
| RESEARCH ARTICLE

Constructing Textual Coherence through Pronominal Systems: A Contrastive Study of Albanian, German, and English

Jetmire ALIU

PhD in Linguistics, University of Tetova, North Macedonia

Corresponding Author: Jetmire ALIU, **E-mail:** jetmire.aliu@unite.edu.mk

| ABSTRACT

This article develops a contrastive analysis of how Albanian, German, and English construct textual coherence through their pronominal and morphosyntactic systems. The study places Arbër Çeliku's (2009, pp. 19–24) distinction between Coherence A (explicit coherence) and Coherence B (implicit coherence) at the center of its theoretical framework, demonstrating that the three languages differ fundamentally in how they encode referentiality, continuity, and discourse progression. Albanian, as a pro-drop language with a highly developed clitic system, realizes coherence largely through implicit and inferential mechanisms, whereas German and English rely primarily on overt pronominal expressions and expletive subject pronouns. Through a tri-lingual mini-corpus (6000 words) and synthetic media-inspired examples based on contemporary Albanian news discourse (2023–2025), the study shows that Albanian systematically omits subjects, encodes referents through morphology and clitics, and avoids expletive subjects in existential and impersonal constructions. German and English, conversely, maintain explicit subject positions and rely heavily on overt pronouns for coherence. These findings argue for a coherence typology in which Albanian exemplifies a prototypical Coherence B system, and German and English codify Coherence A. The article concludes that a modern theory of coherence must account for languages in which implicit mechanisms, not surface markers, carry the essential burden of textual cohesion.

| KEYWORDS

Albanian, German, English, coherence, pronominal systems, Coherence A/B, clitics, pro-drop

| ARTICLE INFORMATION

ACCEPTED: 01 December 2025

PUBLISHED: 22 December 2025

DOI: 10.32996/ijls.2025.5.5.2

1. Introduction

Textual coherence has long been a central concept in linguistics, yet its formulation has often been shaped by languages whose grammatical constraints mandate pronominal and structural markers. Most classical theories of cohesion and coherence developed in the late 20th century implicitly assume languages like English and German, where subjects must be realized overtly, where expletive subject pronouns (*it*, *there*, *es*) are structurally obligatory, and where coherence is often visible through surface links such as pronouns, lexical repetition, and connectors.

However, this architecture does not apply universally. **Albanian**, a morphologically rich pro-drop language, demonstrates that coherence is not necessarily achieved through surface forms. Instead, Albanian relies on:

- zero subjects (subject omission),
- verbal morphology,
- preverbal object clitics,
- inferential reconstruction by the reader,

- the generic *ti* ("you"),
- flexible topicalization patterns.

These features create a system in which coherence is not overtly marked but **inferred**, creating what Çeliku (2009, pp. 19–24) terms **Coherence B**: implicit coherence. By contrast, German and English instantiate what he calls **Coherence A**: explicit coherence marked on the textual surface.

This distinction is not merely terminological; it reflects a deep difference in:

- how languages organize discourse,
- how readers process texts,
- and how textual relations are grammatically encoded.

The purpose of this article is to provide the most comprehensive application to date of Çeliku's Coherence A/B typology to the pronominal systems of Albanian, German, and English. While earlier contrastive studies often limited themselves to grammatical equivalences or translation problems, the present work shows that the three languages embody fundamentally different **coherence architectures**, each with distinct textual behaviors.

The goals of this study are:

1. **To define and refine Coherence A and Coherence B** using recent linguistic theory.
2. **To analyze pronominal coherence markers**: subject pronouns, object pronouns, clitics, zero subjects, expletives in the three languages.
3. **To demonstrate with recent media-inspired examples** how Albanian employs implicit coherence mechanisms in real discourse.
4. **To explain why translation between these languages demands structural and cognitive adjustments.**

To accomplish this, the study integrates:

- a 6000-word tri-lingual mini-corpus,
- updated theoretical insights from discourse linguistics (e.g., Kehler, 2002; Arnold, 2010),
- synthetic but realistic examples modeled on 2023-2025 Albanian media.

The remainder of this article is organized into theoretical, methodological, empirical, and interpretive sections leading to a cohesive argument for a revised cross-linguistic coherence typology.

2. Theoretical Framework

This section presents and elaborates Arbër Çeliku's (ibidem) distinction between **Coherence A** and **Coherence B**, integrating it with current research in discourse analysis, reference theory, and typology.

2.1 Çeliku's Core Distinction: Coherence A vs. Coherence B

In *Herstellung von Textkohärenz im Deutschen und im Albanischen* (2009), Çeliku redefines how coherence should be conceptualized across languages. His fundamental distinction is Coherence A (explicit coherence), which is realized through overt linguistic means, including:

- personal pronouns (*er, sie, es; he, she, it*),
- demonstrative pronouns,
- repeated noun phrases,
- explicit connectors (e.g., *however, therefore, denn, aber*),
- overt subjects in all finite clauses,
- expletive subject pronouns (*es, it, there*) in impersonal or existential clauses.

Languages that follow Coherence A typically possess:

- rigid syntactic requirements for overt subjects,
- limited tolerance for subject omission,
- and a surface-oriented coherence architecture.

German and English clearly belong to this category.

Coherence B (implicit coherence) is achieved through implicit, inferential mechanisms, including:

- zero subjects (pro-drop),
- reliance on verbal morphology for person/number reference,
- preverbal object clitics,
- omission of elements recoverable from context,
- absence of expletive subject pronouns,
- generic *ti* expressing generalized human action,
- contextual and pragmatic inference.

For Çeliku (2009, pp. 19–24), Albanian is a prototypical Coherence B language because:

- it routinely eliminates overt subject pronouns,
- it expresses object reference through clitic morphology,
- it encodes participant continuity implicitly,
- and it constructs coherence through inference rather than textual redundancy.

These characteristics make Albanian radically different from German and English not only in grammar but in *textual logic*.

2.2 Coherence A and B as Cognitive Strategies

Çeliku's typology is more than a grammatical observation; it is a cognitive model.

- Coherence A readers expect explicit cues and rely on visible markers.
- Coherence B readers expect implicitness and rely on verbal morphology + inference.

Thus, Albanian readers tolerate and indeed prefer implicitness, while German and English readers expect overt signals.

2.3 Integration with Modern Discourse Theory

Although Çeliku's model is rooted in Balkan philological tradition, it aligns well with modern theories:

Kehler's (2002) coherence relations

Coherence is maintained by relations such as cause, elaboration, and contrast. Coherence A languages often mark these relations; Coherence B languages often infer them.

Arnold (2010) – Accessibility and pronouns

In Coherence B systems, high accessibility - *zero form*.
In Coherence A systems, high accessibility - *reduced but overt pronoun*.

Typological work (Roberts, 2010)

Pro-drop is not merely syntactic; it entails discourse consequences. Çeliku's model provides those discourse consequences.

2.4 Why Albanian Requires a Coherence B Lens

The Albanian data cannot be explained adequately through traditional cohesion models because:

1. Zero subjects are the norm, not the exception.
2. Clitics express referential structure implicitly, not overtly.
3. No expletive subjects exist; meaning is reconstructed from context.
4. Generic *ti* (you) merges deictic and human-generic reference.
5. Coherence is built through inference, not surface cues.

Therefore, Albanian must be analyzed as a fundamentally implicit coherence system.

3. Methodology

The methodological framework of this study combines contrastive linguistic analysis, discourse-functional interpretation, and corpus-based observation. The aim is not to produce statistical generalizations but to build a qualitative, theory-driven account of how three languages: Albanian, German, and English, construct textual coherence through their pronominal systems.

3.1 Research Questions

The study addresses four central questions:

1. How do Albanian, German, and English differ in their use of pronominal and morphosyntactic devices that contribute to textual coherence?
2. How do these devices map onto Arbër Çeliku's (2009, pp. 19–24) distinction between Coherence A (explicit) and Coherence B (implicit)?
3. How do contemporary Albanian media texts (2023–2025) exemplify implicit coherence mechanisms?
4. What implications follow for coherence theory, typology, and translation studies?

The analysis thus moves from: form, function, coherence architecture and cross-linguistic implications.

3.2 Corpus Design

The study uses a tri-lingual mini-corpus of 6000 words:

- 2000 words of Albanian, drawn from contemporary newspapers and online media (news reports, opinion articles, political commentary).
- 2000 words of German, collected from reputable German newspapers (*Süddeutsche Zeitung*, *Die Zeit*, *FAZ*) and narrative prose.
- 2000 words of English, drawn from international news sources and contemporary narratives.

These texts were selected for:

- Genre comparability (informative prose with narrative elements),
- Recency (to ensure applicability to modern language use),
- Authenticity (professionally edited media and published texts),
- Representativeness (balanced across topics such as politics, society, crime, economy).

Because quoting full copyrighted media texts is restricted, the study uses synthetic but realistic examples that accurately reflect structures observed repeatedly in the corpus. These examples preserve the grammar, coherence patterns, and phrasing typical of Albanian news discourse.

3.3 Annotation Scheme

Each clause in the corpus was manually annotated for:

- **Subject type:**
 - overt NP,
 - overt pronoun,
 - zero subject ($\emptyset$).
- **Object type:**
 - full NP,
 - pronoun,
 - preverbal clitic (e, ia, ua, t'i, ma, ta, etc.).
- **Coherence function:**
 - anaphoric,
 - cataphoric,
 - bridging inference,
 - generic reference.
- **Construction type:**
 - narrative,
 - impersonal,
 - existential,
 - opinion/generic.
- **Presence of expletive subject pronouns** (*es, it, there*) in German/English.

This multi-layered annotation allowed identification of systematic coherence strategies.

3.4 Analytical Procedure

Analysis proceeded in four steps:

1. Identification of coherence devices in each language.
2. Functional explanation of how these devices' structure discourse.
3. Application of Çeliku's Coherence A/B typology to categorize coherence mechanisms.
4. Contrastive interpretation across the three languages.

The combination of corpus evidence, theoretical analysis, and synthetic examples enables a robust account of cross-linguistic coherence systems.

4. Corpus Study

The corpus reveals three robust, cross-linguistically stable tendencies:

4.1 Albanian: Zero Subjects as Default

Across the Albanian sub-corpus:

- 67% of finite clauses with human subjects used zero subjects ($\emptyset$).
- In narrative and news reporting contexts, this rose to 74%.
- Subject pronouns (*ai, ajo*) appeared mostly:
 - when introducing a new referent,
 - for contrastive focus,
 - or when ambiguity could arise.

This demonstrates that Albanian systematically suppresses overt subjects. Subject visibility is governed by discourse context, not by grammar.

4.2 Albanian: Clitics as Referential Anchors

Preverbal clitics occur frequently in object-position contexts:

- 41% of transitive clauses had clitics.
- 22% showed clitic doubling (*e pa atë, ia dha librin atij*).
- Clitics were obligatory when:
 - a pronominal object was topical,
 - anaphoric reference needed reinforcement,
 - the discourse tracked actions of a stable referent over multiple clauses.

Thus, clitics function as micro-cohesive devices, supporting Coherence B.

4.3 German and English: Obligatory Overt Subjects

In German and English:

- Nearly 100% of finite clauses had overt subjects.
- Impersonal and existential constructions used:
 - German: *es*
 - English: *it, there*
- Subject omission was nearly impossible except in:
 - imperatives,
 - coordination ellipsis,
 - colloquial fragments.

This confirms German and English as Coherence A languages where coherence is constructed explicitly.

The contrast between Albanian and the two Germanic languages provides the grounding for the analysis that follows.

5. Analysis

This section presents a contrastive analysis of pronominal coherence mechanisms in Albanian, German, and English, using updated synthetic examples that accurately reflect modern media discourse.

5.1 Zero Subjects vs. Overt Subjects

Example 1 in Albanian:

Një 54-vjeçar me banim në Torino u arrestua mbrëmë në Shijak, pasi kishte tentuar të vidhte dy automjete. Më pas u konstatua se Ø kishte marrë sende me vlerë nga një makinë tjetër. Sipas policisë, Ø do të hetohet në gjendje të lirë.

Here:

- The referent is introduced once (Një 54-vjeçar).
- All later subjects are zero (Ø).
- Verbal morphology handles reference.

This illustrates a canonical Coherence B configuration, where coherence is established implicitly through subject omission and morphological agreement rather than through explicit pronominal realization.

German/English versions:

German:

Ein 54-Jähriger ... wurde festgenommen, nachdem **er** versucht hatte ... Später stellte sich heraus, dass **er** ... Nach Angaben der Polizei wird **er** ...

English:

A 54-year-old man ... was arrested after **he** attempted ... Later it was found that **he** ... According to police, **he** ...

Coherence A in both languages requires subject realization. Without overt *er/he*, the sentences would be ungrammatical.

5.2 Object Clitics vs. Object Pronouns

Example 2 in Albanian:

Policia njoftoi se kishte monitoruar lëvizjet e të dyshuarit për disa javë. Mbrëmë, pranë banesës, **Ø e** arrestoi pa rezistencë. Gjatë kontrollit, **Ø ia** gjeti në çantë paratë dhe dokumentet e vjedhura.

This pattern exhibits the following coherence-relevant characteristics:

- *e* (direct object clitic) = "him"
- *ia* (dative + accusative cluster) = "to him/from him"
- The subject remains zero throughout.
- Clitics obligatorily express referents in an implicit coherence system.

German:

... verhaftete **sie ihn** ... fand **sie bei ihm** ...

English:

... arrested **him** ... found the documents **on him** ...

German and English lack clitic systems and cannot encode referential structure morphologically. They compensate with full pronouns.

5.3 Impersonal and Existential Constructions

Example 3 in Albanian:

Ø U raportua për një shpërthim pranë një objekti industrial. **Ø** U dëmtuan disa automjete, por **Ø** nuk ka të lënduar.

Three successive clauses have no subject at all.

This is permissible because Albanian:

- allows **zero subjects**,
- has no expletive subject pronouns,
- expresses impersonal/causative events through **verb morphology alone**.

German:

Es wurde über eine Explosion ... Dabei wurden mehrere Fahrzeuge beschädigt, aber **es** gibt keine Verletzten.

English:

It was reported that ... Several vehicles were damaged, but **there** were no injuries.

These forms (*es, it, there*) are expletive subject pronouns, structurally necessary in Coherence A languages.

5.4 Generic Pronouns and Discourse Stance

Example 4 in Albanian:

Në këtë vend, **ti** paguan taksa rregullisht, por kur të duhet një shërbim bazë, **ti** pret me orë. Në fund, **ti** e kupton se askush nuk mban përgjegjësi.

In contemporary Albanian opinion and editorial discourse, the second-person singular **ti** frequently appears in a function that extends well beyond its canonical deictic meaning.

In this function, *ti* acts as:

- **a generic human agent whose reference extends to any member of the speech community,**
- **a non-addressive second-person form that does not target an actual interlocutor,**
- **a discourse-pragmatic device that constructs shared experiential alignment between writer and reader,**
- **a stance-marking tool that allows critique without naming specific actors,**
- **a mechanism for transforming individual action into collective experience.**

German:

In diesem Land zahlt **man** regelmäßig Steuern, aber wenn man eine grundlegende Dienstleistung benötigt, wartet **man** stundenlang in der Schlange. Am Ende erkennt **man**, dass niemand Verantwortung übernimmt.

English:

In this country, **you** pay your taxes regularly, but when you need a basic service, **you** end up waiting in line for hours. In the end, **you** realize that no one takes responsibility.

German uses *man*, English uses *you* generically.

Albanian merges:

- deictic "you",
- generic "one/people",
- community stance.

This merging is characteristic of Coherence B systems where implicitness binds speaker and reader.

5.5 Cataphoric and anaphoric coherence

Cataphoric and anaphoric reference behave quite differently in Albanian when compared to German and English, and these differences are revealing for the coherence architectures of the three languages. German and English, as Coherence A systems, easily permit cataphoric reference because overt subject pronouns are structurally obligatory. A clause such as *She is a brilliant woman – Tracy has always impressed her colleagues* or its German equivalent *Sie ist eine kluge Frau — Tracy hat ihre Kollegen stets beeindruckt* is entirely natural, since the languages require the pronoun in subject position regardless of its discourse antecedent. The pronoun can therefore appear before the referent without causing interpretive difficulty; coherence is secured by the transparent, surface-visible pronominal form that signals to the reader that referential specification will follow.

Albanian, by contrast, shows a strong and systematic preference for anaphoric coherence. Once a referent has been introduced, the language typically maintains reference through **zero subjects** rather than overt pronouns. A sentence such as *Tracy është një*

grua shumë e zgjuar. Ø ka impresionuar gjithmonë kolegët e saj ("Tracy is a very intelligent woman. Ø has always impressed her colleagues.") demonstrates the pattern clearly: verbal morphology and contextual accessibility are sufficient to sustain reference, eliminating the need for an overt pronominal subject. Because cataphora would require an explicit pronoun (*ajo/she/sie*), which disrupts the default economy of Albanian discourse, cataphoric constructions appear marked, rhetorically loaded, and comparatively rare. In other words, Albanian's tolerance for cataphora is constrained precisely because its coherence system does not normally rely on overt pronominal realization.

Interpreted through Çeliku's framework, this asymmetry reflects the broader distinction between Coherence A and Coherence B. German and English maintain coherence explicitly through obligatory pronominal forms, which makes both anaphoric and cataphoric reference equally accessible. Albanian, however, constructs coherence implicitly through morphology, inference, and subject omission; as a result, only anaphoric relations align naturally with its coherence system, whereas cataphora remains a marked and secondary option. These contrasts underscore that reference strategies are not merely stylistic choices but manifestations of deeper structural and typological properties governing how coherence is encoded in each language.

6. Discussion

The analysis demonstrates that coherence is shaped fundamentally by the grammatical architecture of each language rather than by universal surface markers. German and English construct coherence through overt pronominal forms and fixed syntactic requirements; Albanian, by contrast, relies on inference, morphology, and subject omission. These patterns confirm the distinction formulated by Çeliku (2009, *ibidem*): German and English exemplify **Coherence A**, where coherence is explicitly encoded, while Albanian aligns with **Coherence B**, where coherence emerges implicitly.

The contrast is particularly evident in subject expression. In German and English, overt subjects—whether referential or expletive—are obligatory, resulting in coherence chains built from repeated pronouns and clear anaphoric or cataphoric links. Albanian, however, maintains reference through zero subjects and verbal morphology; the reader reconstructs coherence internally rather than being guided by surface repetition. This implicit mechanism does not weaken coherence but redistributes it across the interpretive context.

Object reference follows a similar pattern. Albanian clitics encode pronominal information morphologically and obligatorily, creating compact and cohesive structures without the need for full pronouns. German and English, lacking such clitics, rely on overt object pronouns or lexical repetition. Again, coherence in Albanian derives from grammatical economy, while Germanic coherence rests on explicitness.

The absence of expletive subjects in Albanian further strengthens its implicit coherence profile. Impersonal and existential constructions are expressed through verb-centered clauses, whereas English and German introduce syntactic placeholders (*it*, *there*, *es*) that visually mark coherence. Albanian instead relies on contextual inference, consistent with a Coherence B system.

The generic use of *të* provides an additional illustration. While German and English distinguish clearly between generic and deictic pronouns, Albanian employs *të* to construct a shared experiential frame through pragmatic inference. This form of alignment reinforces Albanian's preference for coherence via interpretative engagement rather than overt marking.

Overall, the data show that coherence functions within a broader typological system. German and English foreground coherence on the surface of the text; Albanian embeds it in morphology and discourse expectations. These differences do not signal varying levels of complexity but reflect alternative, systematic strategies for organizing meaning in discourse.

7. Conclusions

This study set out to examine how Albanian, German, and English construct textual coherence through their pronominal and referential systems. The contrastive analysis has shown that coherence is intimately tied to the structural design of each language and cannot be understood independently of its grammatical and pragmatic foundations. While the Discussion section highlighted the operational mechanisms underlying each coherence type, the broader conclusion that emerges is that languages organize reference according to internally coherent principles that reflect long-term typological developments rather than isolated stylistic preferences.

Albanian, with its pro-drop syntax, obligatory clitic system, and context-dependent strategies of reference, illustrates how a language can distribute coherence across morphology and inference rather than surface visibility. This implicit orientation makes Albanian discourse appear highly economical, yet it also requires readers to draw actively on contextual knowledge to track participants and relations. German and English, by contrast, maintain coherence through a stable reliance on overt subjects, pronominal repetition, and syntactic structure. Their coherence is not derived from interpretative work but from the predictability and regularity of explicit markers.

The comparison demonstrates that coherence cannot be treated as a single, universal metric. Instead, it emerges as a typologically variable domain, shaped by the interaction of syntax, morphology, discourse norms, and reader expectations. Albanian's coherence system shows how reference can be maintained even when explicit signals are minimized, while the Germanic systems illustrate how reference can be stabilized through redundancy and formal marking. These divergent but equally effective strategies reveal the adaptability of human languages in organizing discourse.

Furthermore, the integration of recent, media-inspired Albanian data confirms that these coherence patterns are not confined to literary or formal contexts but permeate everyday public communication. This suggests that coherence types are deeply embedded in the communicative culture of each linguistic community, influencing how information is packaged, transmitted, and interpreted.

Ultimately, the study affirms the value of Çeliku's Coherence A/B framework for cross-linguistic research. It provides a productive lens through which to analyze the interplay between explicitness and implicitness in coherence formation. By applying this framework to pronominal systems across three typologically distinct languages, the study highlights the need for coherence theories that account for both overt and inferential strategies. Future research may extend this approach to other pro-drop languages or explore how bilingual speakers navigate these differing coherence architectures in multilingual contexts.

References

- [1] **Arnold, J. E.** (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, 4(4), pp. 187–203. <https://doi.org/10.1111/j.1749-818X.2010.00193.x>
- [2] **Asher, N., & Lascarides, A.** (2003). *Logics of conversation*. Cambridge: Cambridge University Press.
- [3] **Çeliku, A.** (2009). *Herstellung von Textkohärenz im Deutschen und im Albanischen*. Frankfurt am Main: Peter Lang Verlag.
- [4] **Gries, S. Th.** (2021). *Statistics for linguistics with R* (3rd ed.). Berlin: De Gruyter Mouton.
- [5] **Haegeman, L.** (2014). *Thinking syntactically: A guide to argumentation and analysis*. Chichester: Wiley-Blackwell.
- [6] **Kehler, A.** (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
- [7] **McEnery, T., & Hardie, A.** (2019). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.
- [8] **Roberts, I.** (2010). *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press.