

How Far from Genocide? The Semiotic Construction of African Foreign Nationals as Existential Threats in South Africa

Peter Junior Tshetu,^{1*}

Design and Future Learning, Faculty of Humanities, EDUVOS, South Africa

*) Corresponding Author

Email: peterjuniortshetu@gmail.com

DOI: [10.18326/jopr.v8i3.934-971](https://doi.org/10.18326/jopr.v8i3.934-971)

Submission Track:

Received: 14-07-2026

Final Revision: 18-09-2026

Available Online: 01-10-2026

Copyright © 2026 Authors



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

Abstract

South Africa has experienced recurrent violence against African foreign nationals for more than three decades, while organized anti-migrant movements since 2021 have shifted hostile rhetoric from the political margins to the center of public discourse. This study argues that the language used to describe African foreign nationals constitutes dangerous speech that constructs them as existential threats and creates conditions conducive to collective violence. Drawing on Susan Benesch's Dangerous Speech Framework (DSF), Saussurean syntagmatic and associative meaning relations, and George Stanton's staged model of genocide, the study pursues three objectives: to identify the discursive construction of African foreign nationals as existential threats; to determine the prevalence of the five dangerous-speech hallmarks; and to assess South Africa's position on Stanton's genocide continuum. The corpus comprises 2,324 social media posts from *X* and *Facebook*, together with 214 *YouTube* news segments relating to the activities of Operation Dudula, the March and March movement, and the MK Party between 2021 and 2026. The findings reveal the presence of all five dangerous-speech hallmarks, with dehumanization emerging as the most frequent (60.8%) and threats to women and children the least frequent (27.6%). These patterns are realized through identifiable linguistic mechanisms, including derogatory labels, metaphors of disease and infestation, criminality modifiers, and expulsive imperatives. The study introduces the Dangerous Speech-Genocide

Escalation (DSGE) model, which integrates linguistic analysis with the DSF and Stanton's framework. Applied to South Africa, the model indicates that the country is not experiencing genocide but has progressed significantly along the pre-genocidal pathway, highlighting the urgent need for preventative intervention.

Keywords: *Dangerous Speech, African Foreign Nationals Genocide, Anti-migrant discourse*

INTRODUCTION

Xenophobia is a persistent and, at times, deadly feature of post-apartheid South Africa. Although the democratic transition of 1994 promised an inclusive society founded on the constitutional values of equality, dignity and human rights, hostility towards African migrants has become an enduring feature of the social and political landscape (Adonis, 2025). Violent attacks, forced displacement, the looting of migrant-owned businesses and public campaigns demanding the removal of undocumented immigrants have progressively normalised antagonism towards non-nationals of African origin (Cinini and Mkhize, 2021). Over three decades the country has seen at least four major waves of collective anti-migrant violence, in 2008, 2015, 2019 and 2026, alongside a continuous stream of lower-intensity attacks (Crush, 2008; Hayem, 2013; Matsinhe, 2011). This study begins from the premise that these are not merely episodes of social unrest but the visible surface of a linguistic process, and that understanding how language constructs the African foreigner as an existential threat is a precondition for understanding, and preventing, the violence that follows.

Public debate about immigration is deeply polarised. Many South Africans frame campaigns against undocumented migrants as a legitimate response to irregular immigration, rising unemployment, pressure on public services, crime and economic competition, and as a lawful effort to protect scarce national resources (Visage and Castel-Branco, 2026). Migrants, civil-society organisations, human-rights advocates and many scholars counter that such campaigns routinely exceed concerns about immigration enforcement and instead reproduce xenophobic attitudes that indiscriminately target African foreigners regardless of legal status, erasing the distinction between documented and undocumented persons and producing collective suspicion, dehumanisation and violence (Kaziboni et al., 2022; Gordon, 2024). Language occupies the centre of these

competing narratives. Public discourse has become saturated with linguistic expressions that portray foreign nationals as criminals, invaders, parasites and economic threats, and as the source of the country's socio-economic ills. These representations circulate most intensively today across digital platforms, where they construct foreigners as legitimate targets of exclusion while simultaneously legitimising discrimination and violence (Naicker, 2016, p. 46).

Structural drivers: inequality, unemployment and governance failure

Dangerous speech does not travel in a vacuum; its potency depends on context, and the South African context is unusually enabling. The country remains one of the most unequal societies in the world, with official unemployment persistently above thirty per cent and youth joblessness far higher (Shah, 2022, p. 7). Overcrowded clinics, deteriorating schools, housing backlogs and endemic violent crime form the backdrop against which migrants become convenient explanations for structural failure (Crush, 2008; Kaziboni et al., 2022). Public-opinion research is unambiguous that anti-immigrant sentiment is widespread and rising: analysis of the South African Social Attitudes Survey shows the share of adults willing to welcome all immigrants falling steeply since 2003, with the Human Sciences Research Council warning that sentiment has gone “from bad to worse” and that the 2026 local-government elections may act as an accelerant (Human Sciences Research Council, 2026; Gordon, 2022).

Crucially for a linguistic account, relative deprivation and perceived competition, rather than actual contact or objective threat, are the stronger predictors of hostility (Gordon, 2024). This is precisely the pattern a semiotic reading would expect, since dangerous speech works by manufacturing threat perception through the meanings attached to words rather than by reporting a state of the world.

A genealogy of xenophobic violence

The contemporary moment sits atop a long lineage of violence and the value of that lineage for the present argument lies less in its recurrence than in its pattern: each wave is preceded by an identifiable intensification of eliminationist language, and the specific

content of that language, not merely the presence of grievance, tracks the form the subsequent violence takes. In May 2008, attacks that began in Alexandra spread nationwide, leaving more than sixty people dead and tens of thousands displaced, with African foreigners specifically targeted (Matsinhe, 2011; Hayem, 2013). In 2015, remarks attributed to a prominent traditional leader, widely interpreted as telling foreigners to leave, were followed by violence in Durban and beyond, after which the state's Operation Fiela blurred the line between law enforcement and collective punishment (Pollit, 2015): the sequence, public exhortation to leave followed by removal by force, recurs, elaborated and organisationally amplified, in every subsequent wave.

In September 2019, renewed attacks in Johannesburg and Pretoria targeted foreign-owned businesses, again following sustained framing of migrants as economic parasites. In April 2022, the Zimbabwean national Elvis Nyathi, a forty-three-year-old father of four, was beaten and burnt to death by a vigilante mob in Diepsloot after being identified as an “illegal foreigner”, a label that functioned there as a license to act rather than a neutral description; the killing was condemned by the African Commission on Human and Peoples' Rights, the United Nations and the University of Pretoria's Centre for Human Rights, the last explicitly describing it as the culmination of persistent hate speech and propaganda against non-citizens rather than as an isolated or spontaneous act (African Commission on Human and Peoples' Rights, 2022).

This recurrence does more than establish a template for collective violence: it supplies the evidentiary basis for treating language as a cause of escalation rather than a passive index of pre-existing tension. If anti-migrant discourse merely reflected static structural grievance, its intensity should track that grievance and its violent outcomes should be roughly uniform in trigger and form across the four waves; what the record shows instead is a tightening sequence in which each episode begins with a naming or accusation, is licensed by a period of unpunished circulation, culminates in violence, and is followed by a further, more organised round of rhetoric, the impunity feedback loop that the study's Dangerous Speech-Genocide Escalation model, developed below, identifies as

escalation's engine. The Nyathi killing renders the mechanism explicit rather than inferential: the causal path runs from a specific linguistic act, the public identification of a person as an "illegal foreigner", through community uptake of that label, to a specific act of lethal violence attributable to it, not merely coincident with it. It is this demonstrated, mechanism-level chain from rhetoric to killing, repeated and intensifying across four decades of waves, rather than the mere co-occurrence of hostile language and violent episodes, that grounds the study's claim that anti-migrant language functions causally in escalation rather than simply reflecting tensions that would produce violence with or without it.

Organized movements and the mainstreaming of anti-migrant discourse

From 2021, anti-migrant mobilisation acquired durable organisational form. Operation Dudula, whose isiZulu name means "to force out", began as a Soweto-based movement against crime and unemployment before focusing on foreign nationals, marching on shops, clinics and schools and demanding that residents produce identity documents (Human Rights Watch, 2024). It was joined in 2024 by the March and March movement, led by former radio presenter Jacinta Ngobese-Zuma, which issued an ultimatum for undocumented foreigners to leave the country by 30 June 2026, a deadline that spread fear among refugee communities and prompted a large state security response (Daily Maverick, 2026).

Human Rights Watch (2024) documented how candidates in the 2024 general elections scapegoated and demonised foreign nationals, moving anti-migrant rhetoric from the political fringe to the centre; monitoring by Xenowatch recorded scores of incidents across this period. Analysts have noted links between March and March and the uMkhonto weSizwe (MK) Party, whose 2025 manifesto pledged to prioritise South African workers over foreign nationals, illustrating how party-political actors have absorbed and amplified movement discourse (Daily Maverick, 2026). This mainstreaming is decisive: when influential political and institutional figures echo movement claims, the reach and legitimacy, and therefore the dangerousness, of the speech increase sharply.

The June 2026 deadline and the current moment

The immediate context of this study is the discourse surrounding the 30 June 2026 ultimatum and the demonstrations that have accompanied it. In the weeks preceding the deadline, calls for undocumented and, in practice, all African foreigners to leave intensified across *X*, *Facebook* and mainstream broadcast coverage, with the deadline functioning as a shared temporal anchor around which hashtags, rallies and news cycles converged. June 2026, reports documented killings of foreign nationals, the displacement of thousands, and widespread fear among Malawian and Zimbabwean nationals, although march organisers maintained that their campaign targeted only undocumented migrants and criminality (Daily Maverick, 2026). Continued demonstrations against “illegal foreigners” outside home-affairs offices, clinics and informal-trading sites, and their live amplification on social media, make the present moment not a historical case study but an unfolding one. It is this live, platform-mediated discourse, rather than an archive of past statements, that the study analyses.

Institutional recognition: the November 2025 High Court finding

The gravity of this discourse has been recognised by the judiciary. On 4 November 2025, the Gauteng High Court found that Operation Dudula had engaged in intimidation, harassment, incitement to violence and hate speech on the grounds of nationality, social origin and ethnicity, and interdicted the movement and its leaders from such conduct, including hate speech on social media (*Kopanang Africa Against Xenophobia and Others v Operation Dudula and Others*, 2025). The Court characterised the movement’s conduct as a manifestation of “xeno-racism”, directed disproportionately at black African non-nationals, and affirmed that human dignity has no nationality. It further found that the state had failed to implement the National Action Plan to Combat Racism, Racial Discrimination, Xenophobia and Related Intolerance, including its early-warning and rapid-response mechanism. That an independent court, applying the Equality Act and the Regulation of Gatherings Act, reached findings of incitement and hate speech provides external validation of this study’s premise that the discourse under analysis is not benign political

expression. The judgment is taken here as an institutional confirmation of danger already legible in the language itself.

The gap, the purpose and the research questions

Existing scholarship offers rich accounts of xenophobia's structural drivers (Crush, 2008; Kaziboni et al., 2022), its ideological grammar (Matsinhe, 2011; Naicker, 2016) and its measurement through attitude surveys (Gordon, 2022, 2024), while human-rights bodies have documented incitement and toxic rhetoric (Human Rights Watch, 2024). What remains under-developed is a properly linguistic account that specifies the sign-relations through which the African foreigner is constructed as an existential threat, that quantifies the frequency of the resulting dangerous-speech hallmarks, and that translates those findings into a defensible estimate of how close the society stands to atrocity. This study addresses that gap.

The purpose of the study is accordingly to situate the discursive construction of African foreign nationals as existential threats in South Africa using the Dangerous Speech Framework; to detail the frequency of each of its five hallmarks within the discourse; and to position that discourse within Stanton's (2013) staged model of genocide so as to ascertain how far from a genocide South Africa is. From this purpose the study derives a single overarching research question, with three sub-questions: How does the social-media and mainstream-media discourse of Operation Dudula, the March and March movement and the MK Party discursively construct African foreign nationals as existential threats; with what frequency do the five hallmarks of dangerous speech occur; and how far along Stanton's staged model of genocide does this discourse place South Africa?

- a. Through which syntagmatic and associative meaning relations is the African foreigner constructed as an existential threat across *X*, *Facebook* and *YouTube*?
- b. What is the frequency of each of the five Dangerous Speech hallmarks in this corpus, overall and by platform?
- c. When the hallmarks are mapped onto Stanton's ten stages, how far from genocide is South Africa?

Conceptual and Theoretical Framework

The study is built on three articulated components: a linguistic method for describing how meaning is made (syntagmatic and associative analysis); a risk framework for judging whether that meaning is dangerous (the DSF); and a staged model for locating the danger on a trajectory toward atrocity (Stanton). Their integration constitutes the study's theoretical contribution and is set out at the end of this section.

A framework that treats hostile language as a risk factor for violence must also specify what it does not capture, since a theory of dangerous speech that could not distinguish an ordinary, constitutionally protected argument about immigration policy from an incitement to expel or harm a population would be analytically useless and, in a democratic society, dangerous in its own right. The DSF answers this problem structurally rather than lexically: dangerousness is not read off individual words or topics, immigration numbers, border enforcement and the fiscal cost of undocumented migration remain legitimate objects of political debate in any democracy, but off the convergence of message content with the five contextual variables that condition an audience's likely response (Benesch, 2013; Leader Maynard and Benesch, 2016).

A call to tighten enforcement of immigration law, made by a minor commentator with no organisational reach, addressed to an audience with no history of collective violence against migrants, carries negligible risk under the framework regardless of its rhetorical vigour. The same proposition, issued by a mass movement with demonstrated capacity to mobilise crowds, to an audience primed by documented relative-deprivation grievance (Gordon, 2024), in a society with an unpunished record of xenophobic killing (African Commission on Human and Peoples' Rights, 2022), and packaged with a departure deadline, crosses into dangerous speech not because the topic has changed but because the context multiplies the same words' capacity to move an audience toward violence. This consequentialist, context-weighted design is what allows the framework to code content rather than viewpoint: opposition to irregular migration is not, in itself, a hallmark; dehumanising the migrant, reversing victim and aggressor, or issuing a coded

expulsion order are. The distinction tracks, without replicating, doctrines used in comparable democratic jurisdictions to separate protected offensive speech from incitement; South African case law under the Equality Act likewise asks not whether speech criticises a group but whether it advocates hatred that constitutes incitement to cause harm, and the November 2025 Gauteng High Court finding against Operation Dudula turned on exactly this distinction between critique of immigration policy and incitement on the grounds of nationality (*Kopanang Africa Against Xenophobia and Others v Operation Dudula and Others*, 2025).

The framework's reliability nonetheless has limits that this study treats as a methodological constraint rather than a settled matter. Because the DSF is a probabilistic risk instrument rather than a bright-line legal test, its coding decisions retain an interpretive component, and two risks follow. The first is a false-positive risk: strong civic language used loosely, describing a policy failure as a "crisis" or migration as a "flood" in an otherwise neutral report, can superficially resemble a dangerous-speech hallmark without the contextual conditions that make it dangerous.

The second is a false-negative risk: the coded, deniable language documented under the coded-call-to-action hallmark is designed precisely to satisfy legal and normative standards of legitimate debate while functioning, for its intended audience, as an exhortation. To manage the first risk, coding in this study required convergence rather than isolated triggers: an item was coded as exhibiting a hallmark only where a syntagmatic or associative marker co-occurred with corroborating contextual evidence, speaker identity, platform, and proximity to a mobilising event such as the 30 June 2026 deadline, so that a single hostile word in an otherwise policy-focused post was not sufficient on its own to code the item as dangerous speech. To manage the second, the analysis attended explicitly to indirection, agentless imperatives, loyalty tests, coded ultimata, as a distinct hallmark rather than requiring explicit incitement, following Benesch's (2013) own argument that dangerous speech characteristically avoids the language that would trigger legal sanction.

The framework's distinguishing power is accordingly real but bounded: it separates the two categories reliably at the extremes, ordinary policy argument on one side, vermin lexis and departure ultimatums on the other, while acknowledging, and coding conservatively around, the deniable middle ground that democratic societies' own protections for political speech make available to dangerous-speech actors.

A further question remains: does coding conservatively at the deniable middle ground risk reproducing exactly the conditions that let coded incitement succeed, and so under-count the actual risk? The concern is granted rather than deflected. Because convergence coding requires corroborating context before an item is classified as a hallmark, the design is asymmetric by construction: its error, where it errs, runs toward false negatives rather than false positives, so the frequencies reported in Table 3 and Table 4 are properly read as a floor on the corpus's dangerous-speech content rather than a ceiling. Three considerations bound the resulting risk of underestimation rather than eliminate it. First, indirection is coded as its own hallmark precisely so that agentless imperatives and coded ultimatums, the material a purely lexical trigger would miss, are captured through pattern and recurrence rather than through an isolated form of words. Second, the requirement is convergence with context, not certainty of intent: an item recurring across speaker, platform and proximity to the 30 June 2026 deadline is coded even where no single utterance would independently satisfy a legal test for incitement. Third, this asymmetry is the same one the early-warning literature recommends for atrocity-prevention instruments, which tolerate missed detections over false alarms given the respective costs of each (Harff, 2003); a coding design calibrated the other way, toward maximal sensitivity, would inflate the hallmark counts at the price of coding ordinary policy debate as dangerous speech, which would corrode the very distinction the framework exists to preserve. The reported prevalence figures should accordingly be read as conservative, not exhaustive, estimates of the corpus's dangerous-speech content.

Syntagmatic and associative meaning relations

The linguistic core of the study is Saussure's account of the two axes along which any sign acquires value (Saussure, 1916/2011). On the syntagmatic axis, signs derive meaning from their combination with the other signs that precede and follow them in a linear chain; the value of a term is shaped by what it is placed next to. On the associative axis, which Saussure termed *rapports associatifs* and which later structuralists reframed as the paradigmatic axis, a sign derives meaning from the network of other signs it evokes in absentia, the mental family of related terms from which it was selected and against which it is contrasted (Saussure, 1916/2011; Chandler, 2017). Meaning, on this account, is relational and differential: a word means by virtue of its combinations and its selections, not by naming a thing directly.

This apparatus is the natural instrument for analysing dangerous speech, because dehumanisation and threat are achieved precisely through combination and selection rather than through explicit predication. Syntagmatically, the study attends to collocation, the characteristic company a word keeps (Sinclair, 1991): which modifiers attach to *foreigner*, which verbs take *foreigners* as object, which possessive determiners frame the national "we". Associatively, it attends to lexical selection and connotation: what is chosen when *makwerekwere*, *cockroach* or *invasion* is used instead of the neutral alternatives *migrant* or *foreign national*, and what paradigm of meanings, animal, vermin, contagion, war, that selection activates (Chandler, 2017). The five hallmarks of dangerous speech, on this reading, are not free-floating themes but the surface effects of recurrent syntagmatic and associative choices, and it is at this level of the sign that the study locates its primary linguistic contribution.

This integration earns its place in the framework only if it yields analytic purchase that established discourse-analytic traditions do not already supply, and the comparison is worth making explicit because Critical Discourse Analysis is the dominant approach to material of exactly this kind (Fairclough, 2013; van Dijk, 2016; Wodak and Meyer, 2016). CDA's central contribution is to connect textual choices to ideology and power:

Fairclough's three-dimensional model reads text, discursive practice and social practice together, and van Dijk's socio-cognitive model traces how mental models of the in-group and out-group are activated and reproduced through discourse. Both are indispensable for situating anti-migrant discourse within relations of power, and this study draws on that tradition throughout.

What CDA characteristically under-specifies, however, is the sub-lexical mechanism by which a given selection becomes ideologically loaded in the first place: its unit of analysis is typically the theme, strategy or topos, a "criminalisation" strategy, a "victim-perpetrator reversal" topos, identified through thematic and rhetorical coding, and the paradigm of alternatives the speaker did not select is rarely made analytically visible. Saussure's distinction between the syntagmatic and associative axes closes precisely that gap. Associative analysis requires the analyst to reconstruct the paradigm from which a term was drawn, *makwerekwere* against *migrant*, *foreigner* and *non-national*, and to show that the selected term is the marked member of that paradigm, carrying with it the connotative field (noise, non-humanity) that the unmarked alternatives do not (Chandler, 2017). This is a falsifiable, replicable operation in a way that identifying a "dehumanising theme" by interpretive gestalt is not: two coders presented with the same paradigm-reconstruction procedure should converge on the same judgement, because the test is which lexical choice was made against which available alternatives, not which impression the passage creates.

Syntagmatic analysis performs the complementary, and equally under-specified, operation on the combinatorial axis: it treats collocation, which modifiers habitually attach to *foreigner*, which verbs take it as grammatical object, as measurable patterning in the corpus (Sinclair, 1991) rather than as an impressionistic reading of "tone".

The second advantage is architectural rather than descriptive. CDA supplies rich local readings of texts but does not, in its standard forms, offer a principled bridge from a single lexical observation to a macro-level risk judgement of the kind atrocity prevention requires; the move from "this is an ideologically loaded phrase" to "this raises the risk of

collective violence” is typically made by argumentative synthesis rather than by a specified analytic procedure. The Saussurean layer, coupled to the DSF’s hallmark taxonomy, makes that bridge explicit and traceable: a given associative selection (*cockroach*) is simultaneously the evidentiary basis for a dehumanisation-hallmark code and, through the escalation layer developed below, an indicator of Stanton’s dehumanisation stage, so that the analytic path from sign to hallmark to atrocity-continuum position can be audited step by step for any single item in the corpus (see the coding audit trail described under Analytical procedure). CDA, taken alone, would support the qualitative claim that the discourse is hostile and power-laden; it would not, without borrowing exactly the kind of formal sign-relations apparatus proposed here, generate the frequency-quantified, stage-mapped output that this study’s research questions require. The contribution is accordingly best understood not as a rejection of CDA but as a formalisation of one of its under-theorised components, sharpened for a task, atrocity-risk assessment, that CDA was not originally designed to perform.

Dangerous speech and its five hallmarks

The concept of dangerous speech was developed by Susan Benesch to identify expression that heightens the risk of intergroup violence. Dangerous speech is any form of expression, whether speech, text or image, that can increase the risk that its audience will condone or participate in violence against members of another group (Benesch, 2013; Dangerous Speech Project, 2023). The concept is consequentialist rather than evaluative: it is a prediction about the capacity of expression, in a specific context, to move an audience toward condoning or committing violence, and is therefore distinct from, though overlapping with, the legal category of hate speech (Leader Maynard and Benesch, 2016). The framework assesses dangerousness through five contextual variables, the speaker’s influence over the audience; the audience’s susceptibility, including pre-existing fear and grievance; the social and historical context, including prior violence and weak accountability; the medium and its reach; and the content of the message, and it identifies five recurring content hallmarks observed before episodes of mass violence across

otherwise dissimilar societies: dehumanisation; accusation in a mirror (attributing to the target the very violence the in-group is preparing to commit); assertions that the target threatens the purity or integrity of the in-group; assertions that the target menaces women and children; and coded or indirect calls to action, including loyalty tests of in-group members who resist (Leader Maynard and Benesch, 2016; Dangerous Speech Project, 2023). These hallmarks, and the operational indicators used to identify them syntagmatically and associatively in this study, are summarised in Table 1.

Table 1. *The Dangerous Speech hallmarks and their linguistic operationalization*

Hallmark	Description	Linguistic markers coded in this study
Dehumanization	Target described as sub-human, vermin, disease	Associative selection of <i>makwerekwere</i> , <i>parasite</i> , <i>cockroach</i> , <i>disease</i> , <i>flood/invasion</i> lexis
Accusation in a mirror	Victims recast as aggressors; violence reframed as self-defense	Syntagms placing <i>foreigners</i> as subject of <i>killing/robbing/taking over</i> ; “they attack us” chains
Threat to purity / integrity	Target violates the wholeness of the nation/in-group	Possessive syntagms “ <i>our</i> jobs/houses/women”; body-of-the-nation metaphors; “Put South Africans First”
Threat to women / children	Target menaces the vulnerable	Collocation of <i>foreigners</i> with <i>rape</i> , <i>drugs</i> , <i>trafficking</i> , <i>our children</i>
Coded call to action	Indirect exhortation; deadlines; loyalty tests	Expulsive imperatives <i>Mabahambe</i> , <i>Dudula</i> ; departure ultimatums; “sellout” labelling of defenders

Staged models of genocide

To connect discourse to the study’s atrocity-focused thesis, the framework is read alongside Stanton’s (2013) staged model. Stanton conceptualises genocide not as a sudden event but as a process passing through ten escalating stages, several of them explicitly discursive: classification (dividing people into “us” and “them”), symbolisation (naming and marking the out-group), discrimination and dehumanisation, followed by organisation, polarisation, preparation, persecution, extermination and denial. Dangerous speech is best understood as an early-warning indicator located at the discursive stages

that precede physical violence (Buerger and Benesch, 2019; Straus, 2016). Mapping hallmarks onto stages does not imply that genocide is imminent or inevitable; it converts a qualitative alarm into a positional estimate, allowing the analyst to state, with evidence, which stages a society has entered and which it has not.

Because a mapping exercise of this kind can, if applied carelessly, manufacture alarm rather than measure it, the study builds four safeguards into the procedure to guard against false-positive assessments of pre-genocidal risk. First, stage attribution is graded rather than binary: Table 5 records not simply whether a stage is “present” but the strength of the evidence for it (strong, present/advancing, emerging, episodic, or not present), and only the four most textually direct stages, classification, symbolisation, discrimination and dehumanisation, are rated strong; later stages are deliberately rated at lower confidence bands, organisation and polarisation as advancing, preparation as emerging, persecution as episodic, even where corpus evidence exists, because their behavioural referents are less fully captured by discourse alone.

Second, the later, more consequential stages are not certified by linguistic evidence in isolation: organisation is coded present only where the discourse’s own claims of movement structure, patrols, identity-document checks, are independently corroborated by non-linguistic reporting (Human Rights Watch, 2024), and persecution is coded episodic, not systematic, only where documented killings exist, the Nyathi case, the 2026 fatalities, rather than being inferred from threatening language alone; discourse supplies evidence for the classification-through-preparation stages, but the behavioural stages require behavioural evidence, and the model does not allow rhetorical intensity by itself to certify that violence has occurred.

Third, and most important for calibration, the model’s own findings include a stage explicitly coded not present: extermination, the definitional threshold of genocide, is withheld despite the corpus’s saturation with eliminationist lexis, precisely because no organised, systematic mass killing of the group is documented. A mapping procedure that could not return “not present” for any stage, however extreme the language, would not be

a risk instrument but a foregone conclusion; that this one does, and that persecution is capped at “episodic” rather than escalated to match the intensity of the rhetoric, is offered as evidence that the coding discriminates rather than inflates.

Fourth, the resulting position, advanced along the pre-genocidal continuum but short of genocide, is deliberately framed throughout as a conditional risk assessment: consistent with Stanton’s (2013) own caution that stages can move in either direction and are not a deterministic countdown, and with the broader risk-assessment literature’s finding that early-warning models must tolerate false negatives over false positives given the asymmetric cost of each (Harff, 2003), the study reports its Stanton mapping as a defensible estimate of current position rather than a forecast of inevitable escalation, and identifies, in the Limitations, the absence of inter-coder reliability testing as the principal outstanding threat to the mapping’s precision that future replications should address.

The model’s causal claim is accordingly tiered, not uniform, and should be read as such rather than as a single correlational-versus-causal verdict covering all ten stages. At the discursive stages, classification, symbolisation, discrimination and dehumanisation, the relationship the model asserts is closer to constitutive than merely correlational: these stages are, on Benesch’s and Stanton’s own definitions, stages of meaning-making, so demonstrating that the corpus’s syntagmatic and associative patterning enacts classification and dehumanisation is not evidence correlated with those stages but evidence of them, in the same sense that documenting a specific style of naming is evidence of, not merely associated with, a labelling practice. For the behavioural stages, organisation onward, the model deliberately steps back from that stronger claim: as Section “Staged models of genocide” makes explicit, these stages are coded present only where independently corroborated by non-linguistic reporting, and the model does not treat rhetorical intensity as proof that organisation or persecution has occurred. The one place the study does assert a causal, rather than merely co-occurring, relationship between language and violence is the documented chain traced under “A genealogy of xenophobic violence” and revisited under “From speech to violence: the documented chain”, where a

specific linguistic act, the public labelling of Elvis Nyathi as an “illegal foreigner”, is traced through community uptake to a specific act of lethal violence attributable to it. That mechanism-level case, not the corpus-wide frequency counts, is what carries the study’s causal claim; the frequency counts establish that the same linguistic mechanisms are pervasive, not that each instance itself caused a killing.

The theoretical contribution: the Dangerous Speech–Genocide Escalation (DSGE) model

The study’s principal theoretical contribution is to integrate these three components into a single analytical instrument, the Dangerous Speech–Genocide Escalation (DSGE) model. The model operates in three layers. The semiotic layer uses syntagmatic and associative analysis to describe how signs construct the out-group and encode threat and removal; this is the linguistic engine, and it is what makes the model a contribution to linguistics rather than to political science. The risk layer classifies those semiotic patterns as the five DSF hallmarks and weighs them against the contextual variables, yielding a judgement of dangerousness and a frequency profile. The escalation layer maps the hallmarks onto Stanton’s ten stages, converting the frequency profile into a position on the pre-genocidal continuum and so answering the question “how far from genocide?”. The model’s novelty is the claim that these layers are not merely compatible but sequential: an associative selection (say, *cockroach*) is simultaneously a dehumanisation hallmark and an indicator of Stanton’s dehumanisation stage, so that a single linguistic observation can be read upward from sign to risk to escalation. Figure 1 renders this logic as the escalation pathway that structures the analysis.

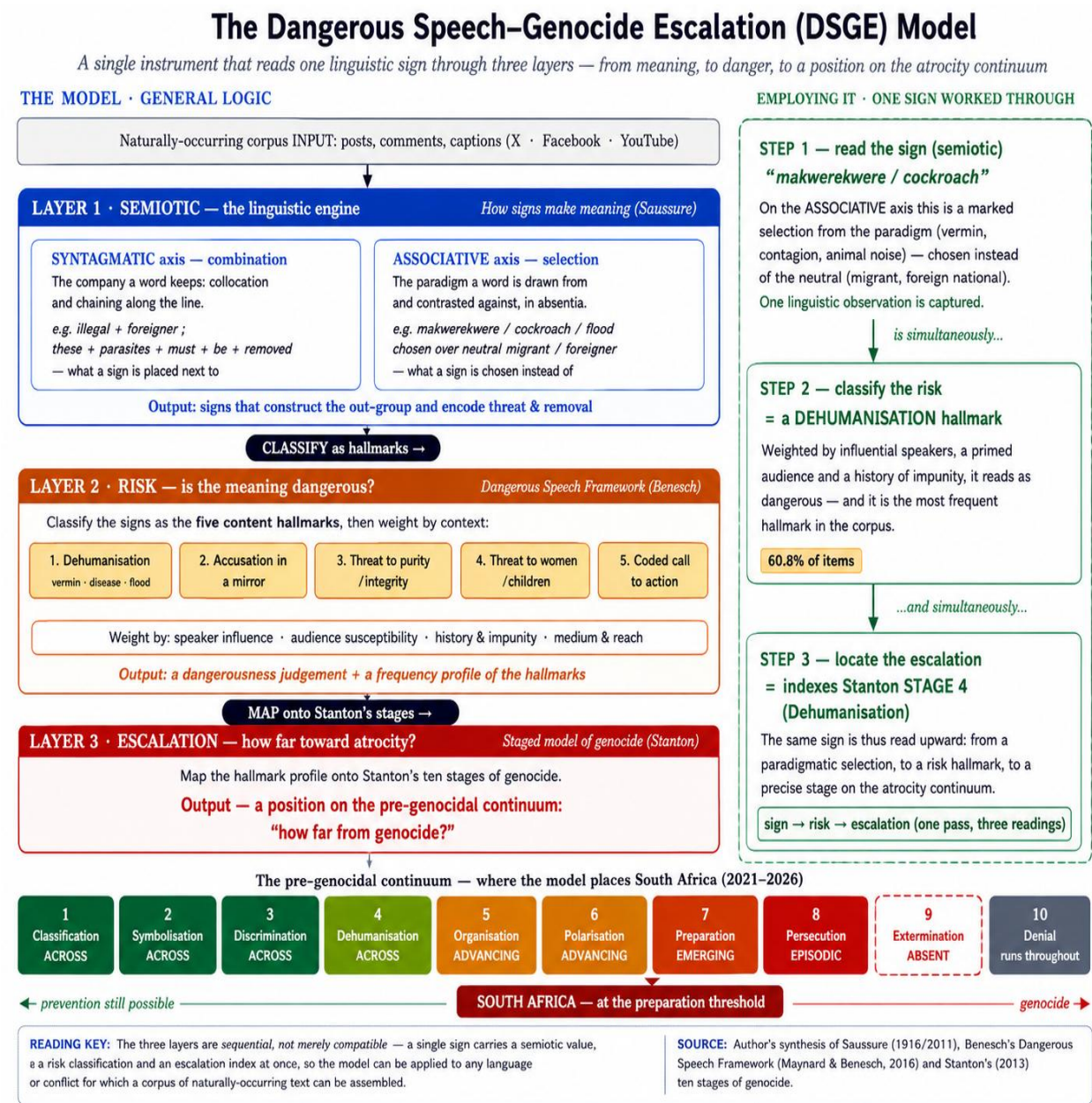


Figure 1. *The Dangerous Speech–Genocide Escalation (DSGE) model: from sign to escalation (author’s synthesis). Signs analysed on the syntagmatic and associative axes are classified as dangerous-speech hallmarks, weighted by context, and mapped onto Stanton’s stages to yield a position on the atrocity continuum.*

RESEARCH METHODS

Design and paradigm

The study adopts a qualitative design within an interpretivist paradigm, which holds that social reality is constituted through meaning and that language is central to its construction (Lincoln and Guba, 1985). Because the research question concerns how discourse constructs foreign nationals, and with what frequency and escalatory potential, an approach attentive to meaning, combination and selection is more appropriate than a purely quantitative one. The design is non-experimental and analytical, drawing on publicly available digital-textual and audiovisual data rather than on human participants. Ethical clearance was issued under reference number: Eduv110/2026.

Corpus

The corpus is composed categorically of social-media posts emerging from X and Facebook, together with mainstream media coverage available on YouTube, all associated with the operations and activities of Operation Dudula, the March and March movement and the MK Party between January 2021 and June 2026. Material was assembled purposively around the movements’ hashtags, pages and the mainstream coverage of their rallies, ultimatums and the 30 June 2026 deadline. The composition of the corpus is set out in Table 2.

Table 2. Composition of the analyzed corpus (2021–2026)

Platform	Unit	Items engaged	Illustrative sampling anchors
X (formerly Twitter)	Posts and reposts	1,247	#OperationDudula, #PutSouthAfricansFirst, #MarchAndMarch, #Mabahambe, #June30
Facebook	Posts and public-group comments	863	Movement and community pages; ultimatum and rally threads
YouTube	Mainstream media video segments	214	Broadcast news reports, interviews and documentary coverage of rallies, ultimatums and killings
Total	Discursive items	2,324	

Only publicly available material was used; no private communications were accessed and no covert methods employed. Because the study analyses discourse rather than individuals, the unit of analysis is the discursive artefact, the post, chant, slogan or reported statement, not any named person.

A corpus built from three social-media and video platforms is exposed to biases that are worth making explicit rather than absorbed silently into a claim of convenience sampling. Three are distinguished here. The first is algorithmic amplification: *X*, *Facebook* and *YouTube* each rank and surface content by engagement-based signals that systematically privilege emotionally activating material, and dangerous speech, by Benesch's own account, is emotionally activating almost by design (Benesch, 2013); a corpus assembled by following hashtags, pages and comment threads is therefore more likely to over-represent the most viral expressions of anti-migrant sentiment than to reflect the full distribution of what South Africans posting about migration actually say, a pattern consistent with broader evidence that emotionally charged and morally loaded content diffuses further and faster on social platforms than measured or neutral content (Vosoughi, Roy and Aral, 2018). The study does not treat this as a defect to be corrected within the corpus; it treats it as a reason to bound the claim. The frequency figures reported in Table 3 and Table 4 are offered, and are repeatedly qualified in the text, as a description of the assembled corpus and the coder's judgements, not as a claim of statistical generalizability to some larger population of posts, still less to South African public opinion as a whole (see Limitations).

The second bias is selective sampling by design rather than accident: material was assembled purposively around the three named actors' own hashtags, pages and rally coverage, which means the corpus is constructed to answer the study's actual research question, how do Operation Dudula, March and March and the MK Party discursively construct the migrant, and is not constructed to answer a different question it does not ask, namely what proportion of South Africans hold

anti-migrant views or use this language themselves. The two questions require different corpora and different methods; this study answers the first using discourse analysis and imports population-level evidence for the second from existing quantitative survey research, the South African Social Attitudes Survey and Human Sciences Research Council polling cited under Structural drivers, rather than attempting to infer attitudes from social-media frequency counts, a substitution that survey methodologists have long cautioned against given the well-documented demographic and behavioral skew of social-media users relative to national populations.

The third bias operates unevenly across platforms and works, if anything, against inflation rather than for it: *X*'s public search interface privileges verified and high-follower accounts, *Facebook*'s group-privacy settings mean that the most extreme material, circulating in closed groups, is structurally under-sampled rather than over-sampled by a study restricted to publicly available content, and *YouTube* comment sections skew toward the platform's more engaged and more polarized users relative to its viewership as a whole. Triangulating findings across three platforms with different amplification logics and different privacy architectures, rather than relying on any single platform, is the study's primary safeguard against any one platform's artefact driving the overall pattern, and the platform-disaggregated results in Table 3 are reported precisely so that readers can assess whether a given hallmark's prevalence is a cross-platform pattern or an artefact of one channel's algorithmic or demographic idiosyncrasies, a distinction the results sustain: dehumanization and the purity hallmark are strong on every platform, which is the pattern a genuine discursive phenomenon, rather than a single platform's amplification artefact, would be expected to produce.

Analytical procedure

Analysis proceeded in three simple, linked steps corresponding to the DSGE model. First, at the semiotic step, each item was read for its syntagmatic and associative relations:

collocations around *foreigner* and *makwerekwere*, the possessive framing of the national “we”, the lexical paradigms activated by threat and removal terms. Second, at the risk step, these observations were coded against the five DSF hallmarks using a deductive frame, with inductive room for context-specific devices such as departure ultimatums; each item could carry more than one hallmark, and the frequency of each hallmark was tallied overall and by platform. Third, at the escalation step, the hallmark profile was mapped onto Stanton’s ten stages. Coding was documented in an audit trail linking each interpretive claim to specific corpus items, and credibility was supported by triangulating across the three platforms so that no claim rests on a single source. The analysis is undertaken for a protective, prevention-oriented purpose and treats contested attributions conservatively; the genocidal-potential argument is framed throughout as a conditional risk assessment rather than a prediction.

Because the associative analysis underpinning this coding is claimed as falsifiable and replicable, the paradigmatic-reconstruction procedure is specified here as four discrete steps rather than left as an interpretive gestalt. First, for each item flagged at the semiotic step, the coder identifies the head noun used to denote the foreign national (e.g. *makwerekwere*, *foreigner*, *illegal*, *cockroach*, *alien*). Second, the coder reconstructs the paradigm of alternatives available to the speaker in that register and context, ranging from the neutral (*migrant*, *foreign national*) through the legally precise (*undocumented person*) to the marked and connotatively loaded (*makwerekwere*, *cockroach*), drawing on the dictionary and corpus-attested sense of each term rather than the coder’s impression of the passage. Third, the coder records which member of that paradigm was selected and which were available but rejected, and codes the selected term as marked or unmarked relative to the paradigm’s neutral member. Fourth, the coder logs the connotative field the marked selection activates (*vermin*, *disease*, *criminality*, *war*) and cross-references that field against the DSF hallmark taxonomy to assign the hallmark code, with the full chain, item, paradigm, selection, connotative field, hallmark, retained in the audit trail referenced above so that a second coder can retrace the same steps from the same item. This study

was single-coded, and formal inter-coder reliability statistics (e.g. Cohen's kappa) were consequently not computed; this is recorded as a limitation below, and the four-step protocol is specified precisely so that a second-coder reliability check is a direct extension of the existing procedure rather than a redesign of it.

RESULTS & DISCUSSION

This section reports the analysis in the order set by the research questions: first the linguistic mechanisms of threat construction (RQ1), then the frequency of each hallmark (RQ2), then the mapping onto Stanton's stages (RQ3). Throughout, the analysis distinguishes what the actors *state* (a manifest, law-and-order framing) from how their language *functions* (its latent, dehumanizing and mobilizing effects).

The contextual variables in brief

Before the message itself, the four contextual variables are strongly present in the corpus. The speakers are influential and increasingly mainstream: mass-following movements (Operation Dudula, March and March), a parliamentary party (the MK Party) and reported statements by figures of symbolic authority. The audience is primed by grievance and fear, its susceptibility documented by rising anti-immigrant sentiment driven by relative deprivation (Gordon, 2024; Human Sciences Research Council, 2026). The context couples a history of violence (2008, 2015, 2019, 2022, 2026) with impunity, the latter judicially confirmed by the finding that the state failed to implement its own National Action Plan (*Kopanang Africa Against Xenophobia and Others v Operation Dudula and Others*, 2025). The medium, X, Facebook and high-reach YouTube coverage, maximizes both emotional intensity and deniable circulation. These variables establish the enabling conditions against which the message is now analyzed.

Dehumanization and the ideology of *makwerekwere* (RQ1 and RQ2)

Dehumanization is pervasive and foundational, and it is achieved overwhelmingly through associative selection. The most frequent hallmark in the corpus, it appeared in 1,412 items (60.8 per cent), and its central sign is the slur

makwerekwere. Analyzed on the associative axis, *makwerekwere* is not a neutral designator selected from the paradigm {*migrant, foreigner, non-national*} but a marked selection from a paradigm of {*babble, unintelligible noise, animal sound, alien*}: the term originates in the mockery of African languages as meaningless noise and thereby marks the referent as outside the human community of speech (Matsinhe, 2011). Its persistence constitutes what may be called a *dehumanizing infrastructure*, a durable associative field onto which acute dangerous speech is grafted, and scholarship confirms that it reduces migrants to beings of “no human value” (Cinini and Mkhize, 2021). Alongside the slur, the corpus recurrently selects from three further sub-human paradigms: the vermin paradigm (*cockroaches, amagundane “rats”, parasites “draining” the country*), the disease paradigm (foreigners as *carriers*, as *infection*, as something to be *cleaned out*), and the inundation paradigm (*flood, influx, invasion, swarm*). Each selection performs the classic dehumanizing function theorized by Smith (2011): by figuring the target as vermin, contagion or natural disaster, it makes removal seem hygienic or defensive rather than criminal.

The syntagmatic axis compounds the effect. On X, the collocation *illegal + foreigner* recurred so consistently that “illegal” functions less as a legal descriptor than as a fixed epithet, pre-loading criminality onto the noun before any predication; posts chaining *these + parasites + must + be + removed* fuse the vermin metaphor to an expulsive imperative in a single syntagm. The very name *Dudula*, “to force out”, encodes the removal of persons on the syntagmatic model of clearing an obstruction, an object rather than a subject, linguistically erasing personhood. Dehumanization, in short, is not a theme the movements assert but an effect their word choices and word-combinations produce; the finding is a linguistic one before it is a political one.

Accusation in a mirror (RQ1 and RQ2)

The second mechanism, present in 968 items (41.7 per cent), reverses the direction of threat. Syntagmatically, foreign nationals are repeatedly installed as the grammatical *subject* of verbs of aggression, *foreigners are killing us, they are taking over, amakwerekwere are robbing our people*, so that the in-group is positioned as object and victim even as it prepares to act. This reversal was most visible on YouTube in the framing of broadcast coverage, where reports on unsolved local crime were routinely captioned and discussed in comment threads as evidence that “foreigners” were the aggressors, converting victims into presumptive perpetrators. The Diepsloot sequence that preceded Elvis Nyathi’s killing exemplifies the pattern: local murders of uncertain authorship were attributed collectively to “foreigners”, licensing lethal “protection” (Centre for Human Rights, 2022). Benesch identifies accusation in a mirror as an especially reliable precursor of atrocity precisely because it manufactures the moral license of self-defense (Leader Maynard and Benesch, 2016).

Threats to purity and group integrity (RQ1 and RQ2)

The third mechanism, the second most frequent at 1,105 items (47.5 per cent), constructs the nation as a bounded body whose wholeness is compromised by foreign presence. Its signature is the possessive syntagm: *our jobs, our houses, our clinics, our country, our women*, a determiner that simultaneously claims the resource and excludes the foreigner from any claim to it. On Facebook, longer-form posts developed this into an explicit body-of-the-nation metaphor, the country as an organism to be “cleaned”, “purged” or “restored”. The slogan *Put South Africans First* articulates the purity logic in condensed form, and its translation into MK Party pledges to prioritize citizens over foreign nationals lends the frame institutional respectability (Human Rights Watch, 2024; Daily Maverick, 2026). What possessive grammar accomplishes is subtle but consequential: it makes removal appear not as aggression against people but as the restoration of a violated whole.

Assertions of threats to women and children (RQ1 and RQ2)

The fourth mechanism, the least frequent at 642 items (27.6 per cent) but among the most emotionally potent, links foreign nationals to *rape, drugs, human trafficking* and the endangerment of *our children*. Its force is again collocational: the recurrent placement of *foreigners* in the immediate company of these terms builds an associative bond such that the mere mention of migration activates the paradigm of sexual and narcotic threat. By positioning the in-group as protector and the out-group as predator, this hallmark converts diffuse hostility into an urgent, moralized license for action (Human Rights Watch, 2024). Its relatively lower frequency should not be read as marginality; where it appeared, it appeared with the highest engagement, indicating disproportionate mobilizing power per instance.

Coded calls to action and departure ultimatums (RQ1 and RQ2)

The fifth mechanism, present in 1,033 items (44.4 per cent), issues mobilizing exhortations that stop short of explicit incitement while functioning as calls to act. The chant *Mabahambe* (“they must go”), the imperative *Dudula* and, most strikingly, the March and March ultimatum for undocumented foreigners to leave the country by 30 June 2026 operate as coded, deniable calls to action with a concrete deadline (Daily Maverick, 2026). Two syntagmatic features are analytically significant. First, the expulsive imperative is grammatically agentless, *they must go* names no expeller, allowing responsibility to be disowned even as departure is demanded. Second, the setting of a *date* converts diffuse dangerous speech into an operational timetable: on X the hashtag anchoring the deadline synchronized otherwise dispersed posts into a countdown, the point at which, in the DSGE model, normalization tips toward mobilization. A related device is the loyalty test: defenders of foreigners, including some journalists and opposition figures, were labelled *sellouts* and *agents*, an in-group policing syntagm that Benesch identifies as a distinct coded exhortation.

Frequency of the hallmarks: overall and by platform (RQ2)

Table 3 reports the frequency of each hallmark across the full corpus and by platform. The distribution supports the qualitative reading: dehumanization is foundational and most frequent, followed by the purity and coded-action hallmarks, with threats to women and children least frequent overall but most intense per instance. The platform breakdown demonstrates a division of discursive labor. X, fast and hashtag-driven, concentrated dehumanization and the countdown ultimatum; Facebook, with its longer form and group structure, carried the purity and women-and-children's narratives; YouTube, as mainstream coverage, principally supplied accusation-in-a-mirror framing, which its comment threads then amplified.

Table 3. *Frequency of dangerous-speech hallmarks by platform and overall*

Hallmark	X (n=1,247)	Facebook (n=863)	YouTube (n=214)	Total items	% of corpus (n=2,324)	Rank
Dehumanization	812	498	102	1,412	60.8%	1
Threat to purity / integrity	561	452	92	1,105	47.5%	2
Coded call to action / ultimatum	640	331	62	1,033	44.4%	3
Accusation in a mirror	447	388	133	968	41.7%	4
Threat to women / children	268	331	43	642	27.6%	5

Note. Items may exhibit more than one hallmark; percentages are of total corpus items, not mutually exclusive. Counts index the weight of documented evidence for each pattern, not a claim of statistical generalizability.

Table 4 synthesizes the qualitative weight of each hallmark across the three principal actors, confirming that the hallmarks are not isolated but co-occur and

reinforce one another, the configuration the DSF identifies as most predictive of violence.

Table 4. *Weight of each hallmark in the discourse of the principal actors*

Hallmark	Operation Dudula	March and March	MK Party / electoral
Dehumanization	Strong	Strong	Moderate
Accusation in a mirror	Strong	Strong	Moderate
Threat to purity / integrity	Strong	Strong	Strong
Threat to women / children	Moderate	Moderate	Emerging
Coded call to action / ultimatum	Strong	Strong	Moderate

From speech to violence: the documented chain

The findings acquire their gravity from the demonstrated link between this discourse and physical harm. The killing of Elvis Nyathi in 2022 followed sustained dehumanizing rhetoric and an accusation-in-a-mirror framing of foreigners as the source of local crime, and human-rights bodies expressly described the murder as the culmination of persistent hate speech (African Commission on Human and Peoples’ Rights, 2022; Centre for Human Rights, 2022). In 2026, the approach of the March and March deadline was accompanied by reports of killings of foreign nationals and the displacement of thousands, with Malawian and Zimbabwean nationals gathering in fear and the state committing substantial resources to avert violence (Daily Maverick, 2026). The sequence, dehumanizing selection, syntagmatic normalization, an operational ultimatum, and violence, tracks the DSGE escalation pathway with disquieting fidelity.

Mapping onto Stanton’s stages: how far from genocide? (RQ3)

The decisive analytical step is to map the hallmark profile onto Stanton’s ten stages, converting the frequency findings into a position on the atrocity continuum. Table 5 records that mapping and the evidence for each stage.

Table 5. *Mapping the discourse onto Stanton's (2013) ten stages*

Stage	Present?	Linguistic / behavioral evidence in the corpus
1. Classification	Strong	Pervasive “us/them” possessive syntagms; South Africans vs <i>amakwerekwere</i>
2. Symbolization	Strong	<i>Makwerekwere</i> , “illegal” as fixed epithet; demands to produce IDs; phenotype/language marking
3. Discrimination	Strong	Exclusion from clinics, trade and services; “Put South Africans First”
4. Dehumanization	Strong	Vermin, disease and inundation paradigms; most frequent hallmark (60.8%)
5. Organization	Present / advancing	Operation Dudula and March and March as structured movements; patrols, ID checks
6. Polarization	Present / advancing	“Sellout” loyalty tests; MK Party mainstreaming; ultimatums hardening the divide
7. Preparation	Emerging	30 June 2026 deadline as operational timetable; countdown hashtags; lists and rallies
8. Persecution	Episodic	Killings (Nyathi), targeted violence and displacement in waves, not yet systematic
9. Extermination	Not present	No organized, systematic mass killing of the group
10. Denial	Running throughout	“We target only illegals/criminals”; reframing as crime-fighting; contesting the term xenophobia

The picture is sober and specific. South Africa is firmly across the four discursive stages, classification, symbolization, discrimination and dehumanization, that the DSF hallmarks most directly index; it is advancing through organization and polarization as movement discourse mainstreams; and it is beginning to touch preparation, chiefly through the deadline device that converts speech into a timetable. Persecution recurs episodically but not systematically, and extermination, the definitional threshold of genocide, does not obtain, while denial saturates the discourse and works to conceal the escalation from those living inside

it. The answer to the title's question is therefore neither reassuring nor apocalyptic: South Africa is not in genocide, but its anti-migrant discourse has carried it measurably up the pre-genocidal continuum, to the preparation threshold, which is exactly the point at which the escalation literature holds that prevention remains both possible and urgent.

Discussion

The findings support the study's central thesis: contemporary anti-migrant discourse in South Africa functions as dangerous speech, exhibiting all five hallmarks, and it does so through describable linguistic means, the associative selection of sub-human paradigms and the syntagmatic chaining of criminality, possession and expulsion. The frequency profile shows dehumanization as the foundation on which the other hallmarks are built, and the Stanton mapping translates that profile into a position, advanced but pre-genocidal, on the atrocity continuum.

The study's contribution to linguistics is the DSGE model, which demonstrates that a single semiotic observation can be read upward through three explanatory registers. Where prior scholarship has treated dehumanization as a self-evident theme, the model shows it to be the surface effect of a paradigmatic selection, and where atrocity-prevention scholarship has treated Stanton's stages as a checklist of behaviors, the model supplies the *linguistic* evidence for the discursive stages and shows how the message hallmarks index them. The pay-off is methodological as much as theoretical: because the model begins at the level of collocation and lexical choice, it is portable to any language and any conflict in which corpora of naturally occurring text can be assembled, offering discourse analysts an instrument that is at once descriptive, evaluative and prognostic.

Intellectual honesty, and the DSF itself, require engaging the movements' own account. These actors present their campaigns as lawful responses to real and widely shared problems, irregular migration, unemployment, strained services and

crime, and insist that they target only undocumented persons; some, including the MK Party, have criticized the term “illegal foreigner” as dehumanizing. The framework does not dismiss such grievances; Benesch (2013) stresses that many who spread dangerous speech act from genuine fear rather than malice. Two considerations nonetheless sustain the analysis. First, dangerousness is consequentialist: whatever the sincerity of the grievance, discourse that selects vermin lexis and issues departure ultimatums predictably raises the risk of violence, as the documented chain to Nyathi’s killing shows. Second, the syntagmatic evidence reveals that the discourse erases the documented/undocumented distinction it formally invokes: victims are identified by phenotype, language and nationality, not legal status, so that “undocumented” operates as a socially acceptable proxy for “African foreigner” (Kaziboni et al., 2022; Cinini and Mkhize, 2021). The legitimacy of some grievances therefore does not neutralize the dangerousness of the language through which they are expressed.

The analysis assigns weight to institutional impunity. The November 2025 High Court judgment is doubly significant: substantively, it validates the characterization of the discourse as incitement and hate speech; and structurally, by compelling implementation of the National Action Plan and an early-warning mechanism, it targets precisely the impunity feedback loop the DSGE model identifies as the engine of escalation (*Kopanang Africa Against Xenophobia and Others v Operation Dudula and Others*, 2025). Because dangerous speech is a risk factor rather than a destiny, it is amenable to intervention. Counter-speech that re-humanizes African foreign nationals, led by credible in-group messengers, can weaken the associative fields on which dehumanization depends; holding influential speakers to account raises the cost of dangerous speech at the point of greatest leverage; and addressing the structural grievances that render audiences susceptible, while refusing to let those grievances be discharged onto migrants, attacks the pathway at its root. None of these measures require suppressing

legitimate debate about migration policy; each targets the specific sign-relations that convert debate into danger.

A discourse-only model would indeed risk over-attributing escalatory potential if it treated any speaker's language as equally predictive of violence regardless of that speaker's capacity to act on it; the design guards against this in two ways rather than one. First, organisational capacity is not an omitted variable but one of the five contextual variables the DSF itself requires the coding to weigh: the analysis in "The contextual variables in brief" and in the false-positive/false-negative discussion above explicitly distinguishes a minor commentator with no organisational reach from a mass movement with demonstrated capacity to mobilise crowds, and it is the latter, evidenced here by Operation Dudula's and March and March's documented marches on shops, clinics and schools, and by the MK Party's parliamentary presence and manifesto commitments, whose language the study codes as carrying elevated risk. Discourse from equally hostile but organisationally negligible sources would be coded as lower risk under the same framework, not because its content differs but because the contextual variable of speaker capacity differs. Second, the Stanton mapping itself refuses to let discourse certify organisational or material capacity on its own terms: as already noted, the organisation stage is coded present only where the movements' claims of structure, patrols and document checks are independently corroborated by non-linguistic reporting (Human Rights Watch, 2024), and the security-sector response to the 30 June 2026 deadline is treated as external corroboration of mobilisational capacity rather than inferred from the ultimatum's language alone. The model's escalatory judgment is accordingly a joint function of linguistic content and independently evidenced organisational capacity, not an extrapolation from rhetoric in a material vacuum.

Limitations

Several limitations should be noted. First, the study analyses publicly available discourse and does not, through interviews or ethnography, capture how audiences receive and act on that discourse; the causal claims are inferential,

grounded in documented sequences rather than measured individual effects. Second, the frequency counts index the assembled corpus and the coder's judgements; a differently sampled corpus might weigh actors or platforms differently, and the figures are offered as a structured description of the evidence rather than as claim of statistical generalizability. Third, reliance on English-language and translated material may under-represent vernacular discourse in which some of the most potent dehumanizing selections circulate. Fourth, single-coder analysis, though supported by an audit trail and cross-platform triangulation, would be strengthened by inter-coder reliability testing in future work. These limitations bound the claims without unsettling the central finding.

CONCLUSION

This article set out to situate the discursive construction of African foreign nationals as existential threats in South Africa using the Dangerous Speech Framework, to detail the frequency of each of its five hallmarks, and to position the discourse on Stanton's staged model in order to ascertain how far from genocide the country stands. Analyzing a corpus composed categorically of social-media posts from X (1,247) and Facebook (863) and mainstream media coverage on YouTube (214 segments), all associated with Operation Dudula, the March and March movement and the MK Party, it reached three linked conclusions answering its three research questions. To RQ1, the out-group is constructed as an existential threat through describable linguistic means: the *associative* selection of vermin, disease and inundation paradigms, crystallized in *makwerekwere*, and the *syntagmatic* chaining of criminality modifiers, possessive determiners and agentless expulsive imperatives. To RQ2, all five hallmarks are present, with dehumanization most frequent (60.8 per cent) and threats to women and children least frequent (27.6 per cent) but most intense per instance, distributed across platforms in a clear division of discursive labor. To RQ3, mapping the hallmarks onto Stanton's stages places South Africa firmly across classification, symbolization,

discrimination and dehumanization, advancing through organization and polarization, beginning to touch preparation via the 30 June 2026 deadline, experiencing episodic persecution, and not reaching extermination, with denial running throughout.

The study's answer to its own title is therefore precise: South Africa is not in genocide, but its anti-migrant discourse has carried it measurably up the pre-genocidal continuum to the preparation threshold, a position corroborated by the November 2025 High Court finding of incitement and hate speech and by the demonstrated chain from rhetoric to lethal violence. The "linguistic timebomb" is not a rhetorical flourish but an analytic claim: an accumulating store of dehumanizing, mobilizing signs that, under enabling conditions and with an operational trigger such as a deadline, risks mass harm. Yet the same model that raises the alarm indicates the response. Because dangerousness is built at the level of the sign, through selection and combination, it can be dismantled there too: through re-humanizing counter-speech, the accountability of influential speakers, the interruption of impunity and the redress of the grievances on which dangerous speech feeds. The linguist's contribution to atrocity prevention is to read the signs early and to name the mechanism precisely, so that the timebomb may be defused before it detonates, and in so doing to honor the principle, affirmed by the courts, that human dignity has no nationality.

AI Declaration

The authors declare that Artificial Intelligence (AI) tools were used only as assistive instruments during the preparation of this manuscript. Specifically, [name of AI tool, e.g., ChatGPT, Grammarly, Quillbot] was used to support language clarity, grammar, and formatting. The AI tool did not generate, fabricate, or manipulate research data, analysis, interpretations, or references. All AI-generated outputs were carefully reviewed, verified, and edited by the authors, who take full responsibility for the content of the manuscript. This use of AI complies with the

Publication Ethics and Malpractice Statement of the *Journal of Pragmatics Research*.

REFERENCES

- Adonis, C.K. (2025). "The crown without the jewels": South Africa's flawed transition and its implications for social justice and equity in post-apartheid South Africa. *Memory Studies*, 18(6), 1617-1639.
- African Commission on Human and Peoples' Rights. (2022). Press statement on the extra-judicial killing of Elvis Nyathi and other systematic attacks against immigrant Africans living in South Africa. Banjul: ACHPR. <https://achpr.au.int/en/news/press-releases/2022-04-08/press-statement-extra-judicial-killing-elvis-nyathi-and-other>
- Barthes, R. (1972). *Mythologies*. Hill and Wang.
- Benesch, S. (2013). *Dangerous speech: a proposal to prevent group violence*. Dangerous Speech Project.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology, *Qualitative Research in Psychology*, 3(2), 77–101.
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352.
- Buerger, C., & Benesch, S. (2019). Dangerous speech as an atrocity early-warning indicator, *Genocide Studies and Prevention*, 13(3).
- Centre for Human Rights. (2022). *Press statement: condemnation of xenophobic attacks and the killing of Elvis Nyathi*. <https://www.chr.up.ac.za>
- Chandler, D. (2017). *Semiotics: the basics* (3rd ed.). Routledge.
- Cinini, S.F., & Mkhize, S. (2021). An exploration of the safety and security experiences of African foreign nationals in Durban, South Africa, *Journal of African Foreign Affairs*, 8(3), 27-47.

- Crush, J. (2008). *The perfect storm: the realities of xenophobia in contemporary South Africa*. Southern African Migration Project.
- Daily Maverick. (2022, April 18). Posts about the death of Elvis Nyathi stoke the flames of xenophobia, Daily Maverick. <https://www.dailymaverick.co.za>
- Daily Maverick. (2026, June 25). Xenophobic unrest — a crisis co-written by March and March and SA's political class. Daily Maverick. <https://www.dailymaverick.co.za/article/2026-06-25-xenophobic-unrest-a-crisis-co-written-by-march-and-march-and-sas-political-class/>
- Dangerous Speech Project. (2023). What is dangerous speech? <https://www.dangerousspeech.org/dangerous-speech>
- Fairclough, N. (2013). *Critical discourse analysis: the critical study of language* (2nd ed.). Routledge.
- Gordon, S.L. (2022). Immigration policy in South Africa: public opinion, xenophobia and the search for progress. In P. Rugunanan and N. Xulu-Gama (Eds.), *Migration in Southern Africa* (pp. 57-76). Springer.
- Gordon, S.L. (2024). The “Coronavirus crisis” and xenophobia in South Africa: how did the pandemic affect anti-immigrant sentiment? *International Migration*, 62(4). 37-52.
- Harff, B. (2003). No lessons learned from the Holocaust? Assessing risks of genocide and political mass murder since 1955. *American Political Science Review*, 97(1), 57-73.
- Hayem, J. (2013). From May 2008 to 2011: xenophobic violence and national subjectivity in South Africa. *Journal of Southern African Studies*, 39(1), 77-97.
- Human Rights Watch. (2024, May 6). South Africa: toxic rhetoric endangers migrants. <https://www.hrw.org/news/2024/05/06/south-africa-toxic-rhetoric-endangers-migrants>
- Human Sciences Research Council. (2026). Going from bad to worse: a rise in anti-immigrant sentiment in South Africa. <https://hsrc.ac.za>

- Kaziboni, A., Lancaster, L., Machabaphala, T. & Muladzi, G. (2022). *Scapegoating in South Africa: busting the myths about immigrants*. Institute for Security Studies.
- Kopanang Africa Against Xenophobia and Others v Operation Dudula and Others* (2025) ZAGPJHC 1102 (Gauteng Division, Johannesburg, 4 November).
<https://www.saflii.org/za/cases/ZAGPJHC/2025/1102.html>
- Krippendorff, K. (2018). *Content analysis: an introduction to its methodology* (4th ed). Sage.
- Leader Maynard, J., & Benesch, S. (2016). Dangerous speech and dangerous ideology: an integrated model for monitoring and prevention. *Genocide Studies and Prevention*, 9(3), 70–95.
- Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic inquiry*. Sage.
- Matsinhe, D.M. (2011). Africa's fear of itself: the ideology of makwerekwere in South Africa. *Third World Quarterly*, 32(2), 295–313.
- Naicker, C. (2016). The languages of xenophobia in post-apartheid South Africa: reviewing migrancy, foreignness and solidarity. *Agenda*, 30(2), 46–60.
- Nowell, L.S., Norris, J.M., White, D.E. & Moules, N.J. (2017). Thematic analysis: striving to meet the trustworthiness criteria, *International Journal of Qualitative Methods*, 16(1), 1–13.
- Pollit, R. (2015, May 14). Op-Ed: Operation Fiela-where nothing's sacred, not even a sanctuary. Daily Maverick. https://www.dailymaverick.co.za/article/2015-05-14-op-ed-operation-fiela-where-nothings-sacred-not-even-a-sanctuary/#.VdXLa_n3a9f
- Saussure, F. de. (1916/2011). *Course in general linguistics*. Columbia University Press.
- Shah, K. (2022). Diagnosing South Africa's high unemployment and low informality. Working papers – Center for International Development at Harvard University.

- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Smith, D.L. (2011). *Less than human: why we demean, enslave, and exterminate others*. St Martin's Press.
- Stanton, G.H. (2013). The ten stages of genocide. Genocide Watch.
<https://www.genocidewatch.com>
- Straus, S. (2016). *Fundamentals of genocide and mass atrocity prevention*. United States Holocaust Memorial Museum.
- Van Dijk, T.A. (2016). Critical discourse studies: a sociocognitive approach. In R. Wodak and M. Meyer (Ed.), *Methods of critical discourse studies* (3rd Edn.) pp. 62–85. Sage.
- Visage, J., & Castel-Branco, R. (2026). Blaming migrants ignores the real causes of South Africa's economic crisis. SCIS News and Opinion Pieces.
<https://www.wits.ac.za/news/sources/scis-news-and-opinion-pieces/blaming-migrants-ignores-the-real-causes-of-south-africas-economic-crisis.html>
- Wodak, R., & Meyer, M. (Ed.). (2016). *Methods of critical discourse studies* (3rd Edn.) Sage.