Stroop-like interference of grammatical and visual number

Experimental evidence from Polish speakers

Piotr Gulgowski
University of Wrocław
piotr.gulgowski@gmail.com

Joanna Błaszczak
University of Wrocław
joanna.blaszczak@gmail.com

Abstract: The current paper presents results of two experiments attempting to replicate with Polish speakers a Stroop-like interference of grammatical number with the counting task, first reported by Berent et al. (2005) for Hebrew. Both experiments tested the influence of the type of number morphology (marked with overt suffix vs. unmarked) of nouns on the strength of the interference effect. Additionally, the second experiment investigated the processing of nouns with a mismatch between grammatical and conceptual number and tested the possible effect of animacy on number interpretation in order to determine the time at which the information about grammatical number is activated. The first experiment showed a significant interaction between the grammatical number and visual numerosity of the counted words and the effect of markedness, with marked singulars producing a bigger congruency effect than unmarked singulars. However, in the second experiment the influence of morphology was reversed and the overall effects were considerably weaker.

Keywords: Stroop effect; grammatical number; numerical cognition; markedness; plural

1. Background

Most human languages possess special means of expressing number contrasts in a systematic way through quantifiers or dedicated number words. Many languages treat number as a special grammatical category. This is a particularly interesting linguistic device, because in languages like English or Polish grammatical number is an obligatory property of nouns forcing the speakers to constantly pay attention to this feature in both production and comprehension. At the same time, we have the ability to mentally represent quantities in a way that seems at least partially independent of linguistic description (numerical cognition). Both grammatical number and numerical cognition are well defined and extensively studied

phenomena and their overlap constitutes an interesting research subject for investigating the interface between language and other cognitive systems.

1.1. Number in language

Languages differ in the distinctions they "carve" in the number field. The situation most familiar to speakers of European languages involves a two-way opposition between singular and plural forms. The former refer to just one entity whereas the latter are used to talk about a set of two or more things. Adopting a wider cross-linguistic perspective reveals a much more intricate picture, with values like dual, trial or paucal number, as described in Corbett's (2000) comprehensive monograph. Languages can also choose different means through which the available grammatical number values can be expressed. The primary number markers in a language may include independent function words, affixes, partial or complete reduplication, stem suppletion or clitics (Corbett 2000).

Grammatical number is primarily a nominal category[1] – it is often reflected in the form of nouns and pronouns and it is relevant for their interpretation. Like the grammatical category of gender or person, number enters the morpho-syntactic relations of agreement. The number value of a noun can affect the form of another element in the sentence (sometimes across several intervening words): determiners (e.g., **this** book/**these** books), verbs (e.g., *The new book of this author* **sells** *well/The new books of this author* **sell** *well*) or adjectives (e.g., Polish adjectives in noun phrases: **ciekawa** książka 'interesting.SG book'/**ciekawe** książki 'interesting.PL books'). Those additional manifestations of number establish syntactic relations between words.

From a psycholinguistic perspective, the processing of grammatical number in natural languages presents a wealth of research problems that still require investigation. Two specific issues are the main subjects of the present study: morphological markedness of number forms and a conflict between grammatical and conceptual number. Additionally, the influence of animacy on grammatical number processing was investigated.

1.2. Markedness of number forms

In languages like English, the expression of number is relatively straightforward – with a handful of exceptions, number is marked as a suffix on

---

[1] For a discussion of verbal number, see Durie (1986) and Corbett (2000).

plural nouns, while singulars are bare stems. For this reason, plural nouns can be characterized as the more marked element of the singular-plural number opposition.[2] The consequences of this fact can be observed in the way language users comprehend and produce subject-verb agreement relations. Agreement received considerable attention from psycholinguists. Much of the research in this area focused on experimentally elicited agreement errors (see, for instance, Bock et al. 2004; Bock & Miller 1991; Dillon et al. 2013; Eberhard 1997; Eberhard et al. 2005; Häussler 2009; Molinaro et al. 2008; Parker & Phillips 2016; Pearlmutter 2000; Pearlmutter et al. 1999; Wagers et al. 2009). Errors of agreement can lead to producing or accepting as correct sentences like (1).

(1)   ***The key**[SG] to <u>the cabinets</u>[PL] **were**[PL] on the table.

This phenomenon, also known as attraction or illusion of grammaticality, arises when the verb agrees in number with a noun other than its subject, often the closest noun in terms of the linear order.[3] In sentence (1), the plural noun *cabinets*, an "attractor" embedded in the subject noun phrase, intervenes between the proper subject head noun (*key*) and the verb. This illusion is characterized by an interesting asymmetry: the effect is stronger for sentences where plural nouns intervene between a singular subject and a verb (like in (1)) than when the number values on the nouns are reversed (Bock & Eberhard 1993; Bock & Miller 1991). Bock and Eberhard (1993) proposed a model in which plural nouns have a specific grammatical feature [Plural] whereas singular nouns, being unmarked, lack any number feature. If no feature can be found for agreement purposes, the subject receives a default morpho-syntactic specification [Singular], but this process can be disrupted by an intervening number feature, giving rise to agreement attraction. One prediction of this model is that if a singular noun does receive a grammatical number specification under some circumstances, the pattern of agreement errors should change. This hypothesis was tested in Eberhard (1997) where countable singular English nouns were preceded by the quantifiers *one, each* or *every* in one condition and by the definite article in another. The quantifiers used in the study can combine only with singular nouns so it was assumed that they should endow the following

---

[2] The concept of markedness is rooted in the theory of oppositions proposed by, among others, Jakobson (1957).

[3] This is not always the case, as illustrated by an example from Wagers et al. (2009, 209):

(i) **The drivers**[PL] <u>who the runner</u>[SG] **wave**[PL] to each morning honk back cheerfully.

noun with a specific number feature [Singular]. Indeed, singular nouns following the quantifiers were more resistant to attraction as subjects and generated more errors as intervening attractors in comparison to singular nouns without singular quantification.

1.3. Grammatical number vs. conceptual number

Linking number forms with number meanings seems like a straightforward and intuitive task, but on a closer inspection it turns out to be problematic. It is true that singular forms refer primarily to single and plural forms to multiple entities, but on some uses this relation does not hold. For instance, in English the grammatically plural pronoun *they* can be used to refer to one person if the speaker wants to avoid specifying the gender. For some nouns, there is a conflict between the value of the word's grammatical number and the conceptual number encoded in its lexical semantics. This conflict is particularly noticeable for three noun classes: *pluralia tantum*, collective nouns and mass nouns.

Pluralia tantum (e.g., *glasses*) are nouns whose grammatical number is always plural, but which can nevertheless refer to a single object. In sentence (2), the word *glasses* most likely refers to a single pair of glasses despite being grammatically plural.

(2)   He put his glasses on.

Collective nouns (e.g., *committee*) refer to a collection with multiple salient members. This conceptual plurality can, in some dialects of English, trigger a plural subject-verb agreement even for grammatically singular collectives (Bock et al. 2006; Humphreys & Bock 2005), as illustrated in sentence (3).

(3)   The committee has/have finally made a decision.

Mass nouns (e.g., *snow*) are nouns typically denoting some quantity of unindividuated substance or abstract concepts. They can be argued to exist outside the conceptual singular/plural opposition taking instead a default grammatical number value (typically singular) for the purposes of agreement and generally resisting pluralization without an accompanying change in meaning or a contextually salient unit of measurement (e.g., *two milks = two bottles of milk*).

Research using elicited agreement errors indicates that the specification of a noun as grammatically plural or singular can be stronger than its conceptual number in the computation of subject-verb agreement. Bock &

Eberhard (1993) and Bock et al. (2001) failed to find increased agreement errors in sentences where a singular collective noun intervened between a singular subject and a verb (e.g., ***the strength*** *of* the army*...*), suggesting that the semantic plurality of a collective noun is not enough to attract number agreement. Similarly, in Bock et al. (2001) *pluralia tantum* (e.g., *groceries*) and bipartite (e.g., *scissors*) attractors led to more instances of plural agreement than singular nouns in this function, despite being conceptually singular. The authors concluded that attraction by a plural local noun is due to its grammatical specification and not to its semantic plurality.[4]

1.4. Grammatical number in lexical access

While subject–verb number agreement mechanisms have been investigated in multiple studies, much less is known about the process of extracting grammatical number values from noun forms on-line during lexical access or the format in which grammatical singularity and plurality are mentally represented. In studies using the technique of elicited agreement errors there is always some delay between the noun carrying the grammatical number feature and the verb on which the effect of the feature is observed. This technique, therefore, does not reveal at which processing stage the representation of grammatical number is activated when a singular or plural noun is encountered.

An attempt to shed some light on these processes was made by Berent et al. (2005), who conducted a series of experiments using a technique based on a numerical Stroop interference. Stroop interference is a name for the difficulty with the simultaneous processing of conflicting information coming from different sources. This psychological phenomenon is named after John Ridley Stroop, one of the early researchers investigating this effect (Jaensch 1929; Jensen & Rohwer 1966; MacLeod 1991; Stroop 1935). A classic Stroop effect can be observed in experiments involving color words. When the color of the font is incongruent with the meaning of the word (e.g., the word *red* written in green font), then naming the font color while ignoring the word's meaning is more difficult than when the font

---

[4] It is possible that different types of agreement are more sensitive to morpho-syntactic or conceptual number. Bock et al. (1999) found a difference between pronominal and verbal agreement in terms of number attraction. In sentences with collective noun subjects, verbs were more likely to follow inflectional agreement (agreeing with the grammatical number) whereas pronouns preferred notional agreement (agreeing with the conceptual number).

color and the meaning are congruent (e.g., the word *red* in red font) or when the second dimension is removed altogether (e.g., geometric shapes displayed in different colors). Stroop-interference experiments demonstrate that some features of stimuli are activated involuntarily, even when they are irrelevant for the task at hand. Participants seem to be unable to ignore the word's identity (its sound or meaning) when they try to simply recognize its visual color. For this reason, Stroop tasks have been used to verify claims about the automaticity of processing in different areas of cognition. A kind of Stroop effect exists for symbolic representations of numbers, like digits (*7*) or numerals (*seven*). Stroop interference requires the existence of two potentially interfering dimensions in stimuli that can be manipulated independently. In number symbols, such dimensions correspond to two types of numerical information: the perceivable numerosity of the items (e.g., three instances of the digit *7*) and the numerical meaning of the number symbols. It has been demonstrated that counting instances of number words or digits presented on a card or a screen takes more time when the visual numerosity is incongruent with the numerical value (e.g., symbol *2* repeated four times: *2 2 2 2*) than in congruent or control conditions (Flowers et al. 1979; Naparstek & Henik 2010; Pavese & Umiltà 1998; Windes 1968).

Berent et al. (2005) tested the possibility that grammatical number can interfere with the ability to count visually presented words, giving rise to a Stroop-like effect. The participants (native Hebrew speakers) were presented with singular and plural Hebrew nouns displayed on a computer screen either once (visually single) or repeated twice (visually double). Meaningless strings of repeated letters were used for control. The participants were asked to assess how many tokens they saw on the screen on each trial. When the morphological number of the word was incongruent with the visual numerosity, the participants' responses were significantly slower than for control items. This effect was obtained only for grammatically plural words, i.e., when a word with a plural suffix was presented as a single token (e.g., *dogs*), the responses were considerably longer than when it was presented as two tokens (e.g., *dogs dogs*). Singulars did not differ significantly from the control. The authors interpreted this Stroop-like interference effect as suggesting that number value is extracted automatically from word forms (morphology) and that the numerical meaning associated with grammatical number and the non-linguistic numerosity of visually presented objects activate identical or very similar conceptual representations. They also concluded that plural nouns are marked both morphologically (suffix) and semantically (specific number meaning: *more*

than one) in contrast to singular nouns, which are unmarked in both respects. This interpretation is consistent with ideas like Horn's division of pragmatic labor. Following observations of earlier linguists, Horn (2001) argues that, given a pair of related items (e.g., a singular and plural noun), a general tendency in natural languages is to give the item marked in terms of form (morphologically complex, e.g., an affix) a marked meaning (specific concept). The unmarked element of the opposition (e.g., no special morphological marking) simply lacks the specific semantic concept in question, instead of being associated with a different specific semantic feature of its own (Horn 2001; see also Farkas & de Swart 2010 for a theoretical discussion).

2. Present study

A Stroop-like interference can be used to gain more insight into different aspects of the processing of grammatical number. The present study used the interference effect in two experiments as a diagnostic tool to investigate the processing of grammatical number in lexical access, in particular the effect of overt and null number morphology (Experiments 1 and 2), a mismatch between grammatical and semantic number ((Experiment 2), as well as the possible influence of the category of animacy (Experiment 2).

Berent et al. (2005) found a clear Stroop-like interference only for plural nouns. They interpreted the findings as evidence that singulars are underspecified for number semantically as well as morphologically. Crucially, in the experiment they describe, all singular forms were suffixless and all plurals were created by attaching a regular or irregular ending to a singular base. The situation in Polish is more complex than in Hebrew, and therefore more interesting from the perspective of the markedness debate. First, number in Polish is inextricably fused with case in a system of noun endings. Second, it is possible to find nominative singular nouns with an overt case/number suffix (e.g., *lekcj-a* 'lesson-NOM.SG' vs. *lekcj-e* 'lesson-NOM.PL') and singular nouns without such an ending (e.g., *czołg-Ø* 'tank-NOM.SG' vs. *czołg-i* 'tank-NOM.PL').[5] Dependent cases of such nouns still receive suffixes indicating case and number explicitly (e.g., *czołg-u*

[5] The presence or absence of an overt case/number suffix of Polish nouns depends primarily on the noun's gender value (Nagórko 2007, 143; Swan 2002, 66; Wiese 2011, 117). Zero suffix in the nominative singular form is predominant for masculine nouns (with some exceptions, e.g., a handful of masculines inflectionally resembling feminines, like *poet-a* 'poet', or neuters, like *dziadzi-o* 'grandpa'), present only in a

'tank-GEN.SG'), so it would be somewhat tricky trying to decide whether given singular nouns are morphologically marked or unmarked for number in general. In any case, Polish offers an opportunity to study the influence of the presence or absence of an overt number suffix on the treatment of specific noun forms.

In Experiment 1 of the present study, suffixed (marked) and suffixless (unmarked) singular nouns were used alongside suffixed plural nouns in a counting task. Experiment 2 used the same method to investigate the processing of nouns with a mismatch between number morphology and number meaning. Specifically, the three types of number-inconsistent nouns described above (*pluralia tantum*, mass and collective nouns) were studied. Those three noun classes were contrasted with "ordinary", prototypically countable singular and plural nouns (e.g., *krow-a* 'cow-NOM.SG', *noż-e* 'knife-NOM.PL'). For grammatically singular nouns, the possible influence of morphological markedness was taken into consideration. Experiment 2 was also an attempt to determine whether the animacy of a noun referent has any effect on the Stroop-like interference between grammatical and visual number. To do this, nouns denoting inanimate objects (e.g., *wiertł-o* 'drill-NOM.SG') and animate beings (e.g., *wron-a* 'crow-NOM.SG') were used as stimuli. A link between animacy and grammatical number has been noted in typological studies. In some languages the singular-plural distinction exists only for a subset of nouns. When this is the case, the nouns that accept a plural marking have usually more animate referents than those that do not, that is, their referents are higher in the animacy hierarchy (Corbett 2000; 2001; Haspelmath 2013; Smith-Stark 1974). Grimm (2012) connects the properties of grammatical number systems across languages with the concept of individuation. He argues that entities higher on the animacy scale are also more individuable and therefore more countable. Additional evidence that animacy can influence number comes from the observation that plural verb agreement with a singular collective subject is usually triggered by animate, and not inanimate, collectives (e.g., *The faculty are in a meeting* vs. **The forest are on fire*), although under certain conditions inanimates can also agree with plural verbs (for examples and discussion, see Levin 2001, Chapter 5). Because animacy is part of the lexical semantics of a word, its possible influence on the Stroop-like interaction can help establish at which processing stage grammatical number gets activated.

---

minority of feminine nouns (i.e., those with the stem ending in a functionally soft consonant, like *wieś* 'village' or *mysz* 'mouse') and absent in neuter nouns.

3. Experiment 1

3.1. Hypothesis and predictions

The main hypothesis tested in Experiment 1 was that the Stroop-like interference between grammatical number and visual numerosity depends on the presence of an overt morphological marker of number on nouns. Congruency effect between grammatical number and visual numerosity was expected for plurals because all plural nouns used in the experiment were marked with a number-encoding ending. Participants should be faster to count plural nouns when 2 tokens are displayed on the screen than when only one token is presented. An opposite congruency effect was expected for singular nouns with a suffix (marked). That is, participants should count marked singulars faster in the visually single than in the visually double condition. Unmarked singulars should differ from both plural nouns and singular nouns with a suffix – they should not present any number congruency effect, or the effect should be significantly smaller than for marked singulars.

3.2. Design

3.2.1. Materials

One hundred Polish nouns were used in the experiment, all in nominative case:

- 50 singular nouns
  - 25 unmarked singulars (e.g., *czołg-Ø* 'tank-NOM.SG')
  - 25 marked singulars (e.g., *lekcj-a* 'lesson-NOM.SG')

- 50 plural nouns created from the same stems as the singular nouns by adding a suffix (e.g., *czołg-i* 'tank-NOM.PL', *lekcj-e* 'lesson-NOM.PL')

Case syncretism is quite common in Polish declensional paradigms and some forms can be ambiguous not only in terms of their case but also their number value (cf. *mysz-y* 'mouse-NOM.PL' or 'mouse-GEN.SG'). No word picked for this experiment was number-ambiguous in this way.

Additionally, following the solution in Berent et al. (2005), 40 strings of repeated letters (e.g., *aaaaaa*) were created. Ten different letters of the Polish alphabet were used (*a, b, c, d, e, g, h, u, o, y*), each appearing in

four strings: two five-letter and two six-letter strings. Repeated letters were used to guarantee that no number-related semantic interpretation could be associated with those items.

Unmarked singulars and marked singulars were matched for the number of letters and surface frequency based on the information from *Narodowy Korpus Języka Polskiego* (Przepiórkowski et al. 2012) using the PELCRA system (Pęzik 2012). Plural nouns were on average slightly longer and less frequent than singulars (see Table 1).

Table 1: Mean letter lengths and surface frequencies (per million) for different item types used in Experiment 1 (SDs in parentheses)

	Letters	Frequency
Unmarked singular	5.12 (0.67)	5.20 (4.42)
Marked singular	5.16 (0.90)	5.20 (4.62)
Marked plural	5.64 (0.92)	4.10 (5.26)
Neutral strings	5.5 (0.50)	NA

There were 140 items in total. Each item appeared both as a single token (e.g., *lekcja*) or repeated twice (e.g., *lekcja lekcja*). Therefore the total number of trials in the experiment was 280. The order of items was pseudo-randomized and two lists were created, one being the exact inverted image of the other. Half of the participants saw one list and the remaining half saw the other list.

3.2.2. Procedure

The experiment started with a greeting message and instructions displayed on the screen. The message explained that the task of the participant was to count the number of words (e.g., *lekcja* 'lesson') or letter strings (e.g., *aaaaaa*) appearing on the screen by pressing the left arrow key when the item was visible as a single token (e.g., *aaaaaa*) or the right arrow key when the item appeared on the screen twice (e.g., *aaaaaa aaaaaa*).

Each trial started with a fixation cross in the middle of the screen. The cross remained visible for 300 ms and was followed by 300 ms of blank screen and then either a single token of an experimental item or an item repeated twice appeared. The item(s) remained on the screen until the participant reacted by pressing the left or the right arrow. If the response was incorrect, there was a message informing the participant about the mistake (*ŹLE!* 'wrong!'). The message disappeared after 400 ms. If the

reaction was correct, there was no feedback. In every case the trial ended with 300 ms of a blank screen before the next trial began. See Figure 1.
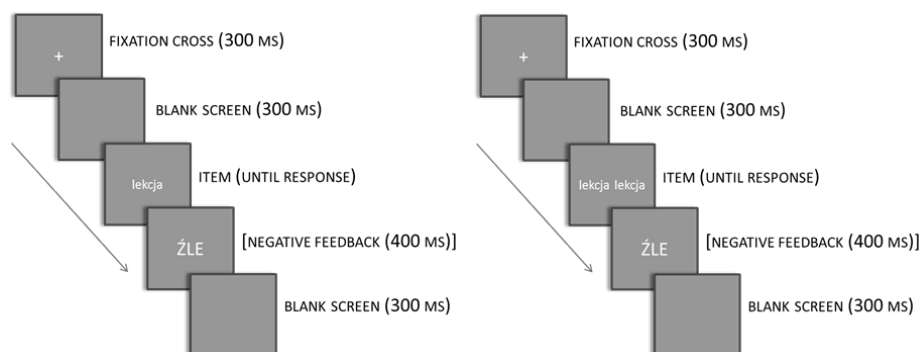

Figure 1: Experimental procedure in Experiment 1

The experiment proper was preceded by a training session that followed the same procedure with the exception that a feedback was given for both incorrect (*ŹLE* 'incorrect') and correct responses (*DOBRZE* 'correct'). There were ten training trials. The training ended with a message informing about the number of correct and incorrect responses, encouraging the participant to ask questions and informing about the possibility of repeating the session. No training item appeared later in the experiment proper.

Halfway through the experiment there was a message informing about a break. The participants could proceed when ready by pressing the space bar. Each experiment session lasted approximately 10–15 minutes. The experiment was designed and presented using the PsychoPy software (Peirce 2007; 2009).

3.2.3. Participants

Thirty one students of the Institute for English Studies of the University of Wrocław (20 women) took part in the experiment in exchange for partial course credit. The participants were all native speakers of Polish and reported no known neurological or reading-related problems. The average age was 20 (SD = 2.13). The oldest participant was 29, and the youngest was 18. Data from one participant had to be removed from the final analysis due to low overall accuracy (below 75%).

3.3. Results

For the reaction time (RT) analysis, the data were cleaned first by removing incorrect responses and then by eliminating the trials with RTs of two standard deviations above or below the mean in each condition for every participant. This resulted in removing 5% of accurate responses, with the outliers roughly equally distributed over subjects and conditions. The remaining trials were subjected to tests performed with the SPSS software (version 22).[6]

Table 2: Mean reaction times and accuracy (percentage of accurate responses) for different item types in Experiment 1 displayed as visually single or double (standard errors in parentheses)

		Visual number					
		Visual 1		Visual 2		Congruency (Visual 1 − Visual 2)	
Item type		RT (ms)	Accuracy	RT (ms)	Accuracy	RT (ms)	Accuracy
Singular		492 (13)	98.8%	499 (15)	97.8%	−7	1.0%
Unmarked singular	*tygrys* 'tiger'	490 (12)	98.7%	493 (14)	97.6%	−3	1.1%
Marked singular	*wieża* 'tower'	494 (13)	98.9%	505 (16)	97.9%	−11	1.0%
Marked plural	*tygrysy* 'tigers'	494 (13)	98.0%	488 (14)	97.2%	6	0.8%
Neutral strings	*yyyyyy*	504 (13)	97.3%	483 (14)	97.0%	21	0.3%

A 3×2 ANOVA was conducted with RT as the dependent variable and the following independent factors:

---

[6] The number-neutral condition (meaningless strings of letters, e.g., *zzzzzz*), contrary to the expectations, produced the greatest difference between the two visual number conditions. Strings of repeated letters turned out to be considerably faster to count when two tokens were displayed on the screen than when they appeared as one token (see Table 2). This, notably, was not the case in the original experiment by Berent et al. (2005). This result makes the strings in the present experiment problematic as a baseline condition to which singular and plural nouns could be compared looking for a congruency effect. For this reason, following a suggestion from an anonymous reviewer, strings of repeated letters were excluded from the analysis.

- Item Type (unmarked singular, marked singular, marked plural)

- Visual Number (visual 1, visual 2)

There was no significant main effect of either Item Type ($F_1(1.92, 55.58) = 2.95$; $p = .06$; $F_2(2, 97) = 2.72$; $p = .07$) or Visual Number ($F_1(1, 29) = 0.34$; $p = .56$; $F_2(1, 97) = 0.67$; $p = .42$).

The interaction between the two factors was significant by subjects ($F_1(1.83, 53.19) = 3.48$; $p = .04$; $\eta_p^2 = .107$) but not by items ($F_2(2, 97) = 2.23$; $p = .11$). An examination of the data revealed the presence of a congruency effect: singular nouns of both types were responded to faster in the visual 1 condition than in the visual 2 condition. The pattern was reversed for plural nouns. To further analyze the nature of this interaction and test the research hypotheses, planned comparisons were computed.
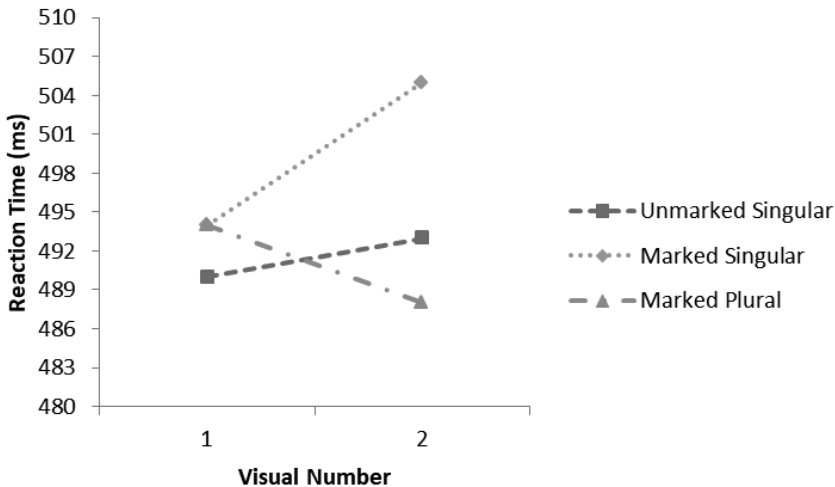

Figure 2: The interaction of Item Type and Visual Number

The first set of comparisons checked the possible Item Type × Visual Number interactions for individual pairs of item types (Table 3). For unmarked singular nouns compared to marked singular nouns the interaction was not statistically significant. It was also not significant for unmarked singulars compared to plurals or for all singulars put together compared to plurals. However, it reached the level of significance (by subjects) for marked singular nouns compared to plurals, indicating that those two item types generated most difference in congruency effects. Results are given in Table 3.

Table 3: Results of planned comparisons testing Item Type × Visual Number interactions for individual pairs of item types (p-values adjusted: Sidak method)

Comparison (Visual 1 vs. Visual 2)	df		t		p	
	Subjects	Items	Subjects	Items	Subjects	Items
Marked singular vs. marked plural	58	97	2.64	2.08	.04*	.15
Unmarked singular vs. marked plural	58	97	1.38	1.03	.53	.77
Unmarked singular vs. marked singular	58	97	−1.251	−0.91	.62	.84
Singular (all) vs. marked plural	58	97	−2.320	−1.91	.09	.22

Another set of comparisons involved checking whether the congruency effect (the difference between average RT in the visual 1 and visual 2 conditions) is significant for individual item types: marked plurals, marked and unmarked singulars and all singulars taken together. None of the comparisons reached the level of statistical significance. Results are given in Table 4.

Table 4: Results of planned comparisons testing the significance of the congruency effect for individual item types (p-values adjusted: Sidak method)

Comparison (Visual 1 vs. Visual 2)	df		t		p	
	Subjects	Items	Subjects	Items	Subjects	Items
Marked plural	55.61	97	0.81	1.19	.89	.66
Marked singular	55.61	97	−1.74	−1.71	.31	.32
Unmarked singular	55.61	97	−0.53	−0.42	.97	.99
Singular (all)	35.73	97	−1.29	−1.51	.60	.44

3.4. Discussion

Experiment 1 managed to replicate the Stroop-like effect reported by Berent et al. (2005) for grammatical number with Polish speakers. As predicted, counting nouns displayed on the screen as one or two tokens was affected by the grammatical number of the counted words – for grammatically plural nouns it took longer for the participants to decide that the word was presented as one token on the screen than when two tokens were displayed and for singular nouns the opposite was true. Singular nouns used in the experiment belonged to two different types: those encoding singular number through an overt suffix and those with no number ending. Morphologically marked singulars produced a bigger congruency effect than unmarked singulars. Only marked singulars differed in terms of congruency effect from plurals, so the data offer some support for the main hypothesis: the strength of a Stroop-like interference between grammatical number and visual numerosity may depend on the presence of an overt morphological marker. It has to be noted that, despite the significant interaction, no congruency effect (a mean reaction time difference between items displayed as visually single or visually plural) for any individual item type proved statistically significant, suggesting that the effect of visual number manipulation is very weak.

The congruency effect for morphologically marked singulars stands in an interesting contrast to the findings of Berent et al. (2005), where (unmarked) singulars did not produce any interference with visual number in the counting task. The outcome of Experiment 1 suggests that singular nouns are more likely to automatically evoke a specific number reading ('exactly one') if they are clearly marked morphologically. The presence of an overt marking may provide an important cue for the parser facilitating the assignment of the number value to the word in contrast to an unmarked singular noun, which may be temporarily perceived as a pure numberless stem requiring an ending. While this form is obviously eventually recognized as a singular nominative noun, the delayed activation of the concept of number might reduce its interaction with the visual number of tokens on the screen in the counting task. This may suggest that instead of talking about semantic unmarkedness and markedness of singular and plural nouns, it might be more accurate to talk about a greater or lesser ease of access to the grammatical number value of a word.[7]

---

[7] An anonymous reviewer suggested the possibility that the inclusion of both unmarked and marked singular nouns in the same experiment highlighted the singular marking.

In the present experiment, number-neutral strings produced the strongest difference between the visual 1 and visual 2 conditions (see Table 2), with RTs for two tokens being considerably shorter than for one token. In other words, strings of repeated letters behaved more plural-like than actual plural nouns. This made the strings problematic as a baseline condition to which singular and plural nouns could be compared and, consequently, they were not included in the analysis (see footnote 6). A possible explanation for this plural-like effect may be that a string of letters is perceived as a plurality of symbols (many letters) rather than a single object (one string), which constitutes a possible confounding factor in the counting task. This possibility was tested in Experiment 2.

4. Experiment 2

4.1. Hypothesis and predictions

Berent et al. (2005) demonstrated that the numerical Stroop interference they found for Hebrew words was not sensitive to the familiarity of the plural nouns or the regularity of their stems. Moreover, in the third experiment described in their paper, the interference effect was present even for nonwords bearing plural morphology, which prompted the authors to conclude that "morphological (and, consequently, semantic) number can be assigned by the grammatical processor even in the absence of lexical information, since nonwords lack such information entirely" (*op.cit.*, 354).

An interesting question in this context is at which point during word processing the information about grammatical number becomes available. Number, as a grammatical category involved in morpho-syntactic processes, may be accessed relatively early, perhaps along with the lemma-level information during the initial structure-building phase (150–300 ms after word onset according to the three-phase model of language comprehension proposed by Friederici 2011). This possibility, coupled with the results of Experiment 1 in the present study, formed the basis for the main hypothesis tested in Experiment 2: grammatical number morphologically marked on a word with an affix is processed automatically in the early stages of lexical analysis, activating the notion of singularity for grammatically singular nouns and the notion of plurality for grammatically plural nouns before the lexical semantics of a noun's stem is fully accessed. Consequently, this should happen even for those nouns whose

notional and grammatical number is inconsistent. If so, plural morphology on *pluralia tantum* words would automatically trigger conceptual plurality (e.g., automatically associating the word *scissors* with something plural) which would then have to be suppressed (assigning a singular interpretation, e.g., *one pair of scissors*). Similarly, grammatically singular mass nouns may activate the concept of singularity and singular collective nouns may initially be understood as denoting a single entity, with the conceptual plurality of the constituent parts receiving more activation at a later processing stage.

The first prediction for Experiment 2 was that it should be possible to replicate the effect of number morphology for marked singulars. Marked singular nouns should differ from both marked plural and unmarked singular nouns, producing significantly shorter reaction times in the visually single than in the visually double condition. Unmarked singular nouns should not present any congruency effect, or the effect should be considerably smaller than for marked singulars.

The second prediction was that singular collective and mass nouns should pattern with ordinary countable singulars. Responses for those nouns should be longer in the visually double than in the visually single condition, provided that they are morphologically marked for number.

Third prediction: *pluralia tantum* nouns should resemble ordinary countable plurals in terms of the Stroop-like congruency effect in the counting task. Responses for those nouns should be longer in the visually single than in the visually double condition.

With respect to the possible effect of the animacy of a noun's referent on the Stroop interference, the following prediction was made: because the activation of grammatical number is expected to precede the access to lexical semantics of the stem, animacy should not affect the Stroop-like interference.

Finally, non-linguistic, visually monolithic stimuli (white rectangles) were used to assess whether a strong effect observed for strings of repeated letters in Experiment 1 was due to the plurality of characters in those strings. If this was the case, the following result should be obtained: white rectangles should produce a minimal congruency effect in comparison to number-neutral words.

4.2. Design

4.2.1. Materials

Eighty prototypically countable nouns were used in the experiment:

- 40 singular nouns (e.g., *rower* 'bike')

- 40 plural nouns (e.g., *młotki* 'hammers')

Half of the countable nouns (20 singulars and 20 plurals) had animate referents, while the remaining half denoted inanimate objects. Singular nouns were additionally divided in terms of number morphology. Of the total set of 40 singular nouns, 20 were without any number suffix (unmarked) and the remaining 20 had an overt number ending (marked). Singulars with a suffix and singulars without a suffix contained an equal number of animate and inanimate words. Examples illustrating the division of singular nouns with respect to animacy and morphology are given in Table 5. All plural nouns had an overt number suffix.

Table 5: Examples of countable singular nouns divided in terms of number morphology and animacy

	Animate	Inanimate
Unmarked	*borsuk-Ø*	*pędzel-Ø*
	'badger-NOM.SG'	'brush-NOM.SG'
Marked	*krow-a*	*łóżk-o*
	'cow-NOM.SG'	'bed-NOM.SG'

Additionally, 82 nouns with a conflict between grammatical and conceptual number were selected:

- 40 mass nouns (e.g., *piasek* 'sand')

- 20 collective nouns (e.g., *stado* 'herd')

- 22 *pluralia tantum* nouns (e.g., *nożyce* 'scissors')

Because a sufficient number of *pluralia tantum* and collective nouns of the right length could not be found, each word from these two types was used twice in the experiment in order to obtain more comparable number of observations for different groups. All mass and collective nouns

were grammatically singular. To check the possibility of the influence of morphological markedness on the interpretation of grammatically singular nouns suggested by the results of Experiment 1, forms with and without a suffix were selected. 12 collective nouns (e.g., *grup-a* 'group-NOM.SG') and 17 mass nouns (e.g., *złot-o* 'gold-NOM.SG') were morphologically marked with a suffix, the remaining collective (e.g., *tłum-Ø* 'crowd-NOM.SG') and mass (e.g., *olej-Ø* 'oil-NOM.SG') nouns were unmarked.

Items were matched for letter length. The average number of letters in all conditions was close to 5.5, except for *pluralia tantum*, for which the number was close to 6 letters. All nouns were in the nominative case. Just like in Experiment 1 no word was number-ambiguous because of case syncretism (see section 3.2.1.).

There were also two groups of items devoid of either grammatical or semantic number:

- 20 number-neutral words, including adverbs (e.g., *żółto* 'in a yellow color'), particles (e.g., *czyż* 'alas'), prepositions (e.g., *przez* 'through') and conjunctions (e.g., *gdyż* 'because')

- 20 white rectangles corresponding roughly in size to the average area of the words used in the experiment

Thus, Experiment 2 included two new neutral control conditions that may better reflect the RT difference in responding to one token vs. two tokens independently of any (in)congruency with morphological number marking or semantic information. Number-neutral words were chosen instead of strings of repeated letters to increase the likelihood of them being treated as a coherent whole (a word) and in that way to improve their suitability as the baseline condition. The inclusion of non-linguistic rectangles was intended to establish whether the visual plurality of letters in words generates its own plurality-related Stroop interference in the counting task, which would account for the unexpected results for number-neutral strings of repeated letters in Experiment 1.

Overall, the experiment used 182 unique words (224 with repetitions) and 20 rectangles. Each item appeared both as a single token on the screen or as two copies. Therefore the total number of trials in the experiment was 488. The trial order was randomized for each participant.

4.2.2. Procedure

The procedure was the same as in Experiment 1 except for the number of breaks (three breaks in Experiment 2). Each experiment session lasted approximately 15–20 minutes. The experiment was designed and presented using the PsychoPy software (Peirce 2007; 2009).

4.2.3. Participants

Thirty two students of the Institute for English Studies of the University of Wrocław (24 women) took part in the experiment in exchange for partial course credit. Participants were all native speakers of Polish and reported no known neurological or reading-related problems. The average age was 22 (SD = 3.47). The oldest participant was 31, and the youngest was 19.

4.3. Results

For the reaction time (RT) analysis, the data were cleaned first by removing incorrect responses and then by eliminating the trials with RTs of 2 standard deviations above or below the mean in each condition for every participant. This resulted in removing 4.7% of accurate responses, with the outliers roughly equally distributed over subjects and conditions. The remaining trials were subjected to tests performed with the SPSS software (Version 22). Average RTs and percentage of correct responses for main experimental conditions are presented in Table 6. The same measures for countable nouns divided according to animacy and morphological marking are given in Table 7.

4.3.1. Replicating the results of Experiment 1

In an attempt to replicate the contrast between overt and null number morphology on singular nouns found in Experiment 1, a 3×2 ANOVA paralleling the ANOVA from section 3.3. was conducted on the countable (singular and plural) nouns using RT as the dependent variable with the following independent factors:

- Item Type (unmarked singular, marked singular, marked plural)

- Visual Number (visual 1, visual 2)

Table 6: Mean reaction times and accuracy (percentage of accurate responses) for different item types in Experiment 2 displayed as visually single or double (standard errors in parentheses)

Item type		Visual 1 RT (ms)	Visual 1 Accuracy	Visual 2 RT (ms)	Visual 2 Accuracy	Congruency (Visual 1 − Visual 2) RT (ms)	Congruency (Visual 1 − Visual 2) Accuracy
Singular		442 (14)	97.97%	444 (15)	97.03%	−2	0.94%
Unmarked	*rower* 'bike'	440 (14)	97.50%	448 (15)	97.03%	−8	0.47%
Marked	*krowa* 'cow'	444 (15)	98.44%	440 (15)	97.03%	4	1.41%
Marked plural	*klucze* 'keys'	447 (16)	97.11%	446 (14)	97.27%	1	−0.16%
Mass		442 (13)	97.73%	438 (14)	96.80%	4	0.94%
Unmarked	*beton* 'concrete'	442 (13)	97.55%	438 (14)	96.46%	4	1.09%
Marked	*złoto* 'gold'	441 (13)	97.97%	437 (15)	97.24%	4	0.73%
Pluralia tantum	*nożyce* 'scissors'	447 (14)	98.01%	441 (13)	96.38%	6	1.63%
Collective		442 (14)	97.11%	437 (13)	96.56%	5	0.55%
Unmarked	*klan* 'clan'	436 (13)	98.43%	440 (13)	94.79%	−4	3.64%
Marked	*załoga* 'crew'	445 (14)	96.54%	435 (13)	97.32%	10	−0.78%
Neutral words	*czyli* 'so'	448 (14)	98.4%	442 (15)	95.6%	6	2.8%

There was no main effect of Item Type ($F_1(2, 62) = 0.903$; $p = .411$; $F_2(2, 77) = 0.630$; $p = .535$) or Visual Number ($F_1(1, 31) = 0.054$; $p = .817$; $F_2(1, 77) = 0.096$; $p = .757$). There was also no significant interaction between the two factors ($F_1(2, 62) = 1.179$; $p = .314$; $F_2(2, 77) = 0.774$; $p = .465$). The RTs and accuracy for each relevant condition are given in Table 6. A visual inspection of the data suggests that only unmarked singulars showed a noticeable tendency towards the expected congruency effect, while marked singulars presented the opposite behavior and plurals showed very little effect of the visual display manipulation (see Figure 3).

Table 7: Mean reaction times and accuracy (percentage of accurate responses) for countable nouns in Experiment 2 divided by animacy and item type and displayed as visually single or double (standard errors in parentheses)

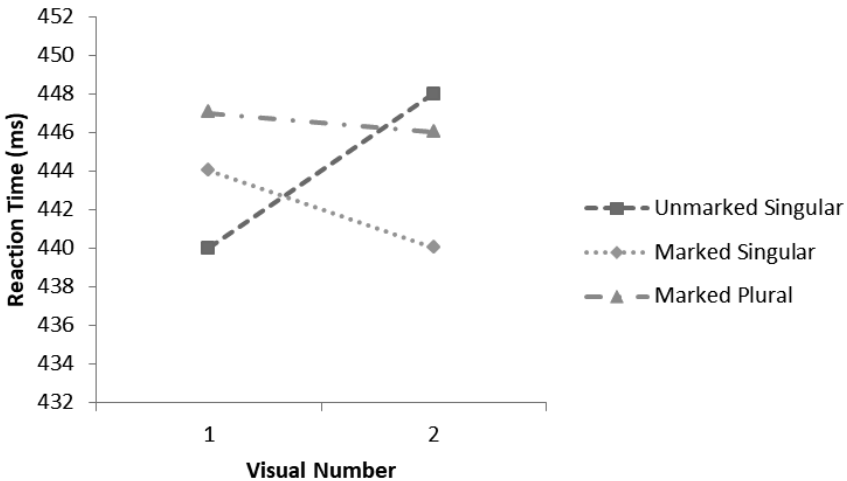|	Item type		Visual number					
			Visual 1		Visual 2		Congruency (Visual 1 − Visual 2)	
			RT (ms)	Accuracy	RT (ms)	Accuracy	RT (ms)	Accuracy
Animate	Unmarked singular	*borsuk* 'badger'	437 (14)	97.18%	444 (16)	96.56%	−7	0.62%
	Marked singular	*mucha* 'fly'	448 (16)	99.37%	442 (16)	97.18%	6	2.19%
	Marked plural	*kruki* 'crows'	452 (18)	98.13%	448 (15)	96.72%	4	1.41%
Inanimate	Unmarked singular	*rower* 'bike'	444 (15)	97.81%	451 (16)	97.5%	−7	0.31%
	Marked singular	*lustro* 'mirror'	441 (16)	97.5%	439 (15)	96.87%	2	0.63%
	Marked plural	*noże* 'knives'	441 (15)	96.09%	444 (14)	97.81%	−3	−1.72%

Figure 3: The interaction of Item Type and Visual Number for countable nouns (unmarked singular, marked singular, marked plural)

4.3.2. Grammatically singular nouns (singular, mass, collective nouns)

To test the possibility that all morphologically marked singular nouns, regardless of their conceptual number, are initially processed in the same way, a $3 \times 2 \times 2$ ANOVA was conducted using RT as the dependent variable with the following independent factors:

- Item Type (singular, mass, collective)

- Morphology (marked, unmarked)

- Visual Number (visual 1, visual 2)

There was no main effect of any factor. The Morphology $\times$ Visual Number interaction was statistically significant by subjects ($F_1(1, 31) = 4.493$; $p = .042$; $\eta_p^2 = .127$) but not by items ($F_2(1, 114) = 2.362$; $p = .127$), indicating that the type of morphology affected the ability to count grammatically singular nouns. Singular nouns with no suffix were on average slightly faster in the visually single than in the visually double condition (a singular congruency effect). In contrast, marked items generated longer reaction times in the visually single than in the visually double condition (see Table 8).

Table 8: Mean reaction times and accuracy (percentage of accurate responses) for marked and unmarked grammatically singular nouns (singular, mass, collective) in Experiment 2 displayed as visually single or double (standard errors in parentheses)

		Visual number					
		Visual 1		Visual 2		Congruency (Visual 1 − Visual 2)	
Item type		RT (ms)	Accuracy	RT (ms)	Accuracy	RT (ms)	Accuracy
Unmarked	*beton* 'concrete'	439 (13)	97.8%	442 (14)	96.1%	−3	1.7%
Marked	*krowa* 'cow'	443 (14)	97.7%	437 (14)	97.2%	6	0.5%

This outcome contradicts the part of the hypothesis concerning the influence of number morphology on the performance in a counting task, which predicted the opposite results for marked and unmarked items. The triple Item Type $\times$ Morphology $\times$ Visual Number interaction was insignificant, suggesting that all three item types were similarly affected by the visual number manipulation (see Figure 4). The complete results of the ANOVA test are given in Table 9 below. The RTs and accuracy for each relevant condition are given in Table 6.
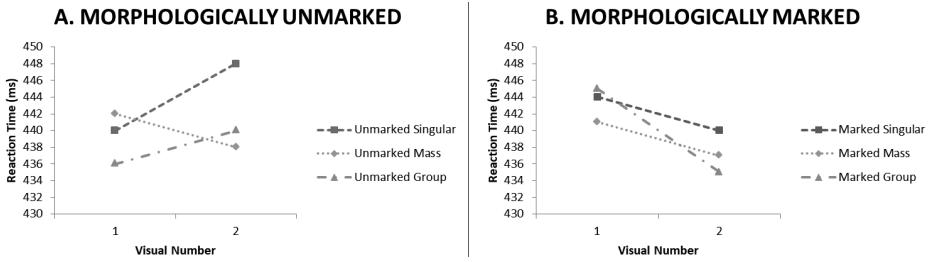
Figure 4: The interaction of Morphology, Item Type and Visual Number for grammatically singular nouns (singular, mass, collective)

Table 9: Results of an Item Type × Morphology × Visual Number ANOVA comparing the mean reaction times for grammatically singular nouns (singular, mass, collective)

Source	df		F		p		Partial eta sq.	
	Subj.	Items	Subj.	Items	Subj.	Items	Subj.	Items
Item Type	2,62	2,114	0.752	0.938	.476	.394	.024	.016
Morphology	1,31	1,114	0.054	0.235	.817	.629	.002	.016
Visual Number	1,31	1,114	0.182	0.124	.672	.725	.006	.001
Item T. × Morph.	2,62	2,114	0.201	0.017	.818	.983	.006	.000
Item T. × Visual N.	2,62	2,114	0.682	0.323	.509	.724	.022	.006
Morph. × Visual N.	1,31	1,114	4.493	2.362	.042*	.127	.127	.020
Item T. × Morph. × Visual N.	2,62	2,114	0.912	0.570	.407	.567	.029	.010

4.3.3. Grammatically plural nouns (marked plural, pluralia tantum)

To find out whether *pluralia tantum* nouns behaved in the counting task like ordinary countable plural nouns, a 2×2 ANOVA was conducted using RT as the dependent variable with the following independent factors:

- Item Type (marked plural, *pluralia tantum*)

- Visual Number (visual 1, visual 2)

There was no main effect of Item Type ($F_1(1,31) = 0.566$; $p = .458$; $F_2(1,82) = 0.731$; $p = .395$) or Visual Number ($F_1(1,31) = 0.246$; $p = .623$; $F_2(1,82) = 0.491$; $p = .485$). The Item Type × Visual Number interaction was also not statistically significant ($F_1(1,31) = 0.432$; $p = .516$; $F_2(1,82) = 0.306$; $p = .581$). The RTs and accuracy for each relevant

condition are given in Table 6. Although *pluralia tantum* and countable plurals did not differ in their interaction with visual number, a very small congruency effect (the difference between the visual 1 and visual 2 conditions) for plurals makes interpreting this test problematic (see Figure 5).
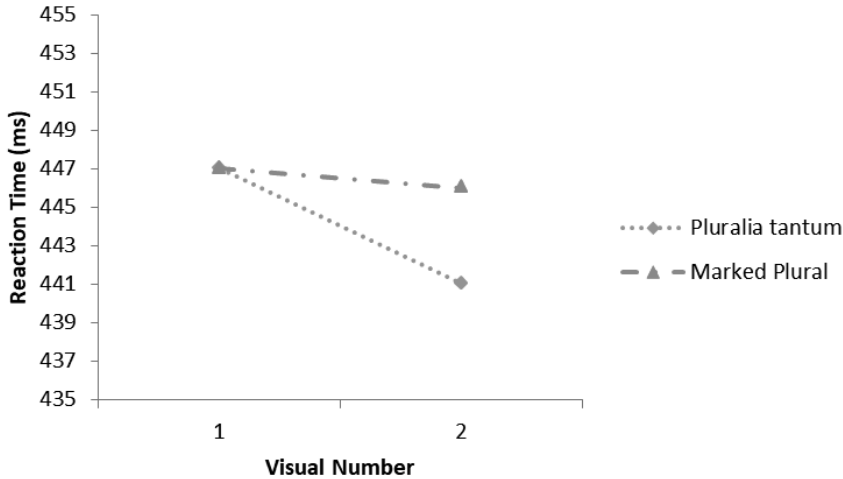

Figure 5: The interaction of Item Type and Visual Number for grammatically plural nouns (marked plural, *pluralia tantum*)

4.3.4. Animacy effect

To check the possible influence of animacy on the performance in the counting task a $3 \times 2 \times 2$ ANOVA was conducted on the countable (singular and plural) nouns only using RTs as the dependent variable with the following independent factors:

- Item Type (unmarked singular, marked singular, marked plural)

- Animacy (animate, inanimate)

- Visual Number (visual 1, visual 2)

None of the main effects or interactions between the factors was statistically significant, indicating that all three item types behaved similarly, regardless of their animacy (see Figure 6). The results of the ANOVA test are given in Table 10. Average RTs and percentage of correct responses are presented in Table 7.
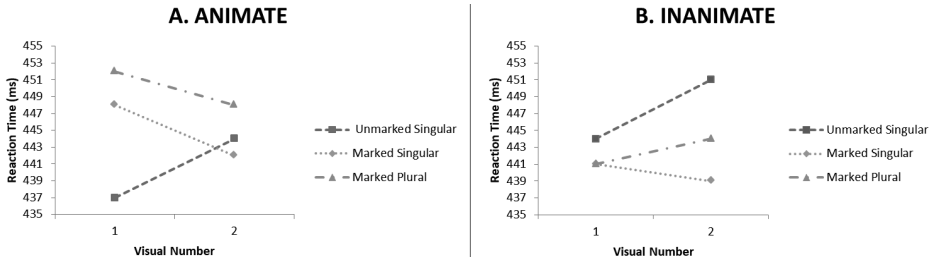
Figure 6: The interaction of Animacy, Item Type and Visual Number for countable nouns (unmarked singular, marked singular, marked plural)

Table 10: Results of an Animacy × Item Type × Visual Number ANOVA comparing the mean reaction times for countable nouns (unmarked singular, marked singular, marked plural)

Source	df		F		p		Partial eta sq.	
	Subj.	Items	Subj.	Items	Subj.	Items	Subj.	Items
Item Type	2,62	2,74	0.795	0.652	.456	.524	.025	.017
Animacy	1,31	1,74	0.252	1.137	.619	.290	.008	.015
Visual Number	1,31	1,74	0.032	0.093	.859	.761	.001	.001
Item T. × Animacy	2,62	2,74	1.719	.1657	.188	.198	.053	.043
Item T. × Visual N.	2,62	2,74	1.033	0.751	.362	.476	.032	.020
Animacy × Visual N.	1,31	1,74	0.397	0.102	.533	.750	.013	.001
Item T. × Animacy × Visual N.	2,62	2,74	0.071	0.185	.932	.831	.002	.005

4.3.5. Rectangles vs. words

In order to test the possibility that the plurality of characters in a string of letters can interfere with the task of counting such strings, a 2×2 ANOVA was conducted on number-neutral items only using RTs as the dependent variable with the following independent factors:

- Item Type (neutral words, rectangles)

- Visual Number (visual 1, visual 2)

There was a main effect of Item Type ($F_1(1, 31) = 60.086$; $p < .001$; $\eta_p^2 = .660$; $F_2(1, 38) = 128.767$; $p < .001$; $\eta_p^2 = .772$). Participants needed on average more time to count rectangles than words. There was also a main effect of Visual Number ($F_1(1, 31) = 11.206$; $p = .002$; $\eta_p^2 =$

.266; $F_2(1, 38) = 21.790$; $p < .001$; $\eta_p^2 = .364$). It took more time for participants to answer that only one token is displayed on the screen than to determine that two tokens are visible.

However, the significant main effects should be approached with caution, because the interaction between the two factors was significant as well ($F_1(1, 31) = 11.001$; $p = .002$; $\eta_p^2 = .262$; $F_2(1, 38) = 13.506$; $p = .001$; $\eta_p^2 = .262$). Manipulating Visual Number produced a bigger effect for rectangles than for words (see Figure 7). Average RTs and percentage of correct responses are presented in Table 11.
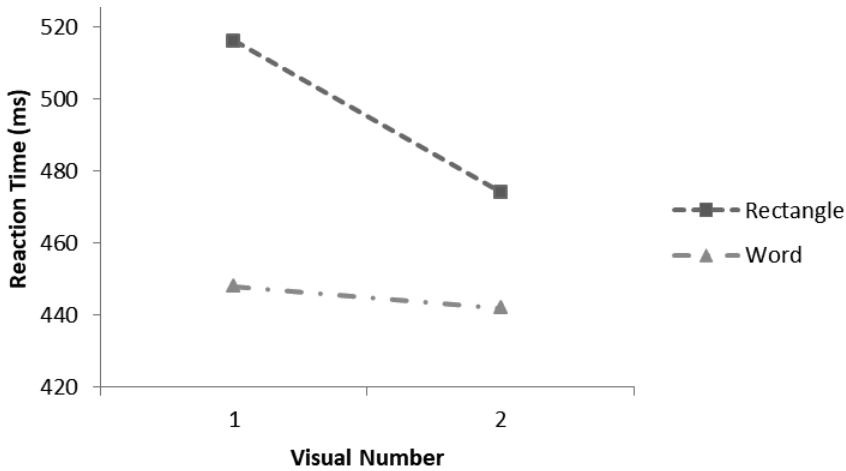

Figure 7: The interaction of Item Type and Visual Number for number-neutral items (rectangles, words)

Table 11: Mean reaction times and accuracy (percentage of accurate responses) for number-neutral items (rectangles and words) in Experiment 2 displayed as visually single or double

Item type		Visual 1		Visual 2		Congruency (Visual 1 − Visual 2)	
		RT (ms)	Accuracy	RT (ms)	Accuracy	RT (ms)	Accuracy
Rectangle		516 (4)	95%	474 (5)	95.9%	42	−0.9%
Word	*przez* 'through'	448 (4)	98.4%	442 (5)	95.6%	6	2.8%

4.3.6. Comparison with number-neutral words

To determine whether any of the major conditions behaved markedly differently from the number-neutral baseline condition, a 9×2 ANOVA was conducted using RT as the dependent variable with the following independent factors:

- Item Type (marked singular, unmarked singular, marked plural, marked mass, unmarked mass, *pluralia tantum*, collective, neutral words)

- Visual Number (visual 1, visual 2)

Number-neutral words were chosen for this test over white rectangles, because they showed significantly less difference between the visual 1 and visual 2 conditions (see section 4.3.5.), which makes them a better baseline condition. There was no significant main effect of Item Type ($F_1(8, 248) = 0.997$; $p = .439$; $F_2(8, 215) = 1.032$; $p = .413$) or Visual Number ($F_1(1, 31) = 0.457$; $p = .504$; $F_2(1, 215) = 0.596$; $p = .441$). The interaction between the two factors was also not significant ($F_1(8, 248) = 0.789$; $p = .613$; $F_2(8, 215) = 0.586$; $p = .789$). The RTs and accuracy for each relevant condition are given in Table 6.

4.4. Discussion

Singular and plural countable nouns in Experiment 2 showed no statistically significant difference in the effects of manipulating their visual numerosity. Only morphologically unmarked singular countable nouns exhibited a trend towards a congruency effect of grammatical and visual number in the expected direction (i.e., they were easier to count when visually single than when visually double). Experiment 2 thus failed to replicate the results of Experiment 1, where a clear congruency effect was found for marked singular nouns and marked plural nouns, but the effect was significantly weaker for unmarked singulars.

In a test limited to grammatically singular (marked and unmarked) nouns, the two-way interaction of morphological markedness with visual number was significant, but the triple interaction with the additional factor of item type was not. This indicates that grammatically singular nouns with a conflict between grammatical and conceptual number (mass and collective nouns) behaved statistically like ordinary countable singular nouns in terms of the Stroop-like interference, which was partly consistent with

the predictions. However, like in the analysis for countable nouns, the effect of morphological marking was contrary to the part of the main hypothesis based on the results of Experiment 1. A singular-number congruency effect (shorter RTs in the visually single condition, longer RTs in the visually double condition) was present only in unmarked nouns, as opposed to the predicted morphologically marked items.

In another test, *pluralia tantum* nouns (grammatically plural, conceptually singular) did not differ from countable plurals. This was in line with the prediction that a conflict between grammatical and conceptual number should not affect the Stroop effect, however the result is difficult to interpret with plural nouns showing virtually no congruency effect between grammatical and visual number.

Finally, no major condition differed statistically from number-neutral words used as a baseline and no significant influence of animacy on the counting task was found.

One possible explanation for the difference between the two experiments presented in this paper could be related to the fact that average RTs in Experiment 1 were approximately 50 ms longer than in Experiment 2. The difference could possibly be due to different lengths of the experiments. A greater number of conditions and trials in Experiment 2 coupled with more breaks may have led to the participants getting better at the counting task. If that was the case, the learning effect might have changed the effects of the interference between grammatical number and visual numerosity in the counting task leading to the null results.

Experiment 2 did manage to answer the question about the multiplicity of characters in letter strings possibly being a source of number interference in a counting task in Experiment 1. There was a significant difference between number-neutral words and visually homogenous rectangles. The difference consisted of a larger advantage for rectangles in the visually double condition, meaning that the general two-item preference observed for number-neutral items in both experiments is not a result of any visual plurality of components. Multiple characters in counted strings do not necessarily lead to an activation of the concept of plurality and, consequently, are not a (major) source of additional Stroop interference in a counting task. The explanation for the observed plural-like behavior of the number-neutral conditions might lie in the fact that the answer 'one' was always made with the left hand while the answer 'two' was always made with the right hand. The "pseudo plural congruency" of number-neutral items may, therefore, reflect a general right-hand preference in the participants, not modulated by any semantic effects.

5. General discussion

The two experiments presented in this paper provided mixed results. Data obtained from Experiment 1 suggest that native speakers of Polish are sensitive to a conflict between the grammatical number and the visual numerosity of stimuli. The conflict manifested itself as slower response times in a counting task when those two values were incongruent (e.g., a singular noun presented as visually double) in comparison with congruent conditions (e.g., a plural noun presented as visually double). This result was a replication of the effect observed by Berent et al. (2005). Additionally, the results of Experiment 1 provide support for the possibility that singular nouns with an overt suffix (e.g., *lekcj-a* 'lesson-NOM.SG') can interfere with the counting task more strongly than unmarked singulars. Experiment 2 was designed to investigate further the processing of overt singular morphology and to test the processing of nouns with a conflict between the grammatical number and the lexical semantics (*pluralia tantum*, collective nouns, mass nouns) in order to determine the time at which the information about grammatical number is activated. The lack of statistical differences between ordinary countable nouns and number-inconsistent nouns, as well as no effect of animacy on the Stroop-like interference, suggest an early automatic extraction of number values blind to lexical semantics. However, very weak congruency effects make any conclusion problematic. Experiment 2 also failed to replicate the effect of morphological markedness from Experiment 1. When all grammatically singular nouns (countable singular, mass, collective) were compared in one test, the result showed a singular-number congruency only for morphologically unmarked nouns, contradicting the findings of Experiment 1. It is possible that different experimental designs put different emphasis on number morphology. This aspect of grammatical number processing still requires more investigation.

References

Berent, Iris, Steven Pinker, Joseph Tzelgov, Uri Bibi and Liat Goldfarb. 2005. Computation of semantic number from morphological informatio. Journal of Memory and Language 53. 342–358.

Bock, Kathryn, Sally Butterfield, Anne Cutler, Kathleen M. Eberhard, J. Cooper Cutting and Karin R. Humphreys. 2006. Number agreement in British and American English: Disagreeing to agree collectively. Language 82. 64–113.

Bock, Kathryn and Kathleen M. Eberhard. 1993. Meaning, sound and syntax in English number agreement. Language and Cognitive Processes 8. 57–99.

Bock, Kathryn, Kathleen M. Eberhard and J. Cooper Cutting. 2004. Producing number agreement: How pronouns equal verbs. Journal of Memory and Language 51. 251–278.

Bock, Kathryn, Kathleen M. Eberhard, J. Cooper Cutting, Antje S. Meyer and Herbert Schriefers. 2001. Some attractions of verb agreement. Cognitive Psychology 43. 83–128.

Bock, Kathryn and Carol A. Miller. 1991. Broken aggrement. Cognitive Psychology 23. 45–93.

Bock, Kathryn, Janet Nicol and J. Cooper Cutting. 1999. The ties that bind: Creating number agreement in speech. Journal of Memory and Language 40. 330–346.

Corbett, Greville G. 2000. Number. Cambridge: Cambridge University Press.

Corbett, Greville G. 2001. Number. In M. Haspelmath, E. König, W. Österreicher and W. Raible (eds.) Language typology and language universals. An international handbook/Sprachtypologie und sprachliche Universalien. Ein internationales Handbuch/ La typologie des langues et les universaux liguistiques. Manuel international (Handbücher zur Sprach- und Kommunikationswissenschaft 20.1). Berlin & New York: Walter de Gruyter. 816–831.

Dillon, Brian, Alan Mishler, Shayne Sloggett and Colin Phillips. 2013. Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. Journal of Memory and Language 69. 85–103.

Durie, Mark. 1986. The grammaticization of number as a verbal category. Annual Meeting of the Berkeley Linguistics Society 12. 355–368.

Eberhard, Kathleen M. 1997. The marked effect of number on subject–verb agreement. Journal of Memory and Language 36. 147–164.

Eberhard, Kathleen M., J. Cooper Cutting and Kathryn Bock. 2005. Making syntax of sense: Number agreement in sentence production. Psychological Review 112. 531–559.

Farkas, Donka and Henriette de Swart. 2010. The semantics and pragmatics of plurals. Semantics and Pragmatics 3. 1–54.

Flowers, John H., Jack L. Warner and Michael L. Polansky. 1979. Response and encoding factors in "ignoring" irrelevant information. Memory & Cognition 7. 86–94.

Friederici, Angela D. 2011. The brain basis of language processing: From structure to function. Physiological Reviews 91. 1357–1392.

Grimm, Scott. 2012. Number and individuation. Doctoral dissertation. Stanford University.

Haspelmath, Martin. 2013. Occurrence of nominal plurality. In M. S. Dryer and M. Haspelmath (eds.) WALS Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Häussler, Jana. 2009. The emergence of attraction errors during sentence comprehension. Doctoral dissertation. University of Konstanz.

Horn, Laurence R. 2001. A natural history of negation. Stanford, CA: Center for the Study of Language and Information.

Humphreys, Karin R. and Kathryn Bock. 2005. Notional number agreement in English. Psychonomic Bulletin & Review 12. 689–695.

Jaensch, Erich Rudolf. 1929. Grundformen menschlichen Seins: (mit Berücksichtigung ihrer Beziehungen zu Biologie und Medizin, zu Kulturphilosophie und Pädagogik). Berlin: O. Elsner Verlagsgesellschaft.

Jakobson, Roman. 1957. Shifters, verbal categories and the Russian verb. Cambridge, MA: Harvard University Press.

Jensen, Arthur R. and William D. Jr. Rohwer. 1966. The stroop color-word test: A review. Acta Psychologica 25. 36–93.

Levin, Magnus. 2001. Agreement with collective nouns in English. Lund: Lund University.

MacLeod, Colin M. 1991. Half a century of research on the Stroop effect: An integrative review. Psychologica Bulletin 163-203. 163–203.

Molinaro, Nicola, Albert Kim, Francesco Vespignani and Remo Job. 2008. Anaphoric agreement violation: An ERP analysis of its interpretation. Cognition 106. 963–974.

Nagórko, Alicja. 2007. Zarys gramatyki polskiej [An outline of Polish grammar]. 7th ed. Warszawa: Wydawnictwo Naukowe PWN.

Naparstek, Sharon and Avishai Henik. 2010. Count me in! On the automaticity of numerosity processing. Journal of Experimental Psychology: Learning, Memory, and Cognition 36. 1053–1059.

Parker, Dan and Colin Phillips. 2016. Negative polarity illusions and the format of hierarchical encodings in memory. Cognition 157. 321–339.

Pavese, Antonella and Carlo Umiltà. 1998. Symbolic distance between numerosity and identity modulates Stroop interference. Journal of Experimental Psychology: Human Perception and Performance 24. 1535–1545.

Pearlmutter, Neal J. 2000. Linear versus hierarchical agreement feature processing in comprehension. Journal of Psycholinguistic Research 29. 89–98.

Pearlmutter, Neal J., Susan M. Garnsey and Kathryn Bock. 1999. Agreement processes in sentence comprehension. Journal of Memory and Language 41. 427–456.

Peirce, Jonathan W. 2007. PsychoPy – Psychophysics software in Python. Journal of Neuroscience Methods 162. 8–13.

Peirce, Jonathan W. 2009. Generating stimuli for neuroscience using PsychoPy. Frontiers in Neuroinformatics 2. 10.

Pęzik, Piotr. 2012. Wyszukiwarka PELCRA dla danych NKJP [The PELCRA search engine for the National Corpus of Polish]. In Przepiórkowski et al. (2012, 253–273).

Przepiórkowski, Adam, Mirosław Bańko, Rafał L. Górski and Barbara Lewandowska-Tomaszczyk (eds.). 2012. Narodowy korpus języka polskiego [National Corpus of Polish]. Warszawa: Wydawnictwo Naukowe PWN.

Smith-Stark, Thomas Cedric. 1974. The plurality split. Chicago Linguistic Society 10. 657–661.

Stroop, J. Ridley. 1935. Studies of inference in serial verbal interactions. Journal of Experimental Psychology 18. 643–662.

Swan, Oscar E. 2002. A grammar of contemporary Polish. Bloomington, IN: Slavica.

Wagers, Matthew W., Ellen F. Lau and Colin Phillips. 2009. Agreement attraction in comprehension: Representations and processes. Journal of Memory and Language 61. 206–237.

Wiese, Bernd. 2011. Optimal specifications: On case marking in Polish. In A. Nolda and O. Teuber (eds.) Syntax and morphology multidimensional. Berlin & Boston: De Gruyter Mouton. 101–127.

Windes, James D. 1968. Reaction time for numerical coding and naming of numerals. ournal of Experimental Psychology 78. 318–322.