

Krzysztof Hwaszcz, Hanna Kędzierska

*Institute of English Studies, University of Wrocław*

## The Rise of an Indefinite Article in Polish: An Appraisal of its Grammaticalisation Stage (Part 2)

### Abstract

The main aim of the reported study is to establish the stage of grammaticalisation of the indefinite article in Polish by contributing the results of a corpus study. We selected and analysed 20,000 sentences containing the word *jeden*. The obtained results demonstrate that the uses of *jeden* as a presentative marker and a specific marker have been both attested, which would suggest that Polish numeral has already reached the specific marker stage. Based on the statistical analysis carried out for the obtained results, a statistically significant increase in the use of *jeden* as an indefinite marker has been revealed. This may be interpreted as evidence for the grammaticalisation phenomena, enhanced by language contacts with article-possessing languages (English and German).

### Keywords

grammaticalisation, indefinite article, corpus study, Polish, language change, numeral ‘one’

### Streszczenie

Artykuł przedstawia wyniki badania korpusowego, którego przedmiotem było określenie etapu gramatyzacji rodzajnika nieokreślonego w języku polskim. W tym celu z Narodowego Korpusu Języka Polskiego wybrano 20 000 zdań zawierających słowo *jeden*, które zostały poddane szczegółowej analizie. Uzyskane wyniki potwierdzają, że *jeden* w połączeniu z rzeczownikiem może pełnić zarówno funkcję słowa wprowadzającego (ang. *presentative marker*), jak i wyznacznika referencji szczegółowej (ang. *specific marker*). Świadczy o tym, że na pięcioetapowej skali zaproponowanej przez Heinego (1997) proces gramatyzacji liczebnika *jeden* przeszedł już etap 3 (dla porównania język angielski jest obecnie na etapie 4). Ponadto zaobserwowano istotny statystycznie wzrost użycia *jeden* w funkcji zaimka nieokreślonego w latach 1992–2011, tym samym potwierdzając istotność procesu gramatyzacji w kontekście kontaktów językowych pomiędzy polskim a językami, w których występują rodzajniki (w szczególności angielskim i niemieckim).

### Słowa kluczowe

proces gramatyzacji, rodzajnik nieokreślony, badanie korpusowe, język polski, zmiany językowe, liczebnik ‘jeden’

In the first part of this paper published in the preceding issue of this journal (Hwaszcz and Kędzierska 2018), we assessed that the grammaticalization of the indefinite article in Polish has reached the stage of a specific marker, with some occasional uses leaning towards the non-specific marker stage, which is the starting point for the corpus-study reported in this part.

## 5. A corpus-based study of the indefinite marker *jeden*<sup>1</sup>

This analysis aims at establishing the pace and orientation of the indefinite article grammaticalisation in Polish. Such an investigation is highly motivated by the number of meanings of *jeden* which have already emerged and which have been presented in details in Sections 3–4. To briefly recapitulate, we have estimated that the grammaticalisation process already reached stage 3 (specific marker), bearing in mind, however, that there exist uses of *jeden* which stretch to stage 4 (in the sense that the speaker may or not be able to identify the referent). The amassed literature on the topic cites a number of illustrative examples, all the same indicating that the number of possible contexts with *jeden* is expanding. We wish to examine whether there is indeed a legitimate reason to believe that there is a growing tendency for the occurrence of *jeden* as an indefinite marker.

### 5.1. Method: Data collection

In order to maximize the similarity of our corpus data to the Polish language which it aims to represent, we collected the distribution of *jeden* in all its paradigms within two independent corpora: NKJP (Janus and Przepiórkowski 2007) and Araneum Polonicum Maius (Benko 2014) and three independent search engines: Pelca (Pęzik 2012), KonText (available at <https://kontext.korpus.cz>) and Monco (available at <http://monco.frazeo.pl>) and then calculated the mean. The outcome of this procedure is illustrated in Table 1 below.

The calculation in Table 1 was used to select the number of sentences with *jeden* in each of its paradigms. These criteria were crafted to mirror the language in such a manner that it would enable us to make precise observations about the extracted sample and then to generalize back to the entire corpus.

We began by selecting the most suitable corpus for our analysis: it was Pelca because out of the three search engines, it is the only one that allows filtering

---

<sup>1</sup> We would like to thank Professor Bożena Rozwadowska and two anonymous reviewers for their insightful and valuable comments on the paper as they led us to a considerable improvement of the work. This work was funded by a grant awarded to Krzysztof Hwaszcz by the Faculty of Letters, University of Wrocław (0420/2570/17).

Table 1. The distribution of *jeden* in all its paradigms in Pelcra, KonText and Monco

Paradigm	Pelcra	Kontext	Monco	Mean
<i>jeden</i>	25%	22%	25%	<b>24%</b>
<i>jedna</i>	11%	10%	11%	<b>11%</b>
<i>jedną</i>	10%	11%	10%	<b>11%</b>
<i>jedne</i>	1%	1%	1%	<b>1%</b>
<i>jednego</i>	13%	13%	13%	<b>13%</b>
<i>jednej</i>	12%	13%	12%	<b>12%</b>
<i>jednemu</i>	1%	0%	0%	<b>1%</b>
<i>jedni</i>	1%	1%	1%	<b>1%</b>
<i>jedno</i>	9%	9%	7%	<b>8%</b>
<i>jednych</i>	1%	1%	1%	<b>1%</b>
<i>jednym</i>	16%	18%	18%	<b>17%</b>
<i>jednymi</i>	0%	0%	0%	<b>0%</b>

by dates, which was essential for measuring the direction of the grammaticalisation process. The threshold of 1,000 sentences containing the word *jeden* was set. Twenty dates (years ranged 1992–2011) have been selected as these have more than 1,000 sentences in the Pelcra corpus. Thus, the overall number of sentences with *jeden* used in the present study amounted to 20,000. The selection of the sentences was randomized so that they would be possibly most dispersed across different texts, disciplines and fields of science. Importantly, a completely uniform distribution of text genres for each year was not always possible: it refers especially to these years in which the number of sentences containing the word *jeden* was only slightly larger than 1,000, which hindered the uniform distribution across different genres and disciplines. Crucially for our study, *jeden* used as an indefinite marker is inherent for a low register of Polish and thus can be most frequently encountered in informal speech. As all the samples we analysed came from the written (and not spoken) corpus, with vast majority of them being the fragments of books, press articles or journals, the register of the analysed texts stayed uniform across the analysed years but it was the one in which the use of *jeden* as an indefinite marker would not be typically observed due to its highly formal nature.

## 5.2. Data division into different meanings of *jeden*

Since the grammaticalisation process of the indefinite article is, to a certain extent, linked with the semantic distribution of the numeral ‘one’ across the language (see Section 2), a meticulous assessment of the meaning in each of the collected sentences had to be implemented. The examples quoted in Section 3, whose aim was to support the particular use of *jeden*, were very illustrative, yet

highly idealised; although they were assessed by language speakers, they were not produced by them. Our data here objectively verify the claims about the different uses of *jeden* in various real-life situations. All in all, we have been able to individualize 10 separate meanings which we divide into three lists: a numeral, an indefinite marker and other functions (such as adjective, intensifier, etc.), exemplified in Sections 5.2.1– 5.2.3 below.

### 5.2.1. *Jeden* as a numeral in the corpus data

*Jeden* as a numeral is its primary function and it is the most frequent one across the board. According to the hypothesis for languages with an incipient indefinite article, a number of instances of *jeden* in this function should decrease over time (with the decrease being relative, not absolute).

The most typical use of *jeden* as a cardinal numeral is presented in (1).

- (1) *W trzyosobowym pokoju był tylko jeden pacjent.*  
 ‘There was only one patient in a room for three people.’

### 5.2.2. *Jeden* as an indefinite marker in the corpus data

Since indefinite markers express referentiality, it is obligatory to always examine the whole context in which *jeden* appears. It is particularly important in the case of examples in (2) to differentiate between indefiniteness marker uses and numeral uses. In the case of the numeral, there is a reference to the noun preceded by *jeden*, e.g., *W mojej pamięci jest jeden obraz – postać taty, gdy groźnie kiwa palcem znad gazety.* ‘There is one image in my memory – a figure of my father when he wiggles his finger up from the newspaper.’ Because *obraz* ‘painting’ is attributed with further reference, it is not just some unidentified memory but it is this particular one. A similar remark applies to the examples in (8) containing the ellipsis, which do not need to have any preceding referents in order to be analysed as indefinite markers.

#### 5.2.2.1. Specific marker

In the examples mentioned in (2) below, *jeden* is used as a specific marker which is not known to the hearer but known to the speaker as he or she narrates the situation. It may be optionally followed by an adjective, as in (2a).

- (2) a. *A był tam jeden sławny aktor, jakiś facet z telewizji.*  
 ‘There was a famous actor there: some guy from the TV.’  
 b. *Na koniec spytałem, które ze stworzeń najbardziej się Panu Bogu udało. A jedna dziewczynka podchodzi do mikrofonu i mówi: „To chyba ja”.*  
 ‘At the end I asked “Which creature given life to by God is the most rewarding/successful?” A girl approached the microphone and said “It was probably me”’

c. Jedna pani, zapatrzona gdzieś w dal niewidzącymi oczyma, uderzyła głową w szklane drzwi.

'A woman, gazing further out with unseeing eyes, hit a glass door with her head.'

It is also noteworthy that *jeden* might sometimes be preceded by the definite pronoun *taki* when used as a specific marker, resulting in structures such as *taki jeden* (3). The use of *jeden* in this meaning is sporadic in Polish: only 0.0016% of the collected corpus sentences contained this use. All of these examples occur in a singular form.<sup>2</sup>

(3) *Miałam wrażenie, że kocha się w takim jednym instruktora, który imponował jej trochę za bardzo.*

'I had the impression that she fell in love with this one instructor, who impressed her too much.'

#### 5.2.2.2. Non-specific marker

In the following example, *jeden* is known neither to the hearer, nor to the speaker.

(4) *Opowiadał też wzruszające historyjki o dzieciach. Na przykład o jednej dziewczynce, o ile się nie mylę ze Lwowa, która, gdy jechał, rzuciła mu razem z innymi dziećmi kwiaty.*

'He also told moving/emotional/heart-touching stories about children. For example, about a girl – from Lvov, if I'm not mistaken – who, together with some other children, was throwing flowers at him.'

#### 5.2.2.3. Plural indefinite marker

The plural form *jedni* 'some' as an indefinite marker appears in the structure *jedni/inni*. The main criterion used to distinguish plural indefinite markers from some of the adjectival uses (see Section 5.2.3.1., and especially uses such as (10a–10b) which are also followed by NPs containing such phrases as *inni* 'others' or *drudzy* 'the second ones') was that in cases such as (5b), *jeden* could be considered a specific marker. Thus, it is possible to substitute it with the indefinite pronoun *pewni*. It is not possible to make a substitution of that kind in the case of the items classified as adjectives in Section 5.2.3.1.

(5) a. *Tam był problem w rozróżnieniu na tych, którzy walczyli z bronią w ręku i których jedni nasi przeciwnicy polityczni uważali za niegodnych miana kombatanów, a my jednak uważaliśmy, że ten tytuł im się należy.*

'The problem involved distinguishing those who had fought with a weapon in their hands and who were not considered combatants by *some our political opponents*. However, we thought that they deserved to be called this way.'

<sup>2</sup> Interestingly, the plural form, i.e., *tacy jedni* also exists, as in *Nie chcę cię martwić ani nic, ale ostatnio kręcili się tu tacy jedni faceci i pytali o twojego ojca.* 'I don't want to worry you or anything but there were these men in the neighbourhood and they asked about your father'. However, this use seems to be extremely rare, at least in the (written) corpus data.

b. *Jedni* nasi przyjaciele mają 11 dzieci, w tym jedno adoptowane, *inni* mają 13, z których aż siedmioro jest adoptowanych.

'Some our friends have 11 children, one of whom is adopted; some others have 13 children, seven of whom are adopted.'

Geist (2011) makes a similar observation for Bulgarian, illustrated in (6).

(6) Včera vidjax edni amerikanci, koito poznavam.  
 Yesterday I-saw one.pl Americans whom I-know  
 'Yesterday I saw some Americans whom I know.'

It is, however, important to note that the distribution of the plural indefinite marker is not uniform from a cross-linguistic viewpoint. In Spanish, for example, an indefinite article *un* (which has been grammaticalised from *uno* 'one') can appear in plural: *unos* (Le Bruyn 2010). The same seems to apply to Portuguese, whose indefinite article *um* (from the numeral *um* 'one') takes the plural form *uns*. Asymmetrically, English and German do not have plural indefinite articles overt in morphology. It has been, however, long debated whether these bare plurals are indeed zero articles because they do not have double-scope reading as singular indefinite markers: compare *I want to buy a car* vs. *I want to buy cars*. The former is ambiguous: it is either one particular car or any car. The latter, on the other hand, has only the reading of any cars in general (Carlson 1977; Diesing 1992; Chierchia 1998; Dayal 2004).<sup>3</sup> Thus, the grammaticalisation path for plural indefiniteness markers can be two-fold: it may take either a path of a plural *one* (as it seems to be in the case of Polish) or it may gradually become eroded to the covert markers (Geist 2011).

#### 5.2.2.4. Elliptical indefinite pronoun

Sometimes *jeden* can be used without a following noun phrase and without a preceding referent. This use of *jeden* is limited specifically to the informal language. The problem with this case is whether an elliptical indefinite pronoun could transform into an indefinite article. Alexiadou and Gengel (2012) argue that it can, supporting their claim with an example from German.

(7) Peter hat viele gebrauchte Autos angesehen und  
 Peter AUX many used car.PL.ACC look-at.PST.3SG AND  
 schließlich ein neu-es gekauft.  
 finally a new.NEUT.SG buy.PST.3SG  
 'Peter has looked at many used cars and finally bought a new one.'

Thus, we classified this use of *jeden* as the one which may prompt the grammaticalisation process. We have additionally observed that the only noun phrases that can be inserted in the place of *jeden* have [+HUMAN] semantic

<sup>3</sup> We would like to thank one of the anonymous reviewers for this clarification.

feature and thus, that they can express only new referents. Consider the examples in (8).

- (8) a. *A tam siedzi Pies i wyje. Jeden [gość], co miał miękkie serce, poszedł go pogłaskać, ale Pies pokąsał go.*  
 ‘There is a howling dog over there. A soft-hearted guy walked up to pet it but it bit him.’
- b. *Jedni [ludzie] jechali do rodziny na wigilijną wieczerzę, do której chcieli zasiąść ze starymi, a ich jawne sentymenty mieszały się z cichym wyrachowaniem, że ceny biletów wzrosły, więc wypadnie to taniej niż przejażdżka autem.*  
 ‘Some [people] were going to spend Christmas Eve with their family and sit at the table with their parents, but their overt sentiments were mixed with covert sophistication that the ticket prices had increased, so using public transport would be cheaper than driving a car.’

### 5.2.3. Other uses of *jeden* in the corpus data

Apart from the numeral and indefinite functions of *jeden*, it has a number of other uses as well, such as an adjective, a pronoun with the ellipsis of a noun phrase, quantifier and intensifier. As for the grammaticalisation of the indefinite article, a number of instances of *jeden* in non-numeral functions should increase over time. This observation would support the claim that the grammaticalisation process is indeed present in the Polish language.

#### 5.2.3.1. Adjective

First, *jeden* may function as a meta-predicative adjective meaning *sam* ‘only’ (9a), *jedyny* ‘alone’ (9b), or *taki sam/identyczny* ‘the same/identical’ (9c) (see Danielewiczowa 2007).

- (9) a. *Przecież wiem dobrze, że Cieślewicz nie jest jeden. Wystawa afiszów w Bibliotheque Nationale pokazuje nam rozwój tej gałęzi sztuki w przeciągu stu ostatnich lat, ale Cieślewicz jeden przekonuje mnie więcej niż ta cała wystawa chaotyczna, wypadkowa, o szalenie nierównym wyborze eksponatów.*  
 ‘After all, I know well that Cieślewicz is not the only one there. The exhibition of placards in Bibliotheque Nationale shows us the development of this branch of art in the last century, but Cieślewicz (alone/as such) persuades me more than this entire exhibition, which is chaotic, accidental and includes a terribly unbalanced selection of pieces of art.’
- b. *Trzy prawdziwe bomby wybuchły w minionym tygodniu tylko w jeden czwartek. W Polsce, jak mi się zdaje, nie toczy się już żadna wojna domowa.*  
 ‘Three genuine bombs exploded in the last week only on Thursday alone. I reckon that there are no civil wars in Poland at the moment.’
- c. *Już od kilku lat „na czasie” są drzewka jednobarwne, monochromatyczne, ozdobione bombkami jednej wielkości.*  
 ‘For several years now, monochromatic trees decorated with baubles of the same size have remained fashionable.’

Interestingly, when *jeden* is followed by an absolute (i.e., non-gradable) adjective (e.g., *wielki/monstrualny* ‘huge/monstrous’) it seems to acquire the adjectival meaning *całkowity* ‘complete’ (10a). This use of *jeden* blends together with the subsequent absolute adjective to further intensify the adjective, cf. *Tam był wielki bałagan* vs *Tam był jeden wielki bałagan*. ‘There was (one) huge mess.’ (10b). When *jeden* is followed by a gradable adjective (e.g., *przeraźliwy* ‘appalling’), its meaning seems to change into *jednolity* ‘uniform/unified’ (10c–10d).

- (10)a. *Jej życie zmieniało się niemal z chwili na chwilę, a wynik tych zmian był jedną wielką niewiadomą.*  
 ‘Her life was rapidly changing and the result of these changes was one big open question.’
- b. *Tyle że całe moje życie to jedna monstrualna pomyłka. Przynajmniej tak to widzę dzisiaj.*  
 ‘It’s just that my whole life is one terrible mistake. At least this is the way I see it today.’
- c. *Mój życiorys mógłby się komuś wydać jednym pasmem dobroczynnych interwencji opatrzości.*  
 ‘One could consider my biography as a uniform stream of benevolent interventions of Divine Providence.’
- d. *Wydał z siebie jeden przeraźliwy okrzyk.*  
 ‘He’s given an appalling cry.’

Finally, we decided to include instances such as (11a–11c) in the category of adjectives. In such sentences, the function of *jeden* is to present the referent in opposition to some other entity or entities. Hence, this sub-type of the adjectival function of *jeden* always comes with the tandems *jeden/drugi* ‘the first one/the second one’ or *jeden/inny* ‘the first one/the other one’. As for *jeden/drugi*, they appear both in singular (11a) or plural (11b), yet in the case of *jeden/inny*, almost all sentences are in the plural form (11c). Also, it is worth mentioning that the use of *jeden* in (11c) has also been classified as an ordinal number in some dictionaries of the Polish language (e.g., Doroszewski 1958, recall the discussion in Section 3.2.).

- (11)a. *Nauczyciel powinien wyczuwać atmosferę, jaka panuje w klasie, aby nie stracić kontroli sytuacji. W jednej klasie można powiedzieć więcej, w drugiej dobrze jest wcześniej zatrzymać się z przekazem informacji, czy zrezygnować z niektórych ćwiczeń.*  
 ‘A teacher should be able to sense the atmosphere in the classroom to maintain the control of the situation. He or she can share more information with one class, while he or she should discuss less or resign from certain exercises with the other.’
- b. *Życie biologiczne polega na zjadaniu osobników jednych gatunków przez drugie i walkach dominacyjnych o prawo do pokarmu w obrębie własnego gatunku.*  
 ‘Biological life consists in consuming specimens of one species by others as well as in fighting for food within their own species.’



c. *Przed oczyma Jasia rozgrywała się jakby tragifarsa, jedni aktorzy schodzili ze sceny, inni się na niej pojawiali.*

‘There was some sort of a tragicomedy taking place right before Johnnie’s eyes; some actors were going off the stage and others were going on the stage.’

### 5.2.3.2. Elliptical pronoun

There are instances of *jeden* in which it is not followed by a noun phrase. However, the noun phrase is introduced in an earlier discourse. We consider such cases as pronouns because of the presence of the preceding referent (12).

(12) *Chłopi od razu zrozumieli, w czym rzecz. Jeden [z nich] powiedział mi: „Chcę pokazać, że Kościół nie takie prześladowania przetrzymał, i dlatego każą wam gadać o historii”.*

‘Peasants got the point straight away. One [of them] told me: *They want to show that the Church endured worse oppressions and therefore they command you to talk about history.*’

### 5.2.3.3. Quantifier

*Jeden* can additionally appear in contexts in which it functions as the neutral quantifier *niektórzy* ‘several (of them)’. The most crucial difference between the neutral quantifier category and the adjectival category (Section 5.2.3.1) is that in the case of quantifiers, *jedni* is not followed by *inni* ‘some’ or *drudzy* ‘second’ in further discourse.

(13) a. *Ludzie przesiadywali na ławkach otoczeni wielkimi torbami pełnymi zakupów, rozmawiali w grupkach, odpoczywali. Jedni zajadali bazarowe przekąski, kolby kukurydzy, szaszłyki na patykach i lody [...].*

People were sitting on the benches, surrounded by bags full of groceries. They were talking and resting. Several [of them] were eating bazaar snacks, corn cobs, shish kebabs and ice-creams [...].

b. *Na świecie żyją różni ludzie... Jedni [z nich] potrafią czytać z gwiazd...*

‘There are different people in the world. Some [of them] can read the stars.’

### 5.2.3.4. Intensifier

When *jeden* is preceded by an insulting noun phrase, its function is to intensify this phrase. It may be used both in singular (14a) or plural (14b). Importantly, the intensifier function does not always involve post position: the word *jeden* in expressions such as *jeden stek bzdur* (lit. ‘one load of nonsense’) in (14c) could also be classified as a type of meta-predicative intensifier.

(14) a. – Łżesz, huncwocie jeden – ojciec dowiedział się o wagarach od jednego z nauczycieli.

‘– You’re lying, you little scamp! – father found out about the truancy from one of the teachers.’

b. *Mamy kłopoty finansowe, kantory i banki nas cisną, Anvashe o nowym modelu wrzeczona trąbią, szerokokadłubowym, do którego, zasrańcy jedni, konia ponoć zdołali upchać, a tu jeszcze ta historia! I to w grudniu, tuż przed świętami!*

‘We’ve got financial problems: banks and exchange offices exert pressure, Anvashe shouts from the rooftops about the new model of a wide-body spindle, into which these suckers were able to squeeze a horse. Not to mention this case! And all this in December, just before Christmas.’

c. *Zeskanowany i przetłumaczony pamiętnik Numeru 5 stanowił dla Hunta jeden stek bzdur i ciąg bełkotu, skreślony drobnym, precyzyjnym, prawie technicznym pismem.*

‘The scanned and translated diary Number 5 – scribbled in a neat, tiny and technical-like inscription – was perceived by Hunt as a load of nonsense and constant gibberish.’

### 5.2.3.5. Peripheral uses

Moreover, *jeden* may occur in fixed expressions, such as *jak jeden mąż* ‘en masse, lit. as one husband’ (15a), *z jednej strony/z drugiej* ‘on the one hand/on the other, lit. from one side/from the second’ (15b), *pod jednym dachem* ‘under the same roof, lit. under one roof’ (15c), *jednym tchem* ‘in the same breath’ (15d).

(15) a. – *Wasze większe miasta – narzekał krasnolud do wtóru skrzekliwych przekleństw papugi – jak jeden mąż pobudowaliście na elfich i naszych fundamentach.*

‘– Your bigger cities – complained the dwarf in the accompaniment of the parrot’s squawking curses – you raised them en masse on the foundations of the elves’ and our ancestors.’ (wszystkie szanujące się krasnoludy znają przecież francuski)

b. *Wygodnie więc byłoby znaleźć takie wielkości, które z jednej strony zawierają wszystkie istotne informacje ukryte w macierzach reprezentacji, z drugiej zaś pozwalają rozpoznać reprezentacje istotnie różniące się od siebie.*

‘It would be convenient to find such values, which, on the one hand, include all essential pieces of information hidden in the representation matrixes, while, on the other, allow for the identification of the significant differences between them.’

c. *Z matką Wilibalda Winklera żyli pod jednym dachem 15 lat. Wreszcie zaczęli budować własny dom, ponieważ chcieli zrobić jej przyjemność.*

‘They had been living under the same roof with Wilibald Winkler’s mother for 15 years. At last they started building their own house because they wanted to make her feel good.’

d. *Kiedy lekarze mówią o wapniu, jednym tchem wyliczają jeszcze fosfor i witaminę D.*

‘When doctors speak about calcium, they speak in the same breath about phosphorus and vitamin D.’

The most problematic cases constitute the structures *jakiś jeden* or *jeden jakiś* ‘some’ because here *jeden* alone does not function as an indefinite marker but it has to be accompanied by *jakiś*, which carries the function of indefiniteness. *Jakiś jeden* constitutes only 0.0001% instances of the sentences in our corpus. We list them below in (16).

(16) a. *Wysłuchiwano sprawozdań przybywających z zagranicy posłów. Niekiedy na naradzie ogniskował uwagę jakiś jeden większy temat. Na końcu zwykle załatwiano sprawy administracyjno-biurowe dotyczące ogółu wydziałów.*

‘The reports of MP’s coming from abroad were heard out. Sometimes, one major issue came into focus during the meeting. At the end, some office-administration issues concerning departments on the whole were dealt with.’

b. *Ze wszystkich rozstajnych dróg jedna jakaś będzie prowadziła!*

‘Out of many, one road has to lead somewhere.’

### 5.3. Analysis of results

In order to verify the hypothesis and estimate whether the tendency to use *jeden* as an indefinite marker has increased in the last years, the number of sentences which belong to the categories described in Sections 5.2 were calculated for each data sample corresponding to years from 1992 and 2011 (see Section 5.1. for the detailed description of the data collection method). The number of observations for each year are presented in Table 2.

Table 2. Different uses of *jeden* in the corpus: summary

Year	Numeral	Indefinite	Other uses	Year	Numeral	Indefinite	Other uses
1992	903	11	86	2002	934	14	52
1993	943	12	45	2003	886	24	90
1994	928	14	58	2004	900	14	86
1995	870	12	118	2005	929	15	56
1996	918	11	71	2006	934	13	53
1997	947	16	37	2007	902	11	87
1998	976	8	16	2008	915	21	64
1999	967	19	14	2009	929	16	55
2000	953	16	31	2010	886	36	78
2001	905	14	81	2011	937	20	43

To check the trends in the three uses of *jeden* between 1992 and 2011, the correlations between the independent variable (YEAR) and the three dependent variables (numbers of *jeden* used as NUMERAL, INDEFINITE or OTHER) were estimated: Pearson correlation coefficients (R) and linear regression equations were calculated for each of the three functions of *jeden* distinguished in Table 2, as changing over the years (see Table 3). Also, plots visualizing the data for *jeden* used as a numeral (Figure 1), an indefinite marker (Figure 2) and other uses of *jeden* (Figure 3) in the investigated time period were created, with trendlines included which depict the tendency of *jeden* to appear in each function over the twenty years.

Table 3. Pearson correlation coefficients (R) and linear regressions for the functions of *jeden*

Function of <i>jeden</i>	Numeral	Indefinite	Other
R	-0.1371	0.5433	0.0173
Linear regression equation	$y = -0.6406x + 2205.3$	$y = 0.5617x - 1108.3$	$y = 0.0789x - 96.963$

On the basis of the obtained measures (Table 3) and the inspection of the plots, a downward trend in the use of *jeden* as a numeral has been observed, though no strong collinearity has been detected between the number of instances in which *jeden* functioned as a NUMERAL and YEAR ( $R < 0.6$ ). Also, no linear correlation was observed between the number of OTHER uses of *jeden* and YEAR, suggesting that the number of instances in which *jeden* is used in these functions (adjective, intensifier, quantifier, elliptical pronoun) did not change significantly over the time period in question. Importantly for our study, when it comes to the use of *jeden* as an indefinite marker, a positive linear correlation has been observed between INDEFITNIE and YEAR variables ( $R = 0.54$ ).

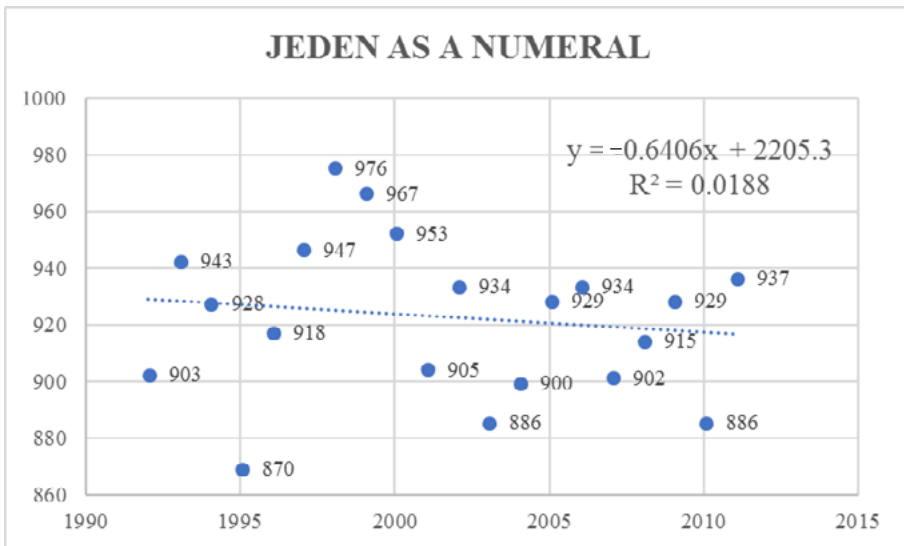


Figure 1. *Jeden* used as a numeral from 1992 until 2011

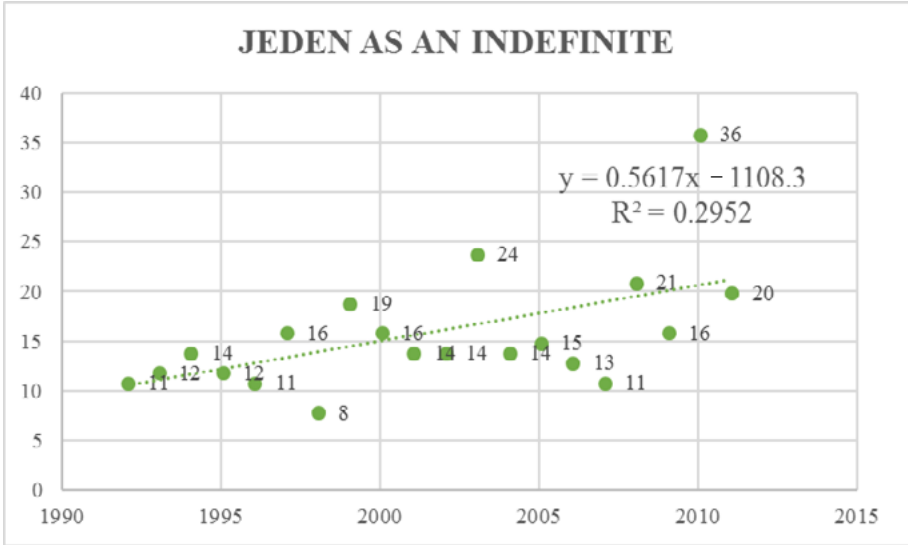


Figure 2. *Jeden* used as an indefinite marker from 1992 until 2011

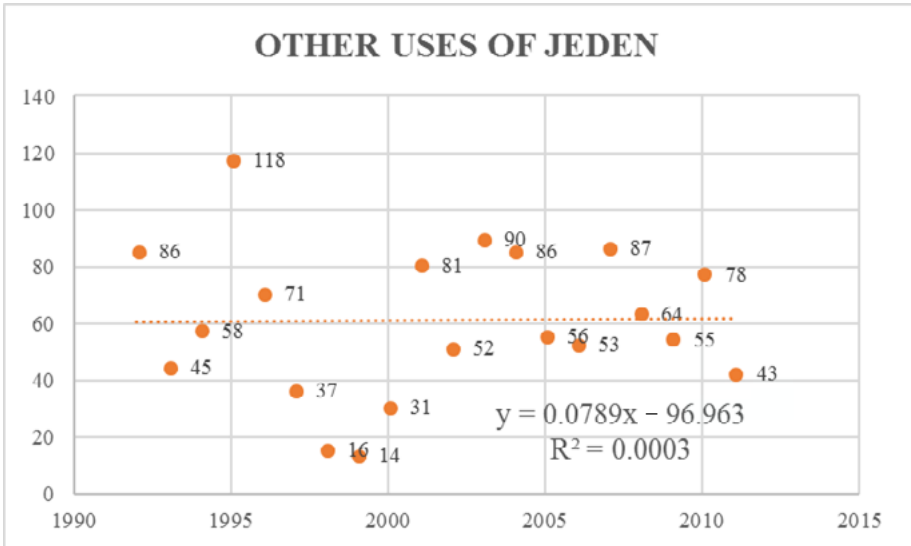


Figure 3. Other uses of *jeden* from 1992 until 2011

In addition, in order to verify whether the function of *jeden* as an indefinite marker has increased significantly between 1992 and 2011, statistical analysis of the data was carried out with the aid of R software (R Development Core Team 2017). The data was divided into four time periods, each of which

covered five consecutive years of the whole investigated time period<sup>4</sup>: TIME PERIOD 1 (1992–1996), TIME PERIOD 2 (1997–2001), TIME PERIOD 3 (2002–2006), TIME PERIOD 4 (2007–2011). Then, a generalized linear model analysis was performed using the *lm* function, available in R software.

Pairwise comparisons between TIME PERIOD 1 and other time periods revealed a statistically significant difference between TIME PERIOD 1 and TIME PERIOD 4 ( $t = 2.474$ ,  $p < 0.05$ ). The differences between TIME PERIOD 1 and the remaining two time periods (i.e. TIME PERIOD 2 and TIME PERIOD 3) were not statistically significant ( $p > 0.05$ ).

## 6. Conclusions

Most of the developments examined in the overview which was presented in the first part of this article (Hwaszcz and Kędzierska 2018) represent early stages of grammaticalisation, aside from Upper Sorbian and perhaps Bulgarian, Macedonian and Polish. Nevertheless, it is by no means obligatory to use the indefinite article in these languages. The status of an indefinite marker in the surveyed languages is open to questions: some of the incipient functions of *one* are used in specific regions or dialects, some are suggested as a possible option, some are preferable only in certain contexts.

We argue that language contact was influential in either starting or boosting the grammaticalisation of indefinite articles in most of the examined cases. A similar stand is taken by experts who investigated the details of the languages in question (e.g., Breu 1994; Löttsch 1996). Furthermore, there exists sociolinguistic evidence that article-less languages (e.g., all Slavic languages in the past) which had contact with article-possessing languages (English, German, French, among others) show a more advanced grammaticalisation process than languages without or with a limited contact of that kind, i.e., Ukrainian and Belarusian (Heine and Kuteva 2006). The illustration of the overview is presented in Map 1 below.

The corpus study conducted on the basis of 20,000 sentences containing the word *jeden* (and embracing the time period from 1992 until 2011) revealed a statistically significant increase in the use of *jeden* as an indefinite marker. To keep the grammatical balance, the use of *jeden* as a numeral decreased for the benefit of the indefinite marker's use as well as other functions of *jeden*. In the latter case, however, the progress was marginal: approximately 0.02,

---

<sup>4</sup> An alternative analysis in which the observations were grouped by five time periods covering four years did not differ with respect to the significance of results from the one presented in the Results section, i.e., statistically-significant differences were observed only between the first time period (i.e. from 1992 until 1995) and last time period (i.e. from 2008 until 2011) ( $t = 3.070$ ,  $p < 0.01$ ).



Map 1. The stages of development of the indefinite article in Slavic languages (with Polish) Source: [https://www.alternatehistory.com/wiki/lib/exe/detail.php?id=blank\\_map\\_directory%3Aall\\_of\\_europe&media=blank\\_map\\_directory:europe1.png](https://www.alternatehistory.com/wiki/lib/exe/detail.php?id=blank_map_directory%3Aall_of_europe&media=blank_map_directory:europe1.png) (accessed 19.07.2018).

which suggests practically no linear correlation. This finding could be explained in terms of the numeral grammaticalisation phenomena, perhaps enhanced by the political and social situation in Poland, with more people learning languages such as German or English since 1989 (after the fall of the Berlin wall), the growing status of English as a *lingua franca* in such fields as business, popular culture, trade and academia, and easier and more accessible travel opportunities.

The outcomes of the research contribute to the current research debate on grammaticalisation, displaying results from one of the least studied Slavic languages in this respect. The results are consistent with the ones found by other linguists who study the grammaticalisation process of indefinite articles.

Finally, we would like to point out the limitations of the study and suggest potential improvements for the future studies on the issue of language change. First and foremost, we tested only written corpora, while spoken corpora (for which there is not enough material to conduct a balanced study) might include more uses and functions of *jeden*. This is because the use of the indefinite marker *jeden* is considered informal or semiformal (Topolińska 1981; Grzegorzczkowska 2001, etc.). Also, it would be vital to investigate the uses of *jeden* in the last five years (i.e. 2012–2017) as our current study did not comprise data from this time period due to its unavailability in (neither written nor spoken) Polish corpora.

## References

- ALEXIADOU Artemis, GENDEL Kirsten (2012). NP ellipsis without focus movement/projections: The role of classifiers. In *Contrast and Positions in Information Structure*. Ivona KUČEROVÁ, Ad NEELEMAN (eds.), 177–205. Cambridge: Cambridge University Press.
- BENKO Vladimír (2014). Aranea: Yet Another Family of (Comparable) Web Corpora. In *TSD 2014, LNAI 8655*. Petr SOJKA, Aleš HORÁK, Ivan KOPEČEK, Karel PALA (eds.), 257–264. Cham: Springer International Publishing.
- BREU Walter (1994). Der Faktor Sprachkontakt in einer dynamischen Typologie des Slavischen. In *Slavistische Linguistic 1993*. Hans Robert MEHLIG (ed.), 41–64. Munich: Otto Sagner.
- CARLSON Gregory N. (1977). A *Unified Analysis of the English Bare Plural*. *Linguistics and Philosophy* 1(3), 413–457.
- CHIERCHIA Gennaro (1998). Plurality of mass nouns and the notion of ‘semantic parameter’. In *Events in Grammar*. Susan Rothstein (ed.), 53–103. Dordrecht: Kluwer Academic Publishers.
- DANIELEWICZOWA Magdalena (2007). Przymiotniki nieprzymiotniki. O pewnym niezwykłym typie wyrażen w języku polskim. *Zbornik Matice Srpske Za Slavistiku* 71–72, 223–236.
- DAYAL Veneeta (2004). Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27(4), 393–450.
- DIESING Molly (1992). *Indefinites: Linguistic Inquiry Monographs 20*. Cambridge MA: The MIT Press.
- DOROSZEWSKI Witold (1958). *Słownik języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.
- GEIST Ljudmila (2011). Bulgarian ‘edin’: The rise of an indefinite article. In *Formal Description of Slavic Languages: The Ninth Conference. Proceedings of FDSL 9*. Uwe JUNGHANN, Dorothee FEHRMANN, Denisa LENERTOVÁ, Hagen PITSCH (eds.), 125–148. Frankfurt am Main: Peter Lang.
- GRZEGORCZYKOWA Renata (2001). *Wprowadzenie do semantyki językoznawczej*. Warszawa: Wydawnictwo Naukowe PWN.
- HEINE Bernd (1997). *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.
- HEINE Bernd, KUTEVA Tania (2006). *The Changing Languages of Europe*. Oxford: Oxford University Press.
- HWASZCZ Krzysztof, KĘDZIERSKA Hanna (2018). The rise of an indefinite article in Polish: An appraisal of its grammaticalisation stage (Part 1). *Studies in Polish Linguistics* 13(2), 93–121.
- JANUS Daniel, PRZEPIÓRKOWSKI Adam (2007). Poliqarp 1.0: Some technical aspects of a linguistic search engine for large corpora. In *The Proceedings of Practical Applications in Language and Computers PALC 2005*. Jacek WALIŃSKI, Krzysztof KREDENS, Stanisław GÓZDŹ-ROSKOWSKI (eds.). Frankfurt-am-Main: Peter Lang.
- LE BRUYN Bert (2010). *Indefinite Articles and Beyond*. Utrecht: LOT.
- LÖTZSCH Ronald (1996). Interferenzbedingte grammatische Konvergenzen und Divergenzen zwischen Sorbisch und Jiddisch. *Sprachtypologie und Universalienforschung* 49(1), 50–59.



- PĘZIK Piotr (2012). Wyszukiwarka PELCRA dla danych NKJP. In *Narodowy Korpus Języka Polskiego*. Adam PRZEPIÓRKOWSKI, Mirosław BAŃKO, Rafał GÓRSKI, Barbara LEWANDOWSKA-TOMASZCZYK (eds.), 253–227. Warszawa: Wydawnictwo Naukowe PWN.
- R Development Core Team (2017). *R: A language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.
- TOPOLIŃSKA Zuzanna (1981). *Remarks on the Slavic Noun Phrase*. Wrocław: Zakład Narodowy im. Ossolińskich, Wydawnictwo Polskiej Akademii Nauk.

Krzysztof Hwaszcz  
Hanna Kędzierska  
Instytut Filologii Angielskiej  
ul. Kuźnicza 22, 50-138 Wrocław  
[krzysztof.hwaszcz(at)gmail.com]  
[hannaewakedzierska(at)gmail.com]

## Appendix 1

The questionnaire contained multiple-choice questions in which any number of answers could be selected. The answers appeared in a randomized order for each respondent.

Apart from original Polish version of the questionnaire, English translations are provided (in italics) below each question. The following English equivalents will be used in the translation of the questionnaire.

- jeden = indefinite marker (translated into *one*)
- jakiś = indefinite non-specific pronoun (translated into *some*)
- pewien = indefinite specific pronoun (translated into *certain*)
- taki = definite pronoun (translated into *this/that* in some contexts and *such as* in others)

The original Polish sentences did not have these elements marked in bold.

1. Jesteś sam/sama w domu. Nagle do drzwi dzwoni nieznajomy mężczyzna. Mężczyzna wręcza Ci paczkę i mówi, że to dla Twojej siostry, po czym pospiesznie odchodzi. Nie masz czasu zapytać go o wyjaśnienia ani przyjrzeć się jego twarzy. Po powrocie siostry do domu chcesz opowiedzieć jej o tym zdarzeniu. Mówisz:

- a) Był tu jakiś mężczyzna i zostawił dla ciebie paczkę.
- b) Był tu  pewien mężczyzna i zostawił dla ciebie paczkę.
- c) Był tu taki mężczyzna i zostawił dla ciebie paczkę.
- d) Był tu jeden mężczyzna i zostawił dla ciebie paczkę.
- e) Inna odpowiedź...

*1. You're alone at home. Suddenly, some stranger rings the doorbell. The man gives you a parcel and says that it's for your sister. Then he leaves. There is no time for you to ask about anything or to remember his face. When your sister comes back home, you want to describe the situation to her. You say:*

- a) *There was some man here and left a parcel for you.*
- b) *There was certain man here and left a parcel for you.*
- c) *There was this man here and left a parcel for you.*
- d) *There was one man here and left a parcel for you.*
- e) *Other answer...*

2. Jesteś sam/sama w domu. Nagle do drzwi dzwoni Janek, Twój dawny kolega ze szkoły. Chłopak wręcza Ci paczkę i mówi, że to dla Twojej siostry. Jesteś zaskoczony/zaskoczona, ponieważ masz pewność, że Twoja siostra nigdy w życiu nie spotkała Janka. Po powrocie siostry do domu chcesz opowiedzieć jej o tym zdarzeniu. Mówisz:

- a) Był tu jakiś chłopak i zostawił dla ciebie paczkę.
- b) Był tu  pewien chłopak i zostawił dla ciebie paczkę.

- c) Był tu taki chłopak i zostawił dla ciebie paczkę.
- d) Był tu jeden chłopak i zostawił dla ciebie paczkę.
- e) Inna odpowiedź...

2. *You're alone at home. Suddenly, your old acquaintance, Janek, rings the doorbell. The boy gives you a parcel and says that it's for your sister. You are surprised because you're sure that your sister has never met Janek. When she comes back home, you want to describe the situation to her. You say:*

- a) There was some guy here and left a parcel for you.
- b) There was certain guy here and left a parcel for you.
- c) There was this guy here and left a parcel for you.
- d) There was one guy here and left a parcel for you.
- e) Other answer...

3. Rozmawiasz z kilkoma osobami na imprezie. Nagle Twoja dobra koleżanka, Marysia, wypowiada następujące zdanie: „Wczoraj odwiedziła mnie jedna nasza przyjaciółka”. Zastanów się chwilę, jak zinterpretowałbyś/zinterpretowałybyś to zdanie. Następnie zaznacz wszystkie odpowiedzi, które uważasz za zgodne z prawdą:

- a) Ty i Marysia na pewno macie wspólną przyjaciółkę.
- b) Ty i Marysia niekoniecznie macie wspólną przyjaciółkę.
- c) Spośród wielu Waszych wspólnych przyjaciółek, tylko jedna odwiedził Marysię wczoraj.
- d) Marysia nie chce, żeby inne osoby w pokoju dowiedziały się, kto ją wczoraj odwiedził.
- e) Oboje wiecie, kogo ma na myśli Marysia, ale nie lubicie tej osoby i sarkastycznie nazywacie ją przyjaciółką.
- f) Marysia żartuje sobie z Ciebie lub się z Tobą droczy.
- g) To zdanie wydaje mi się dziwne. Nie wiem, jak je rozumieć.
- h) Inna odpowiedź...

3. *You're talking to several people at the party. Suddenly, your good friend Marysia says to you: "One of our friends visited me yesterday." Think for a moment how you would interpret the sentence. Mark all the answers that you think are correct.*

- a) *You and Marysia certainly have a common friend*
- b) *You and Marysia do not necessarily have a common friend.*
- c) *Of many of your common friends, only one of them visited Marysia yesterday.*
- d) *Marysia does not want other people at the party to find out who visited her yesterday.*
- e) *You and Marysia know who she is talking about, but you do not like her and you sarcastically call her a friend.*
- f) *Marysia makes fun of you or teases you.*
- g) *This sentence seems strange to me. I do not know how to understand it.*
- h) *Other answer ...*

4. W czasie podróży samochodem postanawiasz włączyć radio. Pierwsze, co słyszysz, to następujące zdanie: „Jeden mężczyzna bierze teraz ślub”. Zastanów się chwilę, jak zinterpretowałbyś/zinterpretowałybyś to zdanie. Następnie zaznacz wszystkie odpowiedzi, które uważasz za zgodne z prawdą:

- a) Istnieje jakiś mężczyzna, który bierze w tym momencie ślub, ale autor wypowiedzi go nie zna.
- b) Istnieje jakiś mężczyzna, który bierze w tym momencie ślub, ale autor wypowiedzi niekoniecznie go zna.
- c) Istnieje jakiś mężczyzna, który bierze w tym momencie ślub i autor wypowiedzi na pewno go zna.
- d) To zdanie wydaje mi się dziwne. Nie wiem, jak je rozumieć
- e) Inna odpowiedź...

4. *While driving a car, you decide to turn on the radio. The first thing you hear is the following sentence: “One man is getting married now.” Think for a moment how you would interpret the sentence. Mark all the answers that you think are correct.*

- a) *There is some man who is getting married at the moment, but the speaker does not know him.*
- b) *There is some man who is getting married at the moment, but the speaker does not necessarily know him.*
- c) *There is some man who is getting married at the moment and the speaker certainly knows him.*
- d) *This sentence seems strange to me. I do not know how to understand them.*
- e) *Other answer ...*

5. Podczas zajęć na uczelni siedząca przed Tobą kobieta odwraca się w Twoim kierunku i prosi: „Czy mógłbyś/mogłabyś pożyczyć mi jeden długopis?” Zastanów się chwilę, jak zinterpretowałbyś/zinterpretowałybyś to zdanie. Następnie zaznacz wszystkie odpowiedzi, które uważasz za zgodne z prawdą:

- a) Kobieta na pewno zauważyła, że masz ze sobą co najmniej kilka długopisów.
- b) Kobieta niekoniecznie musi wiedzieć, czy masz ze sobą zapasowy długopis lub długopisy. Pyta po to, żeby się przekonać.
- c) Gdybyś podał/podała kobiecie dwa długopisy, prośba nie mogłaby zostać uznana za spełnioną.
- d) To zdanie wydaje mi się dziwne. Nie wiem, jak je rozumieć.
- e) Inna odpowiedź...

5. *During the class at the university, the woman in front of you turns back to you and asks: “Could you lend me one pen?”. Think for a moment how you would interpret the sentence. Mark all the answers that you think are correct.*

- a) *The woman certainly noticed that you have at least several pens with you.*
- b) *The woman does not necessarily need to know if you have a spare pen or a few spare pens. She asks to find out.*

- c) *If you gave the woman two pens, the request could not be considered fulfilled.*  
 d) *This sentence seems strange to me. I do not know how to understand them.*  
 e) *Other answer...*

6. Zostałeś/zostałaś poproszona/poproszony przez swoją siostrę o przyprowadzenie z przedszkola siostrzenicy, która ma na imię Zosia i jest ciekawa świata. Idziecie, przechodząc obok trawnika, na którym kwitną kwiaty, Zosia zauważa pszczołę na jednym z nich i pyta: „Dlaczego pszczoła cały czas pracuje?” Odpowiadasz:

- a) Jakaś pszczoła to bardzo pracowity owad.  
 b) Pewna pszczoła to bardzo pracowity owad.  
 c) Taka pszczoła to bardzo pracowity owad.  
 d) Jedna pszczoła to bardzo pracowity owad.  
 e) *Inna odpowiedź...*

6. *You have been asked by your sister to bring a niece from the kindergarten, whose name is Zosia, and who is curious about the world. Going past a lawn where flowers bloom, Zosia notices a bee on one of them and asks: “Why does (-) bee work all the time?”. You answer:*

- a) Some bee is a very hard-working insect.  
 b) Certain bee is a very hard-working insect.  
 c) Such a bee is a very hard-working insect.  
 d) One bee is a very hard-working insect.  
 e) *Other answer...*

7. Widziałeś scenę, w której nieznanemu Ci chłopiec okradł Twojego młodszego brata Marka. Nim zdążyłeś/zdążyłaś zareagować uciekł z pieniędzmi Twojego brata. Zdajesz relację Waszej mamie mówiąc: „Do Marka podbiegł jeden chłopiec”. Dodajesz:

- a) Chłopiec uciekł z jego pieniędzmi, nim zdążyłem cokolwiek zrobić.  
 b) Uciekł chłopiec z jego pieniędzmi, nim zdążyłem cokolwiek zrobić.  
 c) Ten chłopiec uciekł z jego pieniędzmi, nim zdążyłem cokolwiek zrobić.  
 d) Jeden chłopiec uciekł z jego pieniędzmi, nim zdążyłem cokolwiek zrobić.  
 e) *Other answer...*

7. *You saw the scene in which the unknown boy robbed your younger brother Marek. Before you had time to react, he ran away with your brother's money. You describe the situation to your mum saying: “One boy ran up to Marek.”<sup>5</sup> You add (the word order is intentional and acceptable in Polish):*

- a) *Boy ran away with his money before I could do anything.*  
 b) *Ran away boy with his money before I could do anything.*

<sup>5</sup> The original Polish sentence was *Do Marka podbiegł jeden chłopiec.*, lit. ‘To Marek ran one boy.’

- c) *This boy ran away with his money before I could do anything.*
- d) *One boy ran away with his money before I could do anything.*
- e) *Other answer ...*

8. Wchodzisz do biblioteki i dwie nieznanne Ci dziewczyny (jedna blondynka, druga brunetka) rozmawiają ze sobą. Dotarło do Ciebie zdanie blondynki: „Jeden chłopak podszedł do mnie i dał mi tę książkę”. Opierając się na swojej intuicji, przeanalizuj to zdanie pod względem relacji chłopaka z dziewczynami.

- a) Blondynka zna tego chłopaka, a brunetka go nie zna.
- b) Blondynka nie zna tego chłopaka, a brunetka go zna.
- c) Żadna nie zna tego chłopaka.
- d) Obie znają tego chłopaka.
- e) Nie mam pojęcia.
- f) Inna odpowiedź...

8. *You enter the library and two girls you do not know (one blonde, the other one brunette) talk to each other. The blonde girl said: “One boy came up to me and gave me this book.” Based on your own intuition, analyse this sentence in terms of the relationship between the boy and the girls.*

- a) *The blonde girl knows the boy, the brunette one doesn't.*
- b) *The blonde girl doesn't know the boy, the brunette one knows him.*
- c) *None of the girls knows the boy.*
- d) *They both know the boy.*
- e) *I have no idea.*
- f) *Other answer ...*