SCIENTIFIC CULTURE

www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 2.1, (2026), pp. 3065-3077
Open Access. Online & Print

DOI: 10.5281/zenodo.122.126236

SUPRASEGMENTAL PROSODIC FEATURES AS PRIMARY DETERMINANTS OF SARCASM AND IRONY IN SPOKEN ARABIC: TOWARD A PRAGMATICALLY AWARE MACHINE TRANSLATION FRAMEWORK

Saleh Belhassen[1*]

[1]Department of English and Translation, Saudi Electronic University, Riyadh, Saudi Arabia
dr.salehbelhassen@gmail.com

Received: 11/12/2025
Accepted: 02/02/2026

Corresponding Author: Saleh Belhassen
(dr.salehbelhassen@gmail.com)

ABSTRACT

This study establishes suprasegmental prosodic features—intonation, stress, pitch variation, and temporal manipulation—as the primary semiotic system for encoding sarcasm and irony in spoken Arabic. While current Arabic Machine Translation (MT) systems achieve lexical and syntactic accuracy, their pragmatic failure stems from ignoring these acoustic cues, which systematically invert literal meaning. Through systematic field observations across multiple dialects, we develop a fine-grained taxonomy of Arabic sarcastic prosody, formalizing how features like vowel prolongation in "جمييييلة" or strategic pauses in "حكيم قرار ⋯هذا" trigger evaluative reversal. From this taxonomy, we derive three governing linguistic principles—Prosodic Markedness, Evaluative Polarity Reversal, and Contextual Anchoring—which inform a novel computational framework. We propose a Prosodic-Pragmatic Interface Layer (PPIL) architecture, designed to integrate prosodic annotations into MT pipelines to disambiguate intent. To evaluate this, we introduce a human-centric Pragmatic Accuracy (PA) metric. This work provides both the empirical foundation and computational blueprint necessary to advance toward pragmatically aware Arabic MT, bridging a critical gap between acoustic signal and intended meaning.

KEYWORDS: Arabic Prosody, Suprasegmental Features, Sarcasm, Irony, Pragmatic Machine Translation, Computational Pragmatics, Corpus Linguistics.

1. INTRODUCTION

The relentless pursuit of authentic, human-like Machine Translation (MT) has irrevocably shifted from a paradigm of formal equivalence toward one of pragmatic competence. While contemporary neural models demonstrate unprecedented proficiency in syntactic fidelity and lexical selection (Bahdanau et al., 2015; Vaswani et al., 2017), they remain fundamentally deficient in interpreting the nuanced, often unspoken, intentions that govern human communication. This "pragmatic gap" is particularly pronounced and consequential for languages whose rhetorical traditions privilege indirectness, evaluative subtlety, and contextually anchored meaning, with Modern Standard Arabic and its dialectal variants representing a quintessential case (Al-Sayed, 2018; Habash, 2010). In these linguistic ecosystems, the veritable semantic payload of an utterance frequently resides not within the discrete lexicon but in the dynamic, acoustic envelope that carries it—the domain of suprasegmental prosody.

This paper posits that for spoken Arabic, the primary semiotic system for encoding sarcasm and verbal irony—speech acts where evaluative polarity is systematically inverted—is prosodic. Utterances such as the lexically positive "أنت ذكي" (you are smart) or "عمل رائع" (great work) are routinely transformed into potent instruments of criticism or mockery not through lexical substitution, but through the deliberate manipulation of intonational contour, syllabic duration, contrastive stress, and pausal architecture. These suprasegmental features function as a sophisticated, rule-governed meta-commentary, signaling to the listener to suspend a literal interpretation in favor of one predicated on contextual incongruity and shared cultural understanding (Grice, 1975; Wilson & Sperber, 2012). The human cognitive apparatus, honed by sociolinguistic immersion, performs this pragmatic recalibration seamlessly. In stark contrast, state-of-the-art MT and Natural Language Processing (NLP) systems, operating on de-prosodized text, are rendered pragmatically tone-deaf. They are condemned to produce translations that are not merely inaccurate but are socially inappropriate, relationally disruptive, and ultimately fail to convey the speaker's true illocutionary force—a failure that can range from humorous misunderstanding to serious communicative breakdown.

The problem is twofold, stemming from both a theoretical oversight and a practical bottleneck in computational linguistics. Firstly, dominant Anglocentric models of sarcasm detection have disproportionately focused on lexical markers (e.g., "yeah, right"), syntactic constructions, or sentiment clashes within written text, especially from social media corpora (Joshi et al., 2017; Zhang et al., 2021). This orientation has implicitly marginalized the study of prosody, despite its well-documented centrality in spoken irony across languages (Bryant & Fox Tree, 2005; Cheang & Pell, 2008). For Arabic, this marginalization is especially acute, creating a lacuna where the language's own rich rhetorical tradition—'Ilm al-Balāghah—with its explicit concepts like التعريض (taʿrīḍ, indirect criticism) and قلب المدح ذمًا (the reversal of praise into blame), has seldom been computationally formalized (Abdul-Raof, 2006). Secondly, the practical development of prosody-aware systems is hamstrung by the absence of large-scale, annotated multimodal corpora for Arabic that pair acoustic signals with pragmatic intent labels, a resource that is taken for granted in research on English (Bousquet et al., 2023).

Consequently, this study addresses a critical and underexplored intersection: the computational modeling of Arabic sarcasm through its primary real-world manifestation—suprasegmental prosody. We argue that for MT to achieve pragmatic adequacy for Arabic, prosodic features must be elevated from optional paralinguistic metadata to first-class, semantically determinative linguistic signals. Our investigation proceeds through a dual-methodology framework. We first establish a rigorous, empirically grounded qualitative taxonomy of sarcastic prosody in Arabic, derived from systematic field observations across multiple dialects. This taxonomy catalogs the specific acoustic maneuvers—such as the extreme vowel prolongation in "جمييييلة" (beautiful), the emphatic over-stressing in "راااائع" (great), or the artificially high pitch in "!واااو" (wow)that reliably triggers a sarcastic interpretation. We then translate these linguistic insights into a novel computational proposal: a hybrid MT architecture featuring a dedicated Prosodic-Pragmatic Interface Layer (PPIL). This layer is designed to ingest symbolic or continuous prosodic annotations and perform explicit pragmatic disambiguation, thereby guiding the translation engine toward outputs that preserve the speaker's ironic stance, even when this requires a complete pragmatic recalibration of surface lexemes.

By demonstrating how the sarcastic intent of an utterance like "!سرييييع"—produced in response to a losing racehorse—can be computationally detected and rendered pragmatically as "What speed!" (with ironic intent) rather than the literally erroneous "He

is fast!", this research moves beyond mere error correction. It advocates for a paradigm shift toward a culturally and linguistically grounded model of Arabic MT, one capable of approximating the sophisticated inferential processes that underpin human communication. In doing so, it seeks not only to enhance the functional utility of translation technology but also to contribute to the broader project of developing AI that is truly sensitive to the complexities of human language and interaction.

2. LITERATURE REVIEW

2.1. The Nature of Sarcasm and Verbal Irony: A Pragmatic and Cognitive Perspective

The computational modeling of sarcasm and irony necessitates a firm grounding in the theoretical frameworks that define their nature and function in human communication. Classically, these phenomena are situated within the broader domain of non-literal language, where speaker meaning deviates from sentence meaning. The Gricean cooperative principle and its attendant maxims provide a foundational account, positing that irony arises from a blatant and deliberate violation, typically of the maxim of quality, inviting the listener to derive an implicated, often opposite, meaning (Grice, 1975). This "echoic mention" theory was later refined by Relevance Theory (Wilson & Sperber, 2012), which argues that ironic utterances are not mere violations but interpretations of a attributed thought or utterance, expressing a dissociative, critical attitude toward it. This perspective is crucial for understanding Arabic sarcasm, where an utterance like "إِنَّهُ سَرِيييع!" (He is fast!) is not a false statement about a slow horse but a critical echo of an expectation or claim, with the prosody signaling the speaker's dissociative stance.

Cross-linguistically, sarcasm functions as a potent social and rhetorical tool for criticism, humor, and in-group bonding. Its realization, however, varies significantly across languages and modalities. In written English, a substantial body of computational linguistics research has focused on lexical cues (e.g., interjections like "yeah, right"), syntactic constructions (e.g., "as if!"), semantic incongruity within text, and contextual sentiment clashes, often leveraging large-scale social media datasets (Joshi et al., 2017; Zhang et al., 2021). This text-centric paradigm, while productive, has inadvertently marginalized the study of the primary channel for sarcasm in spoken interaction: prosody (Anolli et al., 2002). Research in psycholinguistics and phonetics has robustly established that speakers across languages employ systematic prosodic markers—such as exaggerated pitch range, slowed speech rate, increased intensity, and specific intonational contours (e.g., a "flat" or "singsong" tone) to signal ironic intent (Bryant & Fox Tree, 2005; Cheang & Pell, 2008). This creates a critical disconnect: the dominant computational models are built for a modality (text) where the primary signal (prosody) is absent.

2.2. The Prosody-Pragmatics Interface in Arabic: Bridging Classical Rhetoric and Modern Phonetics

For Arabic, the study of indirect meaning has a deep and sophisticated heritage within the classical Islamic sciences, particularly `Ilm al-Balāghah (the science of eloquence). Concepts central to our inquiry were meticulously categorized by medieval rhetoricians. التعريض (al-taʿrīḍ), or insinuation, refers to expressing criticism indirectly through ostensibly positive or neutral language, a direct precursor to modern notions of verbal irony (Abdul-Raof, 2006). التهكم (al-tahakkum) denotes scornful mockery, often achieved through exaggeration. Most pertinently, the principle of قلب المدح ذمًا–the rhetorical reversal of praise into blame—precisely describes the evaluative polarity inversion that defines sarcastic prosody, as observed in field examples like "عَمَل راااائع!" (Great work!) used to condemn poor performance. While these classical frameworks were not formulated in acoustic terms, they implicitly presuppose a delivery system—the human voice—capable of carrying this meta-pragmatic information. They provide a crucial semantic and functional taxonomy that modern phonetic analysis must acoustically ground.

Modern linguistic studies on Arabic prosody have largely focused on its role in grammar (e.g., interrogative vs. declarative intonation) or metrical systems, with relatively scant attention paid to its pragmatic, attitudinal functions, especially in dialectal speech. Where prosody and pragmatics intersect, research has often been anecdotal or limited to specific constructions. The systematic, corpus-driven investigation of how suprasegmental features like vowel prolongation (e.g., "نااااجح" to imply failure), contrastive stress (e.g., segmented "إبَرا–فو!" for ridicule), or marked intonational contours (e.g., the artificially high pitch in "واااو إنجاز!") conspire to signal sarcasm remains a significant gap in the literature. This study directly addresses this gap by constructing a fine-grained taxonomy derived from empirical observation, thereby building a bridge between the abstract categories of classical balāghah and the concrete, measurable parameters of

contemporary acoustic phonetics.

2.3. Computational Linguistics at the Crossroads: The Neglect of Prosody in Arabic NLP

The field of Arabic Natural Language Processing (NLP) has made remarkable strides in recent decades, achieving high performance in tasks such as morphological analysis, named entity recognition, and sentiment classification (Habash, 2010). Machine Translation for Arabic, particularly into English, has benefited enormously from the neural paradigm, yielding translations of impressive syntactic and lexical fluency (Al-Sayed, 2018). However, as a subset of the wider NLP community, Arabic computational linguistics has largely inherited its focus on textual features, leading to a pronounced "prosodic deficit" in models intended to process or emulate human communication.

A review of the literature reveals a stark asymmetry. For English, there is a growing, if still nascent, body of work on multimodal sarcasm detection that integrates audio features from datasets like MUStARD (Castro et al., 2019). For Arabic, the computational landscape is different. Research on affect and sentiment is overwhelmingly text-based, analyzing written reviews or social media posts (Al-Harbi & Emam, 2021). The few studies that acknowledge sarcasm treat it as a textual classification problem, searching for patterns in lexical choice, diacritics, or emojis. The rich, spoken reality of Arabic sarcasm—where a phrase like "هذا... قرار حكيم" (This is... a wise decision) delivered with a strategic pause conveys profound doubt—is absent from computational models because the requisite multimodal corpora and annotated acoustic-prosodic resources simply do not exist at scale (Bousquet et al., 2023). This lacuna represents both a challenge and an opportunity. It necessitates a foundational, fieldwork-driven approach to data collection and annotation, as undertaken in this study, rather than the application of pre-existing Anglo-centric models.

2.4. Positioning The Present Study: Toward A Grounded, Prosody-First Model

This study distinguishes itself by explicitly rejecting the importation of text-biased assumptions and instead advocating for a prosody-first approach to modeling Arabic sarcasm for MT.

It is positioned at the convergence of three distinct scholarly trajectories:

1. Theoretical Pragmatics & Cognitive Linguistics: It adopts the view of sarcasm as an echoic, dissociative act (Wilson & Sperber, 2012), where prosody is the primary signal of the speaker's critical attitude.
2. Arabic Linguistic Tradition: It seeks to computationally operationalize the insights of classical balāghah concerning indirectness and evaluative reversal, providing them with a formal, acoustic basis.
3. Computational Pragmatics & Multimodal NLP: It addresses a critical bottleneck in Arabic NLP by proposing both a descriptive taxonomy and a computational architecture to integrate prosodic meaning, thereby moving beyond the current text-only paradigm.

By grounding its model in the empirically observed prosodic maneuvers of native speakers—from the elongated vowel in "جمييييلة" to the compound cue of laughter and emphatic stress in "هه... فزنا"—this research does not merely add a feature to existing systems. It argues for a fundamental re-conception of the input to Arabic MT, proposing that for pragmatic adequacy, prosodic features must be elevated from paralinguistic ornamentation to essential, semantically determinative components of the linguistic signal. The proposed Prosodic-Pragmatic Interface Layer (PPIL) is thus not an incremental improvement but a conceptual step toward a machine translation framework that is truly attentive to how meaning is made in spoken Arabic.

3. METHODOLOGY: A DUAL-PHASE, MULTI-MODAL APPROACH TO MODELING PROSODIC SARCASM

This study employs a rigorous, dual-phase methodological framework designed to bridge the chasm between qualitative linguistic observation and quantitative computational modeling. Recognizing that prosodic sarcasm constitutes a complex phenomenon rooted in acoustic realization, pragmatic function, and socio-cultural context, our methodology is intentionally bifurcated. The first phase (Qualitative Elicitation & Taxonomical Construction) is dedicated to meaning discovery, identifying and systematizing the phenomenon in its natural habitat. The second phase (Corpus Engineering & Computational Formalization) focuses on meaning operationalization, translating these insights into annotated data and a formal model suitable for machine learning. This phased approach ensures that the computational model is firmly grounded in empirical linguistic reality, rather than imposing a priori, text-derived assumptions onto spoken Arabic.

3.1. Phase I: Qualitative Elicitation & Taxonomical Construction

The objective of this phase is to develop a comprehensive, empirically grounded taxonomy of suprasegmental features that signal sarcasm in spoken Arabic. This is an inductive, bottom-up process.

3.1.1. Data Collection & Ethical Protocol

Data was gathered through a multi-source strategy to ensure ecological validity and dialectal diversity:

1. **Naturalistic Field Recordings:** With informed consent, spontaneous interactions were recorded in settings where evaluative language and sarcasm are prevalent (e.g., university tutorials, workplace feedback sessions, social gatherings) **across three primary dialect groups: Levantine (Jordan, Lebanon), Gulf (Saudi Arabia, UAE), and Egyptian. The final observed/elicited dataset comprises approximately 450 unique sarcastic utterances, contributed by 87 native speakers (52 male, 35 female) aged 18-65.** This captures contextual incongruity in situ—the essential backdrop against which prosody operates.
2. Broadcast Media Archive: Scripted and unscripted discourse from satirical TV shows, post-match sports commentary, and political talk shows across the Arab world (e.g., MBC, Al Jazeera) was systematically reviewed. This source provides high-quality audio and clear instances of public, performative sarcasm.
3. Elicited Production: To fill gaps in the naturalistic data, short video clips depicting clear failings or incongruous outcomes (e.g., a comedian fumbling a simple task) were shown to native speaker participants, who were asked to provide a verbal reaction. This method elicits controlled yet spontaneous sarcastic utterances like "مُحترف!" (Professional!).

3.1.2. Analysis & Taxonomy Development

Each collected utterance underwent a multi-layered analysis performed by a team of trained linguists and native speakers:

1. Contextual-Pragmatic Tagging: The situational context was documented to establish the incongruity that makes sarcasm relevant (e.g., praising a losing horse).
2. Perceptual Validation: The sarcastic intent of each utterance was validated by at least three independent native listeners. Only utterances with unanimous or strong majority agreement on ironic interpretation were retained.
3. Acoustic-Phonetic Analysis: Using software (e.g., Praat), the selected utterances were analyzed to measure and categorize the prosodic manipulation:

- Temporal Features: Vowel/consonant duration ratios, speaking rate, and pause placement/duration. This quantifies phenomena like the strategic prolongation in "جمييييلة".
- Tonal Features: Fundamental frequency (F0) contours, pitch range, and the shape of nuclear tones (e.g., rising, falling, or complex). This captures the artificially high pitch of "واااو!" or the terminal low fall in "طبعًا كلامك صحيح".
- Dynamic Features: Syllable- and word-level intensity (stress). This identifies emphatic over-stressing, as in "رااائع" or the segmented stress in "فو-بَرا!".

4. Taxonomical Structuring: The analyzed features were categorized into a functional taxonomy (see Section 4), moving from raw acoustic measurements (e.g., +150ms duration) to descriptive linguistic categories (e.g., <PROLONGED>), linked to their pragmatic effect (e.g., Mock Praise).

3.2. Phase Ii: Corpus Engineering & Computational Formalization

The objective of this phase is to construct a resource and a formal model that enable the training and evaluation of prosody-aware MT systems.

3.2.1. The SARCASM (Suprasegmental Arabic Corpus for Sarcasm and Meaning) Corpus Proposal

Based on the Phase I taxonomy, we propose the creation of a novel multimodal corpus. Its architecture is outlined below:

Table 1: Multi-Layer Annotation Scheme for the SARCASM Corpus.

Corpus Layer	Description & Annotation Scheme	Example from Data
1. Raw Audio	High-fidelity recordings of utterances.	.wav file of a speaker saying "أنت ذكي" with elongated /ðiː/.
2. Orthographic Text	Standard Arabic script transcription, including filler words and repetitions.	"إنه سريع"

3. Phonetic Alignment	Forced alignment at phone level (using tools like FAVE or Montreal Forced Aligner) to timestamp each segment.	[ɪ n.nə hʊ] [s ɑ . r iː : ʕ] (with iː : indicating prolonged segment).
4. Acoustic-Prosodic Features	Extracted continuous values: F0 (mean, max, range), duration of target vowels/consonants, intensity peaks, pause duration.	{Target_Word: "سريـــع", Vowel_Duration: 420ms, F0_Max: 280Hz, Pre_Pause: 120ms}
5. Symbolic Prosodic Tags	Discrete labels derived from the taxonomy and acoustic thresholds.	<LEX_POSITIVE> <PROSODY: PROLONGED_VOWEL+> <PRAGMATIC: REVERSAL>
6. Pragmatic Intent Label	Gold-standard label: SARCASMIC, LITERAL, or AMBIGUOUS.	SARCASMIC
7. Contextual Metadata	Dialect (e.g., Levantine, Gulf), speaker-listener relationship (formal, peer), situational description.	{Dialect: Egyptian, Setting: Sports commentary, Context: Comment on last-place runner}

3.2.2. Annotation Protocol & Reliability

Annotation will be performed by linguistically trained native Arabic speakers. A detailed guideline document will define each symbolic tag (<PROLONGED>, <OVERSTRESS>, <IRONIC_HIGH_PITCH>) with acoustic correlates (e.g., <PROLONGED> = vowel duration > 2.5 times the neutral baseline). Inter-annotator agreement will be rigorously measured for pragmatic labels (Krippendorff's α) and symbolic prosodic tags (Cohen's κ), with targets exceeding $\alpha/\kappa > 0.80$ to ensure reliability.

3.2.3. Computational Formalization: From Tags to Pragmatic Operators

The taxonomy and corpus annotations allow us to formalize prosodic features as operators that modify semantic composition.

For instance:

- An utterance U with lexical semantics PRAISE(x) can be modeled as undergoing a pragmatic operator PROS_SARC.
- The meaning derivation becomes: PROS_SARC(PRAISE(x)) → CRITICISM(x).
- In the proposed MT architecture, the Prosodic-Pragmatic Interface Layer (PPIL) learns to associate specific prosodic feature vectors or tag combinations with this PROS_SARC operator. When detected, it instructs the decoder to generate a target language utterance whose pragmatic stance is CRITICISM, not PRAISE, achieving the necessary reversal from "عمل رائع" to an output like "Some 'great' work there."

3.2.4. Experimental Validation Design

To validate the framework, a comparative experiment is designed:

1.Model Variants:

-Baseline: State-of-the-art neural MT model (e.g., Transformer) trained on standard text.

-Pipeline (Symbolic): ASR → Prosody Tagger (trained on SARCASM corpus) → MT system receiving text + symbolic tags (e.g., سريع <PROSODY:SARC>).

-Integrated (Acoustic): End-to-end model where acoustic-prosodic features are fused with text embeddings at the encoder input.

2.Evaluation: Beyond standard metrics (BLEU), the core evaluation is Pragmatic Accuracy (PA), a human-judged metric measuring the preservation of ironic stance in translation (see Section 6.2). A/B testing with bilingual evaluators will determine if the prosody-aware models yield statistically significant ($p < 0.01$) improvements in PA over the baseline for sarcastic utterances, while maintaining performance on literal ones.

This methodology, moving systematically from field observation to formal corpus construction and computational modeling, provides a robust blueprint for teaching machines to "hear" and interpret the sophisticated prosodic code of Arabic sarcasm.

4. TAXONOMY OF SUPRASEGMENTAL SARCASM IN ARABIC

Building on the methodological framework, this section presents the core empirical output: a comprehensive taxonomy of suprasegmental patterns that reliably signal sarcasm in spoken Arabic. This taxonomy is not merely a list of examples, but a structured classification system derived from acoustic analysis and pragmatic validation. It demonstrates the systematic nature of prosodic sarcasm, moving beyond arbitrary vocal variation to identifiable, rule-governed cues. The taxonomy is organized by the primary prosodic feature manipulated, with each category illustrated by canonical examples—including new additio

ns such as "أنت شباب على الدوام" and "برد قارس"—that underscore its breadth and contextual dependency.

Table 2: Prolongation-Based Sarcasm (Temporal Manipulation).

Example	Phonetic Cue & Symbolic Tag	Literal Meaning	Pragmatically Inferred Meaning	Contextual Incongruity Trigger
جمييييلة!	Extreme vowel lengthening of /iː / <PROLONGED+>	"Beautiful!"	"Plain/Unattractive."	Commenting on someone's appearance is contrary to visual evidence.
سرييييع!	Prolonged high vowel /iː / in the adjective <PROLONGED+>	"Fast!"	"Very slow."	Observing a losing racehorse or a slow vehicle.
ناااجح طبعاً.	Marked lengthening of the initial /aː / <PROLONGED>	"He passed, of course."	"He obviously failed."	Discussing a student's performance after a known poor exam result.
بَرد قاارس!	Exaggerated duration of /aː / in "قارس" <PROLONGED+>	"Bitterly cold!"	"Extremely hot."	Describing outdoor weather during a scorching 50°C summer day.

Table 3: Contrastive Stress & Emphatic Accentuation.

Example	Stress Pattern & Symbolic Tag	Literal Meaning	Pragmatic Effect & Stance
عَمَل راائع.	Over-stressing and prolongation on the adjective <OVERSTRESS> <PROLONGED>	"Great work."	**Condemnation:** Mocking the poor quality of the work.
شاااطر يا ولد!	Heavy, emphatic stress on the first syllable <OVERSTRESS>	"Clever, boy!"	**Mockery:** Implying foolishness, often to a child who made a mistake.
بَـرافو! عَلَيْك.	Segmented, staccato delivery with emphasis on "برا" <SEGMENTED_STRESS>	"Bravo to you."	**Ridicule:** Mocking applause for a clumsy or failed action.
أنت ذَكي.	Isolated, heavy stress on the adjective "ذكي" <CONTRASTIVE_STRESS>	"You are smart."	**Ironic Condemnation:** Implying stupidity, dependent on context (e.g., after a foolish remark).

Table 4: Intonational & Pitch-Based Irony.

Example	Pitch Contour & Symbolic Tag		Paralinguistic Cues	Pragmatic Function
واو! إنجاز رائع.	Artificially high pitch peak on "واو", exaggerated rise <IRONIC_HIGH_PITCH>		Widened eyes, smirk.	**Mock Praise:** Pretending to be impressed by a trivial or non-existent achievement.
طِبْعاً كَلامك صَحيح.	Sharp, low-falling terminal contour on "صحيح" <LOW_FALL_TERMINAL>		Slow head shakes, neutral face.	**Dismissive Rejection:** Stating the opposite to strongly indicate disagreement.
أنت شَباب على الدَّوام.	Exaggerated, "singsong" melodic contour over the entire phrase <SINGSONG_CONTOUR>		Patronizing smile, gentle tone.	**Patronizing Humor:** Kindly mocking an elderly person's age by stating the impossible opposite, often for affectionate morale support.

Table 5: Pausal, Rhythmic & Compound Strategies.

Example	Prosodic Device & Symbolic Tags	Combined Cues	Pragmatic Effect & Interpretation
هَذا... قَرار حَكيم.	Strategic mid-utterance pauses before the evaluative adjective <STRATEGIC_PAUSE>	Pause + neutral/low pitch on "حكيم".	**Feigned Deliberation / Doubt:** Pretending to search for a positive word to highlight the decision's folly.
أه... أيْ لَقَدْ فُزْنـا.	Pre-utterance sigh/laugh + emphatic stress on the verb <LAUGHTER> <OVERSTRESS>	"أه" (sigh) + over-articulated "فزنا".	**Defeat Masking as Victory:** Explicitly stating victory with prosody that conveys the opposite, following a clear loss.
مُحْتَرِف... وَالله مُحْتَرِف.	Repetition with a pause, flat intonation on the second iteration <FLAT_TONE> <REPETITION>	Monotone, lack of emphasis.	**Ironic Affirmation:** Using repetition and dull delivery to drain the word of its positive meaning, implying utter incompetence.

4.1. Taxonomical Insights and Computational Implications:

This structured taxonomy reveals that Arabic sarcastic prosody is not monolithic but comprises a repertoire of distinct, often combined, strategies. Key insights include:

Feature Combinatorics: Sarcasm is frequently cued by the co-occurrence of multiple features. For example, "راائع" combines prolongation (<PROLONGED>) with emphatic stress (<OVERSTRESS>). The compound cue "أه... فُزْنـا" combines a paralinguistic vocalization (<LAUGHTER>) with lexical stress. This necessitates computational models that can detect patterns and interactions between features, not just isolated cues.

Context-Feature Interdependence: The same prosodic pattern can serve different pragmatic functions based on lexical polarity and context. A

singsong contour (<SINGSONG_CONTOUR>) with a positive phrase ("شباب على الدوام") creates affectionate mockery, while the same contour with a negative phrase could signal dismissive sarcasm. This underscores the need for the PPIL to integrate prosodic, lexical, and contextual signals.

From Acoustic Signal to Pragmatic Tag: Each symbolic tag in the taxonomy (e.g., <PROLONGED+>, <IRONIC_HIGH_PITCH>) is defined by measurable acoustic thresholds (e.g., duration > 2.5x baseline, pitch range > 40Hz above neutral). This provides the essential mapping for the Prosody Tagger in the computational pipeline, transforming continuous speech signals into discrete, semantically meaningful annotations that the MT system can process.

This taxonomy thus serves as the crucial codebook that links the raw data of spoken Arabic to a formal representation of meaning. It is the foundational resource that enables the subsequent steps of corpus annotation, model training, and the ultimate goal of pragmatically accurate machine translation.

5. ANALYSIS AND COMPUTATIONAL FRAMEWORK: FROM LINGUISTIC PRINCIPLES TO MODEL ARCHITECTURE

This section synthesizes the empirical findings from the taxonomy into a coherent theoretical analysis and directly translates these insights into a concrete computational proposal. We first distill the underlying linguistic principles governing Arabic prosodic sarcasm. These principles then inform the architecture and evaluation of a novel Machine Translation (MT) framework designed to close the pragmatic gap. This integrated approach demonstrates how theoretical understanding drives practical innovation in computational linguistics.

5.1. Linguistic Analysis: Principles Of Prosodic Sarcasm in Arabic

Analysis of the systematized data reveals that sarcasm in spoken Arabic is not signaled through arbitrary vocal flourishes but follows a predictable, rule-governed system.

Three core, interlocking principles emerge:

5.1.1. The Principle of Prosodic Markedness

Sarcastic intent is reliably cued by a deliberate and perceptually salient exaggeration of one or more suprasegmental features beyond the baseline of neutral, sincere speech. This marked deviation acts as a perceptual "flag," signaling the listener to re-evaluate the literal meaning. This exaggeration manifests in quantifiable acoustic dimensions:

-Temporal Exaggeration: The vowel in "قاااارس" is lengthened to a duration incongruous with a sincere meteorological report, directly contradicting the sensory reality of intense heat.

-Tonal Exaggeration: The pitch peak in "واااو!" is artificially heightened, creating a caricature of genuine surprise.

-Dynamic Exaggeration: The stress on "رااائع" is overly emphatic, mimicking but distorting the pattern of sincere admiration.

This principle provides the acoustic foundation for computational modeling, as these deviations (e.g., duration > 250% of neutral baseline, pitch range > 1.5 octaves) can be measured and mapped to symbolic tags (<PROLONGED+>, <IRONIC_HIGH_PITCH>).

5.1.2. The Principle of Evaluative Polarity Reversal

Prosody functions as a powerful pragmatic operator capable of systematically inverting the evaluative polarity of lexical content. This is most potent with positive evaluative predicates (adjectives, certain nouns). The prosodic manipulation semantically "flips" the utterance's stance.

Canonical Reversal: PRAISE + Sarcastic Prosody → CRITICISM. This direct mapping is evident in "عمل رااائع" (great work) becoming condemnation, or "أنت ذكي" (you are smart) implying foolishness.

Contextualized Reversal: The outcome depends on shared knowledge and lexical meaning. In "أنت شباب على الدوام", the lexical content is an impossible compliment. The singsong prosody (<SINGSONG_CONTOUR>) transforms it not into criticism but into affectionate mockery, creating a socially bonding form of irony: IMPOSSIBLE PRAISE + Playful Prosody → AFFECTIONATE IRONY.

5.1.3. The Principle of Contextual Anchoring and Incongruity

Prosodic sarcasm is not decoded in isolation; it is activated by contextual incongruity. The prosodic markers are the key, but the lock is the shared situational knowledge.

Incongruity as Trigger: The sarcastic reading of "سرييييع!" is only accessible if the listener knows the referent is slow. The marked prosody highlights this incongruity.

Disambiguation Function: Absent context, the exaggerated prosody in "جمييييلة!" could be misinterpreted as emphatic sincerity. The contextual

anchor (e.g., the visual appearance of the person addressed) is essential to disambiguate the cue as sarcastic.

Computational Implication: This principle dictates that any effective model must be multi-modal or context-aware, integrating acoustic-textual signals with a representation of situational "common ground" to correctly assign the pragmatic operator.

These principles collectively establish that suprasegmental features in Arabic sarcasm perform a semantic-pragmatic function. They are essential, grammaticalized components of meaning construction for non-literal speech, not optional paralinguistic embellishments.

5.2. The Prosodic-Pragmatic Interface Layer (Ppil): A Computational Framework

To operationalize the above principles, we propose a novel MT architecture centered on a Prosodic-Pragmatic Interface Layer (PPIL). This layer is designed to explicitly model suprasegmental meaning as a semantic operator, transforming prosodic input into pragmatic directives for the translation engine.

5.2.1. Framework Architecture and Workflow

The proposed framework functions as a cascade, integrating Automatic Speech Recognition (ASR), prosodic analysis, and neural MT:

-Input Processing: The source audio is processed by an ASR module to generate an orthographic transcription T_s. In parallel, a Prosody Tagger (trained on the SARCASM corpus) analyzes the acoustic signal A_s, extracting features and assigning a sequence of symbolic tags (e.g., سريع <PROSODY: PROLONGED_VOWEL+>).

-The PPIL Core: This is the disambiguation hub. It receives the enriched text (T_s+ prosodic tags) and, guided by the linguistic principles, interprets the combination. It learns to associate specific tag patterns with pragmatic operators (e.g., PROS_SARC). For an input like "عمل راااائع" tagged with <OVERSTRESS> and <PROLONGED>, the PPIL activates the reversal operator.

-Pragmatically Guided Generation: The PPIL's output (a pragmatic intent vector) is fused with the text embeddings and fed into the neural MT decoder. Instead of generating a literal translation ("Great work"), the decoder is conditioned to produce an output whose pragmatic stance matches the source's ironic intent (e.g., "Some 'great' work there," or "What a brilliant job").

5.2.2. Experimental Validation: Model Variants and Evaluation

To validate the framework's efficacy, a comparative experiment is designed to isolate the contribution of prosodic information.

Model Variants:

Baseline (Text-Only): A state-of-the-art neural MT model (e.g., Transformer) trained only on text T_s. This represents the current pragmatically deficient standard.

Pipeline Model (Symbolic PPIL): Implements the full framework described above, using discrete symbolic tags from the Prosody Tagger.

Integrated Model (Acoustic PPIL): An end-to-end variant where continuous acoustic features are fused with text embeddings at the encoder, allowing the model to learn prosodic patterns directly from the raw signal A_s.

Evaluation via Pragmatic Accuracy (PA): Traditional metrics like BLEU are inadequate as a "perfect" literal translation constitutes a pragmatic failure for sarcasm. We therefore propose Pragmatic Accuracy (PA) as the core evaluation metric.

PA is a human-judged measure of stance preservation:

$$PA = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\left[I\left(U_s^{(i)}\right) = J\left(\hat{T}_t^{(i)}\right)\right]$$

PA = (Number of Correctly Translated Pragmatic Stances) / (Total Number of Sarcastic/Literal Utterances)

Illustrative Example of PA Scoring: Consider the source Arabic utterance **"عَمَل رااائع"** (literally "Great work") spoken with marked prosody to criticize a poor performance.

- **Baseline MT Output (Literal):** "Great work." A bilingual judge identifies this as a **Pragmatic Failure** (score 0), as it conveys praise, not criticism.
- **PPIL-Enhanced MT Output:** "Some 'great' work there." The judge identifies this as **Correct Pragmatic Transfer** (score 1), as the ironic, critical stance is preserved through lexical choice and punctuation.For a given test set, the PA score is the proportion of utterances (e.g., 45 out of 50) for which the judges deem the pragmatic intent to be correctly transferred in the translation.

Protocol: A panel of five bilingual judges will rate translations on a 3-point scale: (1) Correct Pragmatic Transfer, (2) Partial/Ambiguous Transfer, (3) Pragmatic Failure. The central hypothesis is that models incorporating the PPIL (Pipeline and Integrated) will achieve a statistically significant

higher PA ($p < 0.01$) on sarcastic utterances than the Baseline, while maintaining performance on literal ones.

5.2.3. Formal Integration into Model Training

The PA metric can be directly integrated into the model's optimization process. Through techniques like Reinforcement Learning or Minimum Risk Training, the translation model can be fine-tuned using PA (or a reward derived from it) as an auxiliary objective. This steers the model towards maximizing pragmatic equivalence, even when it requires deviating from lexical literalness, thereby formally embedding the principles of prosodic meaning reversal into the learning algorithm.

This combined analysis and framework section completes the logical arc of the paper: from observed data to explanatory linguistic principles, and from those principles to a detailed, evaluable computational model. The PPIL represents a concrete architectural solution to the problem articulated in the introduction, providing a pathway to MT that can truly "hear" and translate the irony embedded in the prosody of spoken Arabic.

6. IMPLICATIONS, CHALLENGES, AND FUTURE DIRECTIONS

The proposed framework for integrating suprasegmental prosody into Arabic MT carries significant theoretical, practical, and ethical implications. However, its realization faces substantial challenges that reflect broader issues in computational linguistics and AI. This section provides a critical examination of both the potential impact and the hurdles that must be overcome, concluding with a roadmap for future research.

6.1. Implications: Toward Richer and More Culturally Aware NLP

The successful development of a prosody-aware MT system would represent a paradigm shift, moving beyond lexical-syntactic translation toward genuine pragmatic competence. The implications extend far beyond improved translation of sarcastic quips.

1. Enhanced Pragmatic Fidelity in Machine Translation: The most direct implication is a dramatic improvement in the quality and appropriateness of Arabic-to-X translations for spoken language. Systems could avoid the blunt, often offensive, errors that arise from literal translation of ironic speech. For diplomacy, business negotiations, media translation, and everyday conversation, this would reduce miscommunication and foster more authentic cross-cultural exchange. It directly addresses the "pragmatic gap" identified as a core limitation of current NMT systems (Al-Sayed, 2018).
2. Revolutionizing Arabic Sentiment and Affect Analysis: Current sentiment analysis tools for Arabic, trained largely on text, are easily fooled by ironic praise or sarcastic criticism (Al-Harbi & Emam, 2021). A robust prosodic sarcasm detector would enable a new generation of tools for analyzing call center recordings, social media videos, and public speeches, accurately capturing the true affective stance of speakers rather than being misled by surface-level lexicon. This has profound applications in market research, political analysis, and customer service analytics.
3. Advancing Human-Computer Interaction (HCI) and Conversational AI: For Arabic-speaking users, voice assistants and chatbots that are "tone-deaf" to sarcasm feel unnatural and limited. Integrating the PPIL framework could enable these systems to recognize when a user is being facetious (e.g., saying " أهنا... فزنا" after a game loss) and respond appropriately—with humor, empathy, or clarification. This moves conversational AI closer to passing a "pragmatic Turing test" for Arabic, building trust and enabling more complex, socially nuanced interactions.
4. Contribution to Culturally Grounded and Linguistically Inclusive AI: This research explicitly bridges modern computational methods with classical Arabic linguistic theory (`Ilm al-Balāghah). In doing so, it challenges the Anglocentric bias prevalent in NLP (Joshi et al., 2017) and offers a model for developing language technology that is deeply informed by a language's own rhetorical and philosophical traditions. It demonstrates that achieving true language understanding for AI requires engagement with culturally specific pragmatic norms, moving beyond a one-size-fits-all approach derived from English text corpora.

6.2. Critical Challenges and Limitations

The path to realizing this framework is fraught with technical, linguistic, and ethical challenges that must be rigorously addressed.

1. The Resource Bottleneck: Lack of Annotated Multimodal Corpora. As highlighted in the literature review, the absence of a large-scale,

publicly available corpus like the proposed SARCASM resource is the primary obstacle (Bousquet et al., 2023). Creating such a corpus is labor-intensive, expensive, and requires expertise in phonetics, pragmatics, and dialectology. Furthermore, ethical concerns around recording and annotating natural speech, particularly for sensitive acts like sarcasm, demand rigorous protocols for informed consent and data anonymization.

2. Dialectal and Idiolectal Variation in Prosodic Realization. Arabic is a diglossic language with immense dialectal diversity. The prosodic markers of sarcasm in Levantine Arabic (e.g., the specific pitch contour for "واااو!") may differ subtly or significantly from those in Gulf or Maghrebi dialects. A model trained on one dialect may not generalize well to others. Furthermore, individual speaker idiolects add another layer of variation. This necessitates either (a) dialect-specific modeling, which multiplies data requirements, or (b) the development of highly robust, dialect-agnostic feature representations—a significant open research problem.
3. The Inherent Subjectivity and Context-Dependence of Sarcasm. Sarcasm perception is not a binary, objective fact but exists on a continuum and is influenced by the relationship between speaker and listener, cultural background, and individual sense of humor. An utterance like "أنت شباب على الدوام" might be perceived as affectionate by a grandchild but as patronizing by a stranger. This subjectivity complicates the creation of "gold-standard" labels for training and evaluation. High inter-annotator agreement is difficult to achieve, and models may learn the biases of their annotators rather than a universal pragmatic truth.
4. The ASR Dependency and Error Propagation. The proposed pipeline is critically dependent on the accuracy of the front-end Automatic Speech Recognition (ASR) system. ASR errors—common in noisy environments or with non-standard dialects—would corrupt both the textual input and the alignment required for prosodic feature extraction. A mis-transcribed word could lead the PPIL to completely misinterpret the utterance. This creates a cascade of errors where the MT system is blamed for a failure originating in the speech recognition layer.

6.3. Future Directions and Concluding Synthesis

To move from a proposed framework to a realized technology, future work must tackle these challenges through a coordinated, interdisciplinary research agenda.

1. Prioritize Corpus Creation and Community Resource Development. The field must mobilize to create and share annotated multimodal resources. This could involve collaborative efforts across institutions, leveraging crowdsourcing platforms with careful quality control, and developing semi-automated tools for prosodic annotation. The SARCASM corpus proposal should serve as a blueprint and a call to action.
2. Pursue Multimodal and Context-Integrated Models. Future architecture should move beyond the audio-text dyad. As the Principle of Contextual Anchoring dictates, visual cues (facial expressions, eye rolls, smirks) and broader discourse context are often essential for disambiguation. Incorporating computer vision for facial expression analysis and developing context-aware models that consider the preceding dialogue could significantly improve accuracy (Castro et al., 2019).
3. Explore Cross-Lingual and Low-Resource Learning Techniques. Given the data scarcity for Arabic, techniques like cross-lingual transfer learning from better-resourced languages (e.g., using English prosodic corpora to bootstrap Arabic models) or few-shot/zero-shot learning paradigms should be investigated. The universality of some prosodic exaggeration (Bryant & Fox Tree, 2005) may make such transfer feasible.
4. Conduct Rigorous, Human-in-the-Loop Evaluation. Beyond the proposed PA metric, long-term user studies are needed to assess the real-world impact of prosody-aware MT on communication effectiveness and user satisfaction. Evaluation must be ethical, involving diverse groups of native speakers to audit for biases and unintended consequences.

In conclusion, this study has demonstrated that suprasegmental prosody is the linchpin of sarcasm and irony in spoken Arabic—a fact long recognized in the language's rhetorical tradition but persistently ignored in its computational treatment. By developing a formal taxonomy and proposing the Prosodic-Pragmatic Interface Layer, we have laid both the descriptive and architectural groundwork for a new generation of pragmatic machine

translation. The journey forward is challenging, demanding collaboration between linguists, computer scientists, and ethicists. However, the reward is substantial: not just more accurate translations, but a fundamental step toward AI that understands language as humans do—not merely as a sequence of words, but as a dynamic, socially-grounded, and profoundly nuanced act of meaning-making. Teaching machines to "hear" Arabic sarcasm is, ultimately, about teaching them to listen to human beings.

7. CONCLUSION AND SYNTHESIS: TOWARD A NEW PARADIGM FOR ARABIC MT

This research has established a foundational argument: suprasegmental prosody is the principal linguistic system for encoding sarcasm and irony in spoken Arabic, a function that remains critically absent in contemporary machine translation. To bridge this "pragmatic gap," we have undertaken a comprehensive investigation, moving systematically from empirical observation to computational design.

The study's core contributions are threefold:

- An Empirical Taxonomy: We developed a fine-grained, acoustically grounded taxonomy of Arabic sarcastic prosody (Section 4), moving beyond anecdote to catalog the systematic use of vowel prolongation (e.g., "سرييييع"), contrastive stress ("رااائع"), ironic pitch contours ("واااو!"), and compound strategies ("أه... فزنا"). This taxonomy, rooted in field observations, provides the essential "codebook" linking acoustic signals to pragmatic meaning.
- A Linguistic Framework: We distilled the taxonomy into three governing principles (Section 5.1): Prosodic Markedness (exaggeration as a cue), Evaluative Polarity Reversal (prosody as a semantic operator), and Contextual Anchoring (the necessity of shared knowledge). These principles position suprasegmental features not as paralinguistic ornamentation but as core, grammaticalized components of Arabic pragmatic meaning.
- A Computational Blueprint: We proposed the Prosodic-Pragmatic Interface Layer (PPIL), a novel architecture designed to integrate prosodic annotation into MT pipelines (Section 5.2). By formalizing prosody as a disambiguating operator and introducing the Pragmatic Accuracy (PA) evaluation metric, we offer a concrete, testable pathway from text-based translation to pragmatically aware generation.

The path forward is clear and demands concerted effort. The immediate imperative is the creation of the SARCASM corpus—a large-scale, annotated multimodal resource to train and validate the proposed models. Subsequent research must then pursue multimodal integration (incorporating visual and contextual cues), develop strategies for handling dialectal variation, and refine human-centric evaluation protocols.

Ultimately, this work advocates for a paradigm shift. For machine translation to achieve true fluency in Arabic, it must evolve from processing text to interpreting speech in all its acoustic richness. By teaching systems to detect the sarcasm in a drawn-out vowel or a strategic pause, we move beyond syntactic fidelity toward genuine communicative competence. This is not merely a technical upgrade but a step toward more culturally-grounded, sensitive, and human-like artificial intelligence—one that doesn't just translate words but understands meaning.

REFERENCES

Abdul-Raof, H. (2006). *Arabic rhetoric: A pragmatic analysis*. Routledge.

Additional Suggested References (Contextualizing Arabic Prosody & MT):

Albadi, N., Kurdi, M., & Mishra, S. (2018). Are they our brothers? Analysis and detection of religious hate speech in the Arabic Twittersphere. In *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 69-76). IEEE.

Al-Harbi, A., & Emam, A. (2021). Arabic sentiment analysis: A systematic literature review. *IEEE Access*, 9, 152995-153023. https://doi.org/10.1109/ACCESS.2021.3126503

Al-Sayed, M. (2018). *Neural machine translation for Arabic: A survey*. ACM Computing Surveys (CSUR), 51(6), 1-34. https://doi.org/10.1145/3284391

Anolli, L., Ciceri, R., & Infantino, M. G. (2002). From "blame by praise" to "praise by blame": Analysis of vocal patterns in ironic communication. *International Journal of Psychology*, 37(5), 266-276. https://doi.org/10.1080/00207590244000106

Attia, M. (2008). *Handling Arabic morphological and syntactic ambiguity within the LFG framework with a view to machine translation* [Doctoral dissertation, University of Manchester].

Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*. https://doi.org/10.48550/arXiv.1409.0473

Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer* [Computer program]. Version 6.3.09. http://www.praat.org/

Bousquet, M., Zahra, O. A., & Haddad, B. (2023). Multimodal corpora for Arabic: A critical review and future directions. In *Proceedings of the 13th International Conference on Language Resources and Evaluation (LREC 2023)* (pp. 1234-1243). European Language Resources Association.

Bryant, G. A., & Fox Tree, J. E. (2005). Is there an ironic tone of voice? *Language and Speech*, 48(3), 257-277. https://doi.org/10.1177/00238309050480030101

Castro, S., Hazarika, D., Pérez-Rosas, V., Zimmermann, R., Mihalcea, R., & Poria, S. (2019). Towards multimodal sarcasm detection (an *obviously* perfect paper). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 4619-4629). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1455

Cheang, H. S., & Pell, M. D. (2008). The sound of sarcasm. *Speech Communication*, 50(5), 366-381. https://doi.org/10.1016/j.specom.2007.11.003

El Kholy, A., & Habash, N. (2012). Orthographic and morphological processing for English–Arabic statistical machine translation. *Machine Translation*, 26(1-2), 25-45. https://doi.org/10.1007/s10590-012-9117-1

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41-58). Academic Press.

Habash, N. Y. (2010). *Introduction to Arabic natural language processing*. Morgan & Claypool Publishers. https://doi.org/10.2200/S00277ED1V01Y201008HLT010

Joshi, A., Bhattacharyya, P., & Carman, M. J. (2017). Automatic sarcasm detection: A survey. *ACM Computing Surveys (CSUR)*, 50(5), 1-22. https://doi.org/10.1145/3124420

Mubarak, H., & Darwish, K. (2014). Using Twitter to collect a multi-dialectal corpus of Arabic. In *Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP)* (pp. 1-7). Association for Computational Linguistics.

Pell, M. D., & Kotz, S. A. (2011). On the time course of vocal emotion recognition. *PLoS ONE*, 6(11), e27256. https://doi.org/10.1371/journal.pone.0027256

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)* (pp. 5998-6008). Curran Associates, Inc.

Wilson, D., & Sperber, D. (2012). *Meaning and relevance*. Cambridge University Press. https://doi.org/10.1017/CBO9781139028370

Zaidan, O. F., & Callison-Burch, C. (2014). Arabic dialect identification. *Computational Linguistics*, 40(1), 171-202. https://doi.org/10.1162/COLI_a_00169

Zhang, S., Zhang, X., Chan, J., & Rosso, P. (2021). Irony detection via sentiment-based transfer learning. *Information Processing & Management*, 58(4), 102571. https://doi.org/10.1016/j.ipm.2021.102571