Journal of Universal Language 4
September 2003, 147-182

On the Universal Neuropsychological Basis of the Syntax of Numerals

Pawel Rutkowski
Warsaw University

Abstract

The aim of this paper is to relate certain properties of linguistic numeral systems to universal principles of human cognition/perception. As has been noted in the literature (c.f., Greenberg 1978; Ifrah 1985; Hurford 1975, 2001), in many unrelated languages all around the world, numerals referring to cardinalities in the range of 'one' to 'four' differ considerably from higher numerals. It seems that the facts in question can be explained only by means of referring to some 'extralinguistic', universal factors—i.e., to factors that are not part of the language viewed as a closed system of purely arbitrary signs. The analysis argued for in the present paper will relate to the neuropsychological work by Cowan (2001). Basing his arguments independent of any linguistic data, Cowan (2001) suggests that there exists something what he calls 'the magical number four', which constrains human perception skills. He presents a wide variety of data on short-term memory capacity limits and shows that, if factors such as rehearsal and long-term memory are not used to combine stimulus items into chunks, a central short-term memory capacity limit averages to contain about four items. This means that the cardinalities 'one' to 'four' can be perceived independently from actual counting. In this paper, I will attempt to show that this universal property of human perception has influenced the development of many numeral systems. Only the lowest numerals seem to be part of the basic human vocabulary (they appear in most languages of the world—c.f., Ifrah 1985, Dixon 1980); numerals higher than 'four' were introduced to languages only when the speakers of those languages developed arithmetic. These different origins have resulted in different morphosyntactic properties.[1]

Keywords: cardinal numerals, the lower/higher numeral dichotomy, Polish, lexical/functional elements, frequency, magical number four, attention, short-term memory capacity

1. Introduction

In this paper, I will attempt to explain why cardinal numerals do not constitute a coherent morphosyntactic class within individual languages. In many languages, we can observe that this class of lexemes is split into two separate morphosyntactic groups: the lower (referring to the cardinalities 'one' to 'four') and the higher ones ('five' and above). Polish (and other Slavic languages) are an exceptionally salient example of such a situation. The fact that the division exists also in languages that are not genetically related to Polish suggests a universal basis of the phenomenon in question. Heine (1997) briefly addresses the issue and gives his explanation related to the frequency of use. This model will be discussed below. However, Linde-Usiekniewicz & Rutkowski (2003) show that Heine's (1997) analysis does not find confirmation in actual frequency data. Therefore, in the present article, I will propose another hypothesis

[1] This article was written while the author was supported by the Foundation for Polish Science (FNP). Thanks are due for this important source of support. The analysis presented here has benefited from the comments received from Jim Hurford, Jadwiga Linde-Usiekniewicz, Clara Molina, Stanisław Puppel, Tomasz Strzałecki, and Kamil Szczegot.

concerning the origin of the split between the lower and the higher numerals. My approach will be grounded on certain experiments done in the fields of cognitive psychology and neurology. Following the line of reasoning presented in Szczegot (2001), my analysis will try to link the linguistic data (i.e., the syntax of numerals) with the universal computational capacity of the human mind. The description of the computational capacity of the mind will be based mostly on Cowan (2001). He refers to many experiments which show that perception spans are conditioned by a limitation on short-term memory to four elements. Cowan (2001) calls this limitation "the magical number four". I will claim that the "magical number four" must have also indirectly conditioned the diachronic development and the modern use of numerals.

2. The Lower/Higher Numeral Dichotomy

This section gives an overview of the syntax of Polish numerals and compares these observations with some basic data from other languages. I will present a possible synchronic analysis of the facts from Polish carried out in the framework of generative linguistics (Veselovská 2001, Rutkowski 2001, Rutkowski & Szczegot 2001). I want to show that the explanatory power of such an analysis has to be limited unless we go beyond the purely linguistic approach and supplement our framework with tools pertaining to other fields of study.

2.1. The Syntax of Polish Cardinal Numerals

Amongst the lexemes that have traditionally been considered cardinal numerals in Polish, we have to distinguish at least two separate syntactic classes. The numerals *jeden* 'one', *dwa* 'two', *trzy* 'three' and *cztery* 'four' behave syntactically like adjectives. They always agree with the noun they quantify in terms of case:

(1) a. trzej lingwiści
three-NOM linguist-PL-NOM
'three linguists'
b. mądrzy lingwiści
wise-NOM linguist-PL-NOM
'wise linguists'
(2) a. trzema lingwistami
three-INSTR linguist-PL-INSTR
'(with) three linguists'
b. mądrymi lingwistami
wise-INSTR linguist-PL-INSTR
'(with) wise linguists'

On the other hand, in structures containing numerals higher than *cztery* 'four', the case marking of the noun depends on the numeral. In certain contexts, the numeral makes the noun assume the so-called "Genitive of Quantification"—GEN(Q) (c.f., Franks 1995). This case marking pattern occurs only in those structures which are syntactic subjects or accusative objects:

(3) a. pięć kobiet/*kobiety
five-NOM woman-PL-GEN/*woman-PL-NOM
śpi
sleep-3-SING
'Five women are sleeping.'
b. on lubi pięć
he like-3-SING five-ACC
kobiet/*kobiety
woman-PL-GEN/*woman-PL-ACC
'He likes five women.'

In other contexts, the Genitive of Quantification is not assigned, which means that the numeral and the noun have the same case marking. The whole phrase is assigned case from outside as shown in the examples below (the verb *ufać* 'to trust' assigns dative case and the preposition *z* 'with' assigns instrumental case).

(4) a. on ufa pięciu
he trust-3-SING five-DAT
politykom/*polityków
politician-PL-DAT/*politician-PL-GEN
'He trusts five politicians.'
b. ona tańczy z pięcioma
she dance-3-SING with five-INSTR
politykami/*polityków
politician-PL-INSTR /*politician-PL-GEN
'She dances with five politicians.'

We can thus observe examples of two different syntactic relations in the same syntactic structure (namely, government in (3) and agreement in (4)).

In order to distinguish the lower numerals from the higher ones Rutkowski (2001) employs the terms *A-numerals* (e.g., *dwa* 'two') and *Q-numerals* (e.g., *pięć* 'five'). This terminological distinction will be used also in the present article.[2]

[2] Rutkowski & Szczegot (2001) distinguish a third class of cardinal numerals in Polish: *N-numerals*, i.e., the numerals that, syntactically, behave like nouns. Only a few lexemes belong to this class: *tysiąc* 'thousand', *milion* 'million', *miliard* 'milliard', etc. (all of them refer to very high cardinalities). They make the quantified noun assume genitive case in all contexts (similarly to nouns such as *grupa* 'group' or *większość* 'majority'). Therefore, some linguists (e.g., Giusti and Leko 1995) analyse N-numerals as belonging to the lexical class of quantity nouns. In the present article, I will omit this group because it does not influence the proposed analysis.

2.2. Lexicality vs. Functionality: A Generative Approach

Many analyses of the syntax of Polish Q-numerals (e.g., Babby 1988, Franks 1995, Przepiórkowski 1996, Veselovská 2001) make reference to the well-established distinction between structural (grammatical) and inherent (lexical, concrete) cases. In the framework of generative syntax, we assume that structural cases depend on the surface syntactic environment (they mark purely syntactic relations between the lexical elements that constitute a given sentence). On the other hand, inherent cases belong to a "deeper" level of the sentence structure (in the so-called "Government-Binding" approach it is referred to as *D-Structure*, as opposed to *S-Structure* — see Chomsky 1981, 1986). Inherent cases seem to be linked to thematic roles, or, in other words, to semantic relations in a given sentence.

Based on Veselovská's (2001) description of numerals in Czech, Rutkowski (2001) and Rutkowski & Szczegot (2001) consider Polish Q-numerals to be functional elements (on a par with lexemes such as determiners for example). This means that, according to Emonds (2000), they are not part of the deep level of syntactic structure. Functional elements, as opposed to lexical (semantic) elements, appear in the syntactic derivation only at a relatively late, surface stage. The above assumption, combined with Chomsky's (1981, 1986) Government-Binding machinery, allows us to offer an explanation for the distribution of the Genitive of Quantification in Polish. As said above, inherent cases are assigned at a deep syntactic level. Therefore, if the Q-numeral enters the derivation only at a surface level, it cannot assign genitive in the context of inherent cases since the noun has already been assigned the inherent case—see Veselovská (2001), Rutkowski (2001), Rutkowski & Szczegot (2001).

My aim here is not to discuss all the intricacies of the above analysis. However, it seems especially important to note that, for independent reasons (connected to the requirements of a particular model of derivation in generative syntax), the crucial assumption is that Q-numerals are functional (i.e., that they are not "regular" lexical items). If we accept this analysis, we are also forced to say that A-numerals are different from Q-numerals precisely because they are regular lexical items, like adjectives or nouns are, for example.

The above analysis offers a principled means of describing the synchronic dichotomy between the two classes of cardinal numerals in Polish. However, it does not show the origin of the dichotomy. If we constrain our tools to the framework of generative grammar, the only answer to the question about the origin of this phenomenon which we can give is that the dichotomy is driven by the lexicon. In other words, it simply happens that some numerals in Polish are marked in the lexicon as functional and others are not. Therefore, the numeral dichotomy should be considered language-specific and unpredictable. However, as pointed out by Nelson & Toivonen (2000), it does not seem to be a coincidence that A-numerals (i.e., the ones that are marked as lexical) form a sequence, and are not randomly distributed within the set of all numerals.

2.3. The Lower/Higher Numeral Dichotomy in Languages Other Than Polish

The distinction between the four lowest numerals and the higher ones is not a phenomenon restricted to Polish. Typological research shows that it can be traced in many unrelated languages (c.f., Greenberg 1978; Ifrah 1985; Hurford 1975, 2001a; Heine 1997). However, the special character of the lowest numerals may manifest itself in many different ways. It is not only syntactic features, such as the process of case assignment in Polish and other Slavic languages (see Giusti & Leko 1995, Heine 1997, Rutkowski 2000) that differentiates the class in question.

In some languages (e.g., in New Guinea), numerals other than 'one' to 'four' have not developed at all (Greenberg 1978). If speakers of such languages have to refer to cardinalities higher than four, they simply say "many". Thus there exist no fine grained distinctions between higher cardinalities. In languages that do have both lower and higher numerals, the latter are very often more complex morphologically and etymologically they are clearly derived from other words, whereas the lower ones are not etymologically transparent. This can be illustrated by the data from Mamvu (spoken in Sudan) given by Heine (1997). In Mamvu, numerals such as *elí qodè juè* 'seven' (literally, 'a hand seizes two') and *jetò jetò* 'eight' (literally, 'four four') are structurally different from morphologically simplex numerals such as *juè* 'two' and *jetò* 'four'. Hurford (1978) states that the lowest numerals are not complex in any of the languages known to him. Thus, a numeral meaning 'two' will never be derived from an expression such as 'one plus one'.

Many other examples of the difference between the two classes of numerals are connected to flexion. Ifrah (1985) notes that, in Latin, only the four lowest numerals had full declension patterns (*unus, duo, tres, quattuor*). In general, idiosyncratic inflectional forms tend to characterise the lowest numerals. This is illustrated in table (5). The table shows the data on morphological realisation of grammatical gender collected by Hurford (2001a). In many languages, only the numerals that refer to numbers 'one' to 'four' change their forms as a result of gender agreement with the nouns quantified (i.e., only these numerals behave like adjectives).

(5) The number of different gender forms: numerals '1'-'10'[3]

Numeral	Greek	Icelandic	Welsh
'1'	3	3	1
'2'	1	3	2

[3] The data in table (5) refer to the maximal number of oppositions. The fact that there are several different gender forms in nominative does not necessarily mean that the same number of oppositions can be observed in all cases. For example, the Polish numeral *dwa* 'two' has three gender forms in the nominative (*dwaj* 'two-masculine', *dwie* 'two-feminine', *dwa* 'two-neuter') but there are no gender oppositions in the genitive, dative or locative cases. Here only one form is observed, *dwóch*.

'3'	2	3	2
'4'	2	3	2
'5'	1	1	1
'6'	1	1	1
'7'	1	1	1
'8'	1	1	1
'9'	1	1	1
'10'	1	1	1

As shown in table (6), the same regularity holds for the morphological realisation of case (Hurford 2001a):

(6) The number of different case forms: numerals '1'-'10'

Numeral	Greek	Icelandic	Albanian
'1'	3	4	3
'2'	1	4	3
'3'	2	4	1
'4'	2	4	3
'5'	1	1	1
'6'	1	1	1
'7'	1	1	1
'8'	1	1	1
'9'	1	1	1
'10'	1	1	1

Data from Polish may also serve as a good example of the correlation between the greater number of different possible case forms and the fact that a numeral is an A-numeral (as opposed to the fewer possible case forms available to Q-numerals). The correlation is shown in (7), where the A-numeral *dwa* 'two' is compared with the Q-numeral *pięć* 'five'.

(7) The number of different case forms: the numerals *dwa* 'two' and *pięć* 'five' (in virile gender)

Numeral	Number of forms	Forms
dwa	5	*dwaj* (NOM, VOC); *dwóch* (NOM[4], GEN, ACC, LOC, VOC); *dwu* (NOM, GEN, DAT, ACC, LOC, VOC); *dwom* (DAT); *dwoma* (INSTR)
pięć	2	*pięciu* (NOM, GEN, DAT, ACC, INSTR, LOC, VOC); *pięcioma* (INSTR)

It is not an impossible task to find many similar examples in languages all over the world (e.g., Hurford 2001a). Moreover, the fact that the lowest numerals ('one' to 'four') have often been perceived as special is reflected also outside the language system *per se*. The special status of the numerals in question is in a way confirmed by the historical development of the notation of numbers. The earliest notations were based on repeating an element such as a dot or a stroke. But this simple technique was never used to numbers greater than four. For example, the notation that developed in Egypt and Crete made use of grouping: the fifth (and the following) strokes were put on a separate line, thus dividing the row of strokes into groups consisting of not more than four elements (see Ifrah 1985):

(8) The notation used in ancient Egypt and Crete:

I	*II*	*III*	*IIII*	*III* *II*	*III* *III*	*IIII* *III*	*IIII* *IIII*
'1'	'2'	'3'	'4'	'5'	'6'	'7'	'8'

Another way to avoid more than four identical elements in a row was to introduce a separate sign for the number 5. That was, for ex-

[4] In many generative analyses (see Franks 1995, Przepiórkowski 1996, Rutkowski 2000) subjects such as *dwóch* are considered to be accusative rather than nominative. The issue is very complicated and it does not influence the analysis presented here.

ample, how the Roman sign *V* developed (the early notation *IIII* was changed for *IV* much later).

(9) The Roman notation:

I	*II*	*III*	*IIII*	*V*	*VI*	*VII*	*VIII*	*VIIII*
'1'	'2'	'3'	'4'	'5'	'6'	'7'	'8'	'9'

Moreover, Ifrah (1985) attempts to link certain cultural phenomena to the special status of the lowest numerals. For example, in Rome, proper names were given only to the first four sons. The younger brothers were simply called *Quintus* ('the fifth'), *Sextus* ('the sixth'), etc. The same principle applied to the names of months in the Roman 304-day calendar (called the 'Romulus Calendar'). It had only 10 months, to which another two were added later on (*Ianuarius* 'January' and *Februarius* 'February'), forming the 12-month calendar we know today. In the 10-month calendar, only the first four months had "real" names (*Martius*, *Aprilis*, *Maius*, *Iunius*). The next ones were referred to as *Quintilis* (later changed to *Iulius* to commemorate Julius Caesar), *Sextilis* (later changed to *Augustus* in commemoration of the next emperor), *September*, *October*, *November*, *December*. All of these names obviously derive from ordinals.

The data cited above may not seem related to the issue of case assignment in structures containing numerals in Polish. However, they suggest that the split between A-numerals and Q-numerals (as argued for in Rutkowski 2001 and Rutkowski & Szczegot 2001) reflects a much broader, possibly universal, phenomenon connected to the way people use and perceive, or rather used to use and perceive, numerals and numbers. This means that the numeral dichotomy in Polish cannot simply be said to be idiosyncratic or merely lexical. If we want to find a plausible and non- framework-specific explanation for the issue at hand, we must go beyond grammar and take a closer look at human numerical competence.

3. An Analysis Based on Frequency

One of very few attempts to give an independent explanation of the numeral dichotomy discussed above has been made by Heine (1997). He has suggested a model based on frequency factors. This will be discussed briefly in the section below. In particular, it will be shown (after Linde-Usiekniewicz & Rutkowski 2003) that Heine's (1997) analysis does not find empirical support.

3.1. Grammaticalisation: Heine's (1997) Model

Drawing on observations made by Greenberg (1978) and Corbett (1978), Heine (1997) notes that one of the most common features of cardinal numerals is that they behave to some extent like nouns and to some extent like adjectives. The fact that numerals are nominal modifiers which often agree with the head noun seems to be an adjective-like property. On the other hand, acting as a case-assigner is rather a noun-like characteristic. As it was said in the previous chapter, the adjective-like behaviour is associated with the lowest numerals.

Heine (1997) tries to explain the mixed adjectival/nominal status of numerals by describing their diachronic development in terms of a model of grammaticalisation. In this model, the semantic change from a more concrete meaning to a more abstract one may result in a change in the categorial status of a given element. For example certain concrete nouns may become abstract adjectives. (See also Heine, Ulrike, & Hünnemeyer 1991). Under the assumption that, in many languages, numerals derive from nouns, Heine (1997) claims that the adjectival characteristics of certain numerals must indicate that their meaning have become more abstract. This is supported by data taken from a number of African languages, in which nouns such as 'hand' and 'man' often become numerals meaning 'five' and 'twenty', Heine (1997). This is in line with a widely-accepted view that the use of an adjective is the most unmarked means of referring to abstract features (Hurford 1987).

Heine (1997) draws a parallel between the above model and the diachronic development of certain colour terms. Colour terms are often derived from nouns denoting fruits or flowers (e.g., the adjective *orange* derives from the name of the fruit). The conceptual shift changes a real object (a fruit) into an abstract quality, which, in a way, is considered independent of the original object. In a similar way, the concrete meaning of a noun such as *hand* changes, as a result of grammaticalisation, into the abstract quality of being ‘five’. This feature can then be applied to other objects (not only fingers) and can be used attributively. Hence, in terms of syntax, it becomes an adjective, modifying other nouns.

The above model seems plausible. However, it also needs to explain why, in languages such as Polish, not all numerals have become syntactic adjectives but only the lowest ones. Heine (1997) assumes that what we actually have is a continuum; the lower the cardinality that a numeral refers to, the more adjective-like the numeral is. Therefore, numerals denoting the highest cardinalities are the most noun-like (see footnote 1 above) and the numerals ‘one’ to ‘four’ resemble adjectives. According to Heine (1997), the factor that shapes this continuum is the frequency of use of the numerals. Many theorists have claimed that frequency influences the process of grammaticalisation (i.e., that high frequency of use is positively correlated with the possibility of a lexical item becoming grammaticalised; see, Hurford 1987, Bloom 2000, and Heine, Ulrike & Hünnemeyer 1991). Thus, Heine (1997) suggests that the fact that the lowest numerals are used much more often than the higher ones results in their undergoing grammaticalisation much earlier. Similarly, not all colour terms become adjectives with the same ease. For example, in Polish the word *indygo* ‘indigo’ is used very rarely so it should not be surprising that it does not have an adjectival paradigm (we could speculate that if it was used more often it would assume an adjectival ending, e.g., *indygowy*).

However, Linde-Usiekniewicz & Rutkowski (2003) show that, if such an analysis is applied to Polish numerals, it does not readily find support in empirical frequency data.

3.2. Synchronic Frequency Data

Linde-Usiekniewicz & Rutkowski (2003) challenge the model proposed by Heine (1997) with data taken from the frequency dictionary by Kurcz *et al.* (1990). They show that the frequency of use of numerals in present-day Polish cannot be said to depend on the cardinality to which a given numeral refers.

As mentioned in the previous section, according to Heine's (1997) model, the numerals 'one' to 'four' became adjectival earlier than other numerals because they are used much more often. From this point of view, the most interesting piece of data that we should have a look at is the difference in frequency between the numerals *cztery* 'four' and *pięć* 'five'. In order to confirm Heine's (1997) hypothesis, we should be able to show that the difference in question is significantly big. However, Linde-Usiekniewicz & Rutkowski (2003) show that the frequency of use of the numeral *pięć* 'five' is actually higher than the frequency of the numeral *cztery* 'four'. The table below (based on Kurcz *et al.* 1990) shows the total number of appearances of numerals denoting the numbers 2-10 in the corpus consisting of 500,000 words.

(10) Total frequency of the cardinals *dwa* 'two' – *dziesięć* 'ten'

Numeral	Frequency
dwa 'two'	936
trzy 'three'	568
cztery 'four'	373
pięć 'five'	431
sześć 'six'	240
siedem 'seven'	164

osiem 'eight'	221
dziewięć 'nine'	135
dziesięć 'ten'	202

Heine's (1997) prediction is right as far as the fact that A-numerals appear in Polish texts more often than, the Q-numeral *siedem* 'seven' for example. But the data in the above table show that the numeral *pięć* 'five' is also one of the most often used cardinals. It seems that numerals referring to the number '5' tend to have a relatively high frequency of use crosslinguistically. For example, Nelson & Toivonen (2000) note that in a frequency dictionary of Spanish (Juillard & Chang-Rodriguez 1964) the numeral *cinco* 'five' has a higher frequency than the numeral *cuatro* 'four'. Moreover, the data discussed by Linde-Usiekniewicz & Rutkowski (2003) show that all numerals connected to the base number '5' are used relatively often. In each series of numerals that can be distinguished in Polish in terms of morphology (*jeden-dziewięć* 'one-nine', *jedenaście-dziewiętnaście* 'eleven-nineteen', *dziesięć-dziewięćdziesiąt* 'ten-ninety' and *sto-dziewięcset* 'one hundred-nine hundred'), the numerals that include *pięć* 'five' appear more often than those that include the numerals *cztery* 'four' or *sześć* 'six'.

(11) Total frequency of cardinals that include the numerals *cztery* 'four', *pięć* 'five' and *sześć* 'six'

Numeral	Frequency	Numeral	Frequency	Numeral	Frequency
czternaście 'fourteen'	46	*czterdzieści* 'forty'	185	*czterysta* 'four hundred'	65
piętnaście 'fifteen'	109	*pięćdziesiąt* 'fifty'	262	*pięćset* 'five hundred'	102
szesnaście 'sixteen'	39	*sześćdziesiąt* 'sixty'	229	*sześćset* 'six hundred'	67

Other data also seem to contradict Heine's (1997) generalisation. For example, the lexemes *dziewięćset* 'nine hundred' (572 appearances in the corpus used by Kurcz *et al.* 1990) and *tysiąc* 'thousand' (1207 appearances) are amongst the most frequently used cardinals despite Heine's (1997) model predicting that they should be used very rarely.

Linde-Usiekniewicz & Rutkowski (2003) interpret the above data as illustrating how extralinguistic factors influence the frequency of use. The fact that numerals derived from the number '5' are used very often is obviously related to its base function in the decimal system that we use (it appears in approximations etc.). In addition, the high frequency of the lexemes *tysiąc* 'thousand' and *dziewięćset* 'nine hundred' seems to be driven by non-linguistic factors. The corpus used by Kurcz *et al.* (1990) was collected in the 1960s. At that time, the numerals meaning 'thousand' and 'nine hundred' must have been used very often in dates (for example *rok tysiąc dziewięćset sześćdziesiąty* meaning 'year nineteen sixty') and while counting money (many goods in Poland cost at least a few thousand zlotys). Pisarek (1972) shows that the use of certain words is in a way "independent" from the speaker's intentions. Instead, it is driven by the social context. We can, for example, predict that in the present century the frequency of use of the numeral *dziewięćset* 'nine hundred' will by much lower than it used to be because it is no longer used in referring to contemporary dates.

Linde-Usiekniewicz & Rutkowski (2003) conclude that Heine's (1997) model can be considered plausible only under the assumption that it applies to early stages of language development, i.e., to a historical period when counting was not as widespread as it is now. The ability to count inevitably results in a great number of numeral data being used in every-day conversations (money, time, calendars). Linde-Usiekniewicz & Rutkowski (2003) illustrate this point further with frequency data concerning colloquial Polish taken from a frequency dictionary by Zgółkowa (1983). The special status of lexemes such as *pięć* 'five' is much less salient there. It could be said that colloquial varieties of language are less exposed to the influence of the civilisation (approximations and dates appear more often in news or scientific texts). However, by no means does it seem justified to claim that, at any stage of the development of numerical competence, the frequency of use of the numerals referring to the numbers 'one' to 'four' could be so much higher than the frequency of use of the other numerals that the difference must have resulted in the syntactic dichotomy described in chapter 2 of this paper. According to Heine (1997), the lowest four numerals must have been used so often that, unlike in the case of the other cardinals, it was possible for their meaning to shift from the concrete to the abstract and for them to obtain the syntactic features of adjectives. Such a claim does not find support in any frequency data, neither in the standard, nor in colloquial varieties of Polish. Moreover, if Heine's (1997) explanation was correct, we should observe crosslinguistic variation in terms of where the border-line between adjective-like and noun-like numerals is. But it seems that, if such a border exists, it usually divides the set of numerals in exactly the same way (namely, distinguishing the lowest four numerals, and not, for example, the lowest nine or thirteen from the rest). If we followed Heine's (1997) analysis, we would have to conclude that it was always the discrepancy between the frequency of use of the numeral meaning 'five' and the numeral meaning 'four' (and no other numerals) that happened to be significant in many unrelated languages. This seems rather improbable.

4. A Psycholinguistic Approach

Szczegot (2001) proposes that the split between A-numerals and Q-numerals in Polish might be driven by neuropsychological processes. Drawing on observations made by cognitive psychologists, he views the dichotomy in question as connected to the limitations of human perception. This line of reasoning will be thoroughly discussed and elaborated on in the rest of this article. The discussion will be illustrated with the neuropsychological data taken from the experiments described by Cowan (2001). I will try to show that such an analysis can be combined with the generative model of the syntax of numerals proposed by Rutkowski (2001) and Rutkowski & Szczegot (2001).

4.1. Limitations on the Human Capacity to Store and Process Information: Cowan (2001)

James (1890) notes that there exists a difference between what could be called primary and secondary memory. The former has a limited capacity, whereas the latter, thanks to complex ways in which the data may be stored and processed, gives an almost unlimited number of different possibilities of memorising things. Miller (1956) follows the same track and describes short-term memory as limited with respect to the number of elements which can be stored in it at the same time. He argues that the limit can be approximated as 7±2 elements. However, Cowan (2001) aims to show that Miller's (1956) "magical number 7" does not account for the facts precisely enough. However, it correctly expresses a general intuition that the short-term-memory-based perception is limited. Nonetheless, Cowan (2001) points out that the perception spans are usually broadened by means of grouping, repeating, reference to the long-term memory, etc. Therefore, the "pure" limit of perception can be noticed only under specific conditions. Let us take a brief look at a few of the necessary restrictions that Cowan (2001) mentions if one is to obtain reliable data. For example, one of the methods that was used in the experiments that he discusses was an 'information overload'. This procedure eliminates repetition during the experiment because a great number of stimuli given at the same time limits the possibility of complex computation. Cowan (2001) also points out that the influence of long-term memory can be minimised by introducing an additional task (that is not part of the experiment, for example the simultaneous repetition of one word during the experiment (Baddeley 1986). Another important condition that was met in the experiments was the elimination of grouping (Cowan 2001). Subjects were not able to associate the elements on which the experiment was based with each other. If they were, greater chunks of information could potentially be created. Cowan (2001) illustrates such a process with the example taken from Miller (1956): the string "fbicbsibmirs" can be remembered much easier if we notice that it could be divided into the well-known abbreviations FBI, CBS, IBM and IRS. This simple grouping results in creating four elements instead of twelve. Further grouping is also possible: FBI and IRS are American governmental institutions, whereas IBM and CBS are big companies. Such associations make the twelve-letter string quite easy to remember. Therefore, in the experiments which were meant to show the limited capacity of human perception, the possibility of grouping had to be eliminated.

Cowan (2001) mentions many experiments that were conducted that kept the above (and many other) conditions in mind. Even the earliest and the least complicated of them show that there exists a clear limitation on our ability to perceive multi-element sets. A brief glance at a group of identical objects is usually enough to say that the group consist of one, two, three or four elements. Ifrah (1985) calls it the "direct perception of number", whereas Ullman (1984) uses the term "visual counting". On the other hand, sets consisting of more than four elements are much more problematic as far as immediate counting is concerned. Without additional time (necessary to make use of more complex mental procedures such as grouping etc.), humans are not able to say if a given set contains, exactly nine or eleven elements (Gelman & Gallistel 1978 and Mandler & Shebo 1982). This is well illustrated in the classic experiment by

Jevons (1871). The experiment was very simple; Jevons threw a handful of beans into a box, had a glance at them for a short while and tried to approximate their number. Then, he counted the exact number of beans and compared the results. Having repeated the trial for more than a thousand times, Jevons (1871) concludes that numbers smaller than five could always be estimated perfectly accurately, whereas it was not so in the case of greater numbers.

It is possible to find many similarly simple facts that confirm Jevons's (1871) observation. For example, telephone numbers are always divided into sequences of no more than four digits. But Cowan (2001) also gives examples of experiments that were far more complex and makes use of modern technologies. For instance, one experiment conducted by Cowan *et al.* (1999) had a form of a computer game. The aim was to compare the name of the picture that appeared in the centre of the screen to the names of four surrounding pictures and indicate (with a mouse click) the name of which of the surrounding pictures rhymed with the central one. The game was played many times and its purpose was not to let subjects concentrate on the lists of digits that they were simultaneously listening to through headphones. From time to time, the rhyming game disappeared from the screen and the subject was asked to recall the digits from the spoken list. Regardless of the length of the spoken list, the average number of items recalled by subjects was about 3.5.

Pylyshyn & Storm (1988) and Yantis (1992) conducted another interesting experiment. The task involved what is often referred to as "multi-object tracking." Before the experiment, a group of small objects was shown to the subject. Some of them flashed several times and then all of them wandered randomly across the screen. When they stopped, the subject was asked to report which objects had flashed. If the number of elements to be tracked was bigger than four out of ten, then the subjects judged the task as extremely difficult.

Many of the experiments mentioned by Cowan (2001) show that, even if subjects are capable of operating with more than four elements, their performance deteriorates significantly as the number of elements involved in the experiment increases from four to five. In an experiment described by Luck & Vogel (1997), subjects were shown an array of one to twelve small coloured squares for 100 milliseconds, followed by a 900-millisecond blank pause and another array of coloured squares which was exactly the same as the first one or differed in the colour of one square. The subjects' task was to say whether the arrays were the same or different. Luck & Vogel (1997) report that the subjects' performance was almost perfect for arrays of one to three squares, slightly worse for four-square arrays, and significantly worse for arrays consisting of more than four elements.

Logan (1988) conducted an experiment in which subjects were asked to say if equations such as "B + 3 = E" (which means that, in the alphabet, E is the third letter after B) are correct. If the number that indicated the distance was higher than four, the performance was much worse than in the case of the lower numbers. The majority of subjects reported that the equations with five were very difficult and that, in order to make the task easier, they tended to memorise the results, whereas the equations with the addends of two, three or four could be solved 'automatically').

As can be easily seen, very similar results were obtained from all the independent experiments mentioned above. Cowan (2001) gives many more examples. Therefore, Cowan (2001) concludes that the focus of attention, or, as Vogel, Luck & Shapiro (1998) call it, the visual working memory must be capacity-limited. The average number of stimulus items that were recalled correctly in the experiments brought together by Cowan (2001) fell within the range of 3-5 items. The "magical number four" means that humans are able to hold only about four chunks in a pure capacity-limited short-term memory, which, according to Cowan (2001), equals the focus of attention. This results in the fact that a human being can concentrate on only a limited number of elements at a time and leads to what Cowan (2001) refers to as scene coherence. It further supports and illustrates Mandler's (1985) example of 'a scene in a park'. The observer of such a scene may be conscious of "four children playing hopscotch, or of children and parents interacting, or of some people playing chess; but a conscious content of a child, a chessplayer, a father, and a carriage is unlikely (unless of course they form their own meaningful scenario)" (Mandler 1985:68). Cowan (2001) notices that a coherent scene can be viewed as limited to four separate elements if there are no special associations between them that could make the observer's awareness of each of them mutually dependent. Thus, coherent scenes are limited because they are formed in the focus of attention.

Cowan (2001) attempts to explain why the limit equals exactly four. He notes that there exist purely mathematical arguments for the idea that the size of four is optimal for working memory. Dirlam (1972) shows that the most efficient basis of a memory search is a chunk of about four items. The average of 3.59 minimises the number of accesses to the information stored in memory. MacGregor (1987) drew a similar conclusion. He calculated that organising list items into higher-level chunks is only efficient with an exhaustive search when there are more than 4 items. Therefore, a one-level search system is more efficient with chunks of the size of four. Thus, the "magical number four" seems to be the optimal number of items that can be beneficially processed in an ungrouped manner.

There are also some facts that seem to indicate that the "magic" of "magical number four" has a neurophysiological basis. Cowan (2001) analyses in detail a few experiments whose aims were to establish a connection between short-term memory and the activity of gamma waves in the brain. For instance, he reports on an experiment in which a set of elements is represented in the short-term memory as a low-frequency, five to twelve Hertz, oscillation, whereas a separate element is stored in a high-frequency (40 Hertz) subcycle of that low-frequency oscillation. The maximal number of elements that may be stored in the short-term memory is, thus, dependent on how many high-frequency subcycles fit into one low-frequency cycle. Cowan (2001) shows that this model can be used to motivate the existence of a memory span of about four elements. If we assume that the low frequency is around 10 Hetz, what we get is the following equation: 40 subcycles/second / 10 cycles/second = 4 subcycles/cycle. Both these neurophysiological facts and the mathematical findings discussed briefly above support Cowan's (2001) claim that the empirically observed capacity limit of the focus of attention is grounded in the computational limitations of the human brain.

4.2. The "Magical Number Four" and the Syntax of Numerals

The idea of the neuropsychological "magical number four" argued for by Cowan (2001) seems to be very insightful as far as the numeral-dichotomy puzzle is concerned (discussed in section 2 above). None of the experiments mentioned by Cowan (2001) referred to linguistic data. Cowan's (2001) aim was to describe a phenomenon that is much broader than the syntax of numerals; namely, the limitations of human perception and attention. However, these independently motivated findings let us interpret the distinction between the lower (lexical) and the higher (functional) numerals which has been proposed by Rutkowski (2001) and Rutkowski & Szczegot (2001) not as an arbitrary and unpredictable idiosyncrasy of Polish, but rather as an effect of universal mental processes.[5]

[5] A similar type of analysis (linking linguistic and neurological data) has been recently proposed by Hurford (2001b). He shows that the basic predicate-argument structure of a sentence in natural languages might be conditioned by the attention limit described as the "magical number four". Such a basic structure consists of maximally four elements (such as subject, object etc.), and, according to Hurford (2001b), seems parallel to the structure of a single scene in our perception (c.f.,

It is beyond any doubt that the ability to count is not encoded genetically. There are societies (for example certain tribes in Australia, the Amazon Forrest and the Murray Islands) that have not developed an abstract idea of number (see Ifrah 1985, Dixon 1980). We can, thus, suppose that, in the past, the same was true of all humans. However, as said in section 2.3, even in societies that do not make use of arithmetic knowledge, people have no problems with referring to the numbers one, two, three and four.

Drawing on the work by Cowan (2001), we can try to link the linguistic numeral data with a universal mental mechanism. Szczegot (2001) points out that, thanks to the basic properties of the short-term memory, the lowest four numerals are easy to distinguish without counting. Therefore, they might be viewed as words denoting a characteristic of a given set (a set consisting of two elements is perceived as being different from a set that consists of three elements, in the same way as a set of red elements is different from a set of green ones, (see also Hurford 2001a). The lowest four numerals might be considered as terms that describe features which are easily perceivable. Ifrah (1985) compares the perceptual mechanism that lets us comprehend the lowest numerals to senses such as hearing or smelling. The fact that the lowest four numerals have appeared at some point in the languages of the world should, therefore, be considered very straightforward (similarly, it seems natural and straightforward that languages have developed the words meaning 'big' or 'small'). Of course, the numerals denoting the cardinalities one to four differ from typical adjectives because they refer to a feature of

---

Mandler's (1985) idea that the capacity of perception limits the number of items that can form a coherent scheme). Therefore, if humans were able to perceive a single scene as consisting of more than four basic elements, the sentence structure could also be more complex. This view is very similar to the one argued for in the present article; syntactic phenomena reflect (philogenetically and ontogenetically) certain mental representations and structures. These representations cannot be considered part of the language faculty. They are rather some sort of mental basis on which constructions that are purely linguistic can be built.

the whole set, and not of one element (apart form the numeral meaning 'one'). But it is worth noticing that, in many languages, similar adjectives exist also outside the counting system, for example, the English adjective *numerous*.

The ease with which a given number can be perceived is precisely something that differentiates the lowest four numerals from the higher ones. In sets consisting of more than five elements, the elements have to be counted, i.e., they are perceived in a different, more complex, way than the one described above. Counting is an abstract operation that requires advanced data processing (e.g., grouping, comparing sets, recalling earlier information that has been previously memorised).

The lowest numbers seem to be distinguishable not only for humans. Hauser, Carey, & Hauser (2000) show that the animal brain can also conceptualise the quantities which we describe with the numerals *one* to *four*. The three authors conducted an experiment with monkeys who appeared to be able to distinguish between two-, three- and four-element sets of fruits. However, the monkeys were not seen to respond to differences between sets consisting of five, six and more fruits. Ifrah (1985) mentions that some birds (nightingales and magpies) also differentiate between the quantities from one to four.[6]

However, Nelson & Toivonen (2000) note that no animals can distinguish quantities on the basis of a counting sequence. They claim that this method is only available to humans, who associate a particular symbol (a numeral) with a given quantity (interpretable only due to the fact that it appears in a sequence). Numbers higher than four cannot be perceived directly so their linguistic representa-

[6] It is necessary to distinguish this type of "counting" from the one that is based on proportions (it is also available for animals). Nelson & Toivonen (2000) give an example of rats which can, in the same way, notice that there is a difference between a one-element set and a two-element one and between an eight-element set and a sixteen-element one.

tions do not refer to anything "real". Therefore, it seems probable that, at early stages of language development, the frequency of use of the numerals higher than *four* was very low. Such a hypothesis cannot, however, be claimed (as it is in Heine 1997) to show the cause of the numeral dichotomy. If the frequency difference had really existed, it would have rather been an effect, and not the cause, of the dichotomy.

As I mentioned earlier, higher numerals refer to certain positions in a conventional sequence. Humans were able to create such sequences because they used fingers or recited sequences of words (similar to today's nursery rhymes—c.f., Ifrah 1985, Hurford 1987). Cowan (2001) reports on an experiment in which the focus of attention was significantly supplemented thanks to the fact that subjects were trained to use their fingers to assist them in remembering some of the information during the task (see also Reisberg, Rappaport, & O'Shaughnessy 1984). This seems similar to the situation in the beginnings of human counting. Ifrah (1985) claims that humans have developed arithmetic precisely because their basic perception was limited to four elements. Such a model of the development of counting is parallel to Dehaene's (1997) assumption that it was only small numbers, or rather quantities, that obtained to mental representations in the early periods of the development of human perception.

When words meaning 'five' or 'seven' appeared in natural languages, their status had to be different from the status of the lowest four numerals. The difference could have been realised as the adjective vs. noun opposition discussed in section 3.1 above. However, it has to be explained why the impact of the "magical number four" can now be seen only in the syntax of some languages. In many other languages, the distinction between the lowest numerals and the other numerals simply does not exist. Szczegot (2001) (similarly Heine 1997) does not attempt to address this question. His reasoning leads to the conclusion that, for some reason, the capacity-limited focus of attention described by Cowan (2001) has more influence on speakers of Polish than, e.g., speakers of English. This does not seem very plausible.

In order to tackle the above issue, we have to assume that the dependence of the syntax of numerals on the properties of the universal capacity-limited focus of attention is a historical phenomenon. In the course of its development, the human mind has become capable of going beyond the limits of the "magical number four". This is why, now, the limitation cannot be easily traced; the experiments brought together by Cowan (2001) were necessarily complicated. One experiment conducted by Ericsson, Chase, & Faloon (1980) shows that humans can, with the use of certain mnemonic strategies, recall strings of more than 80 elements. The subjects in this experiment formed meaningful groups of elements, and then further combined the groups into what can be called "supergroups." Both the group and the supergroup levels were still subject to the limit of the "magical number four" however, so short-term memory capacity could not be said to have increased. However, thanks to hierarchical grouping, the subjects were able to remember strings that would normally seem impossible to remember. This example shows that the "magical number four" has ceased to be a barrier that could not be crossed. In our times, the human brain can process numbers and numerals higher than four with no difficulty. The perceptual difference between the four- and five-element sets have become fuzzy. Therefore, it cannot be considered surprising that in many languages the distinction between the lowest four numerals and the higher numerals has disappeared, for example by making all numerals adjectival, as predicted by Heine's (1997) model.

Keeping the above in mind, we have to consider the linguistic phenomena described in section 2 of this paper to be a result of the fact that some languages are more "conservative" than others. Hurford (2001a) illustrates this with a metaphor borrowed from Wittgenstein's *Philosophical investigations* (Wittgenstein 1953); Language is like an ancient city. In the centre, we find a chaotic maze of buildings from different periods of the city's history. However, the centre is surrounded by suburban districts, which are new, uniform and regular. In natural languages, the lowest numerals seem to belong to the oldest part; the core vocabulary. Their properties show that they must have appeared in languages much earlier than the rest of what we now today consider to be the category 'numerals'. I claim that this category is conditioned by the limitations of human perception. However, the development of more complex computational mechanisms in the brain made the "magical number four" lose its impact on the short-term memory and, indirectly, on the way people perceived numbers. Thus, the distinction between A- and Q-numerals in Polish has to be viewed as an example of the remnant of a universal mental phenomenon that was active a long time ago.

What has been claimed so far means that numerals referring to numbers higher than four appeared in languages together with the knowledge of counting. If we assume that ontogenesis reflects philogenesis, we could try to look for an independent confirmation of the above analysis in the data on the acquisition of numeric competence in children. Nelson & Toivonen (2000) elaborate on experiments which show that children have some "pre-linguistic" awareness of numbers. In the seventh month of life, their ability to perceive numbers is similar to the one noticed in monkeys by Hauser, Carey & Hauser (2000); these very young children were able to distinguish sets consisting of two and three elements (see also Wynn 1995). The first numerals they start to use are, therefore, the lowest numerals.

As early as at the one-word stage of language acquisition, children seem to be able to notice that a given word is a numeral. Nelson & Toivonen (2000) describe this phenomenon as the awareness of the syntactic context of quantification. Many experiments show that children can understand the idea of a numeral being a word that refers to something plural. Later on, they learn to count and use numerals by linking a particular lexeme to a given set (see Bloom

2000). It is only then that children start to use numerals denoting numbers higher than four. The ability to process such numerals requires some basic knowledge of arithmetic. The above observations find support in frequency dictionaries (see Linde-Usiekniewicz & Rutkowski 2003). Zgółkowa & Bułczyńska (1987) show that the numerals that are most frequently used by children are the lowest ones. In contrast to the adult data collected by Kurcz *et al.* (1990), the influence of social factors seems minimal; small children do not usually talk about dates or money. Children learn how to use numerals such as *five*, *fifteen* or *fifty* when they learn mathematics (Ifrah 1981, Nelson & Toivonen 2000). The two cognitive systems (the system of numerals referring to numbers higher than four and arithmetic) seem interdependent; children cannot learn arithmetic without using numerals but they also cannot understand numerals without any reference to the extra-linguistic system of arithmetic. This is why the two systems are learnt at the same time.

I would like to consider the above analysis fills a gap in the syntactic description of numerals presented by Rutkowski (2001) and Rutkowski & Szczegot (2001). The fact that A-numerals (from *jeden* 'one' to *cztery* 'four') have to be treated as lexical (and not functional) elements seems, in the context of Cowan's (2001) findings, much more plausible. Rutkowski (2001) and Rutkowski & Szczegot (2001) assume that lexical elements convey the main semantic content of the sentence. In order to form a grammatical sentence, they have to be in a way supported by functional elements (which add information about tense, reference, syntactic relations within the basic argument-predicate structure, etc.). Lexical elements can be compared to bricks that have to be put together with the use of cement (i.e., functional information). As argued for above, A-numerals denote features that can be perceived sensually. They are part of the basic human vocabulary. Therefore, it seems justified to consider them as being elements that form the semantic core of the sentence in a fashion similar to other adjectives or nouns, other lexical elements).

It might be claimed that numerals that refer to numbers higher than four (Q-numerals) have a completely different status in Polish. Under the assumption that linguistic competence is different from mathematical competence (see Chomsky 1980), Q-numerals must be an intersection of the two systems (see the discussion in Hurford 1987). They are not part of the universal basic vocabulary of natural languages. The way they are interpreted must be different from the way lexical elements such as nouns or adjectives are interpreted. Q-numerals denote something that cannot be understood without the system of arithmetic. The fact that Q-numerals are dependent on the knowledge of counting results in their relatively late appearance in languages (including an ancestor of Polish). In the beginning, they must have been nouns. However, their properties made it possible for them to turn into functional elements.[7] From a diachronic point of view, this happened in Polish only a few centuries ago; Q-numerals lost many of their nominal characteristics, for example they ceased to assign genitive in the context of inherent cases (Linde-Usiekniewicz & Rutkowski 2003). Rutkowski (2002) analyses the diachronic difference between 16th century and Modern Polish as a side-effect of the process of grammaticalisation. According to Roberts & Roussou (1999), many processes that have traditionally been referred to as grammaticalisation involve reanalysis of lexical material as functional material. This sort of reanalysis always leads to structural simplification. Rutkowski (2002) shows that, in 16th-century Polish, an expression that contained a Q-numeral consisted of two separate Determiner Phrases (one of them was headed by the Q-numeral, the other by the quantified noun) whereas modern numeral expressions, as shown in subsection 2.2. of the present paper, are monophrasal. This means that Polish numeral structures have undergone a process

[7] Of course, such an option has not been taken in all languages. Therefore, Rutkowski & Szczegot (2001) view it as a kind of parameter.

of structural simplification. Q-numerals have become part of the extended projection of the quantified noun.

In the model proposed by Rutkowski & Szczegot (2001) (based on Emonds 2000) functional elements have to meet certain criteria: they constitute a closed class (new items are not easily introduced) and their meaning is limited and reduced to some basic oppositions and relations (such as determined vs. undetermined reference in the case of the English determiners *the/a*). Thus, Q-numerals seem to be good candidates for a functional class. The class of numerals is definitely closed; numbers are infinite but numerals are not. The semantic content of Q-numerals, as argued for in this paper, can be considered to be some sort of indices that point to the position of a given number in the counting sequence, i.e., in the arithmetic system. The function of a Q-numeral is to mark a place in which it is necessary to refer to extra-linguistic numeric competence. This line of reasoning finds support in the way numeric competence develops in children, as pointed out above. Children cannot interpret numerals without counting. However, when they hear a Q-numeral they understand that it must be in a context of quantification (Nelson & Toivonen 2000). Therefore, we might assume that, in the syntactic structure of a phrase containing a Q-numeral, only the feature of quantification/plurality (e.g., [+/-Q]) is marked and we interpret it thanks to an index that refers us to another type of competence, which is unavailable for children because they have not yet learnt it.

The exceptional character of Q-numerals is also confirmed by the fact that, in constructions consisting of a few numeral lexemes, internal dependencies and rules of word order are not driven by the syntax or semantics of a given natural language. Instead, they are reflecting arithmetic. In a way they could be thought of as imported to the linguistic structure. Chinese complex numerals seem to provide an exceptionally salient example. In Chinese numeral expressions it is possible to use the word *ling* 'zero' (Hurford 1975).

(12) a. ichian ling ell.shyrsyh
one-thousand zero two-ten-four
'one thousand and twenty-four'
b. erh pai ling erh
two hundred zero two
'two hundred and two'

From the point of view of psychological reality, the word *ling* cannot be interpreted as an element of the semantics of the whole expression. The complex numeral has to be considered a calque of the arithmetic code.

Even very long sequences of numeral lexemes seem to occupy only one slot in the syntactic structure. The relative order of such a sequence with respect to other elements (such as nouns) is always the same, for example, after the determiner but before the noun. The internal structure of the phrase that occupies this numeral position is not dependent on other elements of the sentence. Numeral sequences cannot be separated by words that do not belong to them (Hurford 1987). Even in languages with a relatively free word order (such as Polish) the word order in numeral sequences is always rigid. All of the above observations indicate that numerals are in a way external to the language system proper.

5. Conclussion

The analysis argued for in this paper links the empirically observed distinction between the lowest four numerals and the higher numerals, and certain neuropsychological findings. The linguistic phenomenon is interpreted as driven by "the magical number four", that is to say, the universal limitation of the focus of human attention described in detail by Cowan (2001). The lowest four numerals refer to quantities that can be perceived, in a way, indirectly. Therefore, these numerals appear in natural language much earlier than the higher ones. The higher numerals are dependent on the knowledge of counting. Their "extra-linguistic" properties mark them as syntactically special when they enter vocabularies of natural languages. The fact that they are different from other lexical elements (such as nouns or adjectives) resulted in some languages (e.g., Polish) in moving them to a class of functional elements.

References

Babby, L. 1988. Noun Phrase Internal Case Agreement in Russian. In M. Barlow & C. Ferguson (eds.), *Agreement in Natural Language. Approaches, Theories, Descriptions* 287-304. Stanford: Center for the Study of Language and Information, Stanford University.

Baddeley, A. 1986. *Working Memory*. Oxford: Clarendon Press.

Bloom, P. 2000. *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press.

Chomsky, N. 1980. *Rules and Representations*. Oxford: Basil Blackwell.

__________. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

__________. 1986. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.

Corbett, G. 1978. Universals in the Syntax of Cardinal Numerals. *Lingua* 46, 355-368.

Cowan, N. 1995. *Attention and Memory: An Integrated Framework. (Oxford Psychology Series 26)* New York: Oxford University Press.

________. 2001. The Magical Number 4 in Short-term Memory: A Reconsideration of Mental Storage Capacity. *Behavioral and Brain Sciences* 24, 87-114.

Cowan, N., L. Nugent, E. Elliott, I. Ponomarev, & J. Saults. 1999. The Role of Attention in the Development of Short-term Memory: Age Differences in the Verbal Span of Apprehension. *Child Development* 70, 1082-1097.

Dehaene, S. 1997. *The Number Sense*. New York: Oxford University Press.

Dirlam, D. 1972. Most Efficient Chunk Sizes. *Cognitive Psychology* 3, 355-359.

Dixon, R. 1980. *The Languages of Australia*. Cambridge: Cambridge Uni-

versity Press.
Emonds, J. 2000. *Lexicon and Grammar: The English Syntacticon. (Studies in Generative Grammar 50)* Berlin: Mouton de Gruyter.
Ericsson, K., W. Chase, & S. Faloon. 1980. Acquisition of a Memory Skill. *Science* 208, 1181-1182.
Franks, S. 1995. *Parameters of Slavic Morphosyntax.* New York: Oxford University Press.
Gelman, R. & C. Gallistel. 1978. *The Child's Understanding of Number.* New York: Academic Press.
Giusti, G. & N. Leko. 1995. Definite and Indefinite Quantity Expressions in Bosnian. In R. Benacchio, F. Fici, & L. Gebert (eds.), *Determinatezza e Indeterminatezza Nelle Lingue Slave* 127-145. Padua: Unipress.
Greenberg, J. 1978. Generalizations about Numeral Systems. In J. Greenberg, C. Ferguson, & E. Moravcsik (eds.), *Universals of Human Language 3: Word Structure* 249-295. Stanford: Stanford University Press.
Hauser, M., S. Carey, & L. Hauser. 2000. Spontaneous Number Representation in Semi-free-ranging Rhesus Monkeys. *Proceedings of the Royal Society of London (B): Biological Sciences* 267, 829-833.
Heine, B. 1997. *Cognitive Foundations of Grammar.* Oxford: Oxford University Press.
Heine, B., C. Ulrike, & F. Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework.* Chicago, IL: University of Chicago Press.
Hurford, J. 1975. *The Linguistic Theory of Numerals.* Cambridge: Cambridge University Press.
________. 1987. *Language and Number: The Emergence of a Cognitive System.* Oxford: Basil Blackwell.
________. 2001a. Languages Treat 1-4 Specially: Commentary on Stanislas Dehaene's Précis of The Number Sense. *Mind and Language: Symposium on Numerical Cognition* 69-75.
________. 2001b. The Neural Basis of Predicate-argument Structure. Ms. Edinburgh: University of Edinburgh.
Ifrah, G. 1981. *Histoire Universelle des Chiffres.* Paris: Editions Seghers.
______. 1985. *Les Chiffres ou L'histoire d'une GrandeInvention.* Paris: Editions Robert Laffont.
James, W. 1890. *The Principles of Psychology.* New York: Henry Holt.
Jevons, W. 1871. The Power of Numerical Discrimination. *Nature* 3, 281-282.
Juillard A. & E. Chang-Rodriguez. 1964. *Frequency Dictionary of Spanish Words.* The Hague: Mouton.

Kurcz, I., A. Lewicki, J. Sambor, K. Szafran, & J. Woronczak. 1990. *Słownik Frekwencyjny Polszczyzny Współczesnej*. Kraków: Polska Akademia Nauk, Instytut Języka Polskiego.

Linde-Usiekniewicz, J. & P. Rutkowski. 2003. O Zależności Między Właściwościami Składniowymi i Frekwencyjnymi Liczebników Głównych w Polszczyźnie. In R. Huszcza & J. Linde-Usiekniewicz (eds.), *Prace Językoznawcze Dedykowane Profesor Jadwidze Sambor*. Warszawa: Wydawnictwa Wydziału Polonistyki UW (to appear).

Logan, G. 1988. Toward an Instance Theory of Automatization. *Psychological Review* 95, 492-527.

Luck, S. & E. Vogel. 1997. The Capacity of Visual Working Memory for Features and Conjunctions. *Nature* 390, 279-281.

MacGregor, J. 1987. Short-term Memory Capacity: Limitation or Optimization? *Psychological Review* 94, 107-108.

Mandler, G. 1985. *Cognitive Psychology: An Essay in Cognitive Science*. Hillsdale, NJ: Erlbaum.

Mandler, G. & B. Shebo. 1982. Subitizing: An Analysis of its Component Processes. *Journal of Experimental Psychology: General* 111, 1-22.

Miller, G. 1956. The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review* 63, 81-97.

Nelson, D. & I. Toivonen. 2000. Counting and the Grammar: Case and Numerals in Inari Sami. *Leeds Working Papers in Linguistics and Phonetics* 8, 179-192.

Pisarek, W. 1972. *Frekwencja Wyrazów w Prasie*. Kraków: Ośrodek Badań Prasoznawczych RSW, Prasa.

Przepiórkowski, A. 1996. Case Assignment in Polish: Towards an HPSG Analysis. In C. Grover & E. Vallduví (eds.), *Edinburgh Working Papers in Cognitive Science 12: Studies in HPSG* 191-228. Edinburgh: University of Edinburgh.

Pylyshyn, Z. & R. Storm. 1988. Tracking Multiple Independent Targets: Evidence for a Parallel Tracking Mechanism. *Spatial Vision* 3, 179-197.

Reisberg, D., I. Rappaport, & M. O'Shaughness. 1984. Limits of Working Memory: The Digit-digit Span. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10, 203-221.

Roberts, I. & A. Roussou. 1999. A Formal Approach to "Grammaticalisation". *Linguistics* 37, 1011-1041.

Rutkowski, P. 2000. Składnia Polskich grup Liczebnikowych: Próba Opisu Formalnego. *Poradnik Językowy* 8, 10-28.

__________. 2001. Numeral Phrases in Polish and Estonian. *Proceedings of the 18th Scandinavian Conference of Linguistics* 2, 181-190.

__________. 2002. Numerals as Grammaticalised Nouns: A Generative Approach. *Interlingüística* 13, 317-328.

Rutkowski, P. & K. Szczegot. 2001. On the Syntax of Functional Elements: Numerals, Pronouns and Expressions Indicating Approximation. In A. Przepiórkowski & P. Bański (eds.), *Generative Linguistics in Poland: Syntax and Amorphosyntax* 187-196. Warszawa: IPIPAN.

Szczegot, K. 2001. Dlaczego Liczebniki Niższe Nie są Elementami Funkcjonalnymi?. In S. Ostrowska & P. Rutkowski (eds.), *Szkice Humanistyczne II*, 201-211. Warszawa: CI & KFnrD.

Ullman, S. 1984. Visual Routines. *Cognition* 18, 97-157.

Veselovská, L. 2001. Agreement Patterns of Czech Group Nouns and Quantifiers. In N. Corver & H. van Riemsdijk (eds.), *Semi-lexical Categories: The Function of Content Words and the Content of Function Words*, 273-320. Berlin: Mouton de Gruyter.

Vogel, E., S. Luck, & K. Shapiro. 1998. Electrophysiological Evidence for a Postperceptual Locus of Suppression during the Attentional Blink. *Journal of Experimental Psychology: Human Perception and Performance* 24, 1656-1674.

Wittgenstein, L. 1953. *Philosophische Untersuchungen (Philosophical Investigations)*. Red. G. Anscombe & R. Rhees. Oxford: Basil Blackwell.

Wynn, K. 1995. Origins of Numerical Knowledge. *Mathematical Cognition* 1, 35-60.

Yantis, S. 1992. Multielement Visual Tracking: Attention and Perceptual Organization. *Cognitive Psychology* 24, 295-340.

Zgółkowa, H. 1983. *Słownictwo Współczesnej Polszczyzny Mówionej: Lista Frekwencyjna i Rangowa*. Poznań: UAM.

Zgółkowa, H. & K. Bułczyńska. 1987. *Słownictwo Dzieci w Wieku Przedszkolnym*. Poznań: UAM.