

**Jarosław Wiliński**

University of Siedlce  
Ul. Konarskiego 2, Siedlce, PL-08-110  
<https://orcid.org/0000-0002-3136-6529>  
[jaroslaw.wilinski@uph.edu.pl](mailto:jaroslaw.wilinski@uph.edu.pl)

## *AC-RZECZ BIORĄC* V. *AC-MÓWIĄC* IN POLISH: A CORPUS-DRIVEN STUDY OF TWO QUASI-PARTICIPLE CONSTRUCTIONS

Applying the perspective of Construction Grammar and the corpus-based methodology termed *Distinctive-Collexeme Analysis*, this study aims to ascertain adverbial complements (ACs) that show a marked preference for either the *AC-rzecz biorąc*-construction ‘AC-thing taking-construction’ (e.g. *historycznie rzecz biorąc* ‘historically speaking’; verbatim: ‘historically thing taking’) or the *AC-mówiąc*-construction ‘AC-speaking-construction’ (e.g. *ogólnie mówiąc* ‘generally speaking’) in the Polish language. Drawing upon data retrieved from The National Corpus of Polish (NKJP), the findings reveal that these constructions display a strong tendency to co-occur with particular ACs that convey a range of meanings and serve different functions in discourse. In addition, the study reports that both constructions are predominantly observed in journalism, are used commonly in quasi-spoken texts, occur frequently in literature, and are relatively rare in Internet sources, scientific-didactic texts, and conversational contexts.

### **1. Introduction**

Recently, there has been a notable surge in scholarly publications addressing Polish metatextual expressions, specifically those featuring quasi-participles such as *biorąc* ‘taking’, *mówiąc* ‘speaking’, or *ujmując* ‘expressing’ preceded by adverbs and adverbial phrases (Stępień 2014; Birzer 2015, 2017; Kubicka 2017; Żabowska 2020, among others). These expressions have been analyzed from

four distinct perspectives: as lexical or phraseological units (e.g., Bojałkowska 2010: 16–19), as linguistic units belonging to the class of particles (Kisiel and Żabowska 2011; Grochowski 2002), as products of syntactic transformations (Wróbel 1975: 90–94; Moroz 2010: 267–274), and as outcomes of linguistic operations according to Bogusławski's perspective (Bogusławski and Danielewiczowa 2005: 347).

Some investigations have uncovered the origins of some quasi-participles (see Birzer 2017 for the origin of the participle *mówiąc*) and examined questions surrounding their lexical status (Weiss 2005; Kisiel and Żabowska 2011; Stępień 2014, 2015). Others have differentiated between various variants (Żabowska 2020) and briefly mentioned several specific instances (Wróbel 1975; Weiss 2005; Grochowski 2002; Żabowska 2009; Bojałkowska 2010; Moroz 2010; Kisiel and Żabowska 2011). Further studies have elucidated the general functions and prosodic characteristics of these constructions (Stępień 2014), explored diverse formal and semantic properties of adverbs (Kubicka 2017; Moroz 2007 for formal characteristics), and formulated hypotheses regarding the conditions under which adverbs can function independently (Kleszczowa 2015; Kubicka 2017).

Thus far, to the best of the author's knowledge, no investigation has systematically compared and contrasted the use of adverbial complements in the *AC-rzecz biorąc*-construction with those in the *AC-mówiąc*-construction. Moreover, little consideration has been given to the specific discourse functions of these constructions, their semantic properties, distribution patterns across different registers, quantitative assessment of ACs used in these two participial constructions, statistical validation of their occurrence in these functionally related structures, or the empirical substantiation of previous assumptions and speculations regarding their usage.

Previous research has primarily focused on identifying various adverbs commonly employed in these linguistic patterns, as evidenced by the works of Stępień (2014) and Żabowska (2020). It has also involved comparing [*Adv*] *mówić* with *mówiąc* [*Adv*], a topic explored by Żabowska (2020). Additionally, investigations have centered on elucidating the overarching metatextual function of the quasi-participle *mówiąc* in discourse, as undertaken by Stępień (2014), and the general semantic attributes of adverbial complements associated with this structure, as investigated by Kubicka (2017). Upon scrutinizing paren-

thetical expressions with *mówiąc*, *biorąc*, and *ujmując*, Stepień (2014) observed that the quasi-participle *mówiąc* tends to co-occur with adverbial phrases such as *nawiasem* ‘parenthesis’, *szczerze* ‘sincerely’, *krótko* ‘briefly’, *prawdę* ‘truth’, *delikatnie* ‘delicately’, *inaczej* ‘differently’, *ściśle* ‘strictly’, *ogłędnie* ‘cautiously’, and numerous others. Conversely, the expression *rzecz biorąc* frequently colligates with adverbial phrases like *chronologicznie* ‘chronologically’, *dokładnie* ‘accurately’, *formalnie* ‘formally’, *generalnie* ‘generally’, *historycznie* ‘historically’, *konkretnie* ‘specifically’, *krótko* ‘briefly’, *ogólnie* ‘generally’, and several others.

Żabowska (2020: 27–29) in turn differentiated between various quasi-participles, noting that the [ADV] *mówiąc* pattern is associated with adverbial phrases such as *prościej* ‘more simply’, *najprościej* ‘most simply’, *ogólnie* ‘generally’, *ogólniej* ‘more generally’, *delikatnie* ‘delicately’, *dokładnie* ‘accurately’, *dokładniej* ‘more accurately’, *generalnie* ‘generally’, *ściśle* ‘strictly’, *ściślej* ‘more strictly’, and several others. In contrast, the [ADV] *rzecz biorąc* pattern occurs with expressions like *ściśle* ‘strictly’, *ściślej* ‘more strictly’, *ogólnie* ‘generally’, *historycznie* ‘historically’, *zasadniczo* ‘essentially’, *praktycznie* ‘practically’, *teoretycznie* ‘theoretically’, *z grubsza* ‘roughly’, and *najogólniej* ‘most generally’.

Lastly, Kubicka (2017: 109) systematically categorized adverbial expressions that co-occur with the quasi-participle *mówiąc* into two broad functional categories. The first encompasses adverbials conveying an assessment of communication, such as *szczerze* ‘sincerely’, *otwarcie* ‘openly’, *uczciwie* ‘honestly’, *nieskromnie* ‘immodestly’, *poważnie* ‘seriously’, or *ogłędnie* ‘cautiously’. The second involves adverbials signifying actions performed on communication, thereby altering its content, as exemplified by *krótko* ‘briefly’, *skrótowno* ‘concisely’, *w skrócie* ‘in brief’, *z grubsza* ‘roughly’, *ogólnie* ‘generally’, *generalnie* ‘generally’, *precyzyjnie* ‘precisely’, *ściślej* ‘more strictly’, *dokładniej* ‘more accurately’, or *prościej* ‘more simply’, or adopting a symbolic system, including *inaczej* ‘differently’, *symbolicznie* ‘symbolically’, *obrazowo* ‘pictorially’, *metaforycznie* ‘metaphorically’, *przenośnie* ‘figuratively’, *popularnie* ‘popularly’, *delikatnie* ‘delicately’, *łagodnie* ‘gently’, *brutalnie* ‘brutally’, *elegancko* ‘elegantly’, and various others.

Given that no prior investigation has systematically analyzed and compared the *AC-rzecz biorąc*-construction and the *AC-mówiąc*-construction in terms of their

respective preferences for ACs, a gap in understanding the differences between these structures remains. Moreover, earlier studies did not specifically aim to identify the precise semantic and functional properties of ACs associated with these constructions, nor did they reveal nuanced distributional variations between the two. Thus, there is a need to determine the distribution patterns of ACs across different registers and within these closely-related linguistic structures. Furthermore, it is imperative to compare their frequencies of occurrence, considering the potential existence of subtle variations in their usage within discourse. Consequently, this paper aims to investigate consequential distributional disparities between the two quasi-participle constructions. The goal is to illustrate that, despite the initial appearance of functional and collocational similarity between these constructions (as in *ogólnie mówiąc* and *ogólnie rzecz biorąc*), they indeed exhibit nuanced semantic, functional, and distributional distinctions.

The hypothesis is that the *AC-rzecz biorąc*-construction will show a stronger preference for contexts that emphasize practical or theoretical reasoning, generality, objectivity, and formality. It is expected to be more prevalent in registers such as journalism and academic writing, where there is a need to present information in a logical, precise, or statistical manner. Conversely, the *AC-mówiąc*-construction will be more frequent in contexts that require the expression of truthfulness, informality, and succinctness, such as in quasi-spoken discourse, informal writing, and conversational contexts. This construction is likely to be associated with adverbial complements that convey a personal or subjective tone, introduce asides, or rephrase statements, reflecting its role in facilitating direct and clear communication in less formal settings.

The subsequent sections of this paper are structured as follows. In Section 2, attention is directed toward the theoretical underpinnings governing the semantic interpretation of ACs associated with both constructions. Simultaneously, emphasis is placed on the methodological framework guiding the quantitative analysis of ACs. Specifically, this section elucidates the particulars of the corpus, data, and tools employed, offering an outline of the statistical procedure enacted in this inquiry. In Section 3, the outcomes of the quantitative analysis are integrated with a semantic and functional exploration of ACs, thereby revealing

subtle distributional differences between ostensibly similar constructions. Section 4 summarizes and assesses the results.

## 2. Theoretical and methodological assumptions

This section presents an overview of the theoretical and methodological frameworks applied in the context of this research.

### 2.1. Construction grammar

The investigation is grounded in the theoretical frameworks of usage-based Construction Grammar (Goldberg 2006, 2013). The usage-based grammar model, as articulated by Diessel (2015: 296), posits that grammar constitutes a dynamic system characterized by emergent categories and adaptable constraints, continually undergoing change under the influence of domain-general cognitive processes inherent in language use. This perspective, supported by scholars such as Bybee (2010) and Goldberg (2006), rests on three specific assumptions concerning the nature of linguistic elements and the overall organization of the grammatical system. Firstly, grammatical structure should be examined in relation to simple and complex signs, or constructions, defined as pairings of a specific structural pattern with a particular function or meaning. Secondly, all linguistic signs, encompassing lexical units and grammatical patterns, are interconnected through various links, forming a dynamic network of constructions within grammar. Thirdly, our linguistic knowledge emerges from language use. In other words, the regular occurrence of grammatical units is essential for them to become deeply embedded in a speaker's or listener's linguistic system, which is crucial for their recognition as constructions. This phenomenon can be explained through the concept of *entrenchment* (Schmid 2017), a term first introduced by Ronald Langacker (1987) to describe how new expressions are created and become deeply rooted in a language. For example, the words *red* and *neck* were combined to form the new compound word *redneck*. Over time, this compound was used so frequently that its original components were forgotten. In other words, this phrase evolved into a standard expression, becoming so

firmly entrenched in the lexicon that it turned into a regular, well-established word within the language system.

## 2.2. Methodology

This study employs the *Distinctive-Collexeme Analysis* method (Gries and Stefanowitsch 2004a; Stefanowitsch 2013) to identify ACs that are strongly associated with the *AC-rzecz biorąc*-construction in comparison to the *AC-mówiąc*-construction. This method belongs to a broader group of corpus-linguistic techniques known as *collostructional analysis*, which is designed to examine the relationships between words and grammatical constructions (Stefanowitsch and Gries 2003, 2005; Gries and Stefanowitsch 2004a, b). Collostructional analysis encompasses three closely related methods: *simple collexeme analysis*, *distinctive collexeme analysis*, and *covarying collexeme analysis*. These methods are essentially collocation-based techniques adapted for analyzing grammatical constructions within a rigorously quantitative framework.

While the identification of associations between constructions and the words that occur within them is quantitatively objective, the interpretation of these associations is qualitative and subjective. Specifically, words that are strongly associated with a particular construction can be categorized into semantic classes based on introspective judgment. Collostructional analysis, however, differs from traditional corpus-based language studies (e.g., Biber et al., 1998; Stubbs 2001) in several key ways: it prioritizes the relationship between linguistic form and meaning, typically starting with a grammatical construction and identifying the words associated with this construction; it retrieves linguistic signs comprehensively from the corpus, which often requires manual annotation and analysis of large datasets; it relies exclusively on naturally-occurring language data from representative and balanced corpora; and it applies rigorous quantification and statistical analysis to evaluate association strengths.

Like other collocational methods, collostructional analysis relies on a two-by-two contingency table of co-occurrence frequencies, meaning that any distributional statistic for contingency analysis can be used to measure association strengths. However, practitioners of collostructional analysis have consistently opted to use the Fisher-Yates exact test for this purpose. Specifically, they use

the p-value generated by this test as an indicator of the strength of association—whether a word is strongly attracted to or repelled by a construction: the smaller the p-value, the stronger the association. The Fisher exact test offers a particular advantage in that it is well-suited for data that are unevenly distributed or infrequent in the corpus (see Stefanowitsch and Gries 2003: 9; Gries and Stefanowitsch 2004a: 101).

Thus far, distinctive collexeme analysis has been applied to various alternating pairs of constructions across several languages, as demonstrated in a number of studies (e.g., Wulff 2006; Coleman 2009; Desagulier 2014; Wiliński 2017, 2023). For example, Wulff (2006) explored the variation between the *go-V* and *go-and-V* constructions in English, while Coleman (2009) focused on the dative alternation in Dutch. Desagulier (2014) sought to uncover semantic dimensions within the conceptual structure of four English moderators and provided new insights into the use of degree modifiers. Wiliński (2017) identified nouns that show a significant preference for the *on the brink of-noun* construction over the *on the verge of-noun* construction. In a later study, Wiliński (2023) compared and contrasted the use of nouns within the *sea of-NOUN* construction and the *mountain of-NOUN* construction.

The method described by Gries and Stefanowitsch (2004a) is designed to explore the differences between two semantically or functionally similar constructions by identifying the words that most effectively distinguish between them. Unlike traditional collocation methods, which compare the frequency of a word in a construction to its frequency in the corpus as a whole, this approach compares the frequency of a word in one construction to its frequency in a corresponding slot in a semantically similar construction. Specifically, to calculate a word's association strength—or distinctiveness—four frequencies are needed: the frequency of the word in the construction A, the frequency of the same word in the construction B, and the frequencies of both constructions A and B with words other than the collexeme in question. These frequencies are then entered into a two-by-two table and analyzed using the Fisher exact test (Gries and Stefanowitsch 2004a: 104).

The methodology comprised a sequence of four steps. To elucidate these steps, one can specifically examine the adverb *praktycznie* 'practically' in the AC slot of the constructions under consideration. The distribution of this adverb in both

patterns, along with other frequencies essential for a distinctive-collexeme analysis, is provided in Table 1 below.

The initial step involved searching for these patterns in The National Corpus of Polish (NKJP), extracting their instances in concordance lines, discerning their meanings, and calculating the observed frequencies. The selection of this data source was motivated by the corpus's substantial size and its representative nature. The corpus is well-balanced, exhibiting representativeness of linguistic variations, and encompasses a collection of over 250 million tokens. It includes samples derived from diverse written and spoken Polish sources, such as classical literature, daily newspapers, specialized periodicals, journals, conversation transcripts, as well as ephemeral and internet texts.

Table 1. Contingency table cross-tabulating frequency scores of the constructions under study

<b>Constructions</b>	<b>ACs (<i>praktycznie</i>)</b>	<b>All other ACs</b>	<b>Total</b>
<b>AC-<i>mówiąc</i></b>	Frequency of AC ( <i>praktycznie</i> ) in 'the AC- <i>mówiąc</i> -construction' a= 4 (354.45)	Frequency of all other ACs in 'the AC- <i>mówiąc</i> -construction' b= 10988	Total frequency of 'the AC- <i>mówiąc</i> -construction' x = 10992
<b>AC-<i>rzecz biorąc</i></b>	Frequency of AC ( <i>praktycznie</i> ) in 'the AC- <i>rzecz biorąc</i> -construction' c= 435 (84.55)	Frequency of all other ACs in 'the AC- <i>rzecz biorąc</i> -construction' d = 2187	Total frequency of 'the AC- <i>rzecz biorąc</i> -construction' y = 2622
<b>Total</b>	Total frequency of AC ( <i>praktycznie</i> ) e = 439	Total frequency of all other ACs f = 13175	Total frequency of both constructions z = 13614

The PELCRA corpus search engine facilitated the seamless retrieval of the constructions in question. The extraction process involved entering the combinations *rzecz biorąc* and *mówiąc* into the search engine to accurately query the corpus and identify the desired linguistic patterns. Utilizing a concordance tool and CQL (Corpus Query Language), the software automatically extracted all occurrences of ACs within the span of five words to the left of these searched patterns. This also encompassed occurrences of ACs in the variant form of the AC-

*rzecz biorąc*-construction, where the noun *rzecz* is omitted (as in *ogólnie biorąc* ‘generally taking’). Given the complexity of language data, potential errors in adverbial lemmatization and tagging were anticipated. These could include incorrect tagging of word forms or misidentification of adverbial complements, especially in cases where the linguistic context might lead to ambiguity: *Nie czuję się najlepiej mówiąc o literaturze czy obyczajach starożytnych* (I don’t feel very comfortable talking about literature or customs of ancient times) or *Rzecz oczywista, biorąc pod uwagę to, o czym wcześniej wspomniałem* (It’s obvious, taking into consideration what I mentioned earlier). Therefore, manual correction was essential to ensure the accuracy and reliability of the dataset before further analysis.

The computational tool generated concordance lines, which were subsequently manually scrutinized to identify genuine combinations and ascertain the frequencies of all ACs occurring in the pertinent patterns. The retrieval process was confined to five positions to the left of the participles under examination. Any incorrect combinations that deviated from the specified patterns or suggested co-reference between the participle’s subject and the initial argument of the matrix verb were excluded from subsequent analysis. For example, the expression *często biorąc* meaning ‘often taking’ in the sentence *Poświęca się, często biorąc na siebie zbyt dużo* (He dedicates himself, often taking on too much) was considered a false hit as it indicated a regular use of the active verb participle *biorąc*.

This retrieval process was iterated for every combination, facilitating the computation of observed frequencies for both the patterns and the remaining instances of ACs, resulting in the identification of 10,992 occurrences of the *AC-mówiąc*-construction and 2,622 occurrences of the *AC-rzecz biorąc*-construction (including its reduced variant form). The numerical values (a, x, c, y) in Table 1 were directly extracted from the corpus, while the other values were determined through the operations of addition and subtraction.

In the second stage, these observed frequencies were utilized to compute the expected frequencies of ACs in both constructions. This calculation was executed in Microsoft Excel according to the following procedure: for example, for the adverb *praktycznie* in each pattern, the column total was multiplied by the row total, and the resulting score was divided by the overall table total. As an il-

lustrative example, the expected frequencies for the lemma *praktycznie* in each construction are provided in parentheses (see Table 1). If the observed frequency of the adverb *praktycznie* in the *AC-mówiąc*-construction significantly deviates from the expected, it indicates an association, either attraction or repulsion. In such cases, the adverb is deemed a significantly attracted or repelled collexeme of the *AC-mówiąc*-construction. Similarly, if the observed frequency of the adverb *praktycznie* in the *AC-rzecz biorąc*-construction significantly differs from the expected, it suggests a significant attraction or repulsion, categorizing the adverb as a collexeme of the *AC-rzecz biorąc*-construction.

The third step involved the computation of the strength of association (*collostruction strength*) between ACs and the constructions under investigation. To achieve this, the values (a, b, c, d) from Table 1 were input into a 2-by-2 table and subjected to the Fisher exact test, particularly recommended by statisticians when at least one expected value in the table is less than 5 (Levshina 2015: 29) or when data exhibits significant unevenness and/or infrequency (cf. Gries and Stefanowitsch 2004: 101), and when the total number of observations in all cells is less than 20 (Sheskin 2011: 646).

The resulting p-value from this test was used to assess the collostruction strength of each AC, indicating the degree of attraction to either the *AC-rzecz biorąc*-construction or the *AC-mówiąc*-construction. A smaller p-value implies a higher likelihood that the observed distribution is not due to chance and signals a stronger attraction between an adverbial complement and one of the constructions under scrutiny. The computation of statistical significance employed a built-in function of the R language for statistical computing called *fisher.test*. Subsequently, each p-value (e.g., 1.432667e-316 for *praktycznie*) was transformed into the logarithm to the base 10 (e.g., 315.84), providing a more easily interpretable score than the p-values often expressed in powers of ten. A collostruction strength exceeding 1.301 indicates a strong association of the lexeme with the construction, while a collostruction strength below 1.301 indicates that the lexeme exhibits a weak attraction to the construction. This analysis demonstrates that the adverb *praktycznie* is highly significant for one of the two constructions. However, to ascertain which one, a comparison of the observed frequencies of the adverb needs to be made with the expected frequencies. This comparison reveals that *praktycznie* occurs more frequently than expected in the *AC-rzecz biorąc*-

construction and less frequently than expected in the *AC-mówiąc*-construction. In essence, *praktycznie* is a very strongly distinctive collexeme of the *AC-rzecz biorąc*-construction as opposed to the *AC-mówiąc*-construction.

Lastly, the findings were organized based on the degree of attraction and subsequently assessed qualitatively. They were integrated with a semantic categorization of ACs, providing an in-depth explanation of subtle semantic and functional distinctions between these two constructions. The classification was developed based on the empirical analysis of the corpus data, where specific patterns of co-occurrence between ACs and the constructions *rzecz biorąc* and *mówiąc* were observed. This approach allowed the researcher to identify and group ACs according to their functional and semantic roles within the constructions, rather than imposing predefined categories from pragmatic theory. The goal was to let the data reveal the categories most relevant to these specific constructions.

The ACs were classified according to the functions they serve within the context of the *rzecz biorąc* and *mówiąc* constructions. These functions were identified through a close examination of how the ACs interact with the participles and the broader sentence structure. For example, some ACs were found to frequently introduce general statements, while others were more likely to indicate precision or qualify the truth of a statement. This construction-specific focus allowed the classification to directly reflect the roles that these ACs play in the patterns being studied, providing insights into how speakers use these constructions to achieve particular communicative goals.

The classification was intentionally designed to be flexible and sensitive to the specific context of use, rather than strictly adhering to predefined pragmatic categories. This flexibility was necessary because the corpus data revealed a range of nuanced uses for each AC that might not fit neatly into traditional pragmatic categories. By allowing the classification to emerge from the data, the analysis could account for the diversity of functions that these ACs serve in different contexts, highlighting their pragmatic versatility.

While the classification does not explicitly reference pragmatic notions such as modality, evidentiality, or hedging, these concepts nonetheless inform the analysis. The observed functions of the ACs—such as indicating generality, specificity, or cautiousness—can be seen as related to these broader pragmatic

notions. For example, ACs that introduce precise statements might be implicitly linked to notions of epistemic modality, where the speaker's certainty about the information is emphasized. Similarly, ACs that soften the impact of a statement could be seen as performing a hedging function, even if not explicitly labeled as such in the classification.

The decision to use a data-driven classification approach was motivated by the need to accurately reflect the patterns and nuances observed in the corpus. Traditional pragmatic categories, while useful, might not fully capture the specific ways in which ACs interact with these particular constructions. By deriving the classification from the data itself, the analysis remains closely related to the actual language use observed in the corpus, ensuring that the findings are directly relevant to the constructions under investigation.

In summary, the classification of ACs in this analysis is tailored to the specific findings of the study, reflecting the unique patterns of usage observed in the corpus. While it does not directly reference established pragmatic notions, it is grounded in the empirical data and focuses on capturing the construction-specific functions of the ACs. This approach allows for a flexible, context-sensitive classification that is directly relevant to the linguistic patterns being studied, while still being informed by broader linguistic concepts.

### 3. Results and discussion

An examination of concordance lines facilitated the identification of 10,992 instances of the *AC-mówiąc*-construction and 2,622 instances of the *AC-rzecz biorąc*-construction. Put differently, the former occurred roughly four times more frequently than the latter in the corpus. The computed frequencies, based on the token count of ACs in both constructions, reveal that the *mówiąc* pattern co-occurs with 123 types of ACs, of which 45 types were singular occurrences in the specified construction. In contrast, the *biorąc* pattern associates with 90 types of ACs, 40 of which were used only once in conjunction with this construction. This suggests that numerous ACs exhibit a loose association with both patterns, while others demonstrate a stronger affinity toward one of these constructions. Nevertheless, given the spatial limitations, our emphasis in this

section will predominantly center on the quantitative results pertaining to the 30 ACs that exhibit the highest degrees of distinctiveness within both constructions.

The term *type* in this context refers to distinct lexical items or unique forms of adverbial complements that occur with the construction. In contrast, a *token* refers to the individual instances of these adverbial complements in the corpus, regardless of whether the same word or form is repeated. For example, if the word *praktycznie* ‘practically’ (as an adverbial complement) appears four times with the *AC-mówiąc*-construction, it would count as one type but four tokens. Therefore, the mention of 123 types of ACs means there were 123 different adverbial complements associated with the *AC-mówiąc*-construction, whereas the 10,992 instances refer to the total number of occurrences (tokens) of all these types combined.

Regarding the frequency differences observed between the two constructions, there are several hypotheses that could be proposed. One possibility is that the observed frequency differences are influenced by the nature of the corpus sample, particularly its emphasis on written language. Since both constructions are predominantly observed in journalism and quasi-spoken texts (see Subsection 3.3), the higher frequency of the *AC-mówiąc*-construction could be attributed to its broader applicability and versatility in various written contexts, which are well-represented in the corpus. The *AC-mówiąc*-construction may be more adaptable to the concise and direct communication style often found in journalistic writing, where brevity and clarity are prioritized.

In contrast, the *AC-rzecz biorąc*-construction, which appears to be more specialized and formal, may be less frequently used due to its specific communicative function. This construction is often associated with conveying generality or objectivity, which corresponds with the stylistic demands of journalism but may not be as widely applicable across other registers. The less frequent occurrence of the *AC-rzecz biorąc*-construction could therefore imply that it serves a more specialized role, particularly in contexts where a formal tone or analytical perspective is required, such as in scientific-didactic texts.

Additionally, the distinct distribution patterns across registers suggest that the choice of construction is closely tied to the communicative goals of the speaker

or writer. The *AC-mówiąc*-construction's higher frequency might reflect its utility in a wider range of discourse functions, such as summarizing or emphasizing key points, which are common across multiple registers. On the other hand, the *AC-rzecz biorąc*-construction may be reserved for contexts where a more formal, structured expression is needed, limiting its overall frequency.

In conclusion, the frequency difference between the two constructions could be due to a combination of factors, including the corpus sample's composition and the specialized communicative functions of each construction. Further research into these hypotheses, particularly through a more detailed analysis of the contexts in which these constructions are used, would help clarify the underlying reasons for their differing frequencies.

### 3.1. Findings for the *AC-rzecz biorąc*-construction

The results of this investigation empirically validate the presence of collexemes that differentiate between the *AC-rzecz biorąc*-construction and the *AC-speaking*-construction. Moreover, they bring to light the semantic disparities between these patterns and the semantic constraints they impose on the ACs co-occurring with them. In Table 2 below, the thirty most distinctive collexemes of the first pattern are presented, along with the observed frequencies used for determining the direction of association (whether attracted or repelled) and the strength of association (the distinctiveness of ACs). Additionally, the table provides the expected frequencies for each adverbial complement (a and c), as well as the results of the distinctive-collexeme analysis, encompassing p-values and collostructional strength.

The findings reveal that the five most notable ACs are *praktycznie*, *ogólnie*, *generalnie*, *formalnie*, and *z grubsza*. Their distinctiveness is underscored by remarkably high log transformations: 315.84, 312.87, 168.03, 94.87, and 73.40, respectively. A comparison of observed versus expected frequencies for each AC and the two quasi-participial constructions reveals a pattern: these ACs occur more frequently than anticipated in the *AC-rzecz biorąc*-construction and less frequently than expected in the *AC-mówiąc*-construction. Essentially, they emerge as highly significant and markedly distinctive collexemes in the former in contrast to the latter. Notably, *praktycznie* stands out as the most robust

collexeme for the *AC-rzecz biorąc*-construction, as evidenced by its exceptionally high collostructional strength (315.84) calculated through the Fisher exact test. Additionally, the observed frequency surpasses the expected frequency in this particular pattern, indicating that this lexeme occurs more often than chance would suggest in the construction with the participle *biorąc* and less frequently than predicted in the construction with the participle *mówiąc*. These and the rest of ACs can be grouped into several categories based on their semantic and functional characteristics.

Table 2. The thirty most distinctive collexemes of the *AC-rzecz biorąc*-construction

rank	AC	a	c	e	f	x	y	z	b	d	(a)	(c)	p-values	coll. strength
1.	praktycznie	4	435	439	13175	10992	2622	13614	10988	2187	354.45	84.55	1.432667e-316	315.84
2.	ogólnie	179	634	813	12801	10992	2622	13614	10813	1988	656.42	156.58	1.3564540201e-313	312.87
3.	generalnie	32	379	411	13203	10992	2622	13614	10960	2243	331.84	79.16	9.368498053980195e-169	168.03
4.	formalnie	0	131	131	13483	10992	2622	13614	10992	2491	105.77	25.23	1.3392922637593726e-95	94.87
5.	z grubsza	6	114	120	13494	10992	2622	13614	10986	2508	96.89	23.11	4.0305653459959684e-74	73.40
6.	obiektywnie	4	66	70	13544	10992	2622	13614	10988	2556	56.52	13.48	1.2649806759708686e-42	41.90
7.	statystycznie	0	56	56	13558	10992	2622	13614	10992	2566	45.21	10.79	5.401642474734505e-41	40.27
8.	logicznie	0	55	55	13559	10992	2622	13614	10992	2567	44.41	10.59	2.8531698603270153e-40	39.55
9.	teoretycznie	0	54	54	13560	10992	2622	13614	10992	2568	43.60	10.40	1.5065803468084298e-39	38.82
10.	historycznie	0	43	43	13571	10992	2622	13614	10992	2579	34.72	8.28	1.313021220388576e-31	30.88
11.	ściśle	93	109	202	13412	10992	2622	13614	10899	2513	163.10	38.90	1.798745735564222e-28	27.75
12.	najogólniej	146	113	259	13355	10992	2622	13614	10846	2509	209.12	49.88	1.4404130600785588e-19	18.84
13.	realnie	1	17	18	13596	10992	2622	13614	10991	2605	14.53	3.47	9.757786018733967e-12	11.01
14.	technicznie	0	15	15	13599	10992	2622	13614	10992	2607	12.11	2.89	1.8014867127032362e-11	10.74
15.	zasadniczo	0	12	12	13602	10992	2622	13614	10992	2610	9.69	2.31	2.552239402645162e-9	8.59
16.	dokładnie	37	35	72	13542	10992	2622	13614	10955	2587	58.13	13.87	1.7133650125472063e-8	7.77
17.	przeciętnie	0	9	9	13605	10992	2622	13614	10992	2613	7.27	1.73	3.605801561807026e-7	6.44
18.	racjonalnie	0	8	8	13606	10992	2622	13614	10992	2614	6.46	1.54	0.0000018768376453894393	5.73
19.	ściślej	320	124	444	13170	10992	2622	13614	10672	2498	358.49	85.51	0.00000712311429918414	5.15
20.	psychologicznie	0	6	6	13608	10992	2622	13614	10992	2616	4.84	1.16	0.00005080119851919909	4.29
21.	średnio	0	6	6	13608	10992	2622	13614	10992	2616	4.84	1.16	0.00005080119851919909	4.29
22.	chronologicznie	0	5	5	13609	10992	2622	13614	10992	2617	4.04	0.96	0.00026417787949724883	3.58
23.	politycznie	0	5	5	13609	10992	2622	13614	10992	2617	4.04	0.96	0.00026417787949724883	3.58
24.	ogółem	1	5	6	13608	10992	2622	13614	10991	2617	4.84	1.16	0.0013310612843954767	2.88
25.	prawnie	0	4	4	13610	10992	2622	13614	10992	2618	3.23	0.77	0.0013733617035486198	2.86
26.	ekonomicznie	0	3	3	13611	10992	2622	13614	10992	2619	2.42	0.58	0.007137390663408783	2.15
27.	filozoficznie	0	3	3	13611	10992	2622	13614	10992	2619	2.42	0.58	0.007137390663408783	2.15
28.	fizjologicznie	0	3	3	13611	10992	2622	13614	10992	2619	2.42	0.58	0.007137390663408783	2.15
29.	merytorycznie	0	3	3	13611	10992	2622	13614	10992	2619	2.42	0.58	0.007137390663408783	2.15
30.	naukowo	0	3	3	13611	10992	2622	13614	10992	2619	2.42	0.58	0.007137390663408783	2.15

**a** = Observed frequency of AC (e.g. *praktycznie*) in ‘the *AC-mówiąc*-construction’; **b** = Frequency of all other ACs in ‘the *AC-mówiąc*-construction’; **c** = Observed frequency of ACs (e.g. *praktycznie*) in ‘the *AC-rzecz biorąc*-construction’; **d** = Frequency of all other ACs in ‘the *AC-rzecz biorąc*-construction’; **e** = Total frequency of AC (e.g. *praktycznie*); **f** = Total frequency of all other ACs; **x** = To-

tal frequency of ‘the *AC-mówiąc*-construction’;  $y$  = Total frequency of ‘the *AC-rzecz biorąc*-construction’;  $z$  = Total frequency of both constructions; **(a)** = Expected frequency of AC (e.g. *praktycznie*) in ‘the *AC-mówiąc*-AC-construction’; **(c)** = Expected frequency of AC (e.g. *praktycznie*) in ‘the *AC-rzecz biorąc*-construction’; **p-values** and **collostructional strength** = indexes of statistical significance

### 3.1.1. ACs denoting practical or theoretical reasoning

The first group includes ACs connected with practical or theoretical reasoning. *Praktycznie* suggests a pragmatic perspective on the topic under discussion, making the discussion more accessible and relevant to real-life situations, as in (1). *Realnie* ‘realistically’ in rank 13 also serves to inject a sense of practicality and reality into the discourse. *Teoretycznie* ‘theoretically’ in rank 9 signals a focus on abstract reasoning, suggesting that the upcoming discussion will be rooted in theoretical considerations rather than practical examples, as in (2).

- (1) Biedni ludzie w Afryce, *praktycznie rzecz biorąc*, nie mają żadnych pieniędzy. (Lapidaria IV-VI) ‘Poor people in Africa, *practically speaking*, have no money.’
- (2) *Teoretycznie rzecz biorąc*, o takich działaniach należy informować uczestników rynku. (Polityka) ‘*Theoretically speaking*, one should inform market participants about such activities.’

### 3.1.2. ACs introducing general statements

The second category encompasses ACs that are used to convey information in a general, all-encompassing manner, as exemplified in (3). Notably, *ogólnie* ‘generally’ and *generalnie* ‘generally’, ranked second and third, stand out as the most distinctive collexemes in this category. They are succeeded by *z grubsza* ‘roughly’, *najogólniej* ‘most generally’, and *ogółem* ‘in general’, securing ranks 5, 12, and 24, respectively.

- (3) *Ogólnie biorąc*, lubią swego prezydenta, w każdym razie szanują go. (Polityka) ‘*Generally speaking*, they like their president, or at least they respect him.’

### 3.1.3. ACs introducing domain-specific statements

The predominant category consists of ACs providing information in the context of a specialized field or domain, as in (4). Among the noteworthy lexemes in this group, *statystycznie* ‘statistically’ holds a significant position, securing the seventh rank. Other lexemes falling into this category include *logicznie* ‘logically’ at rank 8, *historycznie* ‘historically’ at rank 10, *technicznie* ‘technically’ at rank 14, *racjonalnie* ‘rationally’ at rank 18, *psychologicznie* ‘psychologically’ at rank 20, along with various others occupying subsequent ranks such as *politycznie* ‘politically’, *prawnie* ‘legally’, *ekonomicznie* ‘economically’, *filozoficznie* ‘philosophically’, *fizjologicznie* ‘physiologically’, and *naukowo* ‘scientifically’.

- (4) *Statystycznie rzecz biorąc*, pijemy 90 litrów piwa rocznie. (Gazeta Pomorska) ‘*Statistically speaking*, we consume 90 liters of beer per year.’

### 3.1.4. ACs introducing logical reasoning

*Logicznie* and *racjonalnie* can also be defined relative to the domain of reasoning or rationality. These adverbs are used to form opinions in a manner that adheres to principles of logic, reason, and objective analysis, as in (5). In the broader context of reasoning or rationality, *obiektywnie* (ranked sixth) can also be interpreted. This term signifies a commitment to presenting information objectively and analytically.

- (5) *Logicznie rzecz biorąc*, nie było powodu, aby rozszerzać Unię. (Ozon) ‘*Logically speaking*, there was no reason to expand the Union.’

### 3.1.5. ACs introducing precise statements

The next group encompasses ACs that emphasize a high degree of accuracy, specificity, and attention to detail. The foremost lexeme in this category is *ściśle*

‘strictly’, securing the eleventh rank. Following closely are *dokładnie* ‘accurately’ and *ściślej* ‘more strictly’ at the sixteenth and nineteenth ranks, respectively. These lexemes are used to provide precise and well-defined information, often within the context of established facts or rules, as in (6).

- (6) *Ściśle rzecz biorąc*, obszar ten jest dwukrotnie lub trzykrotnie większy. (Watykan) ‘*Precisely speaking*, this area is two or three times larger.’

### 3.1.6. ACs conveying a sense of the average situation

The ranking list also contains *przeciętnie* ‘on average’ in rank 17 and *średnio* ‘on average’ in rank 21. These adverbs are used to make statements that reflect the average or typical situation, outcome, or characteristic within a given context, as in (7).

- (7) Dzisiaj polskie przedsiębiorstwa, *przeciętnie rzecz biorąc*, są przygotowane o wiele lepiej do konkurencji na rynku europejskim niż to było w przeszłości[...] (Nowiny Raciborskie nr 16) ‘Today Polish companies, *on average*, are much better prepared for competition in the European market than they were in the past.’

### 3.1.7. Other ACs

In Table 2, we can also find adverbs like *formalnie* ‘formally’, *zasadniczo* ‘fundamentally’, *chronologicznie* ‘chronologically’, and *merytorycznie* ‘substantively’. Notably, *formalnie* occupies the fourth position, making it the most distinctive collxeme in this set. The others are positioned at 15, 22, and 29, respectively. These ACs convey different senses. *Formalnie* implies that the statement following it adheres strictly to formal rules, procedures, or official standards. *Zasadniczo* emphasizes a focus on the core or fundamental aspects of the matter under discussion. *Chronologicznie* guides the audience to interpret the information in the context of a timeline or a specific sequence of events. Lastly, *merytorycznie* suggests a focus on the substantive or meaningful aspects of the topic.

### 3.2. Findings for the *AC-mówiąc*-construction

The results for the *AC-mówiąc*-construction are presented in Table 3. It is noteworthy that *prawdę*, *szczerze*, *krótko*, *nawiasem*, and *inaczej* emerge as the most distinctive collexemes in the *AC-mówiąc*-construction. This distinctiveness is reflected in the remarkably high log transformations resulting from the Fisher exact test for these lexemes: 179.60, 179.10, 174.43, 148.65, and 112.42. Upon comparing the observed values with the expected ones, it is evident that these lexemes occur more frequently than expected by chance in the *AC-mówiąc*-construction, particularly when compared to the pattern involving the participle *biorąc*. Notably, these five lexemes collectively represent approximately 74.38% of all instances of this construction in the corpus, adding an interesting dimension to the analysis. This, in turn, suggests that they should be regarded not only as instances (i.e., constructs) of the *AC-mówiąc*-construction but also as integral constructions, involving pairings of form and meaning in their own right. These, along with other ACs strongly associated with this construction, can be organized into various semantic and functional categories.

Table 3. The thirty most distinctive collexemes of the *AC-mówiąc*-construction

rank	AC	a	c	e	f	x	y	z	b	d	(a)	(c)	p-values	coll. strength
1.	<i>prawdę</i>	1789	0	1789	11825	10992	2622	13614	9203	2622	1 444.45	344.55	2.5677246510968913e-180	179.60
2.	<i>szczerze</i>	1786	0	1786	11828	10992	2622	13614	9206	2622	1 442.02	343.98	8.126527385953148e-180	179.10
3.	<i>krótko</i>	1826	4	1830	11784	10992	2622	13614	9166	2618	1 477.55	352.45	3.673159663704041e-175	174.43
4.	<i>nawiasem</i>	1525	1	1526	12088	10992	2622	13614	9467	2621	1 232.10	293.90	2.2231126952408507e-149	148.65
5.	<i>inaczej</i>	1251	5	1256	12358	10992	2622	13614	9741	2617	1 014.10	241.90	3.7986504170148427e-113	112.42
6.	<i>delikatnie</i>	588	4	592	13022	10992	2622	13614	10404	2618	477.98	114.02	1.4424182682076942e-49	48.84
7.	<i>właściwie</i>	129	2	131	13483	10992	2622	13614	10863	2620	105.77	25.23	4.659300677134602e-10	9.33
8.	<i>między nami</i>	86	0	86	13528	10992	2622	13614	10906	2622	69.44	16.56	1.793341413634901e-8	7.75
9.	<i>najkrócej</i>	150	7	157	13457	10992	2622	13614	10842	2615	126.76	30.24	1.0861104758352385e-7	6.96
10.	<i>ogłędnie</i>	76	1	77	13537	10992	2622	13614	10916	2621	62.17	14.83	0.0000029851841977919546	5.53
11.	<i>łagodnie</i>	59	0	59	13555	10992	2622	13614	10933	2622	47.64	11.36	0.0000064177599741831764	5.19
12.	<i>najprościej</i>	97	7	104	13510	10992	2622	13614	10895	2615	83.97	20.03	0.0004274956448070303	3.37
13.	<i>brzydko</i>	38	0	38	13576	10992	2622	13614	10954	2622	30.68	7.32	0.0005889350894221652	3.23
14.	<i>najdelikatniej</i>	38	0	38	13576	10992	2622	13614	10954	2622	30.68	7.32	0.0005889350894221652	3.23
15.	<i>obrazowo</i>	25	0	25	13589	10992	2622	13614	10967	2622	20.19	4.81	0.008802664135314189	2.06
16.	<i>kolokwialnie</i>	20	0	20	13594	10992	2622	13614	10972	2622	16.15	3.85	0.021521600392520907	1.67
17.	<i>najogłędniej</i>	21	0	21	13593	10992	2622	13614	10971	2622	16.96	4.04	0.022366833869479196	1.65
18.	<i>potocznie</i>	17	0	17	13597	10992	2622	13614	10975	2622	13.73	3.27	0.057361747016375794	1.24
19.	<i>umownie</i>	23	1	24	13590	10992	2622	13614	10969	2621	19.38	4.62	0.06830465913219758	1.17
20.	<i>poważnie</i>	38	3	41	13573	10992	2622	13614	10954	2619	33.10	7.90	0.070804522498863	1.15
21.	<i>nieładnie</i>	15	0	15	13599	10992	2622	13614	10977	2622	12.11	2.89	0.092209087045376	1.04
22.	<i>w skrócie</i>	20	1	21	13593	10992	2622	13614	10972	2621	16.96	4.04	0.10214325759294991	0.99
23.	<i>najłagodniej</i>	11	0	11	13603	10992	2622	13614	10981	2622	8.88	2.12	0.1384665078361616	0.86

24.	lekko	11	0	11	13603	10992	2622	13614	10981	2622	8.88	2.12	0.1384665078361616	0.86
25.	eufemistycznie	12	0	12	13602	10992	2622	13614	10980	2622	9.69	2.31	0.13970466676568294	0.85
26.	po ludzku	12	0	12	13602	10992	2622	13614	10980	2622	9.69	2.31	0.13970466676568294	0.85
27.	popularnie	12	0	12	13602	10992	2622	13614	10980	2622	9.69	2.31	0.13970466676568294	0.85
28.	brutalnie	13	0	13	13601	10992	2622	13614	10979	2622	10.50	2.50	0.14860170836734427	0.83
29.	prościej	13	0	13	13601	10992	2622	13614	10979	2622	10.50	2.50	0.14860170836734427	0.83
30.	uczciwie	47	6	53	13561	10992	2622	13614	10945	2616	42.79	10.21	0.16414763917613628	0.78

### 3.2.1. ACs expressing veracity

The ACs *prawdę* ‘truth’, *szczerze* ‘sincerely’, and *uczciwie* ‘honestly’ are related to veracity. Notably, *prawdę* emerges as the most significant lexeme in this category and holds the distinguished position of the leading collexeme, ranking first in the list. Following closely is *szczerze*, securing the second rank, while *uczciwie* occupies the 30th position. In the discourse, as in (8), their function is pivotal—they establish a direct and forthright tone, fostering a sense of transparency and authenticity in communication. Furthermore, they play a crucial role in building trust between the speaker and the audience, signifying a steadfast commitment to honesty.

(8) *Prawdę mówiąc*, to nic lepszego nie mogło mi się przydarzyć. (Wolna Trybuna) ‘*To tell the truth*, nothing better could have happened to me.’

### 3.2.2. ACs expressing conciseness

The second category is represented by ACs reflecting conciseness. The primary lexeme in this semantic category is *krótko* ‘briefly’, which holds the third position. Additionally, two other lexemes, *najkrócej* ‘most briefly’ and *w skrócie* ‘in short’, represent this category, with the former holding the 9th position and the latter securing the 22nd position. In discourse, they serve to signal a speaker’s intent to provide a condensed summary, ensuring clear, efficient communication, maintaining focus on key points, and facilitating transitions to succinct portions of the discourse, as in (9).

(9) *Krótko mówiąc*: to dzięki niemu obaj żyjemy. (Narrenturm) ‘*In short*: thanks to him, we both are alive.’

### 3.2.3. *ACs introducing an aside*

The term *nawiasem*, situated in the fourth position, signifies ‘by the way’, whereas *między nami*, denoting ‘between us’, holds the eighth position. Both lexemes are used to interject an additional remark or comment that deviates from the main topic, often adopting a more casual or informal tone, as in (10). The phrase *między nami mówiąc* also serves to signal that the information following it is meant to be shared in confidence between the speaker and the listener, as in (11).

(10) *Dotyczy to, nawiasem mówiąc, także twego obecnego położenia. (Bajki robotów)* ‘It concerns, *by the way*, your current situation.’

(11) *Między nami mówiąc: wątpię, żeby im się udało. (Marsz Polonia)* ‘*Between us*, I doubt they will succeed.’

### 3.2.4. *ACs introducing a rephrasing*

The following lexeme in the ranking is *inaczej*, meaning ‘differently’, which holds the fifth position. A comparable meaning is expressed by *innymi słowy* (coll. strength: 0.43), translated as ‘in other words’, positioned at rank 36. Both lexemes serve the purpose of presenting an alternative or clearer expression of the preceding idea, as in (12).

(12) *Inaczej mówiąc, bez cechy planowości nie ma budżetu. (Finanse publiczne)* ‘*In other words*, without the characteristic of planning, there is no budget.’

### 3.2.5. *ACs introducing an understatement*

A different semantic category consists of ACs that introduce milder and more cautious language to diminish the significance or seriousness of a situation, as in (13).

(13) *Delikatnie mówiąc, mija się z prawdą i nie wspomina o kosztach. (Trybuna)* ‘*To put it mildly*, it misses the truth and doesn’t mention the costs.’

Ranked sixth in significance, *delikatnie*, meaning ‘delicately’, stands out as the most prominent lexeme within this category. It is accompanied by *ogłędnie* ‘cautiously’, *łagodnie* ‘mildly’, *najdelikatniej* ‘most delicately’, *najogłędniej* ‘most cautiously’, *najłagodniej* ‘most mildly’, and *lekko* ‘lightly’, positioned in ranks 10, 11, 14, 17, 23, and 24, respectively. Additionally, the term *eufemistycznie* ‘euphemistically’ at rank 25 can be categorized within this group, as it involves the use of softer expressions to convey potentially sensitive topics with enhanced tact and subtlety.

### 3.2.6. ACs expressing precision and clarity

The next category encompasses ACs introducing precision and clarity into the discourse to ensure that the intended message is easily understood by the audience, as in (14). *Właściwie* ‘properly’, which occupies the seventh position, is the most significant lexeme of this category. It is followed by *najprościej* ‘most simply’ and *prościej* ‘more simply’, which are ranked at 12th and 29th, respectively.

- (14) *Właściwie mówiąc*, od dawna chciałem pana poznać. (Kiwony) ‘*Properly speaking*, I’ve wanted to meet you for a while.’

### 3.2.7. ACs conveying negativity and harshness

In Table 5, we also find the adverbs *brzydko* ‘ugly’, *brutalnie* ‘brutally’, and *nieładnie* ‘rudely’, ranked 13, 21, and 28, respectively. These expressions convey a blunt and straightforward assessment, emphasizing the negative aspects of the message being communicated. They all contribute to a tone that is harsh, critical, and not softened in any way, as in (15).

- (15) Po drugie, mają na niego, *brzydko mówiąc*, haka. (Nakielski Czas) ‘*Secondly*, they, *to put it crudely*, have dirt on him.’

### 3.2.8. *ACs introducing informality and everyday language*

The following set encompasses ACs associated with informality and everyday language. Among these ACs, *kolokwialnie* ‘colloquially’ holds the highest significance, securing the 16th position. It is succeeded by *potocznie* ‘informally’, *po ludzku* ‘humanly’, and *popularnie* ‘popularly’, which are ranked at 18, 26, and 27, respectively. These ACs are used to convey information in a relaxed, familiar, and approachable manner, adopting a style characterized by informality and accessibility frequently observed in everyday interactions, as in (16).

- (16) *Po ludzku mówiąc*, im człowiek starszy, tym chętniej poprawia i uzupełnia, co raz napisał. (Polityka nr 2514) ‘*Speaking humanly*, as a person gets older, they are more inclined to revise and add to what they’ve written.’

### 3.2.9. *ACs signaling the style of communication*

Moreover, the list encompasses two lexemes related to the style of communication: *obrazowo* ‘pictorially’ at rank 15 and *umownie* ‘conventionally’ at rank 19. These ACs suggest that the speaker is either employing figurative language to convey a non-literal meaning (as in 17) or using conventional terms in a customary or agreed-upon manner.

- (17) *Obrazowo mówiąc*, płyny te odklejają bakterie, uniemożliwiając im kolonizację, a zatem i atak. (Polityka) ‘*Figuratively speaking*, these liquids detach bacteria, preventing their colonization, and thus, an attack.’

### 3.2.10. *ACs denoting seriousness*

Ultimately, the term *poważnie* ‘seriously’, positioned at rank 20, is used to signal that the information being communicated is to be taken seriously, without levity or frivolity, as in (18).

- (18) Ale *poważnie mówiąc* nie bardzo wiem co to jest kino kultowe. (Trybuna Śląska) ‘But *seriously*, I don’t really know what a cult film is.’

### 3.3. Distribution across registers

The distribution of the *AC rzecz biorąc*-construction and the *AC-mówiąc*-construction follows a similar pattern across various registers in NKJP. In Tables 4 and 5, the observed frequencies show the occurrence of the top twenty ACs collocating with the participles *biorąc* and *mówiąc* within six chosen registers: journalism, literature, spoken conversation, quasi-spoken texts (including parliamentary transcripts), Internet sources (forums, chatrooms, mailing lists, etc.), and academic writing and scientific-didactic textbooks. A closer examination of these frequencies reveals that these constructions are predominantly observed in journalism, are consistently present in quasi-spoken texts, tend to occur quite frequently in literature, are relatively rare in Internet sources and scientific-didactic texts, and are exceptionally infrequent in conversational contexts.

#### 3.3.1. The distribution patterns of the *AC-rzecz biorąc*-construction

The distribution patterns in NKJP, as evident from both tables, reveal notable disparities in the usage of both quasi-participle constructions. For example, in the *AC-rzecz biorąc*- construction, combinations involving terms like *ogólnie* ‘generally’, *z grubsza* ‘roughly’, *ściśle* ‘strictly’, *obiektywnie* ‘objectively’, *statystycznie* ‘statistically’, *logicznie* ‘logically’, *historycznie* ‘historically’, *dokładnie* ‘precisely’, and *realnie* ‘realistically’ are more prevalent in journalism compared to literature or quasi-spoken texts. On the other hand, *praktycznie* ‘practically’, *generalnie* ‘generally’, *formalnie* ‘formally’, *ściślej* ‘strictly’, *najogólniej* ‘most generally’, or *teoretycznie* ‘theoretically’ occur more frequently in quasi-spoken discourse than in journalism and literature. Terms such as *ogólnie*, *generalnie*, *ściślej*, *najogólniej*, and *ściśle* manifest more frequently in scientific and didactic texts than in literature and Internet sources. Additionally, *praktycznie*, *formalnie*, *z grubsza*, *obiektywnie*, *logicznie*, and *teoretycznie* are more frequent in literature than in Internet and scientific-didactic sources.

Table 4. The distribution of the twenty ACs of the *AC-rzecz biorąc*-construction across six registers

Note: <b>jour</b> = journalism; <b>qu-sp</b> = quasi-spoken; <b>lit</b> = literature; <b>Net-int</b> = Internet-interactive; <b>sc-di</b> = scientific-didactic; <b>conv</b> = conversation							
rank	ACs	jour	qu-sp	lit	Net-int	sc-di	conv
1.	ogólnie	172	64	37	37	39	4
2.	praktycznie	37	254	47	0	3	2
3.	generalnie	59	226	4	8	9	1
4.	formalnie	34	35	9	2	3	0
5.	ściślej	23	25	5	1	10	0
6.	z grubsza	24	5	6	2	3	0
7.	najogólniej	25	40	3	1	5	1
8.	ściśle	15	9	7	2	9	0
9.	obiektywnie	19	8	14	0	4	2
10.	statystycznie	28	4	2	4	1	0
11.	logicznie	17	6	9	0	4	2
12.	teoretycznie	13	15	5	2	1	0
13.	historycznie	20	5	2	0	3	0
14.	dokładnie	14	8	3	0	1	0
15.	realnie	6	4	1	0	0	0
16.	technicznie	2	1	5	2	0	0
17.	zasadniczo	5	5	0	0	0	0
18.	konkretnie	6	5	0	0	0	0
19.	przeciętnie	2	3	0	0	0	0
20.	racjonalnie	3	1	2	0	0	0

### 3.3.2. The distribution patterns of the *AC-mówiąc*-construction

Concerning the *AC-mówiąc*-construction, distinctive usage patterns also emerge across various registers. Combinations with *krótko* ‘shortly’, *szczerze* ‘honestly’, *nawiasem* ‘incidentally’, *inaczej* ‘differently’, *delikatnie* ‘delicately/gently’, and various others exhibit a higher frequency in journalism and quasi-spoken discourse compared to literature. Conversely, combinations featuring *prawdę* ‘truth’, *właściwie* ‘actually’, *ściśle* ‘strictly’, and *między nami* ‘between us’ display notably greater frequency in literature than in other linguistic registers. Additionally, terms like *prawdę*, *szczerze*, *nawiasem*, and *delikatnie* tend to

occur more frequently in Internet sources than in scholarly and conversational contexts. Finally, configurations involving *krótko*, *inaczej*, *ściślej* ‘more strictly’, *dokładniej* ‘more accurately’, *najogólniej* ‘most generally’, or *najprościej* ‘most simply’ are relatively more common in scientific discourse compared to Internet sources and conversational language.

Table 5. The distribution of the twenty ACs of the *AC-mówiąc*-construction across six registers

Note: <b>jour</b> = journalism; <b>qu-sp</b> = quasi-spoken; <b>lit</b> = literature; <b>Net-int</b> = Internet-interactive; <b>sc-di</b> = scientific-didactic; <b>conv</b> = conversation							
rank	ACs	jour	qu-sp	lit	Net-int	sc-di	conv
1.	krótko	614	534	244	44	67	7
2.	prawdę	591	187	655	63	20	2
3.	szczerze	684	295	228	334	8	58
4.	nawiasem	635	232	207	86	28	3
5.	inaczej	452	301	84	28	163	1
6.	delikatnie	329	135	46	34	1	0
7.	ściślej	114	59	45	1	28	0
8.	ogólnie	78	38	11	20	11	3
9.	dokładniej	35	19	15	5	8	0
10.	najkrócej	85	30	3	11	5	0
11.	najogólniej	78	31	1	4	15	0
12.	właściwie	5	1	66	0	0	0
13.	najprościej	60	9	6	1	7	0
14.	ściśle	19	15	25	1	5	0
15.	między nami	13	8	47	0	0	1
16.	ogłędnie	46	11	9	2	4	0
17.	łagodnie	33	7	4	1	1	0
18.	uczciwie	13	13	13	1	1	0
19.	brzydko	13	14	1	4	1	4
20.	najdelikatniej	20	12	4	0	0	0

The distinctive preferences observed across different linguistic registers suggest that speakers strategically choose these constructions based on the specific communicative goals associated with each register. For example, the selection of certain adverbial complements in the *AC-rzecz*

*biorąc*-construction, such as those related to generality or objectivity, may correspond to the communicative style and requirements of journalism, while adverbial complements emphasizing practicality or informality may be more suited for quasi-spoken discourse. Similarly, in the case of the *AC-mówiąc*-construction, the varying frequencies of specific adverbial complements across registers indicate that speakers tailor their language choices to the communicative needs of each context. Adverbial complements emphasizing brevity or honesty may correspond with the journalistic style, while those conveying precision or formality might be more prevalent in scientific discourse. This implies that speakers deliberately consider discourse functions and communicative intentions when selecting these constructions in different registers.

#### 4. Concluding remarks

This research contributes significantly to our understanding of the nuanced usage and functions of the *AC-rzecz biorąc*-construction and the *AC-mówiąc*-construction in Polish, shedding new light on their status, patterns of use, functions in discourse, and semantic constraints. It highlights the fact that each construction carries specific semantic implications and constraints that influence the choice of ACs they co-occur with. The *AC-rzecz biorąc*-construction, for instance, does not collocate with certain ACs that convey truthfulness or informality, and it is used less frequently with ACs suggesting brevity or conventionality. Conversely, it is more commonly associated with ACs that emphasize practical or theoretical considerations, general overviews, domain-specific statements, logical or objective descriptions, precision, typical scenarios, formal declarations, and essential aspects of a subject. These tendencies underscore the construction's function in facilitating a focus on practical or theoretical reasoning, domain-specific information, or a presentation of general information in discourse.

In contrast, the *AC-mówiąc*-construction demonstrates the clear avoidance of ACs typically associated with formality, statistics, logic, theory, history, technicality, and norms. Instead, it has a marked preference for ACs that convey

truthfulness, sincerity, confidentiality, gentleness, negativity, informality, everyday speech, and simplicity. These preferences suggest that the *AC-mówiąc*-construction is typically employed to express truth and honesty, provide concise summaries, introduce confidential asides, rephrase content, offer understated remarks, articulate clear and precise statements, and convey both negative and informal tones. The construction's choices in ACs reflect its suitability for conveying personal insights, emphasizing straightforwardness, and introducing an informal or colloquial style to the discourse.

The contrasting preferences of these two constructions highlight their distinct functions and the importance of context in their use. The *AC-rzecz biorąc*-construction is suited to contexts requiring objective, generalized, or theoretical perspectives, while the *AC-mówiąc*-construction is more appropriate for subjective, truthful, or informal expressions. This distinction is crucial for understanding the subtleties of Polish language discourse, as the choice between these constructions can significantly alter the conveyed meaning. The research demonstrates that the subtle semantic differences are not arbitrary but are systematically related to the different discourse functions that each construction performs. This insight is valuable for both linguistic theory and practical applications, such as language teaching, translation, and natural language processing.

Finally, although the analysis in section 3 indicates that both constructions impose specific semantic constraints on the ACs they co-occur with, thereby serving various functions in discourse, the data reveals an overlap in some adverbial complements (ACs) shared by both the *AC rzecz biorąc*-construction and the *AC mówiąc*-construction. Notably, terms such as *ogólnie* 'generally', *generalnie* 'generally', *ściśle* 'strictly', and *dokładnie* 'precisely' appear in both constructions, although their frequency and significance may differ. For example, *ogólnie* and *generalnie* are used in both constructions to introduce general statements, indicating that these terms have a flexible role in the language, applicable to various discourse functions. Similarly, *ściśle* and *dokładnie* serve to introduce precise and accurate statements in both constructions, although they might be more prominent in one construction over the other.

The presence of these overlapping ACs suggests that while each construction has its distinctive set of preferred collocates, there are certain adverbials that are versatile enough to co-occur with both constructions, potentially fulfilling

similar communicative functions across different contexts. This overlap also highlights the fluidity of these constructions in accommodating a range of adverbial complements, depending on the specific discourse needs.

## References

- BIBER, DOUGLAS; CONRAD, SUSAN; REPPEN, RANDI. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press. New York.
- BIRZER, SANDRA. 2015. Generally speaking, connectivity and conversation management combined: The functions of Russian *voobščę govorja* and Polish *ogólnie mówiąc*. *Russian linguistics* 39/1. 81–115. doi.org/10.1007/s11185-014-9143-0.
- BIRZER, SANDRA. 2017. Assessing the role of pattern and matter replication in the development of Polish discourse structuring elements based on non-finite ‘*verba dicendi*’. *Zeitschrift für Slavische Philologie* 73/1. 159–185. doi.org/10.20378/irb-50392.
- BOJAŁKOWSKA, KRYSZYNA. 2010. *Opis składniowy imiesłowów przysłówkowych w współczesnym języku polskim*. Wydawnictwo Naukowe UMK. Toruń.
- BOGUSŁAWSKI, ANDRZEJ; DANIELEWICZOWA, MAGDALENA. 2005. *Verba polona abscondita. Sonda słownikowa III*. UW. Warszawa.
- BYBEE, JOAN. 2010. *Language, Usage, and Cognition*. Cambridge University Press. Cambridge.
- COLLEMAN, TIMOTHY. 2009. Verb disposition in argument structure alternations. A corpus study of the Dutch dative alternation. *Language Sciences* 31. 593–611. doi.org/10.1016/j.langsci.2008.01.001.
- DESAGULIER, GUILLAUME. 2014. Visualizing distances in a set of near-synonyms: Rather, quite, fairly, and pretty. *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*. Ed. Glynn, Dylan; Robinson, Justyna A. John Benjamins. Amsterdam – Philadelphia. 145–178.
- DIESEL, HOLGER. 2015. Usage-based construction grammar. *Handbook of Cognitive Linguistics*. Ed. Dąbrowska, Ewa; Divjak, Dagmar. Mouton de Gruyter. Berlin. 295–321.
- GOLDBERG, ADELE. 2006. *Constructions at Work. The Nature of Generalization in Language*. Oxford University Press. Oxford.
- GOLDBERG, ADELE. 2013. Constructionist approaches to language. *Handbook of Construction Grammar*. Ed. Hoffmann, Thomas; Trousdale, Graeme. Oxford University Press. Oxford. 15–48.

- GRIES, STEFAN TH.; STEFANOWITSCH, ANATOL. 2004a. Extending collocation analysis: A corpus-based perspective on alternations. *International Journal of Corpus Linguistics* 9/1. 97–129. doi.org/10.1075/ijcl.9.1.06gri.
- GRIES, STEFAN TH.; STEFANOWITSCH, ANATOL. 2004b. Co-varying collexemes in the into-causative. *Language, Culture, and Mind*. Ed. Achard, Michel; Kemmer, Suzanne. CSLI. Stanford, CA. 225–236.
- GROCHOWSKI, MACIEJ. 2002. Wielowyrazowe jednostki funkcyjne. Wprowadzenia do problematyki. *Problemy frazeologii europejskiej V*. Ed. Lewicki, Andrzej M. Norbertinum. Lublin. 35–50.
- KISIEL, ANNA; ŻABOWSKA, MAGDALENA. 2011. O zakresie klasy partykuł w Słowniku gniazdowym partykuł polskich. *Polonica* 31. 113–132.
- KLESZCZOWA, KRYSZYNA. 2015. *U źródeł polskich partykuł. Derywacja funkcjonalna, przemiany, zaniki*. Wydawnictwo Uniwersytetu Śląskiego. Katowice.
- KUBICKA, EMILIA. 2017. Jak mówimy jakoś mówiąc? Formalne i semantyczne właściwości adverbialnych uzupełnień quasi-imiesłowowego mówiąc. *LingVaria* 12/23. 99–113. doi.org/10.12797/LV.12.2017.23.07.
- LANGACKER, RONALD W. 1987. *Foundations of Cognitive Grammar: Vol. 1. Theoretical Prerequisites*. Stanford University Press. Stanford, CA.
- LEVSHINA, NATALIA. 2015. *How to Do Linguistics with R: Data Exploration and Statistical Analysis*. John Benjamins. Amsterdam.
- MOROZ, ANDRZEJ. 2010. *Parenteza ze składnikiem czasownikowym we współczesnym języku polskim*. Wydawnictwo Naukowe UMK. Toruń.
- SCHMID, HANS-JÖRG. 2017. *Entrenchment and the Psychology of Language Learning: How We Reorganize and Adapt Linguistic Knowledge*. De Gruyter Mouton. Berlin – Boston.
- SHEKIN, DAVID J. 2011. *Handbook of Parametric and Non-parametric Statistical Procedures*. 5th Edition. Chapman AND Hall/CRC. London.
- STEFANOWITSCH, ANATOL. 2013. Collocation analysis. *Handbook of Construction Grammar*. Ed. Hoffmann, Thomas; Trousdale, Graeme. Oxford University Press. Oxford. 290–306.
- STEFANOWITSCH, ANATOL; GRIES, STEFAN TH. 2003. Collocations: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8. 209–243. doi.org/10.1075/ijcl.8.2.03ste.
- STEFANOWITSCH, ANATOL; GRIES, STEFAN TH. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1/1. 1–43. doi.org/10.1515/cllt.2005.1.1.1.

STĘPIEŃ, MARZENA. 2014. *Wyrażenia parentetyczne w strukturze wypowiedzi – właściwości semantyczne, składniowe, prozodyczne*. Wydział Polonistyki Uniwersytetu Warszawskiego, BEL Studio. Warszawa.

STĘPIEŃ, MARZENA. 2015. To samo czy inne? Właściwości prozodyczne tzw. wyrażen quasi-imiesłowowych. *Sens i brzmienie [Prace językoznawcze Instytutu Filologii Polskiej UKSW, 7]*. Ed. Danielewiczowa, Magdalena; Bilińska, Joanna; Doboszyńska-Markiewicz, Katarzyna; Zaucha, Joanna. Wydawnictwo Uniwersytetu Kardynała Stefana Wyszyńskiego. Warszawa. 145–169.

STUBBS, MICHAEL. 2001. *Words and Phrases: Corpus Studies in Lexical Semantics*. Blackwell Publishers. Massachusetts.

WEISS, DANIEL. 2005. Nowe przyimki o pochodzeniu imiesłowowym? *Przysłówki i przyimki. Studia ze składni i semantyki języka polskiego*. Ed. Grochowski, Maciej. Wydawnictwo Naukowe UMK. Toruń. 177–207.

WILIŃSKI, JAROSŁAW. 2017. On the brink of-noun vs. On the verge of-noun: A distinctive-collexeme analysis. *Research in Language* 15/4. 425–443. doi.org/10.1515/rela-2017-0024.

WILIŃSKI, JAROSŁAW. 2023. A sea of-noun vs. a mountain of-noun: A quantitative corpus-based study of two metaphorical constructions. *Roczniki Humanistyczne* 71/6. 323–337. <https://doi.org/10.18290/rh23716.20>.

WRÓBEL, HENRYK. 1975. *Składnia imiesłowów czynnych we współczesnej polszczyźnie*. Spółka Wydawnicza OD NOWA. Katowice.

WULFF, STEFANIE. 2006. Go-V vs. go-and-V in English: A case of constructional Synonymy? *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*. Ed. Gries, Stefan Th.; Stefanowitsch, Anatol. Mouton de Gruyter. Berlin – New York. 101–126.

ŻABOWSKA, MAGDALENA. 2009. Wyrażenia metatekstowe w funkcji parentezy. *Język z różnych stron widziany*. Ed. Skarżyński, Mirosław; Czelakowska, Anna. Wydział Polonistyki UJ. Kraków. 157–166.

ŻABOWSKA, MAGDALENA. 2020. Szyk wyrażen jako cecha dystynktywna w polu komentarzy metatekstowych: [Adv] mówiąc, \_ vs. mówiąc [Adv], \_ . *Poradnik Językowy* 9. 21–37. doi.org/10.33896/PorJ.2020.9.2.

## Data sources and tools

*The National Corpus of Polish*. Available at: [http://www.nkjp.uni.lodz.pl/index\\_adv.jsp](http://www.nkjp.uni.lodz.pl/index_adv.jsp). [NKJP]

R Core Team (2022). *R: A language and environment for statistical computing*. Available at: <https://www.R-project.org>.

## ***AC-rzecz biorąc tj. AC-mówiąc na poljskom: Korpusno istraživanje dviju kvaziparticipskih konstrukcija***

### *Sažetak*

Ova studija, koristeći se perspektivom konstrukcijske gramatike i metodologiju korpusne analize nazvanom *Distinctive-Collexeme Analysis*, nastoji identificirati priloške dopune (AC) koje pokazuju sklonost prema konstrukciji *AC-rzecz biorąc* ‘AC-stvar uzimajući’ (npr. ‘povijesno gledano’) ili konstrukciji *AC-mówiąc* ‘AC-govoreći’ (npr. ‘općenito govoreći’) u poljskome jeziku. Rezultati utemeljeni na podatcima iz Nacionalnoga korpusa poljskog jezika (NKJP) otkrivaju da se te konstrukcije često pojavljuju s određenim priloškim dopunama koje imaju različita značenja i funkcije u diskursu. Studija također pokazuje da se te konstrukcije najčešće pojavljuju u novinarstvu, kvazigovornim tekstovima i literaturi, dok su rjeđe prisutne u internetskim izvorima, znanstveno-didaktičkim tekstovima i razgovornim kontekstima.

**Keywords:** quasi-participle construction, NKJP, construction grammar, distinctive collexeme analysis, Fisher exact test

**Ključne riječi:** kvaziparticipska konstrukcija, NKJP, gramatika konstrukcije, analiza distinktivnih koleksema, Fisherov egzaktni test