TYPOLOGICAL PARAMETERS FOR THE LEXICO-SYNTACTIC STRUCTURES OF HIGH-DIGIT NUMERICAL EXPRESSIONS[1]

Cheng Luo
Brock University

ABSTRACT

The objective of this paper is to explore the lexico-syntactic structure of high-digit numerical expressions (henceforth HDNEs*), normally with 4 digits and more, and to propose a set of typological parameters to account for crosslinguistic variations in lexicalization sites. The two parameters of* pivot/anchor *and* increment *are proposed to describe and account for lexicalization sites of* HDNEs*, with language-specific value-setting explaining crosslinguistic variation. Support for this analysis includes psycholinguistic evidence from second language acquisition.*

1. INTRODUCTION

This paper examines the crosslinguistic lexico-syntactic structure of high-digit numerical expressions (HDNEs) in order to provide a principled account for their lexicalization sites, whose crosslinguistic variations are demonstrated in italics in Table 1:

Table 1: Sample of crosslinguistic variation in HDNE lexicalization sites

	English	Chinese	Kannada	Lugwere
1,000	(one) *thousand*	(yi) qian	ondu-*sāvira*	lu-kumi
10,000	ten thousand	(yi) *wan*	haththu-săvira	mu-*tulo*
100,000	(one) hundred thousand	shi wan	ondu-*laksha*	mi-tulo i-kumi
1,000,000	(one) *million*	(yi) bai wan	haththu-laksha	ka-*kairi*
10,000,000	ten million	(yi) qian wan	ondu-*kōti*	bu-kairi i-kumi
100,000,000	(one) hundred million	(yi) *yì*	haththu-kōti	bu-kairi chi-kumi
1,000,000,000	(one) *billion*	shi yì	—	*kataʙarika kamo*

As Chomsky (1980: 248f) suggests, the way in which human beings develop the number system may 'shed light on deep and fundamental charac-

1 An earlier version of this work was presented at the 1997 Mid-America Linguistic Conference at Columbus, Missouri, at which I benefited from the remarks and suggestions from the audience. I would like to thank the anonymous reviewer for helpful comments and suggestions. Any mistakes or remaining discrepancies are my sole responsibility.

teristics of the human species.' Piaget (1961: 188) points out that numerical expressions involve 'fundamental cognitive principles [such as] classification and serialization.' As part of a complete description of grammatical systems, numerical expressions provide a relatively small, but usually well-defined and independent subsystem of language that can be described in its own terms and compared with similar systems in other languages (Brainerd 1968). Crosslinguistic studies of numerals (e.g., Greenberg 1978, 1989; Hurford 1975, 1987; Gvozdanovic 1992a) have shown considerable regularity and generalizability in their semantic structure and morphosyntactic properties. However, most studies have either focussed on lower-digit numerals (e.g., Greenberg 1978, 1989; Seiler 1990; Gvozdanovic 1992a; Franks 1994), or investigated single languages or language groups (e.g., Akiner 1983; Bradley 1981; Gerhardt 1987; Shionoya 1990; Krippes 1991; Potet 1992; Oliverio 1993). Thus, despite their insights into the structure of numerical expressions of individual languages, or of lower-digit numerical expressions crosslinguistically, such studies have limitations as to what significant generalizations can be made about the diverse lexico-syntactic properties of HDNEs found in the world's languages.

That HDNEs need to be studied in their own terms is seen from some important distributional differences between HDNEs and their lower-digit counterparts. For example, lexicalization of English numerals occurs at regular intervals by powers of thousand from the 4th digit on, but shows a more idiosyncratic pattern for the lower digits, due in part to their more dense lexicalization within a semantic domain. Given this and other structural differences to be discussed between high- and low-digit numerical expressions, and given the relatively little attention previously paid to HDNEs, an investigation of crosslinguistic HDNE structure will not only provide evidence for a set of parameters governing the lexicalization of HDNEs, but also lend insights into the setting and re-setting of parametric values in first and second language acquisition. In the following discussion, I will first examine Hurford's (1975, 1987) and Greenberg's (1975, 1989) crosslinguistic accounts of HDNEs, and then propose a pair of parameters supported by a crosslinguistic analysis of HDNEs. It will be shown that the proposed parameters provide a plausible, general, yet simple way to describe the lexico-syntactic structure of HDNEs crosslinguistically.

2. PRIOR CROSSLINGUISTIC STUDIES ON HDNEs

2.1 Hurford's exponentiation and multiplication types

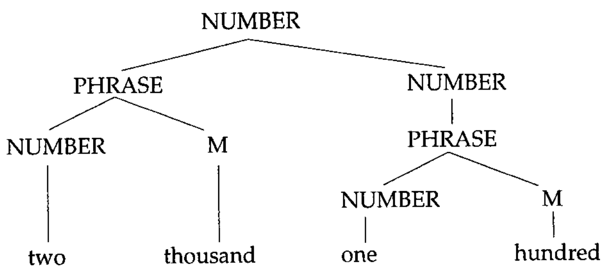
Hurford (1987: 245) proposes three universal syntactic categories for PS-rules in numeral syntax: NUMBER, PHRASE, and M, as in (1). (2) gives an example of a structure generated by (1), where *digit* expands to any of the words *one, two, ...nine*)

(1) a. NUMBER → $\left\{ \begin{array}{l} \text{DIGIT} \\ \text{PHRASE (NUMBER)} \end{array} \right\}$

b. PHRASE → NUMBER M

c. M → *-ty, hundred, thousand, million, billion*

(2)



Of particular interest here is the category M, which functions as base numbers for numerical expressions (Gvozdanovic 1992a: 5), and corresponds to the lexicalization sites in Table 1 except for the numbers 1 to 10 (i.e., DIGITS). In the high-digit domain, Hurford classifies languages into two types: those whose Ms are interpreted by exponentiation and those by multiplication. The exponentiation type is formally defined as in (3) (Hurford 1975: 247), exemplified again by data from English and Chinese in Table 1.

(3) ... the values of the Ms may be arranged into a series $m_1, m_2, ..., m_n$, such that for all adjacent pairs of values m_i, m_j, either $m_j = m_i^2$ or there exists some m_x and some whole number y (y>1), such that $m_i = m_x^y$ and $m_j = m_x^{y+1}$.

According to (3), English Ms *thousand* (m_i), *million* (m_j), and *billion* (m_k) are in an exponential relationship such that $m_j = m_i^2$, and $m_k = m_i^3$. Similarly, Chinese Ms *wan* '10^4' (m_i) and *yì* '10^9' (m_j) also have an exponential relationship of $m_j = m_i^2$. On the other hand, the multiplication type is defined as (4), with (5) as an example from Tamil, a Dravidian language:

(4) a language whose values of the higher-valued Ms can be arranged into a series $m_1, m_2, ..., m_n$, such that for all adjacent triples of values m_i, m_j, m_k there exists some whole number y such that $m_k = m_j \times y$ and $m_j = m_i \times y$ (Hurford 1975: 248).

(5)	Tamil:	ayiram	'1,000'
		laksham	'100,000'
		kōdi	'10,000,000'

In (5), *ayiram* (m_i), *laksham* (m_j) and *kōdi* (m_k) are related not by exponentiation, but by multiplication by 100, such that $m_j = m_i \times 100$, and $m_k = m_j \times 100$. This kind of multiplicative relationship also obtains in HDNEs in Ancient Hawaiian, Yoruba, Ainu, and some other Dravidian languages.

There are several difficulties with Hurford's typology. First, the category M is not explicitly defined, and applies to both the low-digit and the high-digit domains without considering structural differences between them. Further, it does not explicitly specify a digit position at which exponentiation or multiplication starts, thus failing to show whether exponential relationship between lexicalization sites in a language like English starts at the second digit on a decimal basis (*ten, hundred*, and *thousand*) or at the 4th digit on the basis of 1,000 (e.g., *thousand, million, billion*, and *trillion*). Secondly, while multiplication is a very effective way of economizing the expression of higher numbers[2] to avoid inefficiency of total lexicalization (Winter 1992: 18-19), exponentiation is less common in natural languages. Thirdly, Hurford's typology shows limitations in descriptive adequacy when attested against data from other languages. Consider (6) from Lugwere, a Bantu language:

(6)	a.	i-**kumi**	10
	b.	chi-kumi	100
	c.	lu-kumi	1,000
	d.	mu-**tulo**	10,000
	e.	mi-tulo i-kumi	100,000
	f.	ka-**kairi**	1,000,000
	g.	bu-kairi i-kumi	10,000,000
	h.	bu-kairi chi-kumi	100,000,000
	i.	**kataBairika kamo**	1,000,000,000

[2] According to Winter (1992: 18), the optimal solution to efficiency of expression is a combination of addition and multiplication.

As we see in the Lugwere data (6), it is difficult to find any consistent relationship, exponential or multiplicative, between the lexical items *kumi, tulo, kairi,* and *kataʙarika kamo.* In fact, the distinction between exponentiation and multiplication is quite superfluous and unnecessary, and it would be more adequate to describe the structure in terms of multiplicative relationship only, as implied in an increment parameter I will propose.

Finally, Hurford's typology seems to be motivated more by semantic interpretation than by lexico-structural properties. In terms of lexicalization sites, Chinese is as different from English as English is from Kannada (Table 1), but according to Hurford, Chinese is grouped with English as opposed to Kannada. The fact that English and Chinese are rendered non-distinct does not account for some significant psycholinguistic differences by L2 learners in acquiring HDNEs, which I will show later.

To sum up, Hurford's typology of HDNEs has some weaknesses in explicitness, plausibility, and descriptive adequacy. To more adequately describe the structure of HDNEs, we need parameters that not only specify the HDNE lexicalization sites and the syntactic relationship between them in a language, but also account for crosslinguistic variation thereof.

2.2 GREENBERG'S CONCEPT OF *BASE*

Greenberg (1978, 1989) proposes three linguistic procedures for the dimension of enumeration: *atoms*, the set of numerals which receive 'simple lexical representations' (Greenberg 1978: 256), *bases*, and *calculatory operations.* These are functionally related, respectively, to indicativity, iconicity, and predicativity (Seiler 1990). Of these procedures, the *base* is defined as a serialized multiplicand, for example, 10, 100, 1,000 and 1,000,000 (Greenberg 1978: 270). *Bases* function as marks of hierarchical packing, such that languages may have packs of fives (i.e., quinary), tens (i.e., decimal), twenties (i.e., vigesimal), etc. Other properties of *bases* include: (a) that they are nominal, (b) that they are polyvalent, and (c) that they mark turning points (Seiler 1990: 193-6).

Unlike Hurford, Greenberg does not distinguish between exponentiation and multiplication, recognizing only a serialized multiplicative relationship between lexicalization sites of numerical expressions. On the other hand, like Hurford's category M, Greenberg's *base* applies to both lower and higher numerals, overlooking several important structural differences between them, including lexicalization intervals mentioned above. As another difference, we find in English that while *ten* and *hundred* may occur recursively in forming HDNEs, *thousand, million, billion,* etc., can not be so used. Thus, one can say 'ten million ten thousand and ten' but not '*one million million'.

Crosslinguistically, there is also evidence that syntactic processes applicable to lower-digit bases may operate differently for higher-digit ones. Consider the following examples from Chinese involving elliptic forms of numerical expressions:

(7) a. yi bai wu-**shi**
one hundred fifty
'150'

b. yi bai wu

(8) a. yi qian wu **bai**
one thousand five hundred
'1,500'

b. yi qian wu

(9) a. yi wan wu **qian**
one 10^4 five thousand
'15,000'

b. yi wan wu

(10) a. yi yì wu qian **wan**
one 10^8 five thousand 10^4

b. ?*yi yì wu
'150,000,000'

In Chinese, the base *wan* (i.e., 10^4) marks the boundary between lower and higher digits. As can be seen from (7-10), deletion can optionally apply only when the target of deletion is *shi* (10), *bai* (10^2), or *qian* (10^3), but not *wan* (10^4). The inapplicability of deletion to the 5th digit base thus marks a distinct syntactic difference between it and its lower-digit counterparts.

Chinese also has a *ling*-insertion rule to the effect that the word *ling* 'zero' occurs as a place holder to mark a sequence of one or more zero places occurring between two non-zero places (Battistella 1989: 9). As Battistella (1989) reports, *ling*-insertion applies obligatorily to digits below *wan*, but optionally to digits above *wan*, as in (11):

(11) a. san yì (**ling**) er bai shi yi wan yi qian **ling** shi san
three 10^8 zero two hundred eleven 10^4 one thousand zero thirteen
'302,111,013'

b. *san yì ling er bai shi yi wan yi qian __ shi san

In (11a), the place holder *ling* occurs properly both before and after *wan*; however, while the rule applies optionally to the high-digit part of the numerical expression before *wan*, *ling* has to occur obligatorily in the low-digit section after *wan*, as attested in the incorrect (11b).

Another difference between the low- and the high-digit expressions crosslinguistically is that low-digit numerical expressions tend to have less systematic calculatory operations than high-digit ones. At lower digits, many languages have a mixture of bases, such as quinary or vigesimal mixed with decimal bases, or a mixture of calculatory operations, such as progressive operation by addition combined with regressive operation by subtraction (Greenburg 1978, Seiler 1990). For high digits, there tends to be a more uniform pattern of the decimal multiplicands and a more consistent operation of multiplication-addition.

Finally, let's consider the use of *amari* in a quite obsolete native Japanese numerical system called Yamato kotoba (Brainerd & Peng 1968). In this system, *amari* functions as an obligatory connective between two numerical expressions. For example, in (12a), it joins *nana-chi* '7,000' and *iso* '50' on the one hand and *iso* and *mihe* 'three' on the other, which are themselves full-fledged numerical expressions.

(12) a. nana-chi he *amari* iso *amari* mihe
7-thousand CONJ fifty CONJ three
'7,053'

b. futa yorozu *amari* itsu chi-tari
2 10^4 CONJ 5 1000-tari
'25,000'

c. chi i ho yorozu
1000 5 100 10^4
'15,000,000'

In terms of distribution of *amari*, we notice that it occurs only after, as in (12b), but not before, as in (12c), *yorozu* 'ten-thousand', the presumable boundary between low- and high-digit numerical expressions in Japanese. Thus, we may say that the distribution of *amari* is sensitive to the low-/high-digit distinction in Japanese.

To briefly conclude this section, crosslinguistic evidence for the structural differences between low- and high-digit numerical expressions strongly suggests the need to analyze HDNEs independently in terms of a set of parameters capable not only of describing HDNE structures and crosslinguistic variations, but also of capturing these differences as well. In the rest of this paper, I will propose such a set of parameters, and attest it against crosslinguistic data.

3. TYPOLOGICAL PARAMETERS: ANCHOR, PIVOT, AND INCREMENT

Compare English and Mandarin Chinese HDNEs in (13):

(13) 2,354,796,000

a. two **billion** three hundred and fifty four **million** seven hundred and ninety six **thousand**

b. er-shi san yi wu qian si bai qi- shi
twenty three 10^8 five thousand four hundred seven-ty
jiu wan liu qian
nine 10^4 six thousand

The comparison shows several structural properties. First, both languages express some HDNEs periphrastically (e.g., *ten thousand*) but others lexically (e.g., *thousand* and *wan*); however, lexicalization sites differ in the two languages. For example, the 5th digit is a lexicalization site in Chinese but not in English. Secondly, within a periphrastic expression such as *a hundred thousand* in English, the lexical head *thousand* does not recur in other HDNEs as modifier, e.g., **thousand thousand*, whereas *hundred* does (cf. *two hundred million three hundred thousand*). We may call this property *non-recursiveness*. Thirdly, the lexicalization sites are structured in English in such a way that a change in lexeme occurs every three digits starting from the 4th digit (e.g. *thousand, million, billion*). We may call this *3-digit increment*. Chinese, on the other hand, shows a 4-digit increment, i.e., lexeme change occurs every four digits starting from the 5th digit (e.g., *wan, yi*).

For ease of discussion, I will use two terms: *anchor* and *pivot* for the first parameter. An *anchor* is a non-recursive lexeme in any HDNE that can not be a modifier in periphrastic numerical expressions. In English, the anchors are *thousand* (4th digit), *million* (7th), *billion* (10th), *trillion* (13th), etc.; in Chinese, they are *wan* (5th), *yi* (9th), *zhao* (13th), etc.[3] Note that *qian* 'thousand' in Chinese is not an anchor, as it can recur as modifier in higher digit expressions. A *pivot* is the rightmost anchor from which regular increment involving anchor change occurs. By definition, the pivot is *thousand* (4th digit) in English and *wan* (5th) in Chinese. The second parameter is *increment*, defined as the interval at which anchor change occurs.

3 The Chinese anchor *wan*, however, can occasionally be used in periphrastic expressions as an alternative to its more formal counterparts, e.g. *yi wan wan* 'a hundred million' instead of the more formal *yi yi*.

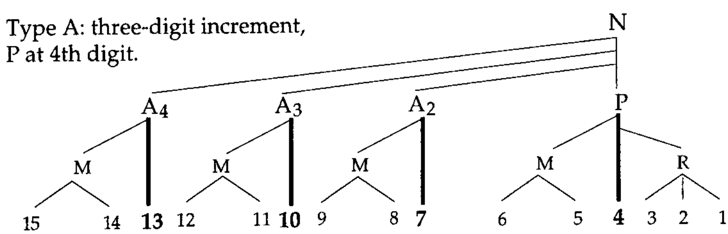
With these parameters, the differences between English and Chinese lexicalization sites for HDNEs can be expressed in terms of different value-setting, as in (14):

(14) While English has the 4th digit as pivot and uses a 3-digit increment which effects change of anchor at the 7th, 10th and 13th digits, Chinese has the 5th digit as pivot and uses a 4-digit increment which effects change of anchor at 9th and 13th digits.

(15) is a formal representation of the two types of lexico-syntactic structures of HDNEs represented by English and Chinese, as stated in (14). In (15), the numbers at the bottom represent digit positions. P[4] marks the pivot position at which increment starts in a language, and together with the As, marks lexicalization sites. The Ms show the respective multiplicative relationships between an anchor and its modifiers. Finally, R represents the digits lower than P, which are not the focus of the present study.

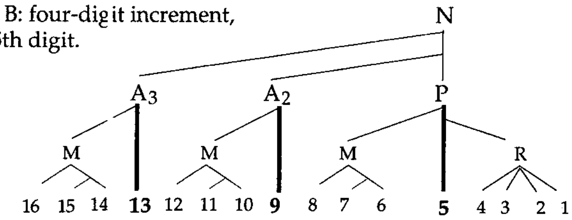
Compared with Hurford's typology, the parameters proposed here more clearly and adequately describe HDNE structures in a language. The pivot is a necessary parameter for marking the starting point for regular increment; and once the value of the pivot is assigned, the increment parameter adequately accounts for lexicalization sites. Secondly, the concept of increment involves only multiplicative relationships between anchors, thus simplifying our typological description. Thirdly, they more adequately account for crosslinguistic variations of HDNE structures from languages like Lugwere as well as those represented by English, Chinese, and Kannada, as will be seen next.

(15) **Typological parameters and syntactic structure of high-digit numerical expressions**



4 P=pivot, N=number, A=anchor, M=modifier, R=residue.

Type B: four-digit increment,
P at 5th digit.



4. CROSSLINGUISTIC EVIDENCE

Table 2 (following References) presents some codified crosslinguistic HDNE data of 14 languages from 5 language families. The languages are divided into 4 groups: the European group, the East Asian group, the Dravidian group and the bantu group. HDNEs are arranged from higher to lower digits. Within each cell, the anchor occurs in the bottom line, and modifiers, if any, occur from top to bottom above the anchor in the order of *ten, hundred,* and *thousand.* Each HDNE should be read from left to right regardless of the linear disalignment. The shaded areas indicate lexicalization sites, with the rightmost one marking the pivot position. In Table 2, crosslinguistic variation in lexicalization sites of HDNEs is accounted for in terms of *pivot* and *increment,* as follows: for the European group: pivot = 10^3; increment = 3-digit; anchors = 10^3, 10^6, 10^9, 10^{12}; for the East Asian group[5]: pivot = 10^4; increment = 4-digit; anchors =10^4; 10^8; 10^{12}; and for the Dravidian group: pivot = 10^3; increment = 2-digit; anchors = 10^3, 10^5, 10^7.

While most groups show regular and consistent increment, Lugwere presents an interesting pattern. As in Chinese, the 5th digit in Lugwere is lexicalized with *mu-tulo.* However, the increment shows a mixed pattern of 2-digit increment between *mu-tulo* 'ten-thousand' and *ka-kairi* 'million', and 3-digit increment between *ka-kairi* 'million' and *kataʙarika* 'billion'. Considering the recurrence pattern of *ikumi* 'ten' especially in the 3rd and 4th digit positions, if we extend the scale rightward to include the 2nd to the 4th digits, we actually have a 3-2-3 increment pattern, with the pivot at the 2nd digit (though not a high digit) and the anchors at the 5th, 7th, and 10th. In a word, what the Lugwere data show is that not all languages have fixed-value increment.

---

5 Diachronically, many of the HDNEs in Chinese, Japanese and Korean were cognates, suggesting that the numerical system originated in Chinese and was borrowed into Japanese and Korean, which accounts for the set of structural features shared by the three languages.

To sum up, crosslinguistic variation in HDNEs can be accounted for in terms of *pivot* and *increment*, and their value-setting. It is suggested that these parameters may be applicable to HDNEs in other languages, which differ from each other in the value-setting of the parameters.

5. PSYCHOLINGUISTIC EVIDENCE

A comparison of Hurford's typology and the typological parameters proposed here shows that they make quite different predictions about second language acquisition, based on language transfer theory of interlanguage studies. According to Hurford's analysis (§2.1), which by treating European languages like English, and Chinese, as belonging to the exponentiation type makes them non-distinct, East Asian English as a Second Language (ESL) learners[6] will learn English HDNEs as easily as, say, Spanish ESL learners. The analysis argued for herein, however, contrasts European and East Asian languages in value setting, and predicts that East Asian ESL learners will have greater difficulty acquiring English HDNEs than their Spanish counterparts, because of differences in processing HDNES.

In order to find out which typology is psycholinguistically more plausible, an experiment (Luo & Wilson 1996) was carried out to test the hypothesis that East Asian ESL learners would encounter greater difficulties than their European counterparts in comprehending or producing English HDNEs as a result of delayed re-setting of typological parameters because of first language (L1) interference. 50 lower-level ESL learner participants (20 Spanish, 30 East Asians) and 16 English native speaker participants performed a listening comprehension task by writing down contextualized HDNEs heard on tape, and an oral production task by promptly uttering HDNEs seen on a computer in Arabic numerals. Errors and processing time for each task were measured, recorded and then converted to performance scores for all the groups.

Statistical results are reported in Table 3 and diagrammatically shown in Figure 2. As the results show, there are significant between-group differences in both comprehension and production of HDNEs. More specifically, while the control group (native English speakers; 138.27 for comprehension and 188.21 for production, respectively) outperformed both the East Asian group (38.88 for comprehension and 100.90 for production, respectively) and the Spanish group (69.33 for comprehension and 137.14 for

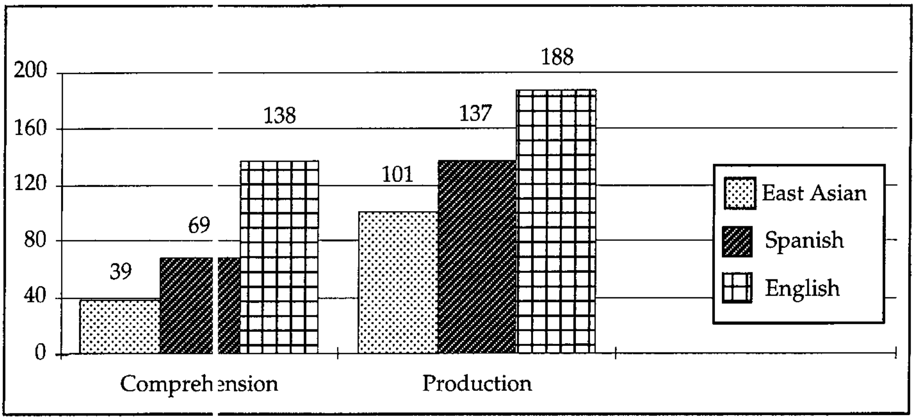
6 The East Asian group refers to Chinese, Japanese, and Korean ESL learners.

production, respectively), the Spanish group significantly outperformed the East Asian group both in production and in comprehension, suggesting presence of negative transfer effect for the latter group. Thus, psycholinguistic evidence in second language acquisition shows greater difficulty in processing target language HDNEs when one's native language HDNEs differ structurally from those of the target language. Such differences are best captrued by *pivot* and *increment* as plausible parameters for the lexico-syntactic description of HDNEs.

Table 3: Descriptive statistics for task performance by participants with different L1s

L1	N	Mean	SD	SE	Variance	Minimum	Maximum
Comprehension:							
English	16	138.27	26.56	6.64	705.43	79.88	163.88
Spanish	20	69.33	34.02	7.61	1157.36	0.00	126.73
East Asian	30	38.88	20.76	3.79	430.98	0.00	91.35
Production:							
English	16	188.21	10.19	2.55	103.84	153.30	195.35
Spanish	20	137.14	34.06	7.62	1160.08	82.03	186.57
East Asian	30	100.90	47.64	8.70	2269.57	34.72	180.68

Figure 1: Comparison of performance on HNDEs by L1



6. CONCLUSION

Based on a critique of Hurford's (1987) typology of HDNEs and Greenberg's parameter of *base*, this paper proposes *pivot/anchor* and *increment* as two parameters for describing and accounting for lexicalization sites of HDNEs. It is suggested that the lexico-syntactic structures of HDNEs in all languages are susceptible to such parameters, whose language-specific value-setting accounts for crosslinguistic variations. Thus, most European languages (e.g., English, Spanish, Slovak, Dutch, German, French) take the 4th digit as the pivot and use a 3-digit increment thereafter, East-Asian languages (Chinese, Japanese and Korean) take a 5th-digit pivot and employ a 4-digit increment, and most Dravidian languages (e.g. Tamil, Malayalam, Kannada, and Tegulu) take the 4th digit as the pivot and use a 2-digit increment. It is also possible for some languages, such as the Bantu language of Lugwere, to adopt a mixed increment system. Psycholinguistic evidence from second language acquisition further supports these parameters as a plausible framework for describing the structure of HDNEs in a most adequate, general and simple way.

Table 2: A crosslinguistic comparison of lexicalization sites...

Digit	13	12	11	10	9
#formative	1000000000000	100000000000	10000000000	1000000000	100000000
European					
English	trillion	hundred billion	ten billion	billion	hundred million
Dutch	biljoen	honderd miljard	tien miljard	miljard	honderd miljoen
German	Billion	hundert Milliarde	zehn Milliarde	Milliarde	hundert Million
French	trillion	cent milliard	dix milliard	milliard	cent million
Spanish	trillon	cien billones	diez billones	billon	cien millones
Slovak	bilion	sto miliard	desat miliard	miliarda	sto miliónov
EastAsian					
Chinese	zhao	qian yì	bai yì	shi yì	yì
Japanese	chō	sen-oku	hyaku-oku	jyu-oku	oku
Korean	cho	chun uk	bac-uk	sip-uk	uk
Dravidian					
Tamil					paththu-kōdi
Malayalam					pathinonnu kōti
Kannada					haththu kōti
Telugu					padhi kōtlu
Bantu					
Lugwere		mi-tulo bu-taʙarika	i-kumi bu-taʙarika	kataʙarika kamo	chi-kumi bu-kairi

...in high-digit numerical expressions

8	7	6	5	4	3	2
10000000	1000000	100000	10000	1000	100	10
European						
ten million	**million**	hundred thousand	ten thousand	**thousand**	hundred	
tien miljoen	**miljoen**	honderd duizend	tien duizend	**duizend**	honderd	
zehn Million	**Million**	hundert tausend	zehn tausend	**tausend**	hundert	
dix million	**million**	cent mille	dix mille	**mille**	cent	
diez millones	**millon**	cien mil	diez mil	**mil**	cien	
desat miliónov	**milión**	sto tisíc	desat tisíc	**tisíc**	sto	
East Asian						
qian wan	bai wan	shi wan	**wan**	qian	bai	
sen-man	hyaku-man	jyu-man	**man**	sen	hyaku	
chun man	bac-man	sip-man	**man**	chun	bac	
Dravidian						
kōdi	paththu-laksham	**laksham**	paththu-āyiram	**āyiram**	nūru	
kōti	pathinonnu laksham	**laksham**	pathinonnu āyiram	**āyiram**	nūru	
kōti	haththu laksha	**laksha**	haththu sāvira	**sāvira**	nūru	
kōti	padhi lakshalu	**laksha**	padhi vēlu	**veyyi**	nūru	
Bantu						
i-kumi bu-kairi	**ka-kairi**	i-kumi mi-tulo	**mu-tulo**	kumi lu-	chi-kumi	**i-kumi**

REFERENCES

AKINER, SHIRIN. 1983. The syntax of the numeral in Byelorussian, compared with Ukrainian, Russian and Polish. *The Slavonic and East European Review* 61,1: 55-68.

BATTISTELLA, EDWIN. 1989. Some notes on Chinese number words. *CUNYForum. Papers in Linguistics* 14, 7-12.

BRADLEY, MAYA. 1981. Numerals in Modern Hebrew. *Working Papers in Linguistics* 7: 87-115. Melbourne, Australia: University. of Melbourne.

BRAINERD, BARRON. 1968. On the syntax of certain classes of numerical expressions. In H.B. Corstius (ed.), *Grammars for Number Names*, Foundations of Language Supplementary Series, Vol. 7, Dordrecht: D. Reidel, 9-40.

BRAINERD, BARRON. & FRED PENG. 1968. A syntactic comparison of Chinese and Japanese numerical expressions. In H.B. Corstius (ed.), *Grammars for Number Names*, Foundations of Language Supplementary Series, Vol. 7, Dordrecht: D. Reidel, 53-81.

CAFLISH, J. 1984. 'A typology of numerical cycles, I: Kartvelian. *The Language Quarterly* 23,1-2: 9-12.

CHAE, W. 1983. A study of numerals and numeral classifier constructions in Korean. *Language Research* 19,1: 19-34.

CHOMSKY, NOAM. 1980. *Rules and Representations*. Oxford: Basil Blackwell.

COMRIE, BERNARD. 1992. Balto-Slavonic. In Jadranka Gvozdanovic (ed.), *Indo-European Numerals*. (*Trends in Linguistics: Studies and Monographs* 57). Berlin: Mouton de Gruyter, 717-833.

CORSTIUS, H. B. (ed.) 1968. *Grammars for Number Names*. Foundations of Language Supplementary Series, Vol. 7. Dordrecht: D. Reidel.

DEHAENE, S & MEHLER, J. 1992. Cross-linguistic regularities in the frequency of number words. *Cognition: International Journal of Cognitive Science* 43,1: 1-29.

FRANKS, S. 1994. Parametric properties of numeral phrases in Slavic. *Natural Language and Linguistic Theory* 12,4: 597-674.

GERHARDT, L. 1987. Some remarks on the numerical systems of Plateau languages. *Afrika und Ubersee: Sprachen Kulturen* 70,1: 19-29.

GREENBERG, JOSEPH H. 1978. Generalizations about numeral systems. In Joseph H. Greenberg (ed.), *Universals of Human Language: Word Structure*. Stanford: Stanford University Press, 249-95.

1989. The internal and external syntax of numerical expressions: explaining language specific rules. *Belgian Journal of Linguistics* 4: 105-118.

GVOZDANOVIC, JADRANKA. 1992a. Remarks on numeral systems. In Jadranka Gvozdanovic (ed.), *Indo-European Numerals*. Berlin: Mouton de Gruyter, 1-10.

(ed.) 1992b. *Indo-European Numerals*. Berlin: Mouton de Gruyter.

HURFORD, JAMES. 1975. *The Linguistic Theory of Numerals*. Cambridge: Cambridge University Press.

1987. *Language and Number*. Oxford: Blackwell.

JUSTUS, C. 1988. Indo-European numerals and numeral systems. In Yoel L. Arbeitman (ed.), *A Linguistic Happening in Memory of Ben Schwartz: Studies in Anatolian, Italic, and Other Indo-European Languages*. Louvain-La-Neuve: Peeters, 521-541.

KOUL, M. & SHARMA, R. 1988. Numeral system in Kashmiri. *Indian Journal of Linguistics Praci Bhasha Vijnan*, 15,2: 43-50.

KRIPPES, K. 1991. A linguistic enigma: the Altaic numerals. *General Linguistics* 31,3-4: 141-52.

LEHMAN, F. K. 1990. Outline of a formal syntax of numerical expressions with special reference to the phenomenon of numeral classifiers. *Linguistics of the Tibeto-Burman Area* 13,1: 89-120.

LUO, CHENG. & WILSON, F. 1996. Acquisition of the ESL high-digit numerical expressions: a reflection on Transfer Theory. In A. Pasquini, L. Rowsell & L. C. Smith (eds.), *Proceedings of the 1996 Annual Conference of the Canadian Linguistic Association*. Calgary, Alberta, (Calgary Working Papers in Linguistics), 221-234.

MACEK, D. 1985. A contrastive study of numerals in English and Serbo-Croatian. In Filipovic, Rudolf (ed.), *Chapters in Serbo Croatian-English Contrastive Grammar*. Zagreb: Faculty of Philosophy., Zagreb University, 135-166.

MANCZAK, W. 1985. Indo-European numerals and the sexagesimal system. In Jacek Fisiak (ed.), *Papers from the 6th International Conference on Historical Linguistics*. Amsterdam: J. Benjamins, 347-352.

MARCOS, M. & FRANCISCO, A. 1992. Numerals and typology in minority languages of China. In André Crochetière, J. Boulanger & C. Ouellon (eds.), *Actes du XVe Congrès International des Linguistes: Les Langues menacées/Endangered Languages: Proceedings of the XVth International Congress of Linguists* Sainte-Foy: Presses Universitaires Laval, 169-72.

MILLER, K. & ZHU, J. 1991. The trouble with teens: Accessing the structure of number names. *Journal of Memory and Language* 30,1: 48-68.

OLIVERIO, G. 1993. Tutelo numeral types. In E. Smith & F. Zephir (eds.), *Proceedings of the 1992 Mid-America Linguistics Conference and Conference on Siouan/Caddoan Languages*. Columbia: University of Missouri-Columbia, 339-346.

PIAGET, JEAN. 1961. La construction des nombres naturels. In Jean Piaget (ed.), *Epistemologie mathématique et psychologie*. Paris: Presses Universitaires de France, 276-299.

PICARD, M. 1986. On the structure of the lower numbers in Pre-PA. *International Journal of American Linguistics* 52,1: 72-77.

POTET, J. 1992. Numeral expressions in Tagalog. *Archipel* 44: 167-82.

RONG, P. 1990. Chinese and English numeral compact expressions: a contrastive study. *Waiguoyu* 3: 1-7, 67.

SEILER, H. 1990. A dimensional view on numeral systems. In W. Croft et al. (eds.), *Studies in Typology and Diachrony*. Amsterdam: Benjamins, 187-208.

SHARMA, D. D. 1985. Numeral system of Tibeto-Himalayan. *Panjab University Research Bulletin (Arts)* 16,2: 75-90.

SHIONOYA, T. 1990. Syntactic properties of Samoan numerals. *Journal of the Linguistic Society of Japan* 97: 18-43.

SIROMONEY, R. 1968. Grammars of number names of certain Dravidian languages. In H.B. Corstius, (ed.), *Grammars for Number Names*. Foundations of Language Supplementary Series, Vol. 7. Dordrecht: D. Reidel, 82-90.

SOYOYE, F. & OLAJUYIGBE, A. 1993. Numerals in Yoruba: an investigation of native speakers' knowledge of a language. *Research in Yoruba Language and Literature* 4: 68-77.

WINTER, WERNER. 1992a. Some thoughts about Indo-European numerals. In Jadranka Gvozdanovic (ed.), *Indo-European Numerals*. Berlin : Mouton de Gruyter, 11-28.

WYNN, K. 1992. Children's acquisition of the number words and the counting system.' *Cognitive Psychology* 24,2: 220-51.