the outlet that carried it and the patterns coded for it. The German wording of fifteen cases was confirmed verbatim against the recording, the official statement or a German press report quoting it. For one of the three faithful cases only a secondary account of the German was available.

The number of outlets indicates how widely a rendering circulated: five of the thirteen are documented in three or more, two in two, and six in one. Where only one outlet is documented, the case shows that a rendering was published and circulated; how far it reached in Greek coverage is left open. One case is a wire translation reproduced by four outlets, which carry one rendering between them.

Three cases carry the experimental work: an unscripted remark at a press conference in 2015, a broadcast interview in 2022, and an institutional policy statement in 2025. They were chosen before the experiments for that spread of discourse types and periods, and because one of the three carries a formal diplomatic response. They are not offered as typical of the corpus, and the experimental results are not generalized to it.

3.2 The translation experiments

3.2.1 Design and prompt conditions

Each of the three German source texts was submitted to two large language models under five prompt conditions, three times per condition, and to two neural machine translation systems. This yields 96 outputs: 90 from the language models (three cases, five conditions, two models, three runs) and six from the NMT systems (three cases, two systems; Vraila 2026).

Prompt wording (Table 1) was fixed before execution and held identical across cases, systems and runs.

Table 1. ***Prompt conditions.***

ID	Label	Prompt as issued
C1	Bare translation	Translate the following text into Greek.
C2	Professional political translator	Translate as a professional translator of political texts.
C3	Newsroom	Translate the following statement for a Greek news website.
C4	Headline task	Write a headline in Greek for a Greek news website based on the following statement.
C5	Journalistic adaptation	Adapt the following statement in Greek for a Greek news audience.

Prompts were issued in English, which is recorded as a parameter of the design, since the language of the instruction may itself condition what the model returns. The source text followed the prompt after a single newline, unquoted, with speaker, date and outlet withheld.

C1 alone asks the language models to do what the NMT systems do, and every comparison between the two architectures is made against it; C2 to C5 vary the task.

3.2.2 Systems, access and execution parameters

Access through each provider's web interface left temperature, top-p, maximum token count, seed and the system prompt at the provider default in every run; no API was used. Web-interface access was chosen for ecological validity, mirroring authentic newsroom practice where journalists interact via default consumer UIs rather than API endpoints. The identifiers in Table 2 are the labels shown in the model selector at execution; such

a label names a product. It does not identify the version or the build that served a given request.

Table 2. ***Systems, access and execution parameters.***

System	Access route	Identifier as recorded	Session dates
GPT-4o	Web interface	GPT-4o (interface label)	13–15 March 2025
GPT-5.3 Instant	Web interface	GPT-5.3 Instant (interface label)	7–8 March 2026
DeepL, free web translator	Web interface	DeepL free web translator	13 March 2025
Google Translate	Web interface	Google Translate web	13 March 2025
GPT-4o	Web interface	GPT-4o (interface label)	13–15 March 2025

The two rounds of language model runs, roughly twelve months apart, used different models, and the change of model is confounded with everything else that altered between March 2025 and March 2026. The second round is therefore reported as a descriptive time point.

The runs were executed by round. For GPT-4o, the first runs were made on 13 March 2025, the second on 14 March and the third on 15 March, with one third run made on 13 March. For GPT-5.3 Instant, the first and second runs were made on 7 March 2026 and the third on 8 March. Each run was issued in a fresh session.

3.2.3 Stability of the machine translation outputs

Each NMT system was run once per case in the main matrix. Two uncoded passes on 8 September 2026 tested what a single run represents: one repeated the submissions within a single sitting, the other set the outputs against those of March 2025.

3.2.4 Coding procedure

The coding scheme is the taxonomy of five patterns, derived from the documented corpus before the experiments were run. Each output received one task type from four values, a binary judgment on each of the five patterns, which may co-occur, and a binary judgment on five modal and register features. A translation renders the propositional content in full. A condensed translation reduces the content without adding to it, and the source syntax remains recoverable. A headline is a nominal or elliptical form from which the source syntax cannot be recovered, and an output meeting that test is classed

as a headline whatever else has shifted. An adaptation asserts something the source does not assert, or drops something it does, so that truth conditions differ. A shift of mood, person or evaluative lexis on its own does not make an output an adaptation; such shifts are recorded in the pattern and modal variables. The first coder also recorded the text span that triggered each code; the second recorded codes alone. The codebook carries definitions, exclusions and five dated amendments; Amendment 5 records a criterion already applied and changes no code (Vraila 2026).

Both coders coded all 96 outputs independently, under different conditions. The second coder is a professional translator of journalistic and political texts working in the German to Greek pair, and had the German source texts alongside the Greek outputs. That access is required by the instrument: three of the five patterns turn on what the output does to material present in the source, as does the weaponization test, which asks whether the rendering adds hostility the German does not carry. The second coder worked blind to the condition, the system and the manuscript with its hypotheses, and outputs were presented under anonymized item numbers in an order shuffled with a fixed seed [random.Random(4242)]. The first coder is the author and coded in the master log, where the condition and the system of every output are visible. The agreement figures therefore compare a blind reading with an informed one.

Calibration followed the first independent round instead of preceding it. Amendments 1 and 2 were made during the first round, as questions arose, and applied to it. Amendments 3 and 4 came out of a calibration session, held after the first round, on the two variables whose agreement fell furthest below that of the rest of the set: Amendment 3 restricted weaponization to hostility added by the rendering, and Amendment 4 set the boundary between translation and adaptation. Task type was then recoded independently by both coders against the amended definition, weaponization was settled by consensus, and every other column was left as first coded, apart from six later changes by the first coder set out in the note to Table 3.

3.2.5 Intercoder agreement

Cohen's κ, computed in spreadsheet formulas, is reported in Table 3 per variable and never pooled, since pooling would let the mechanical variables mask weakness in the interpretive ones.

Table 3. ***Intercoder agreement by variable (Cohen's κ, n = 96).***

Variable	κ	Round	Disagreements
Task type	0.856	2	9
Context isolation	0.907	2	3

Tone modification	0.915	2	4
Narrative-driven translation	0.793	2	10
Weaponization	0.683	1	13
Semantic manipulation	0.915	2	4
Subjunctive preserved	1.000	2	0
Impersonal construction preserved	1.000	2	0
Ironic implicature preserved	0.938	2	3
Legal term rendered with established equivalent	1.000	2	0
Evaluative lexis added	0.957	2	2

Note. Only task type was recoded after calibration; every other variable is compared on its first-round codes. The values for tone modification, narrative-driven translation and semantic manipulation differ from those first computed because three spreadsheet cells overwritten with static values were restored to their formulas. The narrative-driven value also reflects one code the first coder later changed to satisfy the rule that weaponization entails narrative-driven translation, which lowered it from 0.813 to 0.793. The value for evaluative lexis reflects five codes the first coder changed from 1 to 0 after the agreement figures were first computed, on re-examination of outputs in which no added evaluative item could be identified; before the change it stood at 0.957, with two disagreements.

Weaponization is reported at its uncalibrated first-round value ($\kappa = 0.683$), below the threshold of 0.70 set for the study. Disagreements stemmed from distinguishing marked lexical sharpening from outright hostile intent; Amendment 3 resolved this by strictly requiring newly introduced hostility, after which the coders settled remaining discrepancies item by item. Task type, the other variable taken to calibration, stood at 0.762 before recoding.

4 The three cases

The linguistic analysis of the three cases develops an account first given in Vraila and Livas (2025), who examined them as instances of shifting emotional charge in Greek media renderings of German political speech. That analysis is what this article draws from the earlier paper, which has no experimental component. For the 2015 and 2022 cases the German passages used here are fuller than those quoted there, and for the 2022 and 2025 cases the conclusions reached here differ from the earlier ones. The

experimental design, the 96 coded outputs, the codebook and its amendments, and the intercoder agreement procedure are new here.

4.1 Case 1: Schäuble on Varoufakis (2015)

At a press conference on 10 March 2015, answering a question about his meeting with Varoufakis, Schäuble said: "Also, dass er nun plötzlich naiv in Sachen Kommunikation wäre, habe ich ihm gesagt, das ist mir ganz neu. Aber man lernt ja nie aus." (Well, that he should now suddenly be naive in matters of communication, I told him, is quite new to me. But one never stops learning.) The passage is continuous and was submitted to the systems as it stands.

Mega TV broadcast the remark the same day as a subtitled excerpt, cut from the surrounding exchange: «το να είναι τόσο ανόητα αφελής, δεν το περίμενα» (for him to be this foolishly naive was something I had not expected). The Greek ambassador in Berlin delivered a démarche that evening, reported on 11 and 12 March, and German officials answered by referring to what the minister had actually said and denying any intent to offend.

The subtitle removes the reporting frame «habe ich ihm gesagt» (I told him), which placed the judgment in an exchange that had already taken place. It states in the indicative what the German, through the subjunctive «wäre» (would be), puts forward without asserting, and it drops «in Sachen Kommunikation» (in matters of communication), which confined the judgment to one domain. Its degree modifier, «τόσο ανόητα» (so foolishly), has no counterpart in the German, where the predicate is qualified by domain alone. The patterns coded for the subtitle are context isolation, tone modification and semantic manipulation.

Table 4 records the coded outcomes for the six language model outputs in each condition. Under C1, five of the six language model outputs are full translations carrying none of the five patterns. The sixth is an adaptation, «Του είπα ότι είναι εντελώς αφελής στην επικοινωνία, δεν το περίμενα» (I told him he is utterly naive at communication; I had not anticipated it), which delivers the hypothetical as fact and turns a remark about what was new to the speaker into one about what he had anticipated. Under C2, the one condensed rendering drops the closing clause and gives «ganz neu» (quite new) as «με ξενίζει εντελώς» (strikes me as wholly strange). With both changes, the remark stops reporting a claim and starts recording an attitude, and that change to the utterance is why it is coded for context isolation, together with semantic manipulation. The hypothetical, which this rendering carries in a nominalized clause, and the irony that remains are coded as preserved in their own columns.

Under C3, tone modification is coded in five outputs, each time through a degree modifier: «τόσο αφελής», «τόση αφέλεια», «τόσο ανόητα αφελής» (this naive, this much naivety, this foolishly naive).

Of the six headlines, two assign the quality to Varoufakis without comment, two present the remark as an insult, as in «Αφελής επικοινωνιακά ο Βαρουφάκης» - Νέα προσβολή Σόιμπλε ("Varoufakis naive at communication": Schäuble insults him again), and one puts in Schäuble's mouth direct speech the source does not contain. The sixth sets an evaluation of Schäuble against his of Varoufakis: «Η «αφέλεια» του Βαρουφάκη και ο κυνισμός του Σόιμπλε» (Varoufakis's "naivety" and Schäuble's cynicism). Ironic implicature survives in none of the six headlines and in twenty-three of the twenty-four outputs produced under the other four conditions.

Table 4. ***Coded outcomes for Case 1 by condition (n = 6 language model outputs per condition).***

Condition	Task type	Context isolation	Tone modification	Narrative-driven	Weaponization	Semantic manipulation
C1 Bare translation	5 translations, 1 adaptation	0	1	1	1	1
C2 Professional political translator	5 translations, 1 condensed	1	0	0	0	1
C3 Newsroom	3 translations, 3 adaptations	0	5	3	2	3
C4 Headline task	6 headlines	6	6	6	6	6
C5 Journalistic adaptation	6 adaptations	0	4	4	2	6

Note. Under C1, the four coded patterns occurred co-extensively within a single adaptation output.

Of the machine translation outputs, Google Translate returned a translation carrying none of the five patterns and DeepL an adaptation carrying semantic manipulation; ironic implicature was coded as preserved in both.

The wording of the broadcast subtitles recurs in the experimental set. Three outputs use «ανόητα» (foolishly) or its adjectival form, two under C3 and one under C5, and three render the second clause with «περίμενα» (I expected), as the broadcast does, one each under C1, C3 and C5; the output under C5 does both. Where the systems drew this wording from is beyond what the design can show.

4.2 Case 2: Schäuble on sweaters (2022)

The input for this case joins two passages of a Bild TV interview broadcast on 10 October 2022, and the join is unmarked: "Dann zieht man halt einen Pullover an. Oder vielleicht

noch einen zweiten Pullover. Darüber muss man nicht jammern, sondern man muss erkennen: Vieles ist nicht selbstverständlich. Deshalb sollte man tatsächlich immer auch ein paar Kerzen, Streichhölzer und auch eine Taschenlampe zuhause haben." (Then one just puts on a sweater. Or perhaps a second sweater as well. There is no need to complain about it; one has to recognize that much is not to be taken for granted. That is why one really should always keep a few candles, matches and a torch at home.) The sentence to which *Deshalb* (that is why) refers, "Er (der Strom) kann auch mal ausfallen" (It, the electricity, can fail from time to time), lies outside the string, and none of the thirty-two outputs for the case supplies it.

The passage turns on a contrast. *Darüber muss man nicht jammern, sondern man muss erkennen* (there is no need to complain about it; rather, one has to recognize) denies any need to complain and puts the weight on recognition. Its subject, the generic *man* (one), assigns the advice to no one in particular while keeping the speaker inside its scope.

Table 5 records the coded outcomes by condition. Under C1 and C2, all twelve language model outputs are full translations. Each keeps the recognition clause and the generalizing construction, most often as a first-person plural, and each renders the German's denial of need as such, with «δεν χρειάζεται» (there is no need) or «δεν υπάρχει λόγος» (there is no reason): «Δεν χρειάζεται να γκρινιάζουμε γι' αυτό, αλλά να συνειδητοποιούμε ότι πολλά πράγματα δεν είναι αυτονόητα» (We have no need to complain about it; we need to realize that many things cannot be taken for granted).

Under C3, the generalizing construction survives in none of the six outputs, and the Greek addresses the reader: «Δεν χρειάζεται να γκρινιάζετε για αυτό, αλλά να καταλάβετε ότι πολλά πράγματα έπαψαν να είναι αυτονόητα» (You have no need to complain about it; you need to understand that many things are no longer a given). One output delivers the clause as a prohibitive imperative, «Μη γκρινιάζετε» (Do not complain).

Each of the six headlines omits material whose removal changes the force of what remains and adds evaluative wording with no counterpart in the source. GPT-4o returned, for instance, «Σόιμπλε: Να φορέσετε και δεύτερο πουλόβερ και να μη γκρινιάζετε» (Schäuble: Put on a second sweater too, and stop complaining). The three GPT-5.3 Instant headlines place the evaluation in the frame around the quoted words, as in «Αν κρυώνετε, να φοράτε πουλόβερ»: Η κυνική συνταγή Σόιμπλε για τον χειμώνα ("If you are cold, wear a sweater": Schäuble's cynical recipe for winter).

Table 5. ***Coded outcomes for Case 2 by condition (n = 6 language model outputs per condition).***

Condition	Task type	Context isolation	Tone modification	Narrative-driven	Weaponization	Semantic manipulation
C1 Bare translation	6 translations	0	0	0	0	0
C2 Professional political translator	6 translations	0	0	0	0	0
C3 Newsroom	6 adaptations	0	6	6	0	6
C4 Headline task	6 headlines	6	6	6	5	6
C5 Journalistic adaptation	6 adaptations	0	6	6	3	6

DeepL and Google Translate each returned a full translation carrying none of the five patterns.

ERT, Greek public television, published its report on 11 October 2022, and its body stays closer to the German than its headline does. The body keeps the recognition clause, the first-person plural for most of the passage and, because it draws on the whole interview, the causal link made by *Deshalb*. It departs from the German at two points. Its first sentence gives *man* as a third person plural, «Τότε απλά ας φορέσουν ένα πουλόβερ» (Then let them just put on a sweater), and «Δεν πρέπει να γκρινιάζουμε για αυτό» (We must not complain about it) renders *muss man nicht jammern,* which denies a need, as a prohibition.

The headline reads «Συμβουλές Σόιμπλε για τον χειμώνα: Μην γκρινιάζετε, φορέστε πουλόβερ, πάρτε κεριά και σπίρτα» (Schäuble's tips for the winter: Stop complaining, put on sweaters, pick up candles and matches). It turns the body's first-person plural into a second-person imperative and drops the recognition that the German sets against complaint, *sondern man muss erkennen*. An imperative of this kind selects an addressee, and a remark about a condition the speaker shares becomes an instruction to those who bear it. The case also carries semantic manipulation from a separate exchange at the end of the interview. Asked whether he could still laugh, Schäuble answered «Ja klar, lachen befreit und entspannt» (Of course, laughing frees you and relaxes you), which the report gives as «Το γέλιο είναι δωρεάν και χαλαρωτικό» (Laughter is free and relaxing), reading *befreit* as free of charge.

For this case, address in the second-person plural appears in the language model outputs under C3, C4 and C5, and in none produced under C1 or C2. What separates the six headlines from the C3 outputs is the loss of the recognition clause, which survives in every C3 output and in no headline; the same loss separates the ERT headline from

its body. The parallel is one of operation, and nothing in the material links the ERT report to such systems.

4.3 Case 3: Merz on entry bans (2025)

The Merz statement turns on a distinction of legal status. The two consecutive sentences submitted to the systems read: "Es wird ein faktisches Einreiseverbot in die Bundesrepublik Deutschland für alle geben, die nicht über gültige Einreisedokumente verfügen oder die von der europäischen Freizügigkeit Gebrauch machen. Das gilt ausdrücklich auch für Personen mit Schutzanspruch." (A de facto ban on entry into the Federal Republic of Germany will apply to everyone who lacks valid entry documents or who exercises European freedom of movement. This applies expressly to persons with a claim to protection as well.) *Oder* (or) holds apart two classes of person: those without valid entry documents, and those exercising European freedom of movement, an entitlement under EU law for which *Freizügigkeit* is the term of art. Whether an output keeps that second class, and keeps it as an entitlement, is the discriminating test in this case, coded on every output as the legal-term variable.

Table 6 records the coded outcomes by condition, together with the legal-term variable. Under C1 and C2, all twelve language model outputs are full translations, and each keeps both the disjunction and the term, most often with the entitlement made explicit: «ή κάνουν χρήση του δικαιώματος ελεύθερης κυκλοφορίας εντός της Ευρωπαϊκής Ένωσης» (or exercise the right of free movement within the European Union).

Under C3, each of the six outputs adds evaluative wording with no counterpart in the source, and five keep the term, though not always the syntax around it. GPT-4o returned «Θα ισχύσει στην πράξη ντε φάκτο απαγόρευση εισόδου στη Γερμανία για όλους τους παράτυπους μετανάστες που δεν διαθέτουν έγκυρα έγγραφα εισόδου ή κάνουν χρήση της ευρωπαϊκής ελεύθερης κυκλοφορίας» (In practice a de facto ban on entering Germany will apply to all irregular migrants who lack valid entry documents or use European free movement). A single noun phrase now governs both arms of the disjunction, and the people in the second arm are irregular migrants before the sentence reaches their right. Two outputs replace the disjunction with a conjunction, and two render the exercise of the right as its «κατάχρηση» (abuse); one does both and also substitutes a general noun, «ελεύθερη μετακίνηση» (free movement), for the term of art. An output is credited with the term where «ελεύθερη κυκλοφορία» (free circulation) carries a marker of the entitlement, and a merger of the two classes around it is recorded as semantic manipulation.

One GPT-4o headline reads «Μερτς: "Πρακτική απαγόρευση εισόδου" ακόμη και για όσους δικαιούνται προστασία» (Merz: "Practical entry ban" even for those entitled to protection). One GPT-5.3 Instant headline keeps that clause and adds «"Μπλόκο"» ("Blockade") and «Νέοι αυστηροί περιορισμοί στα σύνορα» (New strict restrictions at the border).

Table 6. ***Coded outcomes for Case 3 by condition (n = 6 language model outputs per condition).***

Condition	Task type	Context isolation	Tone modification	Narrative-driven	Weaponization	Semantic manipulation	Legal term rendered
C1 Bare translation	6 translations	0	0	0	0	0	6
C2 Professional political translator	6 translations	0	0	0	0	0	6
C3 Newsroom	6 adaptations	0	6	6	4	6	5
C4 Headline task	6 headlines	6	6	6	6	6	0
C5 Journalistic adaptation	6 adaptations	0	6	6	4	6	2

Merz made the statement in Berlin on 23 January 2025, in a five-point press statement given the day after the attack in Aschaffenburg. It reached Greek readers as one wire translation by ΑΠΕ-ΜΠΕ, the Athens-Macedonian News Agency, and the four outlets that carried it reproduce the quoted passages word for word: ΣΚΑΪ, Protagon and Athina 9.84 with an agency credit, Liberal without one.

In the wire, the two classes of the German become one. The body gives the measure as «"De facto απαγόρευση εισόδου" σε παράνομους μετανάστες, ακόμη και εάν βρίσκονται σε καθεστώς προστασίας» ("De facto entry ban" on illegal migrants, even where they hold protection status). The people the German places in the second arm of the disjunction, who exercise a right, are absorbed into the category of illegality. ΣΚΑΪ's headline names the same group «παράτυπους μετανάστες» (irregular migrants). Later in the report the wire fills the second arm with different content: «Όποιος δεν μπορεί να παρουσιάσει έγγραφα ή έχει ήδη εισέλθει στην Ευρωπαϊκή Ένωση από άλλη χώρα» (Anyone who cannot produce documents or has already entered the European Union from another country). Prior entry through another country takes the place of the entitlement the German names.

Twenty-one of the thirty-two outputs for this case render *Freizügigkeit* with an established Greek equivalent: all twelve language model outputs under C1 and C2, both machine translation outputs, five under C3, two under C5 and none of the six headlines. The wire, a translation of the body of a press statement, loses the term. The labels «παράτυποι μετανάστες» (irregular migrants) and «παράνομοι μετανάστες»

(illegal migrants), counted in any case form, appear in seven of the thirty language model outputs for this case: three under C3, two under C4 and two under C5, and none under C1 or C2.

5 A taxonomy of translational distortion

Five patterns recur across the documented cases of distortion. Each is defined by the operation performed on the source rather than by its effect on a reader, and the categories may co-occur. The definitions are those applied in coding the experimental outputs, and each pattern is illustrated from a case outside the three used in the experiments.

5.1 Context isolation

Material present in the source is removed, and its removal alters what the statement does. The test is the effect of the omission on the utterance, whatever other markers of the source survive it: where the statement still does what the German did, the pattern is not coded.

Merkel, speaking at Meschede on 17 May 2011, said: *Wir können nicht eine Währung haben und der eine kriegt ganz viel Urlaub und der andere ganz wenig* (We cannot have one currency in which one side gets a great deal of holiday and the other very little). Greek coverage in Τα Νέα, Πρώτο Θέμα and in.gr rendered the remark as «Κόψτε τις διακοπές και τις συντάξεις» (Cut the holidays and the pensions) and «Οι Έλληνες κάθονται με τα λεφτά μας» (The Greeks are sitting around on our money). The premise of a shared currency is gone, and with it the comparison of two parties within that currency. The Greek reader receives an instruction addressed to a national audience, where the German stated a condition under which a monetary union can work.

5.2 Tone modification

Evaluative intensity diverges from the source through lexical choice while the propositional content stands. Sharpening and flattening both count. Every case in the corpus also carries semantic manipulation, and the example shows the two patterns in one rendering.

Schäuble, in a Deutschlandfunk interview of 1 February 2012, said: *Wir können nur Hilfe leisten, wenn es nicht ein Fass ohne Boden ist* (We can only give help if it is not a bottomless barrel). news247.gr carried «Σε ένα βαρέλι δίχως πάτο, κάθε βοήθεια είναι μάταιη» (In a barrel with no bottom, all help is futile). The image survives intact. The German

attaches a condition to assistance, and the Greek turns it into a general claim that assistance is futile, which is the propositional shift. The evaluative weight sits in «μάταιη» (futile), which has no counterpart in the German.

5.3 Narrative-driven translation

The output aligns with a recurrent framing template in the receiving discourse that cannot be derived from the source. The coding requires documentary evidence that the template is in circulation.

Kauder said on 15 November 2011: *Jetzt auf einmal wird in Europa Deutsch gesprochen* (Now, all of a sudden, German is being spoken in Europe). Greek coverage in Το Βήμα, Ελευθεροτυπία, in.gr and newsit.gr carried both «Η Ευρώπη μιλάει τώρα γερμανικά» (Europe now speaks German) and «Η Ευρώπη υποτάσσεται στη Γερμανία» (Europe submits to Germany). The first stays with the German claim about which language now sets the terms of European argument. The second recasts it as the subordination of Europe to Germany, a reading the source does not supply, aligned with an antagonistic frame already established in Greek political discourse of the period (Stavrakakis and Katsambekis 2014; Katsambekis 2016).

5.4 Weaponization

The statement is reconstructed as attributing hostile intent or contempt, with the hostility introduced or amplified by the rendering. Hostility the German already carries, conveyed as it stands, is coded zero. Naming a person, or ascribing a negative quality to one, is neither necessary nor sufficient. Because weaponization relies on embedding the statement within an antagonistic receiving frame, it methodologically entails narrative-driven translation.

In an interview with Handelsblatt in 2022, Lindner said: *Nach der Pandemie sind die Schuldenquoten so hoch, dass diese Vorgabe bestimmte Länder objektiv überfordern würde* (Debt ratios after the pandemic are high enough that this requirement would, objectively, ask too much of certain countries). A signed opinion column on reporter.gr appeared under a headline set in quotation marks and attributed to him: «Εμπρός... πίσω στη λιτότητα» (Onward... back to austerity). The German weighs, in the conditional, what a fiscal requirement would do to countries it leaves unnamed. The Greek gives him a program of austerity, a word absent from the interview, and the quotation marks present it as his own. That attribution is where the hostility enters. The Greek text is commentary rather than a newsroom translation, which sets the case apart from the others in the corpus; it is also analyzed in Vraila and Livas (2025).

5.5 Semantic manipulation

Propositional content is added or removed, so that the truth conditions of the output differ from those of the source.

Friedrich, in *Der Spiegel* 9/2012 of 25 February 2012, said: *Ich rede nicht davon, Griechenland rauszuschmeißen, sondern Anreize für einen Austritt zu schaffen* (I am not talking about throwing Greece out, but about creating incentives for it to leave). news247.gr and Το Βήμα ran «Την Ελλάδα εκτός ευρώ θέλει ο Γερμανός υπουργός Εσωτερικών» (The German interior minister wants Greece out of the euro). The German separates expulsion from the creation of incentives and places the speaker on one side of that separation. The Greek attributes to him the position his sentence sets aside.

5.6 Incidence across the corpus

The classifications in Table A1, given in the appendix, were made by the author alone; the reliability figures apply to the experimental outputs.

Counts are given over the thirteen documented departures, since the three faithful cases carry no pattern and would depress every figure if counted.

Semantic manipulation is coded in all thirteen cases and does not discriminate within the corpus; the other four patterns record what accompanied it. Narrative-driven translation is coded in eleven cases and weaponization in seven. Tone modification is coded in five, context isolation in four. Individual cases carry between one and four patterns. Weaponization is coded in four of the seven departures from 2011 and 2012 and in two of the five from 2022 onward, all five of which carry narrative-driven translation.

5.7 Renderings that hold

Three cases in the corpus are faithful. Merkel's remark on multiculturalism of 16 October 2010, Gabriel's on Tsipras of 5 July 2015, and Merkel's *Wir schaffen das* (We can do this) of 31 August 2015 each reach Greek with their lexical and pragmatic force intact, the idiom in the Gabriel case included. Each was carried by an outlet that also carries departures in the corpus. The two from 2015 fall within the period the thirteen span, and the 2010 case precedes it.

6 Results

6.1 What the conditions can separate

Three accounts of where a departure comes from are available, and each predicts a different distribution of departures across the five conditions. They were formulated after the runs of March 2025 and March 2026 and serve as alternative readings of the results; none was set down in advance as a prediction.

On the first account, the dispositions belong to the model, and Greek journalistic convention shapes what it returns whatever it is asked. Incidence would then be similar under every condition, C1 included.

On the second, the repertoire is available to the model but dormant, and the wording of the instruction decides whether it is drawn on. This account expects incidence near zero under C1 and C2 and a rise once the instruction names a Greek news outlet.

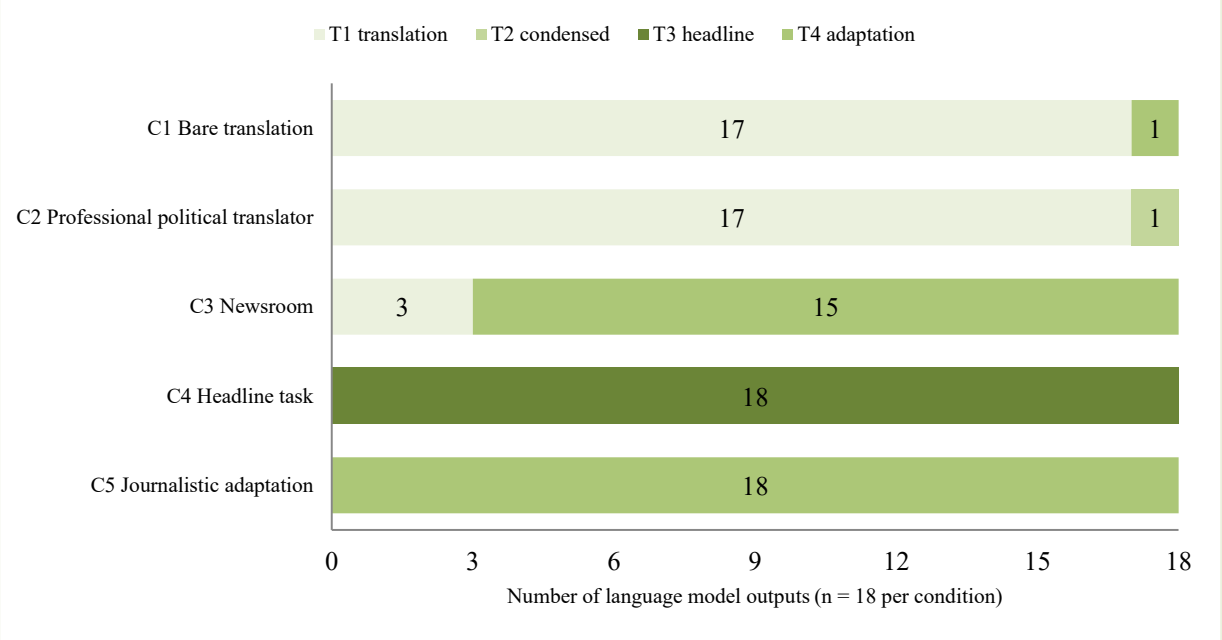
The third account holds that an instruction mentioning a news website is read as a request for a news text. In its strong form, C3 would return headlines, as C4 does. In its weaker form, C3 would return adaptations of the kind requested by name under C5.

6.2 Task type

The counts reported are the first coder's. The blind coder's are given where the two differ on a comparison the three accounts depend on.

Figure 1 classifies the language model outputs by the task they perform. Every output takes one of the four task types, so each bar sums to eighteen.

Figure 1. ***Task type by condition, language model outputs (n = 18 per condition).***



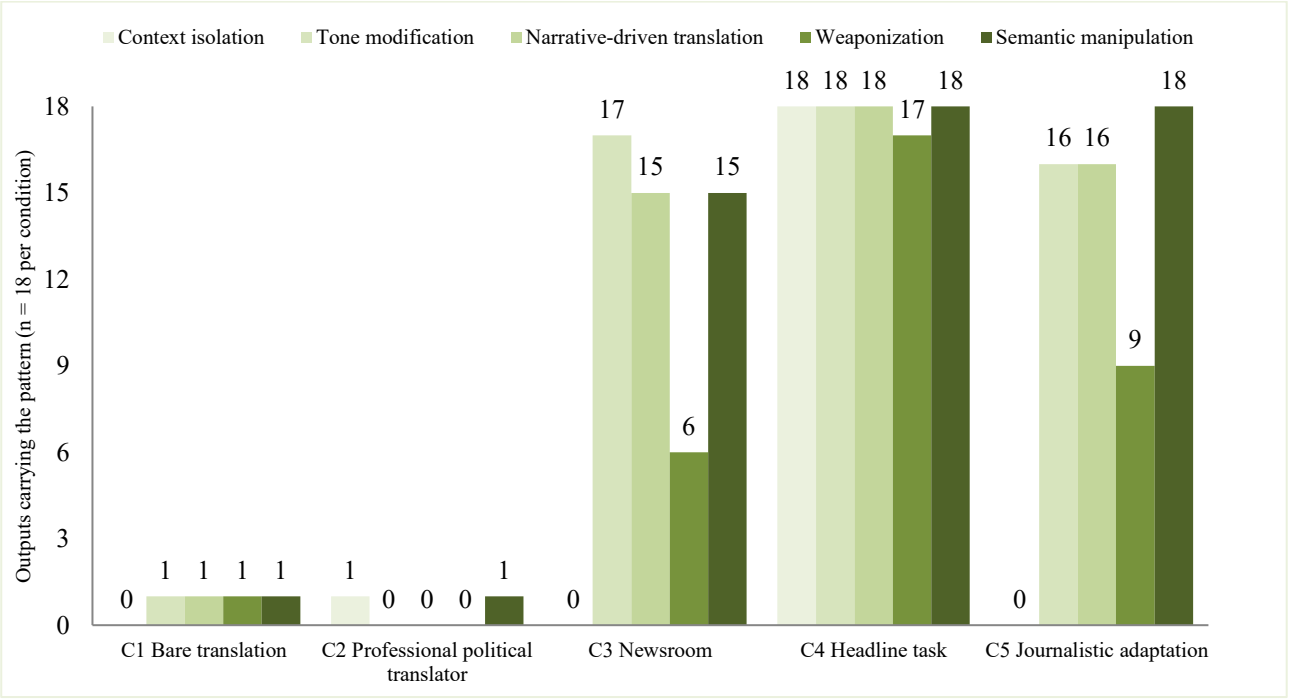
C3 and C4 divide completely: all eighteen outputs under C4 are headlines and none under C3 is. The blind coder counts seventeen headlines under C4 and none under C3. This excludes the strong form of the third account. The weaker form survives. C3 returns fifteen adaptations against one under C1, by either coder's count, and C5 returns eighteen by the first coder's count and fourteen by the blind coder's, who classified two of the other four as translations and two as condensed renderings.

Of the six NMT outputs, read against C1, five are full translations and one is an adaptation.

6.3 Pattern incidence

An output can carry more than one pattern, and each bar in Figure 2 is read against the eighteen outputs of its condition.

Figure 2. ***Pattern incidence by condition, language model outputs*** **(n = 18 per condition).**



Note. Hatching marks weaponization, the one variable reported below the agreement threshold of 0.70.

Under C1 and C2, seventeen of eighteen outputs are clean in each condition. The exception under C1 carries four of the five patterns and the exception under C2 carries two; both come from Case 1 and from GPT-4o. The blind coder finds sixteen clean outputs under C2, coding tone modification and semantic manipulation in a second Case 1 output from GPT-4o. Disaggregating outputs reveals consistent prompt sensitivity across both systems: GPT-4o and GPT-5.3 Instant alike preserved fidelity under C1–C2 and shifted toward adaptation under C3, justifying their aggregated reporting.

Weaponization is coded in six of the eighteen C3 outputs, against fifteen to seventeen for tone modification, narrative-driven translation and semantic manipulation.

The C4 and C5 bars record compliance with the instruction and serve as a benchmark for permitted deviation. Context isolation is coded in all eighteen headlines, a count that describes the genre more than the system.

Read as a benchmark, C5 falls closer to C3 than to C4. C4 reaches seventeen or eighteen on all five patterns. C5 records context isolation in none of its outputs by the first coder's count and in three by the blind coder's, and it stays within three outputs of C3 on every other variable. Here the second and third accounts are hardest to keep apart, and these data leave the choice between them open. The first account requires incidence under C1

and C2 at the level found under C3; the observed incidence there is one output in eighteen in each.

6.4 Variability across runs of one cell

Each of the thirty language model cells returned three different strings across its three runs. For each cell, pattern agreement is the proportion of the five patterns on which all three runs carry the same code. It averages 0.84 across the thirty cells. The coded profile thus varies less than the wording does, and least steadily in Case 1: six of its ten cells fall short of full agreement, against three of ten for Case 2 and two of ten for Case 3. The lowest value, 0.2, is returned by two cells. One of them is the C1 cell for Case 1 on GPT-4o, which holds the single patterned output under C1; its other two runs are clean.

6.5 Stability of the machine translation outputs

Within a single sitting, each of the six submissions returned the same string on resubmission. Set against the outputs collected in March 2025, five of the six differed. That comparison spans eighteen months and concerns change over time.

7 Conclusions

The same instructions, applied to the same passages, return full translations in seventeen of eighteen runs under C1 and under C2, and adaptations in fifteen of eighteen under C3. Of the technological conditions the third research question asks about, the instruction is the one this design isolates, and a newsroom that uses these systems makes an editorial choice when it words the request.

The patterns were derived from renderings by journalists and could be applied to machine outputs with agreement at or above the threshold on every variable except weaponization. In each of the three cases examined closely, the largest departure occurs where the text is excerpted, headlined or relayed, and the faithful renderings in the corpus come from outlets that also carry departures. Both observations point to the stage of production more than to the outlet, although three faithful cases can only suggest this. A wire extends the reach of a departure: the agency rendering of the Merz statement reached four outlets as one text.

The temporal evidence is thinner. The weaponized renderings of 2011 and 2012 are consistent with the antagonistic frame then in circulation (Stavrakakis and Katsambekis 2014), but a purposive corpus registers such distributions without measuring them. Machine translation systems also change, and a statement retranslated from the archive need not match an earlier machine rendering of it.

Two checks follow from the material, and neither requires a judgment about intent. A term lost in transfer, as *Freizügigkeit* was in the wire, is caught by comparing legal and institutional terms with their established equivalents, for which a translation obtained under a bare instruction can serve as reference. A statement rebuilt around a frame is caught by comparing modality, person and evaluative lexis with the source. For newsrooms, machine output is a draft whose instruction, system and date belong in the record, and requesting translation and headline as separate steps lets the headline be written against a full rendering. A check at a news agency reaches furthest, since a wire passes to every outlet that takes it.

The corpus is a purposive sample within one language pair, and the experiments cover three source texts. The reported counts are those of a coder who could see the condition of each output. The ten disagreements with the blind coder on narrative-driven translation all run toward the pattern; on tone modification, context isolation and semantic manipulation most run the other way. Weaponization, reported at its first-round value, has the lowest independent agreement of the coded variables, and its boundary with the neighboring categories is not cleanly drawn. Task type and semantic manipulation share a criterion, a shift in propositional content, and are not independent measurements. Language pairs without the asymmetry of the crisis years, and instructions that name an outlet's editorial standards, would show how far the finding reaches; how these tools enter editorial routine is a question for work with journalists.

7.1 Data availability

The data supporting this study are deposited in Zenodo (Vraila 2026, https://doi.org/10.5281/zenodo.22863810). The deposit holds the master log of the 96 coded outputs, with one row per output giving the source text, the prompt as issued, the system and its recorded identifier, the date, the unedited output and both coders' codes. It also holds the corpus of sixteen documented cases, the codebook with its dated amendments, the agreement spreadsheet with the formulas from which Cohen's κ was computed, and a README.

Vraila, Stavroula Paraskevi. 2026. Political Translation in the AI Era: Journalists as Dual Agents and Systemic Vulnerabilities in Cross-Linguistic Political Communication. In: L10N Journal. 1(5), pp. 39-61.

Bibliography

Baker, Mona. 2006. Translation and Conflict: A Narrative Account. London and New York: Routledge.

Bielsa, Esperança and Bassnett, Susan. 2009. Translation in Global News. London and New York: Routledge.

Conway, Kyle, Gramaccia, Julie Alice, Scholz, Nikita and Averbeck, Téana. 2025. Measuring Lexical Distance Between Parallel Corpora: The Case of AI-generated News Translation. In: Perspectives. Advance online publication. https://doi.org/10.1080/0907676X.2025.2590066

DeepL. 2025. DeepL Translator (March 2025). https://www.deepl.com/translator.

Entman, Robert M. 1993. Framing: Toward Clarification of a Fractured Paradigm. In: Journal of Communication. 43(4): pp. 51-58.

Freeth, Peter J. and Treviño, Rafael (eds.). 2024. Beyond the Translator's Invisibility: Critical Reflections and New Perspectives. Leuven: Leuven University Press. https://doi.org/10.11116/9789461665454.

Google. 2025. Google Translate (March 2025). https://translate.google.com/.

Hagar, Nick and Diakopoulos, Nicholas. 2019. Optimizing Content with A/B Headline Testing: Changing Newsroom Practices. In: Media and Communication. 7(1): pp. 117-127. https://doi.org/10.17645/mac.v7i1.1801.

House, Juliane. 2015. Translation Quality Assessment: Past and Present. London and New York: Routledge.

Kalantari, Esmaeil. 2022. Journalistic Translation: A Gate at Which Journalism Studies and Translation Studies Meet. In: Journalism. 23(7): pp. 1411-1429. https://doi.org/10.1177/14648849221074516.

Katsambekis, Giorgos. 2016. Radical Left Populism in Contemporary Greece: Syriza's Trajectory from Minoritarian Opposition to Power. In: Constellations. 23(3): pp. 391-403. https://doi.org/10.1111/1467-8675.12234.

Munday, Jeremy. 2012. Evaluation in Translation: Critical Points of Translator Decision-Making. London and New York: Routledge.

OpenAI. 2025. GPT-4o (March 2025). https://openai.com/.

OpenAI. 2026. GPT-5.3 Instant (March 2026). https://openai.com/.

Paullada, Amandalynne and Haroutunian, Levon. 2026. Language Justice and Machine Translation on Social Media. In: Modern Languages Open. 0(1): Article 7. https://doi.org/10.3828/mlo.v0i0.466.

Prates, Marcelo O. R., Avelar, Pedro H. and Lamb, Luis C. 2020. Assessing Gender Bias in Machine Translation: A Case Study with Google Translate. In: Neural Computing and Applications. 32: pp. 6363-6381.

Schäffner, Christina. 2004. Political Discourse Analysis from the Point of View of Translation Studies. In: Journal of Language and Politics. 3(1): pp. 117-150.

Schäffner, Christina. 2012. Unknown Agents in Translated Political Discourse. In: Target. 24(1): pp. 103-125.

Stavrakakis, Yannis and Katsambekis, Giorgos. 2014. Left-Wing Populism in the European Periphery: The Case of SYRIZA. In: Journal of Political Ideologies. 19(2): pp. 119-142. https://doi.org/10.1080/13569317.2014.909266.

Triantafyllou, Sotirios, Panagopoulos, Andreas M. and Kapos, Panagiotis. 2025. AI Pioneers and Stragglers in Greece: Challenges, Gaps, and Opportunities for Journalists and Media. In: Societies. 15(8): Article 209. https://doi.org/10.3390/soc15080209.

Vraila, Stavroula Paraskevi. 2026. Political Translation in the AI Era: Journalists as Dual Agents and Systemic Vulnerabilities in Cross-Linguistic Political Communication. In: L10N Journal. 1(5), pp. 39-61.

Valdeón, Roberto A. 2023. On the Cross-Disciplinary Conundrum: The Conceptualization of Translation in Translation and Journalism Studies. In: Translation Studies. 16(2): pp. 244-260. https://doi.org/10.1080/14781700.2022.2162573.

Valdeón, Roberto A. 2025. Journalistic Translation Research Twenty-Five Years On: Methods, Topics and Prospects. In: Perspectives. 33(4): pp. 635-653. https://doi.org/10.1080/0907676X.2025.2516364.

Venuti, Lawrence. 1995. The Translator's Invisibility: A History of Translation. London and New York: Routledge.

Vraila, Stavroula Paraskevi. 2026. Political translation in the AI era: Experimental Dataset and Coding Log (Version v1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.22863810

Vraila, Stavroula Paraskevi and Livas, Sotiris. 2025. Translation, Emotion, and Ideological Reconstruction: Translating Political Discourse in Journalistic Contexts. In: Köksal, Onur, Dolmacı, Mustafa and Ramizoğlu, Ruşen (eds.). 2025. 5th International Congress of Language and Translation Studies: Full-Text Book. Konya: Çizgi Kitabevi. pp. 799-826.