

Research Article

Constructing Collective Agency through “We”: A Corpus-Assisted Analysis of Modality in ASEAN Member States' UN General Debate Speeches

Titi Christiana^{1*}, Yondi², and Berdinata Massang³,

1 Sekolah Tinggi Pastoral Kateketik Santo Fransiskus Asisi Semarang

2 Sekolah Tinggi Teologi Injili Arastamar (SETIA) Ngabang

3 Institut Agama Kristen Negeri Manado, Indonesia

* Corresponding Author: Titi Christiana

Abstract The first-person plural pronoun “we” is a flexible linguistic resource in diplomatic discourse because it can represent collective actors at various scales. Although previous research has discussed political “we”, modality, solidarity, and identity in United Nations discourse, there is still limited understanding of how referential scale and modal meaning interact in constructing collective agency in individual ASEAN member states. This study uses Corpus-Assisted Discourse Studies (CADS) to analyze 92 United Nations General Debate statements from ten ASEAN member states during the period 2015–2024. The corpus consists of 159,857 word tokens and yields 2,361 attributable occurrences of “we”, 738 validated modalized-“we” constructions, and 643 actionive agency-core constructions. Referential scope is classified as national, regional, coalitional, global, generic, or ambiguous, while contextual modal meanings are classified as obligation/necessity, capacity/ability, commitment/volition, and epistemic/prediction. The results show that national and global “we” dominate resolved references, while explicitly regional ASEAN “we” is very rare. Obligation/necessity comprises 60.84% of modalized constructions. Global “we” is highly concentrated on obligation (73.28%), while national “we” shows a much greater commitment orientation (29.52% compared to 6.34% for global “we”). The association between national/global reference and modal function is significant, $\chi^2(3) = 73.27$, $p < .001$, Cramér's $V = .325$. These findings suggest that collective agency is scalar: global collectivities are primarily constructed as actors bearing shared obligations, while national collectivities are more readily constructed as actors making commitments to their own actions.

Received: 15th February 2026

Revised: 16th March 2026

Accepted: 15th April 2026

Published: 30th June 2026

Curr. Ver.: 30th June 2026

Keywords: collective agency; “we”; modality; diplomatic discourse; ASEAN; United Nations General Debate; corpus-assisted discourse studies

1. Introduction

Political speech in international forums is not only a means of conveying a state's position, but also a linguistic space where political actors define who they represent, with whom they build common ground, and what actions are positioned as shared obligations, capabilities, or commitments. The United Nations General Debate (UNGD) provides a particularly productive context for observing these processes because member states deliver diplomatic statements in a relatively comparable and longitudinally available institutional setting. The United Nations General Debate Corpus (UNGDC), originally developed to understand state preferences through text-as-data and later extended to the contemporary period, provides one of the most comprehensive corpus sources for this purpose (Baturu et al., 2017; Jankin et al., 2025).

The UNGD has been used to study various forms of political self-presentation, state preference, legitimacy, and changes in international discourse. (Kentikelenis & Voeten, 2021), for example, use UN speeches to explore legitimacy challenges to the liberal international



Copyright: © 2026 by the authors.

Submitted for possible open

access publication under the

terms and conditions of the

Creative Commons Attribution

(CC BY SA) license

([https://creativecommons.org/li](https://creativecommons.org/licenses/by-sa/4.0/)

[censes/by-sa/4.0/](https://creativecommons.org/licenses/by-sa/4.0/))

order, while (Chelotti et al., 2022) utilize General Debate statements to examine the possible socialization effects of intergovernmental organizations on state preferences. (Adams & Mitrani, 2024) demonstrate how states conduct international self-presentation through multiple collective "we's," while (Hecht & Steffek, 2024) demonstrate that General Debate discourse can capture long-term changes in the salience of global terminology. Recent research by (Hecht, 2026) also uses the UNGD to explore changes in normative status dimensions. In linguistics, (Brun-Mercer, 2021) shows that corpus-assisted analysis of General Debate addresses can reveal systematic representational patterns that are difficult to discern through close reading of a small number of speeches.

One of the simplest and most flexible linguistic resources in political discourse is the first-person plural pronoun "we." Unlike "I," which deictically usually has a relatively stable referent, "we" can represent a variety of collectivities. Diplomatic speakers can use "we" to refer to their own government or country, a regional grouping, a coalition of states, the international community, or humanity as a whole. Therefore, the presence of "we" does not necessarily indicate the identity of the collective actor being constructed.

Previous research has shown that first-person plurals play a crucial role in the construction of political and institutional identities. (Guliashvili, 2023) demonstrated the deictic function of "we" in national identity construction in presidential discourse. (R.-Y. Liu, 2024) demonstrated that institutional actors can activate different referential identities through first-person plurals depending on the local interactional context. On the other hand, (Rehbein et al., 2021) methodologically demonstrated that first-person plural referents need to be disambiguated because a single pronoun can refer to various types of political collectivities. (Williams & Wright, 2024) also demonstrated that ambiguity in "we" can be related to how responsibility is distributed in political communication. Even in interpreter-mediated institutional discourse, the choice of singular or plural person reference can change according to institutional and contextual constraints (N. Liu, 2023).

In diplomatic discourse, the referential problem becomes more complex because a speaker can activate multiple political communities in the same speech. (McEntee-Atalianis & Vessey, 2025) show how person deixis can connect diplomatic actors with layered political affiliations. Therefore, the binary distinction between inclusive and exclusive "we" is not always sufficient. Political speakers can move from "national" "we" to "regional," "coalitional," or "global" "we" without changing grammatical person. Thus, the key issue is not simply whether "we" is inclusive or exclusive, but at what scale of collectivity the pronoun operates.

However, identifying the referent only answers the question of who is constructed as the collective subject. To explain how the collective is positioned as an actor, it is necessary to consider the linguistic resources that connect "we" with potential action. In this study, this relationship is analyzed through modality. Modal expressions such as "must," "should," "need to," "can," and "will" allow speakers to position actions as obligatory, necessary, possible, achievable, or intended. (Palmer, 2001) distinguishes epistemic meanings from event-oriented meanings such as deontic and dynamic modality, while (Halliday & Matthiessen, 2014) show that modality connects interpersonal meaning with propositions and proposals.

Political discourse research shows that modal forms cannot be classified solely based on their surface appearance. "Must," for example, can express obligation but also inference; "will" can express commitment or prediction; and "can" can indicate capacity or specific epistemic functions. Therefore, contextual interpretation is crucial (Almazán Ruiz & Orrequia-Barea, 2024; Orrequia-Barea & Almazán Ruiz, 2021; Y. Zhang & Cheung, 2022).

The relationship between "we" and modality itself is not entirely unexplored. (Legiansyah et al., 2025) demonstrated how "we" and modal expressions work together in the ideological construction of speeches at the Asian-African Conference. More closely related to the UN context, (Gao & Wang, 2024) demonstrated the relationship between "inclusive global-we," solidarity discourse, and collective obligations in Chinese speeches and conference interpreting at the UN General Debate. Therefore, the novelty of this research cannot be based on the claim that "we" and "modality" have never been studied together.

A more defensible gap emerges when the relationship is placed within ASEAN member states and analyzed through referential disambiguation. ASEAN collective identity has become an important theme in regional studies. (Rattanaseevee, 2023) shows that ASEAN identity has varying levels of salience between elite/diplomatic communities and wider publics. (Aras et al., 2025) also included ASEAN in a comparative study of solidarity

narratives in the UN General Assembly. In ASEAN institutional discourse, (C. Zhang et al., 2025) found consistent use of the collective pronoun "we" in Chairman's Statements and linked it to cooperation, progress, and ASEAN identity.

However, ASEAN as an institutional speaker differs from individual ASEAN member states speaking independently before the UN. When Indonesia, Malaysia, Singapore, Thailand, Vietnam, or other members deliver a General Debate statement, it cannot be assumed that "we" always means ASEAN. Speakers can shift from the national "we" to the global "we," or activate specific regional or coalitional memberships. Therefore, this research gap is not an absence gap, but a configuration gap: it remains insufficiently explained how the referential scope of "we," contextual modality, and predicated action pattern together when individual ASEAN member states speak in the same diplomatic arena.

This research approaches this issue through the concept of collective agency. (Shteynberg et al., 2022) distinguish collective identity as a representation of who a group is from collective agency as the positioning of the collective self as a subject capable of caring, desiring, or acting. (Lan & Mehta, 2022) demonstrate that the we-perspective can be a linguistic resource in the construction of collective agency. This research develops this idea by not considering all occurrences of "we" as evidence of agency. Collective agency is treated as a configuration when an identifiable collective referent is positioned through modality toward prospective or purposive action.

Based on this gap, the research proposes three research questions:

RQ1.What referential scopes of "we" are constructed in the UN General Debate speeches of ASEAN member states between 2015 and 2024, and how do their distributions vary across countries and the two study periods?

RQ2.How are different modal words associated with the referential scope of "we" in ASEAN member states' UN General Debate speeches?

RQ3.How do recurring combinations of "we" reference, modality, and predicated action discursively construct collective agency in ASEAN member states' diplomatic discourse?

This research contributes on three levels. Linguistically, it shifts attention from the frequency of first-person plurals to contextual disambiguation of referential scale. Conceptually, it integrates reference and modality to explain how differently scaled collectivities are positioned as actors. Empirically, it shifts focus from ASEAN as an institutional voice to individual member states in a common global diplomatic setting.

2. Literature Review and Theoretical Framework

2.1 First-Person Plural Reference and Political Collectivity

Personal pronouns are an important resource in political discourse because they provide an economical mechanism for establishing boundaries between speakers, institutions, audiences, and broader communities. First-person plurals have a special degree of flexibility because the lexical form "we" does not explicitly specify who is included in its referent. Referential interpretation depends on the antecedent, discourse context, speaker role, institutional setting, and surrounding propositions.

(Rehbein et al., 2021) demonstrate this complexity through systematic disambiguation of first-person plurals in parliamentary debates. (Williams & Wright, 2024) develop a contextually transparent procedure for identifying the referent "we" and demonstrate that ambiguity can function in the allocation of political responsibility. This finding is important because it demonstrates that referential ambiguity should not be addressed by forcing every occurrence into a specific category.

The political "we" also has an identitarian function. (Guliashvili, 2023) demonstrates the use of the deictic "we" in the construction of national identity, while (R.-Y. Liu, 2024) shows that institutional identity can be negotiated through shifting referents from the first-person plural. (Adams & Mitrani, 2024), in their analysis of the United Nations General Debate, show that states can present themselves through various collective "we" and multiple others. In other words, "we" provides a resource for moving speakers between levels of political belonging.

In diplomatic interactions, this flexibility can be more complex. (McEntee-Atalianis & Vessey, 2025) show that person deixis helps connect speakers with layered political affiliations. (N. Liu, 2023), through institutional interpreting, also shows that person-number

choices are related to representational roles and communicative conditions. These findings suggest that a simple analytic distinction between inclusive and exclusive “we” is not necessarily sufficient for multilateral diplomacy.

Therefore, this study uses the concept of referential scale. National “we” constructs the state as a collective subject; regional “we” constructs ASEAN or regional collectivity; coalitional “we” refers to narrower transnational groupings such as developing countries or small states; while global “we” encompasses the broader international community or humanity. Generic and ambiguous references are maintained as separate categories. This approach does not assume that the topic determines the referent. Sentences discussing ASEAN do not necessarily use regional “we”, just as sentences discussing global problems do not necessarily have a global referent. Referential classification must rely on evidence regarding the pronoun's membership, not the topic being discussed.

The separation between topic and referent is also relevant to political responsibility. A country can discuss climate change, international law, or multilateral reform while still using the national “we” to emphasize its own actions. Conversely, speakers can expand “we” to global collectivity even without explicitly mentioning terms like international community, as long as the antecedent and surrounding discourse indicate clear inclusion. Therefore, referential interpretation requires an evidence hierarchy that considers explicit naming, antecedent continuity, possessive forms like *our country* or *our world*, and changes in subject position within the paragraph. This principle aligns with (Rehbein et al., 2021) and (Williams & Wright, 2024), who suggest that referential resolution in political discourse should be treated as a contextual task, not a lexical lookup. In this study, the decision to retain AMB when evidence is insufficient was deliberate to avoid over-resolution that could artificially alter the distributions of NAT, REG, COAL, and GLOB.

2.2 Modalities in Political and Diplomatic Discourse

Modality provides linguistic means for placing propositions or events on dimensions such as possibility, necessity, obligation, capacity, and commitment. (Palmer, 2001) distinguishes epistemic modality, which relates to the speaker's judgment about the status of a proposition, from deontic and dynamic meanings, which relate to events and potential actions. (Halliday & Matthiessen, 2014), in Systemic Functional Linguistics, also demonstrate the importance of modality in connecting interpersonal meaning to proposals and propositions.

In political communication, modality plays a strategic role as speakers continually position policies, future actions, responsibilities, and predictions at varying degrees of commitment. (Orrequia-Barea & Almazán Ruiz, 2021) demonstrate how modality contributed to the reshaping of Brexit discourse, while (Almazán Ruiz & Orrequia-Barea, 2024) expand on this focus through a corpus analysis of lexical modals. (Y. Zhang & Cheung, 2022) also demonstrate the importance of contextual modal analysis in governmental press conference interpreting.

These findings support this study's decision not to use deterministic form-to-function mapping. The form “must” is not automatically coded as “obligation,” “will” is not automatically coded as “commitment,” and “can” is not automatically coded as “capacity.” All forms are examined in a concordance context. Thus, the study attempts to capture modal meaning rather than just modal form.

Four functional categories are used. OBL covers obligation and necessity; CAP covers capacity and ability; COM covers commitment, volition, or intended future action; and EPI covers inference and prediction. These categories represent operational reductions directed at research questions and do not claim to replace the entire theoretical system of modality.

This operational reduction is important because the research does not aim to produce a complete inventory of all modal meanings in UN discourse, but rather to examine how collective actors are positioned toward action. In this context, the distinction between OBL, CAP, and COM has direct analytical value. OBL positions the collective as an actor facing necessity or normative requirement; CAP represents resources or the possibility to act; while COM represents willingness or intention to undertake future action. EPI is retained because some of the same surface forms can only assess the likelihood of a state of affairs. The separation of EPI prevents predictions such as “we will face further challenges” from being

treated the same as commitments such as "we will continue our efforts." Thus, contextual modality coding serves as a filter between grammatical modalization and claims about agency.

2.3 From Collective Identity to Collective Agency

Collective identity and collective agency are related but not identical. (Shteynberg et al., 2022) emphasizes the distinction between the collective self as an object of identification and the collective self as an acting subject. Identity primarily answers the question of who we are, while agency is related to the collective subject which is directed towards intention, responsibility, capability and action.

(Lan & Mehta, 2022) demonstrate how narrative positioning and the we-perspective can build collective agency in group interactions. However, the use of "we" alone is not sufficient to infer agency. Constructions such as "we remember our history" generate collective subjectivity, but differ from "we must act," "we can achieve," or "we will continue," which more clearly position the collective subject in relation to potential action.

This study therefore uses a three-layer analytical model:

WHO — Referential scope of "we"

NAT / REG / COAL / GLOB / GEN

↓

HOW — Modal orientation

OBL / CAP / COM / EPI

↓

WHAT — Predicated action

for example work, strengthen, continue, ensure, support, or address

↓

Discursive construction of collective agency

This model allows for different types of agency to be distinguished based on scale. National and global collectivities can both appear as grammatical "we", but the modal orientation and action attached to them may differ.

2.4 ASEAN Identity and UN Diplomatic Discourse

ASEAN provides an interesting case because its regional collectivity lies between the state and the international community. (Rattanasavee, 2023) shows that the construction of ASEAN identity remains heterogeneous and has varying degrees of institutionalization at the elite and societal levels. (Aras et al., 2025) suggest that ASEAN can also emerge as a distinctive regional grouping within broader solidarity narratives in UN discourse.

In institutional ASEAN texts, the regional pronoun "we" can be particularly salient. (C. Zhang et al., 2025), through a diachronic corpus analysis of ASEAN Chairman's Statements and language-policy materials, found consistent use of the collective pronoun "we" and linked it to cooperation and ASEAN identity. This is logical because this genre institutionally grants ASEAN the authority to speak as ASEAN.

UNGD member-state speeches create different communicative configurations. Each state speaks based on its own sovereign membership. Thus, regional membership is only one of several possible collective alignments. This study leverages this difference to examine whether ASEAN becomes a salient referential scale when member states speak alone, or whether diplomatic discourse shifts more between national and global collectivities.

More broadly, UNGD-based research has demonstrated that the arena is sensitive to changes in discursive and representational patterns. (Kentikelenis & Voeten, 2021) demonstrate shifts in legitimacy discourse; (Chelotti et al., 2022) use the General Debate to understand state preference socialization; (Hecht & Steffek, 2024) demonstrate the relationship between UN conferences and the long-term salience of global terminology; and (Hecht, 2026) identifies changes in normative status dimensions through a combination of automated and manually coded analysis. This literature supports the use of the UNGD not simply as a collection of speeches, but as a comparable site of international self-positioning.

Based on this literature, the research gap in this study does not lie in the absence of research on "we", modality, ASEAN, or the UNGD. The gap lies in the specific configuration between referential scale and modal positioning. This study therefore examines whether changes in "who we are" are also related to changes in "how we are positioned to act."

3. Method

3.1 Research Design and Corpus

The research used Corpus-Assisted Discourse Studies (CADS), which combines quantitative corpus analysis with contextual qualitative interpretation. CADS was used because the research objective was not only to identify the frequency of forms but also to explain how recurring linguistic patterns operate in the surrounding discourse. This approach classically relies on movement between quantitative pattern identification and close contextual reading (Partington et al., 2013). Recent developments in CADS also emphasize systematic integration between corpus evidence and interpretive categorization (Gillings et al., 2023; Heinrich & Evert, 2024).

Data are from the United Nations General Debate Corpus (UNGDC) (Baturo et al., 2017; Jankin et al., 2025). The sampling frame includes Brunei Darussalam, Cambodia, Indonesia, Lao PDR, Malaysia, Myanmar, the Philippines, Singapore, Thailand, and Vietnam during 2015–2024. These countries constitute the ten members of ASEAN throughout the study period. Timor-Leste is not included because it only gained full ASEAN membership on October 26, 2025, after the study period ended (Association of Southeast Asian Nations, 2025).

The combination of ten countries and ten years theoretically yields 100 country-year slots. The raw corpus audit identified 92 available country-session documents. Eight unavailable slots were not imputed using speeches from other genres. The final corpus consists of 159,857 word tokens, with 48 documents covering 2015–2019 and 44 documents covering 2020–2024.

Because document count and speech length differ between countries, comparison uses normalized frequency:

$$NF = (f / N) \times 10,000$$

With f as the frequency of linguistic features and N as the total word tokens in the subcorpus. Temporal comparisons use two periods: 2015–2019 and 2020–2024. Because the corpus is not a perfectly balanced panel, the temporal analysis is read corpus-normalized and descriptive. As a sensitivity check, general temporal trends are also examined on the balanced seven-state subset with complete annual coverage.

3.2 Data Extraction and Referential Coding

Case-insensitive whole-word extraction yielded 2,401 raw occurrences of "we." Each occurrence was examined for voice attribution. Forty tokens contained in direct quotations, attributed speech, titles, slogans, or quoted institutional material were excluded because they did not represent the voice of the state statement. The RQ1 dataset thus consisted of 2,361 attributable occurrences.

Referential analysis uses a sentence as the initial context, followed by previous/next sentences and paragraph context when necessary. Coding evidence includes explicit self-identification, antecedents, named collectivities, possessive/deictic cues, and referential continuity.

Six categories are used: NAT for speaker/state/national collectivity; REG for ASEAN or an identifiable regional collectivity; COAL for an identifiable transnational coalition; GLOB for international community, broad Member-State collectivity, world community, or humanity; GEN for generalized non-specific human reference; and AMB when the referent remains defensible.

Referential classification is performed independently of modality. Modal forms are not used to help determine scope, so the analysis of RQ2 is not circular. Mentioning an institution also does not automatically determine the referent. "We support ASEAN's efforts," for example, can still be considered NAT if the antecedent indicates the state speaker, while "we in ASEAN" is explicit evidence for REG.

A total of 285 of the 2,361 attributable tokens remained classified as AMB. These tokens were not forced into another category. This approach is consistent with literature emphasizing intrinsic ambiguity and the need for contextual procedures in pronoun resolution (Rehbein et al., 2021; Williams & Wright, 2024).

3.3 Modality and Collective Agency Coding

Modal extraction begins with high-recall screening of the central modals must, should, can, could, will, would, shall, may, and might, as well as quasi-modals such as need to, have to, ought to, and be able to. Automatic extraction only produces candidates.

Each candidate was then validated to ensure the modal actually governs the predicate whose actor is "we." For example, "We believe that states must cooperate" is not considered WE + MUST because "states" is the actor of "must cooperate." After structural and contextual validation, 738 modalized-we constructions were obtained.

Modal meaning is classified into OBL, CAP, COM, and EPI. Coding depends on contextual function. "Must" in "We must strengthen cooperation" is OBL, while the inferential "must" can be EPI. "Will" can be COM when expressing intended conduct, but EPI when principally predicting future conditions. This distinction follows the principle of contextual modal interpretation (Halliday & Matthiessen, 2014; Palmer, 2001; Y. Zhang & Cheung, 2022).

For RQ3, collective agency was restricted to constructions where (1) we had an identifiable collective referent, (2) the modality included OBL, CAP, or COM, and (3) the predicate indicated prospective or purposive action. Epistemic, predominantly stative, and metacognitive constructions were excluded from the actionive core. The final dataset consisted of 643 actionive collective-agency constructions. Predicates were then dilemmatized to identify recurring actions.

3.4 Data Analysis and Coding Integrity

RQ1 was analyzed using raw and normalized frequencies, proportional referential distributions, cross-country comparisons, and comparisons between two study periods. Percentages of resolved scopes were calculated using a denominator that excluded AMB.

RQ2 was analyzed through cross-tabulation between referential scope and contextual modal function. Pearson's chi-square was used only when expected counts met assumptions, with Cramér's V as the effect-size measure. Because REG and COAL had small cell counts, the main inferential analysis focused on NAT versus GLOB.

RQ3 uses concordance-based interpretation of recurring combinations of reference, modality, and predicate. Quantitative results determine salient configurations, while interpretation is returned to the textual context, following the iterative logic of CADS.

Coding integrity was maintained through explicit rules, context preservation, conflict auditing, retention of AMB, and independence between referential and modality coding. Automatic classification served as a retrieval aid, not a final adjudicator. The study does not claim formal inter-coder reliability because final coding was not performed by multiple independent coders; instead, auditability was maintained through token-level coding records and explicit decision rules.

The corpus uses official English UNGD texts. Because translated and untranslated diplomatic discourse can have linguistic differences (Shen & Kotze, 2025), findings are not treated as the personal linguistic style of individual leaders. Analytical claims are limited to official English diplomatic discourse attributed to ASEAN member states in UN General Debate records.

4. Results

4.1 Corpus Distribution and Frequency of "We"

The final corpus contains 92 documents and 159,857 word tokens. Raw extraction yielded 2,401 occurrences of "we," equivalent to 150.20 occurrences per 10,000 words.

Table 1. Corpus composition and raw frequency of we

Country	Documents	Word tokens	Raw we	we/10,000
Brunei Darussalam	7	8,386	138	164.56
Cambodia	10	15,009	172	114.60

Country	Document s	Word tokens	Raw we	we/10,000
Indonesia	10	12,385	239	192.98
Lao PDR	10	16,386	144	87.88
Malaysia	10	21,222	346	163.04
Myanmar	6	13,730	174	126.73
Philippines	10	22,704	425	187.19
Singapore	9	17,178	323	188.03
Thailand	10	17,573	267	151.94
Vietnam	10	15,284	173	113.19
Total	92	159,857	2,401	150.20

Indonesia had the highest normalized frequency (192.98), followed by Singapore (188.03), and the Philippines (187.19), while Lao PDR had the lowest rate (87.88). Frequency only indicates how often a pronoun is used and does not explain its referential meaning.

Temporally, the overall “we” rate remained relatively stable: 151.49 per 10,000 words in 2015–2019 and 148.71 in 2020–2024. In the balanced seven-state subset, the trend also declined slightly, from 150.21 to 142.39. Thus, the more substantive change is seen in the composition of references, not in the overall “we” frequency.

4.2 Referential Scopes of “We”

After voice-attribution exclusions, 2,361 tokens were analyzed for RQ1.

Table 2. Referential distribution of attributable we

Scope	n	% attributable	% resolved
NAT	1,236	52.35	59.54
REG	34	1.44	1.64
COAL	51	2.16	2.46
GLOB	728	30.83	35.07
GEN	27	1.14	1.30
AMB	285	12.07	—
Total	2,361	100.00	—

NAT is the largest referential category, followed by GLOB. Among resolved references, the two together account for over 94% of tokens. Explicit regional ASEAN references account for only 1.64% of resolved tokens.

Cross-country distributions show variation. Myanmar has the largest national share of resolved tokens, around 83.8%, despite the country's corpus containing only six documents. Indonesia is the most balanced case, with around 44.7% NAT and 47.7% GLOB. Singapore and Lao PDR also have substantial global proportions.

These variations indicate that the dominance of NAT in the aggregate corpus is not identical to each country's profile. Brunei Darussalam, Cambodia, the Philippines, Thailand, and Vietnam continue to show national reference as the majority pattern, but with varying proportions of global reference. Indonesia is the only subcorpus where GLOB slightly exceeds NAT in the resolved data, while Singapore and Lao PDR show a global share approaching two-fifths. Because corpus size and annual coverage are not uniform, these differences are not used to rank each country's "global orientation." The country-level results

only serve to demonstrate that the aggregate ASEAN pattern is composed of heterogeneous national profiles.

A comparison of the periods shows a clearer redistribution. In 2015–2019, there were 733 NAT and 349 GLOB occurrences, while 2020–2024 produced 503 NAT and 379 GLOB occurrences. In the resolved data, NAT decreased from 64.47% to 53.57%, while GLOB increased from 30.69% to 40.36%.

Normalized NAT frequency decreased from 85.81 to 67.57 per 10,000 words, while GLOB increased from 40.86 to 50.91. The distribution of resolved scopes differed significantly between periods, $\chi^2(4) = 31.98$, $p < .001$, Cramér's $V = .124$. The small effect size suggests a modest redistribution rather than a major transformation.

4.3 Referential Scope and Modal Function

Structural and contextual validation yielded 738 modalized-we constructions.

Table 3. Contextual modal functions

Modal function	n	%
OBL	449	60.84
STAMP	122	16.53
COM	129	17.48
EPI	38	5.15
Total	738	100.00

OBL dominates the overall dataset. However, the distribution differs sharply between referential scopes.

Table 4. Referential scope $\times$ modal function

Scope	OBL	STAMP	COM	EPI	Total
NAT	158	57	98	19	332
REG	3	1	4	0	8
COAL	19	2	3	1	25
GLOB	266	59	23	15	363
GEN	3	3	1	3	10
Total	449	122	129	38	738

On NAT, OBL covers 47.59%, COM 29.52%, CAP 17.17%, and EPI 5.72%. In GLOB, OBL increased to 73.28%, while COM was only 6.34%.

For NAT and GLOB, the association between scope and modal function was significant, $\chi^2(3) = 73.27$, $p < .001$, Cramér's $V = .325$, indicating a moderate association. NAT $\times$ COM was substantially overrepresented, while GLOB $\times$ COM was underrepresented. In contrast, GLOB $\times$ OBL was above the expected frequency.

Representative GLOB $\times$ OBL construction found in Indonesia 2015:

"We must work together in realizing peace and prosperity for our people."

While NAT $\times$ COM appeared in Malaysia 2015:

"We will continue to seek justice through other legal options..."

REG and COAL are too small for inferential comparison. REG produces only eight modalized constructions, while COAL produces 25.

Temporally, OBL increased descriptively from approximately 55.7% in 2015–2019 to 65.3% in 2020–2024. However, the overall association period $\times$ modality did not reach the conventional significance threshold, $\chi^2(3) = 7.23$, $p = .065$, Cramér's $V = .099$. Therefore, the results are treated as descriptive tendencies.

4.4 Predicated Action and Collective Agency

After epistemic, stative, and primarily metacognitive constructs were removed, 643 actionive agency-core constructs were retained for RQ3.

The largest configurations are GLOB $\times$ OBL (240), NAT $\times$ OBL (148), NAT $\times$ COM (93), NAT $\times$ CAP (56), GLOB $\times$ CAP (55), and GLOB $\times$ COM (19). GLOB $\times$ OBL repeats with predicates such as work, ensure, strengthen, build, stop, and address. This pattern positions the broad international collectivity as an actor directed towards cooperative problem-solving or shared responsibilities.

NAT $\times$ COM has a different profile. The predicate "continue" is particularly prominent, appearing 16 times as the action head in this configuration. Forms like "we will continue" position the state as an actor, declaring a continuation of future conduct.

Explicit regional collective agency is very limited. One example is in Lao PDR 2022, when the antecedent established the ASEAN collectivity and the following construction used the phrase "we should continue to engage Myanmar." Coalitional agency also appears in a limited set of cases, such as the G20, developing countries, small states, and other identifiable transnational groupings.

Overall, the findings show two dominant agency configurations: global collective reference linked to obligation-oriented action and national collective reference with a substantially stronger orientation toward commitment.

5. Discussion

5.1 National and Global Scales Dominate Diplomatic "We"

The first finding shows that diplomatic discourse in ASEAN member-state UNGD discourse primarily operates between national and global scales. National references account for almost 60% of resolved occurrences, and global references for around 35%. Explicit ASEAN-regional references account for only 1.64%.

The dominance of the national scale can be understood from the institutional speaking position in the UNGD, namely the sovereign Member State. Speakers have a primary mandate to represent the country, even when discussing regional or global questions. This finding is consistent with (R.-Y. Liu, 2024), who showed that the institutional identity of the state is context-dependent, and (Adams & Mitrani, 2024), who showed that states can manage multiple collective identifications in the UN General Debate.

However, GLOB, which accounts for more than a third of resolved tokens, demonstrates that diplomatic discourse is not confined to national self-reference. Speakers repeatedly expand the collective actor to the international community, Member States, the world community, or humanity. This aligns with (McEntee-Atalianis & Vessey, 2025) view that diplomatic person deixis can connect speakers to layered political communities.

The concept that emerged from these findings is discursive rescaling. The grammatical form "we" remains the same, but the size and membership of the collective subject change. Speakers can shift from "we = our state" to "we = international community" without any explicit grammatical change other than a contextual shift. Referential scope is therefore an essential part of the political meaning of pronouns.

Cross-country variation makes it clear that rescaling is not an identical pattern across ASEAN. Indonesia, for example, exhibits a relative balance between national and global references, while Myanmar is heavily concentrated on national references. These differences should not be read as a direct measure of internationalism, regionalism, or policy preference. Pronoun distribution is influenced by rhetorical agendas, genre constraints, speaker positions, and the issues raised in a given year. However, this heterogeneity emphasizes that treating ASEAN states as a single, homogeneous discourse community will eliminate crucial variation. Aggregate analysis therefore needs to be read as a shared corpus-level tendency, while country profiles reveal the range of strategies that shape that tendency.

5.2 Why Is Explicit ASEAN "We" So Rare?

The low REG is one of the most interesting findings, but it should not be interpreted as evidence of a weak ASEAN identity. (Rattanaseevee, 2023) actually shows that ASEAN collective identity is quite salient in diplomatic and elite environments.

A more useful contrast emerges when the results of this study are compared with those of (C. Zhang et al., 2025). In the ASEAN Chairman's Statements, the collective "we" is consistent and plays a role in the construction of ASEAN identity. The difference lies in the institutional speaking position. Chairman's Statements are institutionally authorized to speak as ASEAN; UNGD statements speak as individual sovereign states.

Thus, ASEAN identity can be available as a political resource without having to become the default self-referential in every forum. The regional "we" becomes more accessible when the genre gives speakers the authority to speak on behalf of the region. Conversely, in UNGD member-state speeches, national and global scales are more readily available. The rarity of REG is more accurately understood as evidence that linguistic activation of collective identity is institutionally conditioned.

This finding also provides important nuances to the literature on regional identity. While much identity research begins with the existence of symbols, norms, institutions, or joint declarations, this corpus demonstrates how identity must be activated in actual linguistic choices. A speaker can support ASEAN, mention ASEAN centrality, or discuss regional cooperation without making ASEAN the referent of "we." In other words, thematic salience and pronominal identification are not identical. This distinction helps avoid the inference that every positive discussion about ASEAN is evidence of regional self-reference. For linguistic research, a more appropriate measure is whether speakers truly include themselves and others within the same collective subject. Because REG rarely appears in the UNGD, regional collectivity appears to function as a strategically available layer rather than a default deictic center. This opens up the possibility of comparative research between member-state speeches and ASEAN-authored genres to examine how changes in institutional authorship directly alter the pronominal scale.

5.3 Global "We" and Shared Obligations

The main finding of RQ2 is the dominance of OBL in the global "we". A total of 73.28% of modalized GLOB constructions fall into the obligation/necessity category. This pattern is consistent with (Gao & Wang, 2024), who found a relationship between the inclusive global "we" and the construction of collective obligations in UN discourse.

This study extends these findings comparatively. The GLOB $\times$ OBL relationship emerges in aggregate across all ten ASEAN member states. Furthermore, referential scope is determined independently of modal form, so global classification does not derive from the presence of must or should itself.

Constructions like "we must work together," "we need to ensure," or "we should strengthen" perform two tasks simultaneously. First, we expand the membership of the collective subject. Second, the modality places the collective under normative or practical necessity. Global collectivity is thus not simply constructed as belonging together; it is constructed as a subject that bears shared responsibility.

Action-predicate results reinforce this interpretation. GLOB $\times$ OBL frequently appears with the terms work, ensure, strengthen, build, and address. These predicates are oriented toward collective problem management. Within (Shteynberg et al., 2022) framework, the discourse moves from collective identification to collective subjectivity oriented toward action.

A comparison with (Gao & Wang, 2024) helps clarify this contribution. Their study demonstrated that global-we and obligation can be resources for solidarity in a specific national/interpretive context. The present results show that similar associations persist when the unit is expanded to ten member states, when we are resolved referentially one by one, and when the national reference is placed as a comparison category. Therefore, the important finding is not only that global-we frequently co-occurs with must or need to, but that propensity toward obligation increases as the referential scale moves from national to global. The measurable differences between NAT and GLOB transform the previous observation into a relational finding: a collective actor's modal profile is related to the scale of actorhood being constructed.

However, the linguistic construction of obligation should not be equated with actual solidarity or implementation. (Aras et al., 2025) demonstrate broader solidarity narratives in regional groupings within UN discourse, but the statement "we must act" remains a discursive

positioning. This study measures how responsibility is represented, not whether coordinated action actually occurs.

5.4 National “We” and Self-Commitment

National “we” show a different profile. OBL remains the largest function, but COM reaches 29.52%, almost five times the proportion of COM in GLOB. The chi-square residual pattern confirms that $\text{NAT} \times \text{COM}$ is substantially overrepresented.

One of the most plausible explanations is discursive authority. State representatives have relatively limited authority to declare what their country will or shall do. Speakers can self-commit to the future conduct of the collective they institutionally represent. In contrast, individual state speakers have much more limited authority to promise what the entire international community will do.

The dominance of the predicate continue in $\text{NAT} \times \text{COM}$ supports this explanation. We will continue... connects past/present state action with promised continuity. This construction establishes a national actor with a temporal trajectory and the capacity to commit itself to future conduct.

This difference can be summarized as NATIONAL WE $\rightarrow$ stronger language of self-commitment, while GLOBAL WE $\rightarrow$ stronger language of shared obligation. This relationship is one of the research's main contributions because it shows that changing referential scope is not simply an identity shift. The scale of the referent is related to the type of agency that speakers can plausibly claim or allocate.

5.5 Scalar Collective Agency

The overall results support the conceptualization of collective agency as a scalar discursive configuration. (Lan & Mehta, 2022) demonstrated the importance of the we-perspective in constructing collective agency, but this study suggests that the we-perspective needs to be differentiated based on its referential scale.

The resulting model is WHO — collective scale + HOW — modal orientation + WHAT — predicated action = mode of collective agency. National and global groups are not just differently sized groups. They acquire different modal profiles. National collectivity is more available for self-commitment, while global collectivity is more available as a locus of shared obligation.

Thus, referential rescaling also constitutes a redistribution of responsibility and commitment. When the speaker shifts from the national to the global “we”, agency is not lost but changes form: from an actor relatively capable of promising its own conduct to a broader actor entrusted with normative responsibility.

This distinction can be interpreted as a question of the relationship between scale and authority. The national “we” has a clearer institutional principal: speakers speak as representatives of states that can, at least discursively, make commitments regarding their own policies or conduct. The global “we” has a much more diffuse membership. No single national speaker has the authority to guarantee the actions of the entire international community. Therefore, obligation is a modal solution that is pragmatically more available than commitment: speakers can state what all parties must do without claiming the authority to promise that all parties will do it. The $\text{NAT} \times \text{COM}$ and $\text{GLOB} \times \text{OBL}$ findings can thus be understood not simply as frequency differences, but as manifestations of unequal commitment authority on different scales of collective reference.

This framework also helps distinguish collective agency from collective efficacy or actual capability. CAP does appear in the corpus, but it does not dominate either NAT or GLOB. This means that diplomatic discourse more often regulates agency through normative requirements and self-commitment than through explicit claims that the collectivity is capable of doing something. This finding serves as a reminder that “agency” in this study is a representational relationship between subject and action, not a psychological measurement of perceived group efficacy. By limiting these claims, the scalar concept of collective agency remains close to the available linguistic evidence.

REG and COAL demonstrate that the scale is not binary. Political subjects can move through national, regional, coalitional, and even global scales. However, the empirical distribution is highly unequal. National and global positions are dominant, while regional and coalitional positions are available more selectively.

5.6 Temporal Redistribution

Temporal findings show a decrease in NAT and an increase in GLOB in the 2020–2024 period. Because the overall frequency of “we” is relatively stable, this change does not mean diplomatic speakers have become more “collective.” What has changed is the composition of the collectivity.

These results are interesting when placed alongside studies on discursive change in the UNGD. (Hecht & Steffek, 2024) show that international terminology can experience long-term salience shifts following global institutional developments, while (Hecht, 2026) shows that social-identification patterns in the General Debate change as the normative international agenda evolves.

However, this study does not claim a causal explanation. The second period did include the pandemic, geopolitical tensions, food and energy pressures, and intensified discussions on climate and multilateral cooperation. However, corpus comparison alone is insufficient to conclude that any one of these events caused the increase in global “we”. The effect size of the referential period association is also small. Therefore, the most defensible result is a modest shift toward broader global reference, not a transformation of ASEAN diplomatic identity.

Modality showed a similar trend but weaker evidence. OBL increased from approximately 55.7% to 65.3%, but the period $\times$ modality test had a $p = .065$. Therefore, this change is only a descriptive tendency.

This caution is important because the 2015–2019 versus 2020–2024 periodization is an analytical convenience balanced in duration, not a natural historical boundary that demonstrates regime change in diplomatic language. In addition to the pandemic, the second period encompasses speaker changes, domestic political transitions, international conflicts, and a different UN agenda. Each of these factors can influence lexical and referential choices. The current research design allows for identification of temporal redistribution but does not perform a causal decomposition of the sources of these changes. Future research could examine annual turning points, link scope shifts to topic distributions, or compare event-specific subcorpora. In this article, maintaining claims at the level of modest redistribution keeps interpretation commensurate with the effect sizes found and prevents longitudinal results from being burdened with untested causal narratives.

5.7 Contributions and Limitations

This research makes three contributions. First, theoretically, it adds a referential scale to the conceptualization of collective agency. Collective agency is not only a matter of whether a group speaks as an actor, but also the scale at which that actor is constructed.

Second, empirically, research shows that individual ASEAN states in the UNGD primarily move between national and global collective positions. The rarity of explicit ASEAN discourse contrasts with institutional ASEAN discourse and demonstrates the importance of speaking position in the linguistic activation of regional identity.

Third, methodologically, the study demonstrates the need to separate referential coding from modality. If we must be assumed to be global because it is seen as a collective obligation, and then the global reference is tested against the obligation, the result becomes circular. Independent scope adjudication in this study avoids this problem. CADS is also appropriate because automated frequency helps identify systematic patterns, while contextual examination is necessary for referential disambiguation and semantic modality classification (Gillings et al., 2023; Heinrich & Evert, 2024; Partington et al., 2013).

The decision to retain 285 AMB tokens also represents a methodological contribution. In corpus research, the drive to obtain a fully classified dataset can result in false precision when the true categories depend on discourse context that doesn't provide sufficient antecedents. By retaining AMB, the aggregate percentage resolved reference becomes slightly more conservative, but the boundary between observed evidence and analyst inference becomes clearer. A similar principle applies to modal extraction: automatic screening is used to achieve high recall, while the inclusion of 738 validated constructions requires verification that the modal is truly present in the predicate whose actor is “we”. The combination of retrieval automation and contextual adjudication allows analysis to be scaled without treating machine rules as a substitute for interpretation.

Several limitations should be noted. First, the corpus is not fully balanced because eight country-year documents were unavailable. Normalization and a balanced-subset sensitivity check mitigate but do not eliminate the consequences of unequal coverage. Second, 285 “we” tokens remain AMB. This retention increases caution but also reduces the amount of fully resolved data. Future research could use multiple independent coders to measure inter-coder reliability.

Third, official English UNGD texts can include translated or interpreter-mediated discourse. (Shen & Kotze, 2025) show that translated and untranslated UNGD discourse can have linguistic differences. Therefore, the findings apply to official English diplomatic texts and should not be interpreted as direct personal speech style. Fourth, the operationalization of collective agency focuses on modality and action predicates. Agency can also be realized through lexical resources such as pledge, commit, intend, seek, or remain determined without modal auxiliaries. Future research could integrate lexical stance, speech acts, transitivity, or evaluative language.

Finally, linguistic agency is not synonymous with actual political behavior. Global obligations can be articulated without resulting in coordinated action, and national commitments in speech do not automatically translate into implementation. The scope of this research is discursive construction, not policy-compliance measurement.

6. Conclusion

This study analyzes how the referential scope of “we”, modality, and predicated action interact in the construction of collective agency in the UN General Debate discourse of ten ASEAN member states during 2015–2024.

The results of RQ1 show that national and global references are the two dominant collective scales. National “we” covers 59.54% of resolved occurrences, while global “we” covers 35.07%. Explicit ASEAN-regional “we” only accounts for 1.64%. Thus, ASEAN membership does not automatically make ASEAN a default collective self when member states speak for themselves in global diplomatic forums.

RQ2 shows that obligation/necessity is the dominant modal orientation, accounting for 60.84% of validated modalized constructs. However, modality is not distributed uniformly across referential scopes. Global “we” is strongly associated with obligation, while national “we” has a much stronger commitment orientation. The NAT/GLOB $\times$ modal function association is significant and has a moderate effect size.

RQ3 shows that these differences are also reflected in action predicates. Global obligation often directs broader collectivity toward joint actions such as work, ensure, strengthen, and address. National commitment, on the other hand, is strongly characterized by constructs such as “we will continue,” which position the state as an actor capable of self-committing to future conduct.

Based on these results, the study proposes the concept of scalar collective agency. Collective agency is linguistically formed through the configuration of referential scale, modal orientation, and predicated action. In this corpus, global collectivity is primarily constructed as an actor obliged to act, while national collectivity is more readily constructed as an actor committed to acting.

The findings also confirm that collective identity and linguistic activation are not identical. A low regional “we” does not prove a weakness in ASEAN identity; it shows that the institutional context influences which collective scale is activated. ASEAN can speak strongly as “we” in institutional ASEAN discourse, while individual member states move more between the national self and the global collectivity at the UN.

Thus, “we” in diplomatic discourse is not simply a pronoun of belonging. It becomes a linguistic mechanism for shifting the scale of political subjects and distributing responsibility, capability, and commitment among various levels of collectivity.

More broadly, these results demonstrate the benefits of combining corpus-level comparison with referential close reading in the study of diplomatic language. This approach allows recurring patterns to be quantitatively examined without detaching pronouns and modalities from the context that determines their meaning. The same framework can be applied to other regional groups, other institutional genres, or specific crisis periods to test whether the relationship between scale, obligation, and commitment is unique to ASEAN-UNGG or represents a more general pattern in multilateral diplomacy. Therefore, this study

provides not only findings about ASEAN member states but also a replicable analytical framework for examining how political collectivities are shaped and given agency through language.

References

- Adams, T., & Mitrani, M. (2024). Juggling identities: Identification, collective memory, and practices of self-presentation in the United Nations General Debate. *The British Journal of Politics and International Relations*, 26(2), 277–298. <https://doi.org/10.1177/13691481231156906>
- Almazán Ruiz, E., & Orrequia-Barea, A. (2024). A corpus study of Brexit political discourse: Exploring modality through lexical modals. *Miscelánea: A Journal of English and American Studies*, 70, 67–92. https://doi.org/10.26754/ojs_misc/mj.20249837
- Aras, B., Dagci, K., & Al-Khulaifi, K. A. (2025). The heartbeat of unity: Exploring solidarity in the Trans-Atlantic, Asia, and Muslim world. *Journal of Peacebuilding & Development*, 20(1), 69–85. <https://doi.org/10.1177/15423166251317941>
- Association of Southeast Asian Nations. (2025). *Chairman's Statement of the 47th ASEAN Summit, Kuala Lumpur, Malaysia, 26 October 2025*.
- Baturo, A., Dasandi, N., & Mikhaylov, S. J. (2017). Understanding state preferences with text as data: Introducing the UN General Debate corpus. *Research & Politics*, 4(2), 1–9. <https://doi.org/10.1177/2053168017712821>
- Brun-Mercer, N. (2021). Women and men in the United Nations: A corpus analysis of General Debate addresses. *Discourse & Society*, 32(4), 443–462. <https://doi.org/10.1177/0957926521992145>
- Chelotti, N., Dasandi, N., & Mikhaylov, S. J. (2022). Do intergovernmental organizations have a socialization effect on member state preferences? Evidence from the UN General Debate. *International Studies Quarterly*, 66(1), sqab069. <https://doi.org/10.1093/isq/sqab069>
- Gao, F., & Wang, B. (2024). Reconstructing solidarity discourse through conference interpreting at the general debates of the United Nations (2008–2021). *The Translator*, 30(4), 553–570. <https://doi.org/10.1080/13556509.2024.2367275>
- Gillings, M., Mautner, G., & Baker, P. (2023). *Corpus-Assisted Discourse Studies*. Cambridge University Press. <https://doi.org/10.1017/9781009168144>
- Guliashvili, N. (2023). National identity revisited: Deictic WE in President Zelenskyy's speeches on Russia-Ukraine war. *Journal of Language and Politics*, 22(6), 779–801. <https://doi.org/10.1075/jlp.22106.gul>
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's Introduction to Functional Grammar* (4th ed.). Routledge.
- Hecht, C. (2026). Dynamics of salient normative status dimensions and issues in a changing international order. *Global Constitutionalism*, 15(1), 33–61. <https://doi.org/10.1017/S2045381725000073>
- Hecht, C., & Steffek, J. (2024). Salient discourses in international society: When and how have United Nations global conferences acted as catalysts? *Journal of International Relations and Development*, 27(3), 283–314. <https://doi.org/10.1057/s41268-024-00324-7>
- Heinrich, P., & Evert, S. (2024). Operationalising the hermeneutic grouping process in corpus-assisted discourse studies. *Proceedings of the 4th Workshop on Computational Linguistics for the Political and Social Sciences*, 33–44.
- Jankin, S., Baturo, A., & Dasandi, N. (2025). Words to unite nations: The complete United Nations General Debate Corpus, 1946--present. *Journal of Peace Research*, 62(4), 1339–1351. <https://doi.org/10.1177/00223433241275335>
- Kentikelenis, A., & Voeten, E. (2021). Legitimacy challenges to the liberal world order: Evidence from United Nations speeches, 1970–2018. *The Review of International Organizations*, 16, 721–754. <https://doi.org/10.1007/s11558-020-09404-y>
- Lan, L., & Mehta, C. M. (2022). Constructing collective agency through narrative positioning in group meetings within a Chinese professional team. *Language & Communication*, 87, 179–190. <https://doi.org/10.1016/j.langcom.2022.07.008>
- Legiansyah, G., Darmayanti, N., & Citraesmana, E. (2025). Examining “we” and modality as ideological practice in opening speeches of Asian-African Conference: A critical discourse analysis. *Law and Humanities Quarterly Reviews*, 4(3), 1–17. <https://doi.org/10.31014/aior.1996.04.03.151>
- Liu, N. (2023). Speaking in the first-person singular or plural: A multifactorial, speech corpus-based analysis of institutional interpreters. *Interpreting*, 25(2), 239–273. <https://doi.org/10.1075/intp.00088.liu>
- Liu, R.-Y. (2024). Constituting institutional identity in political discourse: The use of the first-person plural pronoun in China's press conferences. *Language in Society*, 53(3), 421–444. <https://doi.org/10.1017/S0047404523000386>
- McEntee-Atalianis, L., & Vessey, R. (2025). Putting negotiation on a ‘principal-ed’ footing: A corpus-informed discourse analysis of person deixis in diplomatic debates. *Journal of Pragmatics*, 247, 42–56. <https://doi.org/10.1016/j.pragma.2025.07.013>
- Orrequia-Barea, A., & Almazán Ruiz, E. (2021). Modelling through modality: (Re)shaping Brexit. *ES Review. Spanish Journal of English Studies*, 42, 127–153. <https://doi.org/10.24197/ersjes.42.2021.127-153>
- Palmer, F. R. (2001). *Mood and Modality* (2nd ed.). Cambridge University Press.
- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*. John Benjamins. <https://doi.org/10.1075/sci.55>
- Rattanasavee, P. (2023). Understanding and prospect of identity construction in ASEAN integration. *International Studies*, 60(4), 379–400. <https://doi.org/10.1177/00208817231213500>
- Rehbein, I., Ruppenhofer, J., & Bernauer, J. (2021). Who is we? Disambiguating the referents of first person plural pronouns in parliamentary debates. *Proceedings of the 17th Conference on Natural Language Processing (KONVENS 2021)*, 147–158.
- Shen, L., & Kotze, H. (2025). Linguistic indicators differentiating translated and untranslated diplomatic discourse: A diachronic analysis of the United Nations General Debate (1946–2022). *Lingua*, 320, 103933. <https://doi.org/10.1016/j.lingua.2025.103933>
- Shteynberg, G., Hirsh, J. B., Garthoff, J., & Bentley, R. A. (2022). Agency and identity in the collective self. *Personality and Social Psychology Review*, 26(1), 35–56. <https://doi.org/10.1177/10888683211065921>

- Williams, J., & Wright, D. (2024). Ambiguity, responsibility and political action in the UK daily COVID-19 briefings. *Critical Discourse Studies*, 21(1), 76–91. <https://doi.org/10.1080/17405904.2022.2110132>
- Zhang, C., Zhao, R., & Huang, Y. (2025). The evolution of ASEAN's language policies: A diachronic analysis of official documents and website. *PLOS ONE*, 20(1), e0315076. <https://doi.org/10.1371/journal.pone.0315076>
- Zhang, Y., & Cheung, A. K. F. (2022). A corpus-based study of modal verbs in Chinese--English governmental press conference interpreting. *Frontiers in Psychology*, 13, 1065077. <https://doi.org/10.3389/fpsyg.2022.1065077>