

The Competition of ARB Constructions in Polish*

Maria Katarzyna Prenner and Daniel Bunčić

Abstract: Polish has three quasi-synonymous impersonal constructions: the *-no/-to* construction, the reflexive impersonal, and the 3PL impersonal. This raises the question of what the differences between them are and how one of them is selected. This paper presents the results of an acceptability judgment test, which is informed by an explorative corpus study and examines the following factors: colloquial vs. neutral register; perfective vs. imperfective aspect; present tense vs. preterite; and generic vs. specific reading. The main findings are that the 3PL impersonal turned out to be better in colloquial discourse and the reflexive impersonal is much more acceptable with the imperfective than with the perfective aspect. Furthermore, the corpus data and the acceptability judgments show numerically that both reflexive and 3PL impersonal have a certain tendency towards present tense and generic reading, and that the *-no/-to* construction is more typical with the perfective than the imperfective aspect.

1. Introduction

The Polish language has several impersonal constructions. This paper is concerned with three of them: the *-no/-to* construction (Polish *bezosobnik*, cf. e.g., Fellerer 2008) as in (1), the reflexive impersonal (*się* construction) as in (2), and the impersonal third-person plural as in (3).

- (1) Wrócono do swoich domów.
return_{PST,IMPRS}¹ to one's houses
'People returned to their homes.' (Wiemer 1995: 314)

* Research for this article was funded by the German Research Foundation (DFG), project ID 281511265, SFB 1252 "Prominence in Language".

¹ The following abbreviations, which follow the Leipzig Glossing Rules where possible, are used in the glosses and diagrams: 1 'first person'; 2 'second person'; 3 'third person'; ACC 'accusative'; AUGM 'augmentative'; COND 'conditional'; F 'feminine'; FUT 'future'; GEN 'genitive'; IMPRS 'impersonal'; INSTR 'instrumental case'; M 'masculine'; N 'neuter'; NEG 'negative particle'; NVIR 'non-virile (= not masculine-human)'; PL 'plural'; PRS 'present'; PST 'past, preterite'; REFL 'reflexive (pronoun)'; SG 'singular'; VIR 'virile (= masculine-human)'.

- (2) Czyta **się** książkę.
 read_{PRS.3SG} REFL book_{ACC}
 ‘One reads a book.’ (Wiese 1973: 625)
- (3) Znowu **podnieśli** cenę paliwa.
 again raise_{PST.3PL.VIR} price_{ACC} fuel_{GEN}
 ‘They have raised the price of fuel again.’ (Kibort 2008: 263)

Gast and van der Auwera (2013: 123) subsume these expressions under the category of “human impersonal pronouns” in the wider sense. More specifically, all three constructions belong to a category of impersonals called “*arb* constructions” (Cinque 1988: 544) or “arbs” and defined as “constructions with arbitrary interpretations” (Malamud 2013: 1).² (This definition excludes some of the “human impersonal pronouns”, see (8–11) on p. 204) Although these constructions and their arbitrary nature have been extensively investigated (cf. e.g., Cabredo Hofherr 2003; Mel’čuk 1974; Pađučeva 2012), the proposed analyses are diverse, and the semantics of ARBs is still far from being thoroughly understood (Malamud 2013: 2).

The quasi-synonymity of these three ARB constructions raises the obvious question of the choice of expression: When does a Polish native speaker choose which of these constructions? The extant literature on these expressions does not give sufficient answers to this question. Therefore, this paper attempts to clarify the differences between the three Polish ARBs on the basis of a corpus analysis and an acceptability judgment test. Specifically, we will investigate the effects of register, tense, aspect, and generic vs. specific reading of the implicit subject.

The problem addressed here arose in the context of a project within the Collaborative Research Center 1252 at Cologne University, “Prominence in Language”. This project is based on the finding that ARBs are not equally good with all verbs, which can be explained by assuming a prominence relation in the sense of Himmelmann and Primus 2015, according to which the constructions are deemed more acceptable the more agentive the verbs are (cf. Bunčić 2018, 2019, 2020). However, in order to be able to determine this effect (and in very practical terms, to construct valid test items), we have to know the influence of other factors on the grammaticality and/or acceptability (cf. Haider 2019) of these constructions.

² Note that Cinque (1988: 529) used “arb” for a certain class of meanings and “*arb* constructions” (ibid. 544) as a cover term for constructions that can have these meanings (in this case, the Italian *si* and 3PL impersonal), whereas we follow Malamud (2013: 1) in calling the constructions themselves ARBs and using the terminology introduced in (4–7) for the meanings these constructions can have (cf. also Egerland 2003: 76, fn. 4 on differences in terminology).

In the following, a short definition of the category of ARBs (§2) will be given. On the basis of what is known so far about Polish ARBs, we can then make predictions about the possible factors influencing the choice of construction (§3). The predictions will be tested in an explorative corpus study (§4) and an acceptability judgment test (§5), the results of which will be discussed in detail (§6).

2. Definition of ARBs

All three constructions have a demoted subject/agent (and are thus impersonal constructions in the sense of Siewierska 2008a: 116) with reduced referentiality, which is why Malchukov and Ogawa (2011: 44) classify them as “R-impersonals”—a category that, however, also includes sentences like *It dawn*s (ibid. 25). Gast and van der Auwera (2013: 124) describe this reduction of referentiality as “impersonalization”, by which they mean “the process of filling an argument position of a predicate with a variable ranging over sets of human participants without establishing a referential link to any entity from the universe of discourse”. A number of different meanings can be associated with such expressions, which have been classified with varying degrees of granularity; see (4–8):

- (4) a. *neopredelënno-ličnye predloženiya* ‘indefinite-personal sentences’
 b. *obobščënno-ličnye predloženiya* ‘generalized-personal sentences’
 (Vinogradov 1954: 5–12)³
- (5) a. quasi-existential
 b. quasi-universal (Cinque 1988: 545)
- (6) a. specific (“1st person singular, ‘I’”)
 b. arbitrary (“a non-specific group of individuals”)
 c. generic (“a quasi-universal set of individuals”)
 (Egerland 2003: 76)
- (7) a. specific existential reading (temporally anchored)
 b. vague existential reading (not temporally anchored)
 c. inferred existential reading (inferred from a result)
 d. corporate reading (predicates with a designated subject)
 e. universal reading (licensed by a locative)
 (Cabredo Hofherr 2003: 83)

³ Note, however, that apart from this semantic definition, these terms are also used to distinguish different forms. Thus, Padučeva (2012: 27) uses the term *neopredelënno-ličnye predloženiya* to refer to the Russian 3PL impersonal regardless of its meaning.

- (8) a. episodic; existential, vague
 b. episodic; existential, plural, indefinite
 c. episodic; existential, plural, definite
 d. generic; universal, external
 e. generic; universal, internal
 f. modal; universal, internal
 g. non-assertive; universal, internal

(Gast and van der Auwera 2013: 140f)

However, the “universal”/“generic” readings can easily be achieved with many other expressions as well (cf. Bauer this volume.), e.g., with *człowiek* ‘man, human being’ as in (9), with *ludzie* ‘people’ as in (10), with the second person as in (11), or with the first-person plural as in (12).⁴

- (9) W takim domiszczu **człowiek** spodziewa się
 in such house_{AUGM} man expect REFL
 większych rzeczy.
 greater things

‘In such a huge house one expects greater things.’

(ParaSol; Polish original)

- (10) **Ludzie** nigdy nie dowiedzieli się, dokąd.
 people never NEG found.out REFL whither

‘Nobody ever found out where to.’

(ParaSol; translation from German)

- (11) Prawdy nie **dojdiesz**, **choćbyś** i tam był.
 truth NEG reach_{FUT.2SG} even.if_{2SG} also there be_{PST}

‘One will not learn the truth even if one has been there.’

(ParaSol; Polish original)

- (12) wszystko, **cośmy** raz przeżyli, miałoby się
 all that_{1PL} once experienced must.COND REFL

kiedyś powtórzyć
 someday repeat

‘everything one once experienced would have to recur someday’

(ParaSol; translation from Czech)

⁴ Examples (9–12) were retrieved from the ParaSol corpus, a parallel corpus of Slavic and other languages (von Waldenfels and Meyer 2006–).

None of the expressions in (9–12) can be interpreted in any of the non-universal readings associated with ARBs (i.e., Vinogradov's reading from (4a), Cinque's reading from (5a), Egerland's reading from (6b), Cabredo Hofherr's readings from (7a–d), Gast and van der Auwera's readings from (8a–c)). Therefore, it seems sensible not to treat these expressions as ARBs and to restrict the definition of ARBs to those constructions that can have non-universal (i.e., existential), or arbitrary, readings.

3. Predictions

The three Polish ARB constructions are characterized by important differences and constraints. For example, the *-no/-to* construction and the 3PL impersonal trigger virile (plural) marking in agreement, while the reflexive can occur with any agreement, depending on the implicit referent. See, for example, (13) and (14):

- (13) Pracowano jako nauczyciele / *nauczycielki / *nauczyciel /
 work_{PST.IMPRS} as teacher_{PL.VIR} teacher_{PL.NVIR} teacher_{SG.M}
 *nauczycielka.
 teacher_{SG.F}
 'People worked as teachers.' (Kibort 2008: 267)

- (14) Pracowało się jako nauczyciele / nauczycielki /
 work_{PST.3SG.N} REFL as teacher_{PL.VIR} teacher_{PL.NVIR}
 nauczyciel / nauczycielka.
 teacher_{SG.M} teacher_{SG.F}
 'People worked as (female) teachers. / One worked as a (female)
 teacher.' (Kibort 2008: 273)

Furthermore, the *-no/-to* impersonal is restricted to the preterite and conditional, whereas the reflexive and the 3PL impersonal can be used in all tenses (preterite, present, future; cf. (2), (14), (19), (22)). This might mean that in certain contexts in which the *-no/-to* construction would be preferred to the other two constructions in the preterite, these constructions are perfectly acceptable in the present tense. We therefore predict an effect of tense on the 3PL and reflexive impersonal.

Although in the Slavic languages verbal aspect is such a pervasive category that it plays a role in almost all areas of grammar, so far it has not been considered as a factor influencing the acceptability of ARBs, and no studies have been carried out to investigate this. We will therefore investigate whether aspect has an effect on ARBs.

In the literature, the 3PL impersonal is often described as having “a strongly colloquial flavour” (Siewierska 2008b: 22). This leads to the simple assumption that the 3PL impersonal should be rated better in informal registers. In more formal registers, however, there seems to be a certain preference for the *-no/-to* construction. In the literature, this construction is sometimes described as “stylistically neutral” (Siewierska 2008b: 22) and sometimes as formal or bookish (Bartnicka and Lehmann 2004: 522; Puzynina 1993: 40; Skibicki 2016: 420). A possible explanation for this is that the use of *-no/-to* is perceived as more detached, perhaps because it mainly occurs in newspaper articles where the main function is the “objectivization” of information and where the irrelevance of the agent is emphasized (Laskowski 1984: 147). For this reason, the *-no/-to* construction might be considered less suitable for everyday communication. On this basis, we predict that the *-no/-to* impersonal should be rated better in a formal register. The reflexive impersonal, however, is described as stylistically neutral, which is why we do not expect any differences here.

As seen in Section 2, ARBs can express both universal (generic) and non-universal (specific, arbitrary) readings. The morphosyntactic differences between the three constructions might cause them to express one reading more easily than another, which in the situation of competition among the constructions might have an effect on their choice and acceptability.

We therefore predict that preferences for one of the three constructions might depend on the following factors: register, aspect, tense, and universal vs. non-universal reading.

4. Explorative Corpus Study

Before designing an experiment with native speakers, we conducted an explorative corpus study to see if there are any significant effects of these four factors on the frequency of the ARBs.

However, the problems associated with finding the three constructions in a monolingual Polish corpus like the NKJP (National Corpus of Polish; cf. Przepiórkowski et al. 2012) are very unevenly distributed. Thus, the *-no/-to* construction is completely unambiguous (since the form of the passive participle from which it originally derived now ends exclusively in *-nel/-te*), so that a search for *pracowano* ‘one worked’, for example, does not yield any false positives or false negatives.⁵ The reflexive impersonal is homonymous with a range of other constructions with the reflexive pronoun: real reflexives, recip-

⁵ The only exceptions are purely coincidental homonyms. For example, a search for the rare *padano* ‘one fell’ in the NKJP yielded a great number of false positives, all of which were preceded by *grana* and often capitalized to *Grana Padano* (a kind of Italian cheese).

rocals, anticausatives, reflexiva tantum (or inherent reflexives (15)), middles (16), and antipassives (17).⁶

- (15) Maria wyspała się.
 Maria out.sleep_{PST.3SG.F} REFL

‘Maria had a good sleep.’

(Fehrmann, Junghanns, and Lenertová 2010: 207)

- (16) Te samochody prowadzą się łatwo.
 these cars drive_{3PL} REFL easily

‘These cars are easy to drive.’

(Rivero and Sheppard 2003: 93)

- (17) Chłopiec rzucił się kamnieniami.
 boy throw_{PST.3SG} REFL stones_{INSTR}

‘The boy was throwing stones.’

(Janic 2013: 161)

A restriction to the neuter preterite form immediately adjacent to the reflexive pronoun (e.g., *pracowało się* ‘one worked’) can considerably reduce the number of false positives at the expense of missing all the instances in the present, future, and conditional, and where *się* is separated from the verb by other words. Finally, it is virtually impossible to find the 3PL impersonal in a corpus because it is completely homonymous with the regular (personal) third-person plural with pro-drop, which of course is more frequent by several magnitudes.

Consequently, we decided to use a Polish-German parallel corpus and look for the impersonal pronoun *man* as well as passives in the German version in order to avoid any construction bias that a direct search for the Polish constructions would have. For this explorative study, the ParaSol corpus (von Waldenfels and Meyer 2006–; cf. von Waldenfels 2006) seemed sufficient; furthermore, we restricted the analysis to the only two Polish original texts with a German translation in the corpus in order to exclude effects that foreign texts might have on the choice of construction by a Polish translator.⁷

In the corpus we found many constructions corresponding to German *man* or the German passive that are not ARBs—e.g., the personal passive, personal reflexive, infinitival constructions, verbal nouns, *można* ‘one can’, *trzeba*

⁶ Cf. Rivero and Sheppard 2003: 99. Note, however, that in contrast to other Slavic (and Romance) languages there is no reflexive personal passive in Polish.

⁷ The search terms used were "man" for the impersonal pronoun and "w[ieu]rd.*[]"*(ge|abge|ange|aufge|ausge|be|beige|darge|einge|er|hinge|hinter|nachge|niederge|über|um|unter|ver|vorge|wegge|wider|zer|zuge|zusammenge).*(en|t)" within s for the passive. Only the texts *Kongres futurologiczny* (1971) by Stanisław Lem and *Opowieści galicyjskie* (1995) by Andrzej Stasiuk were selected.

'one must', *widać* 'one can see', *słychać* 'one can hear', *czuć* 'one can feel', *człowiek* 'human being, man', *ludzie* 'people', *kto* 'who(ever)', *ktoś* 'someone', *nikt* 'nobody', etc. However, since the aim of the study was not a contrastive comparison, we discarded these hits and included in our analysis only the three ARB constructions.

All in all, the little corpus search yielded 118 ARBs, of which roughly half ($n = 61$) were reflexive impersonals, a third ($n = 36$) were *-no/-to* constructions, and a sixth ($n = 21$) 3PL impersonals. Their distribution over tenses and aspects is illustrated in Figures 1 and 2 below. The diagrams show that the reflexive impersonal is mainly used in the present tense (at a ratio of 49 : 12 in relation to all other forms) and that the most frequent ARB in the present tense is the reflexive (49 : 9). In the perfective preterite, the vast majority of the instances is covered by the *-no/-to* construction (24 : 5), which in turn is used considerably more often in the perfective than in the imperfective aspect (25 : 11). The main field of competition between the three ARBs seems to be the imperfective preterite, and the 3PL impersonal is the one that does not show clear preferences for any tense (preterite 10 : 9 present) or aspect (IPFV 7 : 5 PFV, not counting the present tense).

As our sample of the corpus contains only written texts, it is hardly possible to get any significant data on register from the small number of ARBs we found. We therefore refrained from a corpus analysis of register. As to the question of universal vs. non-universal readings, in many cases it was hard to tell, even with the context given in the corpus, which reading is to be preferred, as in (18) on the opposite page. Therefore, we did not conduct a

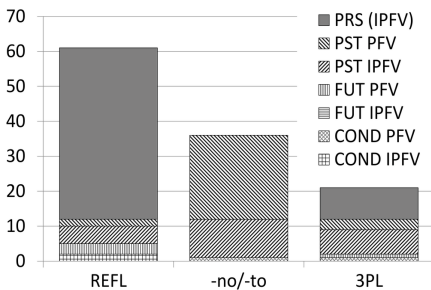


Figure 1. Tense and aspect against ARBs

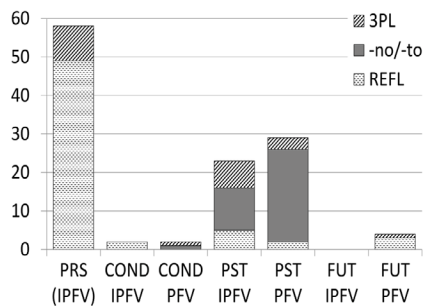


Figure 2. ARBs against tense and aspect

quantitative analysis. However, our qualitative analysis gives us the impression that in the imperfective preterite, which Figure 2 identifies as the main “battleground” of the constructions, the reflexive typically has a universal reading as in (19), whereas the *-no/-to* construction often has a non-universal reading as in (20).

- (18) O polskiej mafii mówiono i pisano już w połowie lat 90.
 ‘Polish mafia was talked and written about as early as the mid-nineties.’
- (19) Wtedy pisało się na maszynach do pisania.
 ‘At that time people wrote on typewriters.’
- (20) W 1980 roku zadłużenie Polski sięgało 11,5 mld zł, a nie 30 mld zł, jak pisano.
 ‘In 1980, the debt of Poland reached 11.5 billion złoty, and not 30 billion złoty as was written.’

5. Acceptability Judgment Test

The results of the corpus analysis made it seem worthwhile to verify all four factors—register, aspect, tense, and universal vs. non-universal reading—in an acceptability judgment test.

5.1. Test Items

As Bunčić (2018, 2019) has shown, at least the *-no/-to* construction and the reflexive impersonal are more acceptable with a more prominent agent than with a less prominent agent. We therefore selected 20 highly agentive Polish verbs⁸—i.e., verbs entailing Dowty’s (1991) agentivity features [+volition], [+sentience], and [+movement]—thus keeping the level of agentivity effects across all test items constant. Apart from that, all verbs are transitive (which has a similar effect because subjects of transitive verbs are more prominent agents than subjects of intransitive verbs). These verbs were embedded in sentences with uniform syntactic behavior; they all start with a prepositional

⁸ The tested verbs are the following: *dać* ‘give’; *dodać* ‘add’; *napisać* ‘write’; *oddać* ‘give back’; *odwrócić* ‘turn’; *pić* ‘drink’; *położyć* ‘lay’; *przestawić* ‘move’; *przynieść* ‘bring’; *rozbić* ‘break’; *rozwiązać* ‘solve, untie’; *ściągnąć* ‘take off, take down’; *stawić* ‘put’; *ugotować* ‘cook’; *wykopać* ‘dig’; *wypić* ‘drink up’; *wziąć* ‘take’; *zbudować* ‘build’; *złożyć* ‘fold, put together’; *związać* ‘tie (up)’.

phrase, followed by the verb in the particular impersonal construction and a direct object, and end with a subordinate clause, as in (21–23):

- (21) W moim samochodzie **rozbito** (smash_{PST.IMPRS}) szyby, bo mam niemiecką rejestrację.
 ‘The windows of my car were smashed in because I have a German license plate.’
- (22) W meksykańskich restauracjach **gotuje się** (cook_{3SG.REFL}) dania, które są bardzo smaczne.
 ‘In Mexican restaurants they cook dishes that are very tasty.’
- (23) Na wfie w gimnazjum **rozwiązali** (untangle_{3PL.PST}) skakanki, ponieważ jakiś głuptas zrobił żart i je wszystkie poplątał.
 ‘In PE at the high school, they were untangling the jumping ropes because some moron had made a practical joke and knotted them all up.’

We did not test any reflexive verbs, since in Polish the reflexive impersonal cannot be formed from inherently reflexive verbs (Siewierska 1988: 265).

The factors to be tested were varied as follows: The tense was either present or preterite (the future was not tested). The aspect of the verb was switched between perfective (only in the preterite) and imperfective. For register, some items were created to contain words marked as colloquial, like (23), whereas the rest was in a neutral register. For the universal/non-universal reading, we only created test items with a clearly generic reading, as in (22), or with a clearly specific reading, as in (21) or (23). (In the following we will therefore refer to this variable as *genericity*.) This set of variables allows for 6×3 different combinations in the preterite and 4×2 combinations in the present tense, since the *-no/-to* impersonal and the perfective aspect only occur in the preterite (see Table 1) and the generic reading does not readily agree with the perfective aspect.

5.2. Design of the Questionnaires

The test items were distributed over five different questionnaires, each containing 16–19 test items and an equal number of fillers (positive and negative control items), all of which were pseudo-randomized. Due to the multiplicity of factors to be examined, it was not possible to represent all factors equally in all the questionnaires. For example, if a questionnaire is balanced for tense (50% present, 50% preterite), it cannot at the same time also be balanced for aspect and include items in the imperfective preterite. Therefore, each of the

Table 1. Possible combinations of factors across ARBs

Tense	Reference	Aspect	Register	ARB
preterite	generic	imperfective	neutral	<i>-no/-to</i> , reflexive, 3PL
preterite	generic	imperfective	colloquial	<i>-no/-to</i> , reflexive, 3PL
preterite	specific	imperfective	neutral	<i>-no/-to</i> , reflexive, 3PL
preterite	specific	imperfective	colloquial	<i>-no/-to</i> , reflexive, 3PL
preterite	specific	perfective	neutral	<i>-no/-to</i> , reflexive, 3PL
preterite	specific	perfective	colloquial	<i>-no/-to</i> , reflexive, 3PL
present	generic	imperfective	neutral	reflexive, 3PL
present	generic	imperfective	colloquial	reflexive, 3PL
present	specific	imperfective	neutral	reflexive, 3PL
present	specific	imperfective	colloquial	reflexive, 3PL

questionnaires was created to examine a certain factor while keeping other factors constant, so that as a fallback strategy it would be possible to analyze each questionnaire separately. However, in the end we decided to evaluate all test items from all five questionnaires together and take care of confounding factors with statistical methods.

The participants were asked to rate the acceptability of each item on a six-point Likert scale ranging from “- - -” (completely unacceptable) to “+ + +” (completely acceptable). For the statistical analysis, the rating categories were converted into an equidistant numerical scale from -1 to +1 (i.e., -1.0, -0.6, -0.2, +0.2, +0.6, +1.0). We decided to choose an even number of rating categories to force participants to at least indicate a tendency in their rating and avoid indifferent “I don’t know” answers. It was also possible to give a free-text comment on every test item.

Together with some questions about the sociolinguistic background of the test participants, the five questionnaires were made available online via sosci-survey.de. After the introductory questions were answered, an urn-drawing mechanism decided which of the five questionnaires would be shown. The link to the online test was sent to colleagues in Poland, Germany, and Austria with the request to distribute it among their students, friends, and colleagues.

5.3. Participants

The five questionnaires were filled out by a total of 298 people. Since six people responded in the negative regarding whether Polish was their native language, and another three did not give any answer, we only used the re-

maining 289 questionnaires. Each of the five individual questionnaires was answered by 53 to 65 people.

Of the test participants, 70% were female and 30% male, while 96% were between 19 and 60 years old. The 289 participants represent all 16 Polish voivodeships, with a 26% majority from Mazovia; 13% live outside Poland (and were not excluded from the analysis since they indicated Polish as their native language). An overwhelming 81% majority holds an academic degree, and another 16% are currently enrolled at a university, so that the level of education is clearly not representative of the population, which should be kept in mind because it might have an influence on the test results.

6. Results

Statistics were conducted in R, version 3.5.1, using the *lme4* library (Bates et al. 2015) to perform a linear mixed-effects model (LMEM) of the acceptability of Polish ARBs under the influence of tense, aspect, register, and genericity. Depending on which factor we tested for the acceptability of the construction, we entered the remaining factors as fixed effects. The variability of subjects and items was taken into account by including them as random intercepts. Throughout the paper, we present *p*-values that are considered significant at the $\alpha = 0.001$ level. For the visualization of the data, we use boxplots, which are based on the five-number summary of the dataset. The whiskers indicate the total value range of the dataset (which in our large datasets generally includes both the minimum and the maximum value). The box encompasses the second and third quartiles of the dataset. The black horizontal line, which separates the quartiles, denotes the median of the dataset, while the circle plots the arithmetic mean. In general, the larger the box, the greater the dispersion of the data.

6.1. Register

As is apparent from Figure 3 (opposite), the reflexive impersonal scored notably better in informal items, but unexpectedly, *-no/-to* is not rated better in formal style, although in grammars it is usually treated as formal. A different picture emerges if we take a closer look at the ARBs in the present tense (see Figure 4). Here the 3PL impersonal is rated significantly better in informal style ($p < 0.001$), and there are no significant differences for the reflexive impersonal. These results confirm our prediction that the 3PL impersonal is marked as colloquial, whereas the reflexive impersonal is applicable in all styles, even though it was formulated independently from tense.

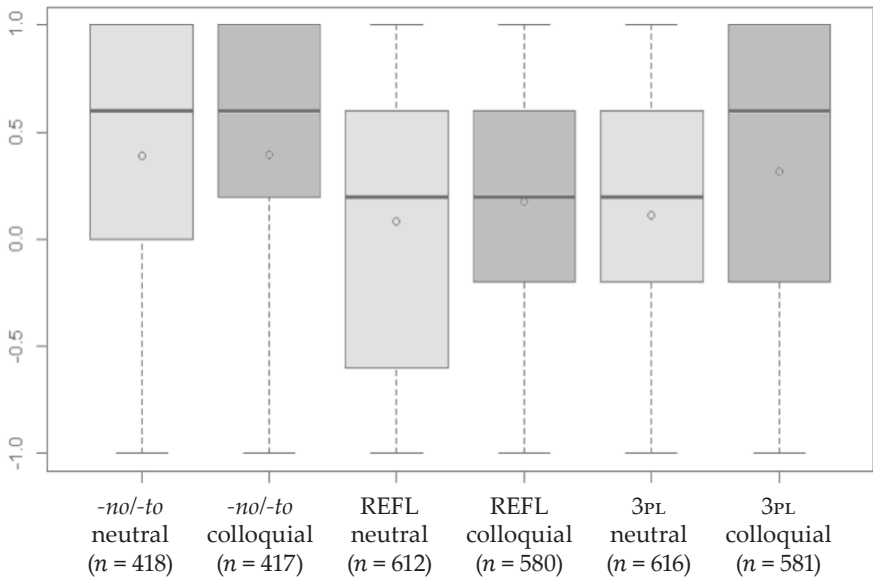


Figure 3. Colloquial vs. neutral register

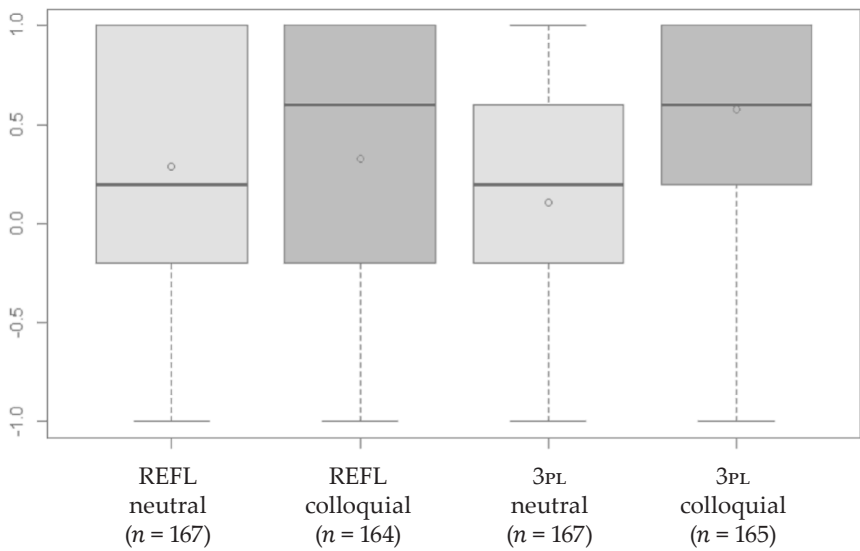


Figure 4. Colloquial vs. neutral register (present tense only)

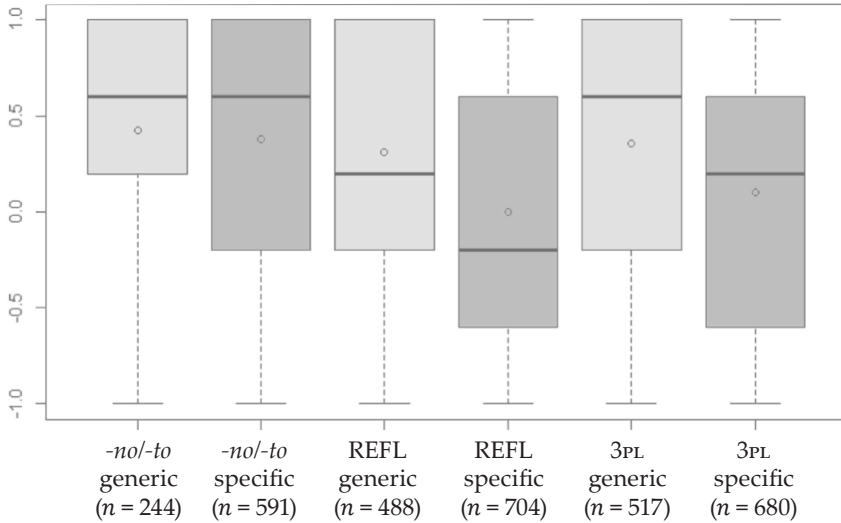


Figure 5. Generic vs. specific reading

6.2. Genericity

Visual inspection of Figure 5 above suggests that the reflexive impersonal is rated considerably better with generic than with specific readings. This is in line with our prediction, which was based on the results from the corpus query where there was a preference for *-no/-to* with an arbitrary reading versus the reflexive impersonal with a generic reading. However, after applying the mixed-effects model with confounding factors, the differences turned out to be statistically insignificant.

6.3. Aspect

Acceptability ratings for ARBs as a function of verbal aspect are shown in Figure 6 on the opposite page. The statistical analysis proves that the reflexive impersonal is significantly more acceptable in the imperfective than in the perfective aspect ($p < 0.001$). This confirms our assumption that aspect might have an influence on the acceptability of particular ARBs.

However, the perfective aspect causes the situation to be read as specific, and Figure 5 has shown that specific readings render the reflexive impersonal (at least numerically) less acceptable. So, in order to isolate the effect of aspect from that of genericity, we excluded the items with generic reference. The results are shown in Figure 7. Even with specific reference, the effect of aspect

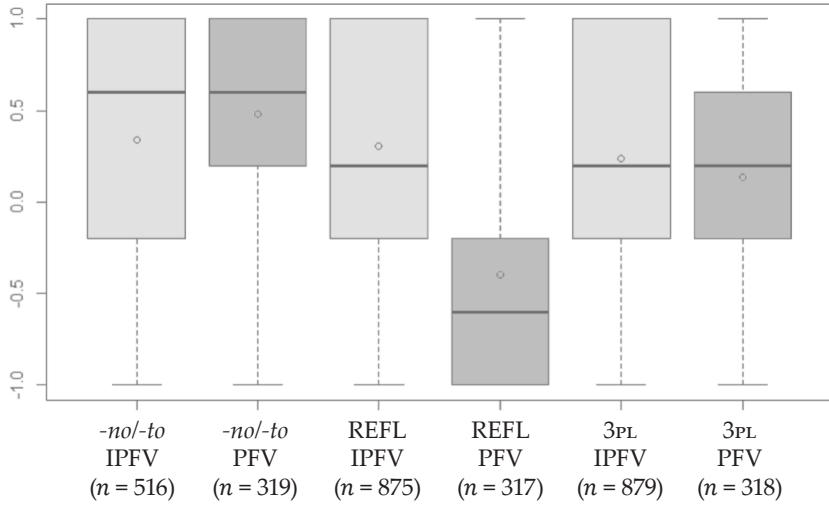


Figure 6. Imperfective vs. perfective aspect

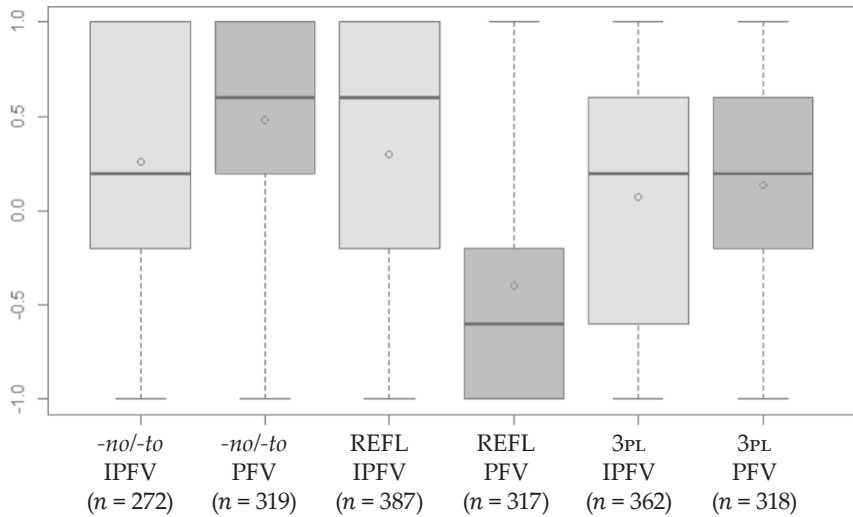


Figure 7. Imperfective vs. perfective aspect (specific only)

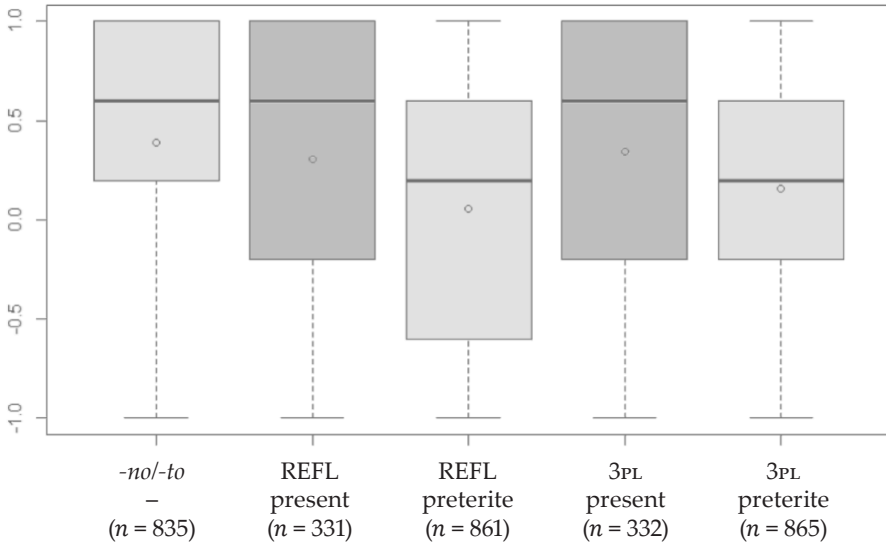


Figure 8. Present vs. preterite

on the reflexive impersonal is still significant ($p < 0.001$). The reverse effect on the *-no/-to* construction, which is better in the perfective aspect, is more pronounced here, but still not significant.

6.4. Tense

The ratings for the test items depending on tense are shown in Figure 8 above. The reflexive and the 3PL impersonal are numerically more acceptable in the present tense than in the preterite. This would corroborate our prediction that these two constructions are less suitable in the preterite, where they are in competition with the *-no/-to* impersonal. However, the statistical analysis reveals no significant effects.

7. Conclusion

It is evident that the three competing Polish ARB constructions are selected on the basis of a rather complex multifactorial situation. While the four possible factors we examined render the situation almost too complex for an acceptability judgment test, one can easily think of further factors that might play a role: e.g., clusivity (cf. Krzek 2015), verb class (transitive, unaccusative, unergative, reflexive, etc.), negation, definiteness or specificity of the object, etc. How-

ever, we have found unambiguous evidence that the following factors play a role in the acceptability of the Polish ARBs. As expected, the 3PL impersonal is better in colloquial discourse, but the *-no/-to* construction shows no significant effect of register. The reflexive impersonal is much more acceptable with the imperfective than with the perfective aspect. The effects of specific vs. generic reading (where both reflexive and 3PL impersonal seem to prefer the generic reading) and of tense (where the same two ARBs seem to prefer present tense), as well as the slight tendency of the *-no/-to* construction towards the perfective aspect, are too small to be significant.

The curious fact that the *-no/-to* construction turned out to be insensitive to register might be explained by the unusually high level of education of our test participants. It is probably true that such highly educated speakers use the *-no/-to* impersonal even in everyday speech. In this case, the classification of this construction as “formal” or “bookish” by the grammars is actually based on a sociolect rather than a register.

Sources

[NKJP] *National corpus of Polish*. Available at: nkjp.pl. Last accessed 1 February 2022.

ParaSol, a corpus of Slavic and other languages. (2006–) Ruprecht von Waldenfels and Roland Meyer. Available at: www.parasolcorpus.org. Last accessed 30 November 2019.

References

- Bartnicka, Barbara and Volkmar Lehmann. (2004) “Der Satz”. Barbara Bartnicka, Björn Hansen, Wojtek Klemm, Volkmar Lehmann, and Hanna Satkiewicz, eds. *Grammatik des Polnischen*. Munich: Otto Sagner, 487–591.
- Bates, Douglas, Martin Mächler, Benjamin M. Bolker, and Steven C. Walker. (2015) “Fitting linear mixed-effects models using lme4”. *Journal of statistical software* 67(1): 1–48. DOI 10.18637/jss.v067.i01.
- Bauer, Anastasia. (this volume) “Impersonalization strategies in Slavic: A contrastive corpus study based on the data from six Slavic languages”. *Journal of Slavic linguistics* 29(2): 123–78.
- Bunčić, Daniel. (2018) “Impersonal constructions in Slavic languages and the agentivity of the verb”. Sebastian Kempgen, Monika Wingender, and Norbert Franz, eds. *Deutsche Beiträge zum 16. Internationalen Slavistenkongress, Belgrad 2018*. Wiesbaden: Harrassowitz, 101–10.
- . (2019) “Agent prominence in the Polish *-no/-to* construction”. Anastasia Bauer and Daniel Bunčić, eds. *Linguistische Beiträge zur Slavistik*. Vol. 24.

- JungslavistInnen-Treffen in Köln, 17.–19. September 2015*. Frankfurt am Main: Peter Lang, 63–76.
- Bunčić, Daniel. (2020) “Agent prominence and movement in Russian third person plural impersonals (‘indefinite-personal sentences’)”. *Die Welt der Slaven* 65(2): 297–311.
- Cabredo Hofherr, Patricia. (2003) “Arbitrary readings of 3pl pronominals”. Matthias Weisgerber, ed. *Proceedings of Sinn und Bedeutung (SuB) 7*. Konstanz: Universität Konstanz, 81–94.
- Cinque, Guglielmo. (1988) “On *si* constructions and the theory of arb”. *Linguistic inquiry* 19: 521–83.
- Dowty, David R. (1991) “Thematic proto-roles and argument selection”. *Language* 67(3): 547–619.
- Egerland, Verner. (2003) “Impersonal pronouns in Scandinavian and Romance”. *Working papers in Scandinavian syntax* 71: 75–102.
- Fehrmann, Dorothea, Uwe Junghanns, and Denisa Lenertová. (2010) “Two reflexive markers in Slavic”. *Russian linguistics* 34: 203–38.
- Fellerer, Jan. (2008) “Kilka uwag o płaskiej koncepcji gramatyki języka polskiego jako obcego”. Władysław T. Miodunka and Anna Seretny, eds. *W poszukiwaniu nowych rozwiązań: Dydaktyka języka polskiego jako obcego u progu XXI wieku*. Cracow: Wydawnictwo Uniwersytetu Jagiellońskiego, 91–101.
- Gast, Volker and Johan van der Auwera. (2013) “Towards a distributional typology of human impersonal pronouns, based on data from European languages”. Dik Bakker and Martin Haspelmath, eds. *Languages across boundaries: Studies in memory of Anna Siewierska*. Berlin: de Gruyter Mouton, 119–58.
- Haider, Hubert. (2019) “Grammatical rules are discrete, not weighted, and not vulnerable”. Ken Ramshøj Christensen, Henrik Jørgensen, and Johanna L. Wood, eds. *The sign of the V—Papers in honour of Sten Vikner*. Aarhus: Aarhus University, 205–26. DOI 10.7146/aul.348.96.
- Himmelfmann, Nikolaus P. and Beatrice Primus. (2015) “Prominence beyond prosody”. Amedeo De Dominicis, ed. *pS-prominenceS: Prominences in Linguistics. Proceedings of the international conference*. Viterbo: DISUCOM Press, 38–58.
- Janic, Katarzyna. (2013) *L’antipassif dans les langues accusatives*. Ph.D. thesis. Lyon: Université Lumière Lyon 2.
- Kibort, Anna. (2008) “Impersonals in Polish: An LFG perspective”. *Transactions of the Philological Society* 106(2): 246–89.
- Krzek, Małgorzata. (2015) “Clusivity of the null subject pronoun in the Polish *-no/-to* construction” [slides of a presentation]. Available at: <https://docgo.net/clusivity-of-the-null-subject-pronoun-in-the-polish-no-to-construction>. Last accessed 1 February 2022.

- Laskowski, Roman. (1984) "Kategorie morfologiczne języka polskiego—charakterystyka funkcjonalna". Renata Grzegorzycowa, Roman Laskowski, and Henryk Wróbel, eds. *Morfologia (Gramatyka współczesnego języka polskiego 2)*. Warsaw: PWN, 121–70.
- Malamud, Sophia A. (2013) "(In)definiteness-driven typology of arbitrary items". *Lingua* 126: 1–31.
- Malchukov, Andrej and Akio Ogawa. (2011) "Towards a typology of impersonal constructions: A semantic map approach". Andrej Malchukov and Anna Siewierska, eds. *Impersonal constructions: A cross-linguistic perspective*. Amsterdam: John Benjamins, 19–56.
- Mel'čuk, Igor' A. (1974) "O sintaksičeskom nule". Aleksandr A. Xoldovič, ed. *Tipologija passivnyx konstrukcij: Diatezy i zalogi*. Leningrad: Nauka, 343–61.
- Padučeva, Elena V. (2012) "Neopredelennno-ličnoe predloženie i ego podrazumevaemyj sub'ekt". *Voprosy jazykoznanija* 2012(1): 27–41.
- Przepiórkowski, Adam, Mirosław Bańko, Rafał L. Górski, and Barbara Lewandowska-Tomaszczyk, eds. (2012) *Narodowy korpus języka polskiego*. Warsaw: Wydawnictwo naukowe PWN.
- Puzynina, Jadwiga. (1993) "Die sogenannten 'unbestimmt-persönlichen' Formen in der polnischen Sprache". Gerd Hentschel and Roman Laskowski, eds. *Studies in Polish morphology and syntax*. Munich: Otto Sagner, 31–61.
- Rivero, María Luisa and Milena Milojević Sheppard. (2003) "Indefinite reflexive clitics in Slavic: Polish and Slovenian". *Natural language & linguistic theory* 21: 89–155.
- Siewierska, Anna. (1988) "The passive in Slavic". Masayoshi Shibatani, ed. *Passive and voice*. Amsterdam: John Benjamins, 243–89.
- . (2008a) "Introduction: Impersonalization from a subject-centred vs. agent-centred perspective". *Transactions of the philological society* 106(2): 115–37.
- . (2008b) "Ways of impersonalizing: Pronominal vs. verbal strategies". María de los Ángeles Gómez González, J. Lachlan Mackenzie, and Elsa González Álvarez, eds. *Current trends in contrastive linguistics: Functional and cognitive perspectives*. Amsterdam: John Benjamins, 3–26.
- Skibicki, Monika. (2016) *Polnische Grammatik*. 2nd edition. Hamburg: Buske.
- Vinogradov, Viktor V. (1954) *Grammatika russkogo jazyka*. Vol. 2. Part 2. Moscow: Izdatel'stvo Akademii nauk SSSR.
- von Waldenfels, Ruprecht. (2006) "Compiling a parallel corpus of Slavic languages: Text strategies, tools and the question of lemmatization in alignment". Bernhard Brehmer, Vladislava Ždanova, and Rafał Zimny, eds. *Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV)*. Vol. 9. Munich: Sagner, 123–38.
- Wiemer, Björn. (1995) "Zur Funktion von Sätzen ohne grammatisches Subjekt im Polnischen und Russischen (Auszüge aus einer vergleichenden Corpu-

sanalyse zu einer ausgewählten Textsorte)”. *Wiener slawistischer Almanach* 35: 311–29.

Wiese, Eva. (1973) “Zur Analyse polnischer syntaktischer Konstruktionen aus dem Bereich der ‘unpersönlichen’ oder ‘unbestimmt-persönlichen’ Sätze”. *Zeitschrift für Slawistik* 5: 625–33.

Maria Katarzyna Prenner
Slavic Department
Giessen University
Giessen, Germany
maria.prenner@slavistik.uni-giessen.de

Daniel Bunčić
Slavic Department
University of Cologne
Cologne, Germany
daniel.buncic@uni-koeln.de