Measure words and classifiers

Palavras de medida e classificadores

Jenny Doetjes*

Abstract

This paper investigates the relation between measure words and mensural classifiers in numeral classifier languages. Based on data from three numeral classifier languages (Mandarin, Mokilese and Taba), the paper gives some preliminary evidence that measure words can be both classifier-like and noun-like in numeral classifier languages. This observation is discussed in the light of Rothstein's (2009, 2011) distinction between measuring and counting, Krifka's (1995) numeral based analysis of numeral classifier languages as kinds. Crucially, if the measure words are treated as nouns, one has to take into account that the atomic entities corresponding to units of measurement typically overlap. This is problematic for the type of interpretation that Chierchia (1997) assigns to kinds, as the kinds corresponding to different units of time would be indistinguishable. Other approaches will need a non-overlap condition on counting structures.

Resumo

Este trabalho investiga a relação entre palavras de medida e classificadores de medida em línguas de classificadores numerais. Baseado em dados de três línguas com classificadores numerais (Mandarim, Mokilese e Taba), o artigo apresenta evidências preliminares de que palavras de medida podem ser tanto classificadores quanto nomes em línguas de classificadores numerais. Essa observação é discutida à luz da distinção entre medir e contar feita em Rothstein (2009, 2011), na análise de línguas de classificadores numerais baseada em número de Krifka (1995) e na proposta de Chierchia (1997) de tratar nomes, em línguas de classificadores, como kinds. Fundamentalmente, se as palavras de medidas são tratadas como nomes, deve-se levar em conta que entidades atômicas que correspondem a unidades de medida tipicamente se sobrepõem. Isso é problemático para o tipo de interpretação que Chierchia (1997) atribui a kinds, já que kinds, correspondendo a diferentes unidades de tempo seriam indistintos. Outras abordagens precisarão de uma condição de não sobreposição para estruturas de contagem.

^{*} Leiden University Centre for Linguistics

Introduction

In numeral classifier languages, the use of a numeral triggers insertion of a classifier, independently of the count-mass status of the noun, as illustrated in (1) for Mandarin. Classifiers are often seen as the counterparts of measure words in languages such as English.

(1)	a. liăng	zhī	bĭ	
				[Mandarin]
	two	CL ^{branch} pen		
	'two pens'			
	b. liăng	jìn	mĭ	
	two	CL^{half_kilo}	rice	
	'two pound			

In this paper I will investigate the relation between measure words and classifiers, and I will show that this relation is more complex than one might expect on the basis of examples such as (1b), where the measure word is realized as a classifier. I will give some preliminary evidence that measure words can be classifier-like or noun-like in classifier languages.

The structure of this paper is as follows. Section 2 provides some background on classifiers and measure words, and addresses the theoretical question of why classifiers are needed in languages such as Mandarin. Section 3

is devoted to measure words in three classifier languages – Mandarin, Mokilese and Taba – and section 4 discusses the main observations in the light of theoretical accounts of numeral classifiers and the status of nouns in classifier languages. Section 5 concludes the paper.

1. Background

Numeral classifier languages come in many different types (GILL, 2005; GRINEVALD, 2005). In some cases the classifiers are obligatory, e.g. in Mandarin, and in others they are optional, e.g. Indonesian. In some languages they are affixal and constitute an obligatory ending on numerals, which function as bound morphemes, e.g. Mokilese (HARRISON; ALBERT, 1976), while in others they can function as independent morphemes that also occur in the absence of numerals, (e.g. Chinese, Cheng and Sybesma, 2005). Also, some languages have a rather large inventory of classifiers, while others have only one or a few (see Grinevald, 2005 for examples).

Setting aside these differences, a basic distinction is often made between sortal and mensural classifiers (see for instance Grinevald, 2005). This distinction is based on the fact that some classifiers name units that are independently present in the denotation of the noun, while other classifiers do not. For the first type, the term sortal classifier is used, and the second type is called a mensural classifier.1 The example in (1a) illustrates the concept of a sortal classifier: the classifier zhī 'cl^{branch}' is used with a number of nouns, including the noun bi 'pen'. As indicated by (GRINEVALD, 2005, p. 1020), sortal classifiers "specify units (not quantity) in terms of which the referent of the head noun may be counted [...]." They constitute a small, closed class. The classifier is semantically redundant in the sense that the unit of counting it introduces is independently provided by the meaning of the noun (CHAO, 1968; CHENG; SYBESMA, 1998, 2005, 1999; DOETJES, 1997, 2012; GRINEVALD, 2005; LI et al., 2009; ZHANG, 2013). This can be illustrated by the fact that one sortal classifier usually functions as a default, which can be used whenever there is no more specific classifier available, and which is usually overgeneralized by children (see Erbaugh, 1992; Hu, 1993 for Mandarin). This classifier clearly does not contain information about the units that are counted, illustrating the semantically redundant nature of the classifier.

Mensural classifiers, on the other hand, are used to measure units that are different from the default unit associated to a count meaning of a noun (GRINEVALD, 2005). This can be illustrated on the basis of the example in (1b). The classifier jin 'CL^{half_kilo}' could be replaced by other classifiers corresponding

¹ In a more refined classification, one can also distinguish event classifiers, group classifiers, and taxonomic classifiers (see for instance Bender and Siegel 2004).

to other units in which the rice could be measured. The classifier thus fully determines the unit of counting. Mensural classifiers are usually compatible with nouns independently of whether these nouns provide default units of counting or not. When looking at classifier languages, languages vary substantially in terms of their inventory of mensural classifiers. Some languages possess a large class of mensural classifiers, while other classifier languages do not. Mensural classifiers come in various types. On the one hand, there are classifiers that correspond to abstract units of measurement, such as jìn 'CL^{half_kilo}'. In other cases, the classifier corresponds to a container noun in English, as illustrated in (2).

(2) liăng xiāng (de) shū

[Mandarin]

two box de book 'two boxes of books'

Depending on the presence or absence of de, two different meanings are available (LI; ROTHSTEIN, 2012; SYBESMA, 1992). If de is present, what counts is the quantity of books, and the books do not need to be contained in two actual boxes. Following Rothstein (2009, 2011), Li and Rothstein analyse this distinction in terms of measuring (the reading with de) and counting (the reading without de), claiming that these two readings correspond to two different structures. In cases where de is inserted, the classifier first forms a constituent with the numeral. The unit liang xiang then combines with the noun 'book', resulting in a reading in which the quantity of books corresponds to a quantity that can fill two boxes (i.e. two boxes-worth of books). In cases where the classifier first combines with the noun, one obtains the reading in which actual boxes containing books are counted. Even though the distinction between sortal and mensural classifiers is not made in this context, a typical sortal classifier will mostly figure in counting structures, while a typical mensural classifier, such as a unit of weight, will normally appear in measuring structures. Expressions such as box will easily show up with both types of reading, depending on the context.

The difference that is usually made between sortal classifiers and mensural classifiers is similar to the one that can be made in a non-classifier language such as English. Collective mass nouns (furniture) have a count meaning while they have a mass syntax. A structure such as the one in (3a) is, at least to some extent, comparable to sortal classifier structures in Mandarin, while the measure word pounds in (3b) shows a similarity with the mensural classifier jìn 'clhalf_kilo' in Mandarin.

- (3) a. two pieces of furniture
 - b. two pounds of apples/flour

As illustrated for the Mandarin classifier xiāng 'box' in (2), one can distinguish between two types of readings for container nouns (ROTHSTEIN, 2009; SELKIRK, 1977). Whereas two glasses of wine refers to actual glasses filled with wine in (4a), it does not in (4b), resulting in a measure reading similar to the one of the abstract measure word pounds in (3b):

- (4) a. John put two glasses of wine on the table
 - b. John put two glasses of wine in the cheese fondue

According to Greenberg (1972), a system with numeral classifiers can emerge from a system with structures of the type in (3). According to Ikoro (1994), this is what happened in Kana, a Niger-Congo language spoken in Nigeria.

From the point of view of the formal semantic literature, the link between classifiers and measure words has also been made, two of which will be discussed here. According to Chierchia (1997) mass nouns are kind-denoting expressions, whereas classifiers and measure words are inserted to make the kind-denoting noun compatible with a numeral.2 Chierchia defines kinds as functions from worlds into pluralities. In a given world w, a kind will denote the sum of all instances of that kind in w (cf. CHIERCHIA 1997, p. 349). In classifier languages all nouns are kind-denoting, resulting in the obligatory insertion of classifiers in the context of numerals. In a non-classifier language such as English, mass nouns are kinds, and typically trigger the insertion of a measure word, as illustrated in (3b). The function of the classifier or the measure word is to turn a kind-denoting expression into a predicate, which enables the classifier-noun complex to be combined with a numeral. As plural count nouns denote predicates, they can directly combine with numerals.

Krifka (1995), on the other hand, proposes that the classifier first combines with the numeral. According to his analysis, the main difference between classifier languages and non-classifier languages is that numerals in non-classifier languages incorporate an abstract unit of counting, which makes it possible to combine them with countable predicates. Note that Krifka assumes that the noun bear, on a par with its Chinese counterpart xíong 'bear', denotes the bear-kind Ursus. The plural marker is seen as an agreement marker and plays no role in the semantics. The derivation of the noun phrase three bears is given in (5) (KRIFKA, 1995:406), where Ri is the realization relation that relates kinds to their specimens in situation i, and OUi is an operator that takes a kind,

² This is a simplification: classifiers and measure words are also inserted in the context of certain quantity expressions (e.g. "many" in English and "ji" 'a few' in Mandarin) and some other expressions (e.g. demonstratives in the case of Mandarin).

yielding a measure function that measures the number of specimens of that kind in situation i: 3, 4

(5)	a.	[[bear]] = Ursus
	b.	$\llbracket \text{three} \rrbracket = \lambda y \lambda i \lambda x [\text{Ri}(x,y) \& \text{OUi}(y)(x) = 3]$
	с.	$[[three]]([[bear]]) = \lambda i \lambda x [Ri(x, Ursus) & OUi(Ursus)(x) = 3]$

As the numeral in (5b) incorporates an operator for the object or kind unit, the meaning of three bears is either three types of bears or three individual bears: in a situation i, three bears denotes a set of sums of three specimens of the kind Ursus. In a classifier language, on the other hand, the numeral does not incorporate a measure word, as illustrated in (6):

a. [[xíong]] = Ursus

(6)

- b. [san] = 3
- c. $[[zh\bar{i}]] = \lambda n \lambda y \lambda i \lambda x [Ri(x,y) \& OUi(y)(x) = n]$
- d. $[[san zhi]] = [[zhi]]([[san]]) = \lambda y \lambda i \lambda x [Ri(x,y) & OUi(y)(x) = 3]$
- e. $[san zhi]([xiong]) = \lambda \lambda x[Ri(x,Ursus) & OUi(Ursus)(x) = 3]$

In this case, the classifier is inserted because of the numeral, which can only be interpreted in relation to a noun when the classifier is present.

Recently, Bale and Coon (2014) argued on the basis of Mi'gmaq and Chol that the classifier systems in these languages offer evidence in favor of this second type of approach. In both of these languages, some numerals take classifiers, while others do not. This is easily accounted for if one assumes that the numerals that do not take classifiers are similar to the English numeral three above, while the ones that take classifiers are similar to Mandarin sān.

In the rest of this paper, I will examine the status of measure words in several classifier languages. According to a naïve view, one might assume that measure words in classifier languages are realized as classifiers rather than as ordinary nouns. That is, if one combines a measure word with a numeral, the measure word will be realized as a classifier, and there will be no other classifier present. However, I will argue on the basis of data from Mandarin (LI; THOMPSON, 1981), Mokilese (HARRISON; ALBERT, 1976) and Taba

³ The definition of $[[_{Num} three]$] is actually slightly more complex, as it generalizes over counting individuals and counting subkinds. Instead of the relation $R_i RT_i$ is used, which combines the realization relation with a taxonomic relation. Similarly, OU_i is replaced by OKU_i which is a measure function that measures either the number of specimens or the number of subkinds of a kind.

⁴ Krifka also offers an alternative to this analysis, which accounts for the distinction between English and Mandarin without assuming that the numerals are distinct. In this view, the numeral *three* has the same interpretation as Mandarin: $s\bar{a}n$ ($[three]] = [s\bar{a}n]] = 3$). In order to combine with mass nouns, a measure word is inserted. However, in this case, count nouns would have a different denotation, and "incorporate" a classifier, the result being a relational noun with a number argument ($[bear]] = \lambda n \lambda i \lambda x [RT_i(x, Ursus) \& OKU_i(Ursus)(x) = n]$). I will not consider this approach here, as it has no implications for the analysis of classifier languages.

(BOWDEN, 2001), that this view is too simple. This will lead to a discussion on the consequences of the data for the theories discussed above.

2. Measure words in classifier languages

In the most familiar classifier languages discussed in the literature, measure words are assumed to belong to the class of classifiers. Examples from Mandarin are given above in (1b) for an abstract measure word and (2b) for a measure word derived from a container expression. Grinevald (2005) gives examples of several classifier languages with large inventories of measure words that function as mensural classifiers, including Thai, Burmese and Vietnamese.

In some cases, classifiers occur without nouns. This is usually the case for measure words referring to times and currencies. The Mandarin examples in (7) show that the word tiān 'day' behaves like a classifier rather than as a normal noun (LI; THOMPSON, 1981:105): 5, 6

(7) a. sān tiān

b.

[Mandarin] three day 'three days' 'sān ge tiān three CL^{general}day

Given that ordinary nouns do not directly combine with numerals in numeral classifier languages, it is quite plausible that these words have a similar status as classifiers. From a semantic point of view, this is also expected: expressions such as tiān 'day' in (7) indicate abstract units of measurement, on a par with words corresponding to units of length or weight. If measure words are realized as classifiers in the language, we expect these words to be combined with nouns directly.

However, measure words that occur without nouns do not all behave in the same way. Whereas tiān 'day' behaves like a classifier in the sense that it directly combines with a numeral, Li and Thompson (1981:169) mention in a footnote that "[t]here are a few nouns denoting measures of time which do require classifiers." The examples they give are yuè 'month', xīngqi 'week' and zhōngtou 'hour'. Contrary to tiān 'day' above, these measure words require

⁵ Similar data can be found in Kana. According to Ikoro (1996:98), measure words such $z\dot{u}\dot{a}$ 'year', $s\dot{2}$ 'time' and $k\dot{2}b\dot{2}$ 'kobo' (a unit of currency) do not take classifiers, unlike other nouns. 6 In examples such as (7b), Li and Thompson use a hyphen between the numeral and the following sortal classifier. I use the same notation as in (1), without a hyphen (cf. for instance Cheng and Sybesma, 2015).

insertion of a classifier in the same contexts as ordinary nouns do (Li and Thompson 1981:169):

(8)	a.	liǎng	ge	yuè	b.	*liǎng	yuè
							[Mandarin]
		two	CL ^{genera}	^l month		two	month
		'two n	nonths'				
	с.	jĭ	ge	zhōngtou	d.	*jĭ	zhōngtou
		a few	CL ^{genera}	hour		a few	hour
		'a few	hours'				

The data show that the class of measure words is not a homogeneous class: some behave grammatically as classifiers and others as nouns.7 This implies that measure words can be ordinary nouns in a numeral classifier language. In what follows I will discuss two other languages in which measure words realize as ordinary nouns, and require the insertion of a sortal classifier.

The first of these langauges is Mokilese, an Austronesian language spoken on the Mokil atoll, part of Micronesia (HARRISON; ALBERT, 1976). Mokilese has a limited set of classifiers that are obligatorily present as suffixes on numerals. The numeral stems are bound morphemes and do not occur without a suffix that corresponds to one of the classifiers in the language. The data in (9) exemplify the use of the three sortal classifiers in the language: a general classifier (-w), a classifier for animate nouns (-men), and a long object classifier (-pas).8

(9)	а.	mwumw	jiluw/	jilmen	[Mokilese]
		fish 'three fish'	three-cl ^{general}	three-CL ^{animate}	[MOKIIese]
	Ь.	jeri roahm child two-c			
		'two children	,		
	с.	suhkoa	rahpas		
		tree	two-cL ^{long object}	ct	
		'two trees'			

⁷ The fact that measure words do not behave as a single class is reminiscent of the behavior of measure words in some Germanic languages, such as Dutch and German. In these languages, some measure words typically lack plural marking when combined with, for instance numerals, even though this is not allowed for ordinary nouns (GEERTS et al. 1984, p. 297). As in the Mandarin case, it depends on lexical properties of the measure word whether this is possible or not. In Dutch, one can say *twee uur* 'two hour' without plural marking, but not **twee dag* 'two day' (this should be *twee dagen* 'two days'). The question of whether these two observations are related is an issue for further research.

298

⁸ In the original examples, there are usually no word to word glosses, but only translations. The word to word glosses are mine, and are based on the information in the grammar.

The classifier -kij indicates that the units of counting are non-default units, as illustrated by the examples in (10).9

(10)	а.	peipa riahkij	
			[Mokilese]
		paper two-cl ^{piece/part}	
		'two scraps of paper'	
	b.	peipa riahpas	
		paper two-cl ^{long} object	
		'two sheets of paper'	
	с.	adroau riahkij	
		egg two-cl ^{piece/part}	
		'two pieces of egg'	
	d.	adrouau riahpas	
		egg two-cl ^{general}	
		'two eggs'	

Next to these items, which Harrison and Albert (1976) call classifiers, there are two more types of elements that can function as suffixes to a numeral base. The first are the words for ten, hundred and thousand, which are used to form complex numerals, see (11).

(11)	war	jilu-ijek	rah-pas	
			[Mokilese]	
	canoe	three-ten	two-cl ^{long object}	299
	'thirty	three canoes'		

In the second place, the morphemes -pak 'time' and -pwong 'day' are also used as suffixes on a numeral base, as illustrated in the examples in (12):

(12)	a.	Riapak	pahw	waluw			
				[Mok	ilese]		
		two-time	four-cl ^{general}	eight-cl ^{general}			
		'Two times	four is eight'				
	b.	Kama pirin	kijoula	Mwoakilloa	jilpwong		
		we fut	travel-to	Mokil	three-day		
		'We will leave for Mokil in three days'					

Interestingly, though, pak 'time' can also be used as a free morpheme, as illustrated in example (13).

⁹ Mokilese has one more element that is similar to classifiers, which is the plural marker – pwi. This marker is obligatorily present on indefinite plural nouns, and is similar to the plural 'classifier' $xi\bar{e}$ in Mandarin, see Doetjes 2012 for a discussion.

(13)	Woallo	ne	kijoula	Pohnpei	pak	jiluw	
						[Mokilese]	
	man-that already		rtravel.to	Ponape	time	$two-cL^{general}$	
	'That man has traveled to Ponape three times'						

The use of one and the same item as a suffix on a numeral base and as a noun is also found for loap 'half'. In cases where loap combines with oa-'one', loap is a suffix, as in oaloap in armaj in Mwoakilloa 'half of the people of Mokil'. However, in loap raw 'half two-cl^{general}', 'two halves', the word is used as a noun. In this last example, there might well be a relation between the use of the classifier with the fact that two halves easily has a counting meaning in the sense of Rothstein (2009, 2011), corresponding to two actual halves rather than the overall quantity corresponding to two halves. One might want to argue that this is also the reason for the use of pak in a noun position in the case of (13): in this example, the difference between measuring (counting the number of times something happened) and counting (specifying that there were three different occasions that are counted) is not easy to make. One could argue then that if the measure word is used as a noun, a counting interpretation is obtained.

A final example of a measure word is given the example in (14), taken from Harrison and Albert (1976:106).

(14)	Kidoahng	ngoahi ap	pas	pinjellen	apel		
					[Mokilese]		
	give	me on	e CL ^{long object}	pencil	and		
	jiluw	poaun in	koahpi	hen			
	$three{\textbf{-}}{CL}^{general}$	pound of	coffee				
	'Give me one pencil and three pounds of coffee'						
		(sai	id in a store	, for example)			

300

As in example (13), the measure word in (14), poaun 'pound', is not used as a suffix on the numeral base, but combined with a numeral plus the general classifier. The use of this word as a suffix on a numeral base is not reported, suggesting that this measure word may only surface as a noun. However, given the example and the context that the authors give for its use, it is not clear whether the example has a counting reading and/or a measuring reading in the sense of Rothstein (2009, 2011). The situation of a store indicated by Harrison and Albert (1976) would allow for both types of readings: three packages of one pound each (counting) or an amount of three pounds (measuring).

To conclude, Mokilese measure words are not systematically treated as classifiers, and also occur as nouns. In the case of poaun 'pound' this may well be the only possibility, while pak 'time' appears either a classifier or a noun. The data from Harrison and Albert suggest that the use of the classifier as a noun

might indicate a counting rather than a measuring reading, but whether this is systematically the case is an issue for further research.

Another example of a numeral classifier language in which measure words are realized as nouns is Taba (Austronesian, Indonesia, Bowden, 2001). In Taba, the classifier is realized either as a proclitic (indicated by =) or as a prefix (indicated by -) on the numeral, as illustrated in (15) (see Bowden 2001: 240, 243):10

(15)	a.	Ι	not		yan	sishot	
							[Taba]
		i	n=ot		yan	sis=hot	
		3sg	3sg=ta	ake	fish	CL ^{2-9animals} =four	
		'He ca	aught fo	our fish	,		
	b.	amplo	р	pwon	am		
		amp l	ор	p-woi	nam		
		envelo	ope	CL ^{gener}	al-six		
		'six envelopes [of A4 size]'					

In Taba, some measure words are realized as classifiers, but others are realized as nouns. The language even has a specific classifier that is used for measure word, as illustrated for the measure word liter in (16), taken from Bowden (2001:253).

(16) liter halu

liter ha=lu liter CL=two 'two liters'

Other words that take the classifier ha= are meter 'meter' and song 'handspan', as well as units of time and container nouns. If the classifier ha= is used, the container noun is used as a measure word, but if the general classifier is used, the noun denotes an object, as illustrated by ember 'bucket' in the pair in (17).11 (18) illustrates the measure word use of blek susu 'milk can' with the classifier ha= in a full sentence (BOWDEN, 2001, p. 397).

Revista Letras, Curitiba, UFPR, n. 96, pp.291-308, jul./dez. 2017. ISSN 2236-0999 (versão eletrônica) 301

[Taba]

¹⁰ In the original glosses, Bowden uses CLASS for classifier, and in a few cases adds a semantic description of the classifier between brackets. In this paper, I use the format used in the examples of Mandarin above, where CL stands for classifier. The superscripts are based on Bowden's characterizations (Bowden 2001: 243). In cases where it is not clear what the superscript should be, no superscript is given (cf. ha=, which is sometimes described as 'intervals of measurement' and sometimes as 'times'). As for the other abbreviations used in the Taba examples: sg = singular, POSS = possessive, APPL = applictive and DETR = detransitivising.

¹¹ It is not clear how the 'counting' reading of Rothstein (2009, 2011) is realized. Given the translation 'bucketful' for *ember* in combination with ha=, it might be the case that in the counting reading, the general classifier has to be used, but no relevant examples are given.

(17) a. ember halu

[Taba]

ember ha=lu bucket CL=two 'two bucketfuls (of something)' b. ember plu ember p=lu bucket CL^{general}=two 'two buckets'

(18) Bonci ni dungko duga blek susu haso bonci ni dumik-o duga blek susu ha=so peanut 3sg.poss be finished-APPL (remainder) only can milk cl=one 'The remainder of the peanuts make up just one milk-can full.'

The classifier ha= is also used in temporal contexts as illustrated in (19a), taken from Bowden (2001: 220) and in complex numerals as in (19b), from Bowden (2001: 246):12

(19) a. Taplod haso nak

b.

[Taba]

ta-plod ha=so nak detr-erupt CL=one again 'It (the mountain) erupted once more.' mapin matyo halu mapin mat=yo ha=lu woman CL^{human}=ten CL=two 'twenty women'

Besides the special classifiers for 'intervals of measurement', in Bowden's terms, speakers also use the general classifier p- with measure words, which Bowden takes to be an influence of Malay. Words such as jam 'hour, *ngan* 'day' and *pait* 'month' require *p*- rather than *ha*=.

302

¹² Bowden glosses *yo* sometimes as *ten* and sometimes as CLASS, for classifier. The reason for this is, presumably, that the general classifier p- is absent in the context of *yo*, as well as *utin* 'hundred' and *calan* 'thousand'. This is due, according to Bowden, to the fact that p- is a true prefix and can only have scope at the word level, that is, at the level of the numeral root. Given that *yo*, *utin* and *calan* are necessarily combined with *haso* 'one', *halu* 'two' etc. (*yo haso* 'ten', *yo halu* 'twenty'), they only permit the use of classifiers with phrasal scope. This is a very interesting observation, as it shows that the use of a classifier can be blocked by morphological constraints: if no appropriate classifier is available in the lexical system of the language, the classifier can be left out. If one wants to claim that the classifier is inserted for semantic reasons, one has to assume (or rather stipulate) that in these cases p- is replaced by a null classifier that may have phrasal scope and that only occurs when no overt classifier is available. I will leave this issue aside for reasons of space.

On the other hand, Bowden observes that there are also measure words that function as classifiers. This is the case for lof 'armspan', tonat 'ten armspans'. An example is given in (20) (BOWDEN, 2001, p. 261):13

(20)	Kurusi	ni	mlongan	lof	SO	lo	lomo
						[Taba]	
	Kurusi	ni	mlongan	lof	SO	lo	lomo
	chair	3sg.poss	length	CL ^{arms}	oan one	and	other
	'The chair is over a handspan long. '						

The examples show that measure words are also not a homogeneous category in Taba: in some cases they behave as mensural classifiers, and combine directly with a numeral, while in other cases a classifier is inserted. In most cases, this is the classifier ha=, but sometimes the general classifier p- is used. Contrary to what might be speculated about Mokilese, the distinction between a classifier position and a noun position is not exploited to distinguish between counting and measuring: the noun uses of the classifiers are clear cases of measure words in their measuring reading. In this respect, the data are more similar to what is observed for the isolated examples in Mandarin at the beginning of this section.

The examples in this section show that measure words may be realized as classifiers, but also in positions where we normally find ordinary nouns. Moreover, even in a single language, there seems to be variation in this respect. As illustrated by the examples of Taba, classifier languages may have a specific classifier for measure words, even though this language also permits the use of a general classifier in the context of certain measure words.

3. Consequences for our view on classifiers

If measure words can be classified nouns themselves, what does this mean for our view on classifiers and the reasons why they are inserted? In what follows, I will first discuss the data in relation to Rothstein's (2009, 2011) distinction between measuring and counting. Then I will turn to the question as to why classifiers are inserted, and the semantic status of nouns that co-occur with classifiers.

Within the framework of Rothstein, one might expect that measure words with a measuring interpretation behave as classifiers, while counting uses of container words would be more noun-like. Whereas the counting uses can <u>be seen as nouns</u> with a meaning that is similar to that of well-behaved count 13 The grammar does not give examples where the measure words (whether behaving as a noun or as a classifier) form a constituent with a 'measured' noun (cf. (1)). The glosses suggest that combinations are possible (see for instance (17a)), but no concrete examples are given.

nouns (showing no possible overlap between atoms), this is not true for measure words on their measuring reading: two atomic stretches of time corresponding to one hour may overlap. The data discussed in this paper show, however, that both types of measure words (ones that have a counting interpretation and ones that have a measuring interpretation) may trigger insertion of a classifier. Even though the data from Mokilese are not conclusive in this respect, the data from Mandarin and Taba illustrate that measure words on their measure reading may need the insertion of a classifier. In what follows, I will explore some of the consequences of the assumption that this indicates that the measure word in these cases should be treated as a noun (cf. for instance Bale, 2006 for English).14

According to Chierchia (1997), nouns in classifier languages have a kind interpretation. The function of the classifier is to turn the noun into a predicate, so that it can be interpreted in combination with a numeral. Recall that for Chierchia, kinds are functions from worlds into pluralities: in a given world w, a kind will denote the sum of all instances of that kind in w (cf. CHIERCHIA, 1997, p. 349). In an expression such as ji ge zhongtou 'a few hours' in (8b), the noun zhongtou should thus be analyzed as a kind-denoting expression. But what does this mean for this noun? If for a given world w, the kind is defined as the sum of all instances of that kind in w, the noun zhongtou should denote in a given world w the sum of all instances of hours in w. This interpretation is not without problems, as it seems hard to make sure that the denotation of zhongtou would be different from the denotation of other measure words of time such as xīngqi 'week', which also requires insertion of a classifier: the sum of all possible stretches of time in a given world that are weeks and the sum of all possible stretches of time that are hours both correspond both to the sum of all possible stretches of time in a given world and cannot be distinguished. Even though the analysis of kinds proposed by Chierchia is attractive, it is problematic if expressions such as zhongtou and other measure words are treated as kind denoting expressions.

Note that the problem of Chierchia's account is not shared by other approaches in which the classifier changes the meaning of the noun in such a way that it becomes compatible with a numeral. In a syntactically motivated account, Cheng and Sybesma (1999) argue that classifiers turn number-neutral predicates into singular predicates.15 Under the assumption that the numeral needs a singular predicate (cf. IONIN; MATUSHANSKY, 2006), the need of the classifier is not only syntactically motivated, but also semantically motivated.

Let us now turn to the analysis of Krifka (1995). The data in this paper do not seem to offer evidence against his type of approach, even though the

¹⁴ I do not want to defend the claim that this assumption is a necessary one: there may be different ways to explain the data. However, given the parallel distribution of the nouns and the measure words, the simplest assumption is that the measure word and the noun have the same type of semantics in this context.

¹⁵ Their arguments are based, among other things, on the meaning of bare CL-N sequences in the absence of numerals.

observation that measure words do not need to be realized as classifiers is interesting from this perspective. Within this approach, the overlap problem discussed above needs to be addressed for independent reasons, as quite generally, the units that are introduced by classifiers may not overlap. Krifka (1995, p. 400) illustrates this on the basis of the example in (21).

(21) sān qún (de) xióng

[Mandarin]

three herd de bear 'three herds of bears'

Under the assumption that classifiers, such as herd in this example, are additive measure functions, a non-overlap condition is introduced by the classifier, given that for additive measure functions sum formation is only possible for non-overlapping entities. If this same assumption is made for sortal classifiers, such as the general classifier ge, there is no overlap problem related to counting in examples such as jĭ ge zhōngtou 'a few hours'.

The meaning of the noun that occurs with a classifier is a separate issue. For Chierchia, the kind denotation of nouns that occur with numeral classifiers is a crucial part of the analysis as a whole, as is the specific interpretation of what it means to be a kind. However, for an approach to numerals along the lines of Krifka (1995), in which the classifier is a seen as a measure expression that is needed by certain types of numerals, the status of the noun as either a kind or a predicate is an independent question. Even though Krifka assumes that nouns that occur with classifiers are kind-denoting, the classifier introduces the function R, relating kinds to instances of the kind, so nothing hinges on this assumption.16

4. Conclusions

This paper discussed some preliminary data showing that measure words do not necessarily form a homogeneous category in classifier languages. Besides being used as mensural classifiers, there are also measure words that require insertion of a sortal classifier, and as such behave in a similar way as nouns with a count interpretation. As shown, Mandarin, Taba and Mokilese all offer examples of measure words that take a sortal classifier when combined with a numeral or another count quantity expression that triggers classifier insertion. In the case of Mandarin, the data are restricted to a few expressions that alternate with classifiers that combine with a numeral without a noun (in particular measure words expressing units of time). Some of these measure words require the insertion of

¹⁶ See Bale and Coon (2014), who adopt Krifka's analysis for numerals and assume for the sake of simplicity that the noun that co-occurs with the classifier has a set denotation.

the general classifier ge, while others behave as classifiers themselves (cf. (7) and (8)). In Mokilese, some measure words occur both with a classifier and without, while other measure words seem to require the insertion of a classifier. In this language, it might be the case that the presence of a classifier is a reflection of a counting interpretation in the sense of Rothstein (2009, 2011), but this needs to be checked with more data. Finally, in Taba, a particular classifier is used with measure words. Besides this, the language also has some measure words that take the general classifier instead, and some measure words are used as classifiers themselves.

From a theoretical point of view, the idea that measure words can be nouns in classifier languages raises a number of issues that have to do with the fact that the atomic entities corresponding to units of measurement typically overlap: two stretches of time corresponding to an hour may correspond to two hours, but this is obviously not necessarily the case. This is problematic for the type of interpretation that Chierchia assigns to kinds, as the kinds corresponding to different units of time would be indistinguishable. In other approaches, a non-overlap condition is needed in counting structures.

The data discussed in this paper raise a number of questions for further research. In the first place, what are the properties of measure words in nonclassifier languages? Do these also fall into different classes? A first indication that this might be the case comes from the fact that certain measure words, but not others, fail to be marked for plural in the context of numerals (as in twee jaar 'two year' vs. *twee dag*(en) 'two days'). A second question is whether certain types of measure words in classifier languages behave like nouns more than others. For instance, it might be the case that measure words such as zhōngtou 'hour', which are used in the absence of "a measured noun", have this type of property more often than others. Another question is whether the noun/classifier distinction is regularly exploited in order to express a difference between measuring and counting. Answering these questions requires comparative data from a much larger sample of languages.

Acknowledgements.

This paper is based on part of a presentation I gave at the XI Workshop on Formal Linguistics (November 2017, Curitiba, Brazil). I would like to thank the organizers and the audience for comments and stimulating discussion. A special thank you to Kate Bellamy for proofreading the prefinal version of this paper.

References

BALE, A. *The universal scale and the semantics of comparison*. Ph.D. Thesis, McGill University, Montreal. 2006.

BALE, A.; COON, J. Classifiers Are for Numerals, Not for Nouns: Consequences for the Mass/Count Distinction. *Linguistic Inquiry*, v. 45, n. 4, p. 695–707, 2014.

BENDER, E. M.; SIEGEL, M. Implementing the syntax of Japanese numeral classifiers. In: KEH-YIH Su et al. (Eds.). *International Conference on Natural Language Processing*, p. 626-635. 2004.

BOWDEN, J. F. *Taba: description of a South Halmahera language*. Canberra: Pacific Linguistics, 2001.

CHAO, Y. R. *A grammar of spoken Chinese*. Berkeley: University of California Press. 1968.

CHENG, L.; SYBESMA, R. Yi-wan Tang, yi-ge Tang: classifiers and massifiers. *The Tsing Hua Journal of Chinese Studies*, v. 28, 385-412, 1998.

CHENG, L., SYBESMA, R. Classifiers in four varieties of Chinese. In: CINQUE, G.; KAYNE, R. S. (Eds.). *The Oxford handbook of comparative syntax*. Oxford: Oxford University Press. 2005. p. 259–292.

CHENG, L.; SYBESMA, R. Mandarin. In: KISS, T.; ALEXIADOU, A. (Eds.). *Syntax – Theory and Analysis*. An international handbook. Berlin: Mouton de Gruyter. 2015. p. 1518–1559.

CHENG, L. L.-S.; SYBESMA, R. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry*, v. 30, p. 509-542, 1999.

DOETJES, J. Quantifiers and selection. On the distribution of quantifying expressions in *French*, *Dutch and English*. The Hague: HAG, 1997.

DOETJES, J. Count/mass distinctions across languages. In: MAIENBORN, C.; von HEUSINGER, K.; PORTNER, P. (Eds.). *Semantics: An International Handbook of Natural Language Meaning*. Berlin: De Gruyter. 2012. p. 2559–2580.

ERBAUGH, M. S. The acquisition of Mandarin. In: SLOBIN, D. I. (Ed.). *The crosslinguistic study of language acquisition* Vol. 3. Hillsdale, NJ: Lawrence Erlbaum Associates. 1992. p. 373–455.

GEERTS, G.; HAESERYN, W.; de ROOIJ, J.; van den TOORN, M.C. *Algemene Nederlandse Spraakkunst*. Groningen/Leuven: Wolters-Noordhoff/Wolters. 1984.

GILL, D. Numeral classifiers. In: HASPELMATH, M.; DRYER, M. S.; GIL, D.; COMRIE, B. (Eds.). *The world atlas of language structures*. Oxford: Oxford University Press. 2005. p. 226–229.

GREENBERG, J. Numeral classifiers and substantival number: problems in the

genesis of a linguistic type. Stanford papers on language universals, n. 9, p. 1-39, 1972.

GRINEVALD, C. Classifiers. In: LEHMANN, C.; BOOIJ, G. MUGDAN, J. (Eds.). *Morphology: a handbook on inflection and word formation*. Berlin: Walter de Gruyter. 2005. p. 1016-1031.

HARRISON, S.; ALBERT, S. Mokilese reference grammar. Honolulu: University Press of Hawaii. 1976.

HU, Q. The acquisition of Chinese classifiers by young Mandarin speaking children. Ph.D. Thesis. Boston University, Boston. 1993.

IKORO, S. Numeral classifiers in Kana. Journal of African languages and linguistics, n. 15, p. 7-28, 1994.

IONIN, T.; MATUSHANSKY, O. The composition of complex cardinals. Journal of Semantics, n. 23, p. 315-360, 2006.

KRIFKA, M. Common nouns: a contrastive analysis of Chinese and English. In: CARLSON, G.; PELLETIER, F. (Eds.). The generic book. Chicago: The University of Chicago Press, 1995. p. 398-412.

LI, C. N.; THOMPSON, S. A. Mandarin Chinese: a functional reference grammar. Berkeley: University of California Press. 1981.

LI, P.; DUNHAM, Y.; CAREY, S. Of substance: the nature of language effects on entity construal. Cognitive Psychology, v. 58, p. 487-524, 2009

LI, X.; ROTHSTEIN, S. Measure readings of Mandarin classifier phrases and the particle de. Language and Linguistics, v. 13, p. 693-741, 2012

> ROTHSTEIN, S. Individuating and measure readings of classified constructions: evidence from modern Hebrew. Brill's Annual of Afroasiatic Languages and Linguistics, v. 1, p. 106-145, 2009.

ROTHSTEIN, S. Numbers: counting, measuring and classifying. Paper presented at Sinn und Bedeutung 16, Utrecht University, September 6-9, 2011.

SELKIRK, E. Some remarks on noun phrase structure. In: CULICOVER, P. W., WASOW, T., AKMAJIAN, A. (Eds.). *Formal Syntax*. London: Academic Press. 1977. p. 285-316.

SYBESMA, R. Causatives and accomplishments. The case of chinese ba. Ph.D. Thesis, Leiden University, Leiden, 1992.

ZHANG, N. N. Classifier structures in Mandarin Chinese. Berlin: De Gruyter. 2013.

Submetido em: 15-08-2017 Aceito em: 15-08-2017

308