

How “General” are General Classifiers? With Special Reference to *ko* and *tsu* in Japanese

Author(s): David A. Zubin and Mitsuaki Shimojo

*Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Semantic Typology and Semantic Universals* (1993), pp. 490-502

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

---

*The Annual Proceedings of the Berkeley Linguistics Society* is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

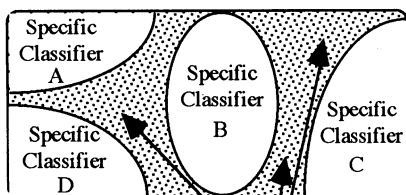
# How "general" are general classifiers? with special reference to *ko* and *tsu* in Japanese

David A. Zubin & Mitsuaki Shimojo\*

Department of Linguistics and Center for Cognitive Science  
SUNY at Buffalo

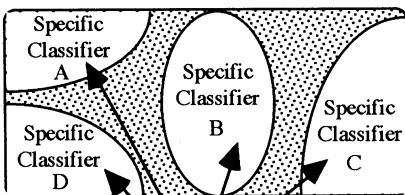
Two principal models form the background to the study of numeral classifier and other semantically motivated<sup>1</sup> nominal classification (NC) systems: (a) a lexical model in which the inventory of categories is viewed as a list of lexical items derived from nouns, and preserving much of the rich and somewhat idiosyncratic meaning of these source items; and (b) a grammatical model in which this inventory, or at least some part of it, is viewed as a grammatical system in which the individual forms have highly abstract and relational meanings, and behave as a system of alternatives. The first model--characterized by a "lexical" particularism in which each form has its own particular history and is related only loosely and unsystematically to other neighboring choices in the language--can be traced to turn of the century lexicalists such as Gilliéron (1918), and is reflected in some current work of the grammaticalization movement (e.g. Bybee (1985)). The second model is characteristic of the Saussurean extremism of Roman Jakobson (e.g. 1936) and other Pragueans who saw, and continue to see grammar and lexicon as highly structured systems of choices in which "tout se tient." In the middle of these extremes falls the centrist linguistic doctrine that grammatical meanings form structured systems while lexical meanings do not. Numeral classifier and other semantically motivated NC inventories fall squarely in the middle between these poles, resulting in controversy over their lexical vs grammatical status. In support of the system analysis is the fact that NC inventories ranging from the noun classes of Dyirbal to the verb-stem classifiers of Navajo to the numeral classifiers of Mandarin to the gender classes of German seem to have "general" members, giving the inventory its systemic structure by picking up the semantic "residue" (Dixon, 1982) of nouns which are semantically incompatible with more specific categories, i.e. by playing the role of globally applicable alternative choice. But the terms "general" and "default" classifier have come to be used in the literature to cover a variety of semantically and pragmatically distinct situations.

**Figure 1a:** complement function - remainder category for entities outside the semantic domains of specific classifiers



Complement Classifier

**Figure 1b:** default function - substitutes for any specific classifier within pragmatically specialized contexts.



Default Classifier

This paper will attempt to decompose the notion of "general" classifier into more precisely definable distributional patterns, and will briefly survey general classifiers in

several typologically diverse NC inventory types before turning to the results of experimental and survey research on the two categories *ko* and *tsu* in Japanese (Shimojo, 1993a,b), research that will pin down their semantic and pragmatic properties more precisely than has been done in the past.<sup>2</sup>

Under the "general classifier" rubric at least three theoretically and empirically distinct phenomena may be distinguished. In many cases these functions are exhibited by the same classifier category, although this is not necessarily the case. A classifier acting in COMPLEMENT FUNCTION (Fig 1a) is a remainder category for nouns and/or referents outside the semantic domains of the other classifiers in the inventory, or at least only in the semantic periphery of these other categories. A classifier in DEFAULT FUNCTION (Fig 1b) can be substituted for other classifiers under specialized pragmatic conditions in which the speaker abandons the available precision of a semantically specific classifier in favor of a semantically neutral, and hence globally available one. Here there is no issue of semantic complementation, since the *default* classifier is used even though the noun or referent fulfills the summary description (Smith & Medin) of a specific classifier.

Yet a third type is UNSPECIFIED REFERENT FUNCTION, in effect when the speaker has little or no information about a referent and builds a referring NP around a non-specific expression as in "there's *something scary* in that house." In addition to exploring the functions displayed by "general classifiers" we will raise the issue of whether the *ko* or *tsu* categories have core meaning,<sup>3</sup> the presence of which is at least potentially in conflict with the general classifier functions outlined above.

**Table 1:** a survey of "general" classifier functions. "Form" = the general classifier category in question. In many cases the core semantics of the general category is problematic.

a. complement function

language	characteristics of system	form	core semantics
Diari	2 gender classes marked on pn/det	<i>nu</i>	masc
Kala Lagau Ya	2 gender classes marked on pronoun	<i>na</i>	fem
Dyirbal	4 gender classes marked on loc. det	<i>bala</i>	inanimate
German	3 gender classes marked on pn/det	(none) <sup>4</sup>	--
Panare	about 21 genitive classifiers <sup>5</sup>	<i>iyu</i>	alienable poss.
Navajo	about 14 predicate classifiers <sup>6</sup>	<i>si'a</i>	3D
Yidiny	about 20 "generic" classifiers <sup>7</sup>	<i>wirra</i>	hand-size instr obj

b. default function

English	unlimited lexical choices in individuating construction (a/the X of Y)	<i>piece</i>	shape-neutral concrete indiv.
Mandarin	100+ numeral classifiers, about 20 in frequent use	<i>ge</i>	human
Chichewa	7 paired sg-pl noun classes, marked throughout NP, pns, and in predicate	<i>-a</i> <i>-zi</i>	human inanimate

c. unspecified referent function

German	3 gender classes marked on pn/det	<i>neuter</i>	inanimate
Navajo	about 14 classes expressed by verb stem	<i>sila</i>	1D flex
Japanese	numeral classifiers	<i>tsu</i>	inanimate

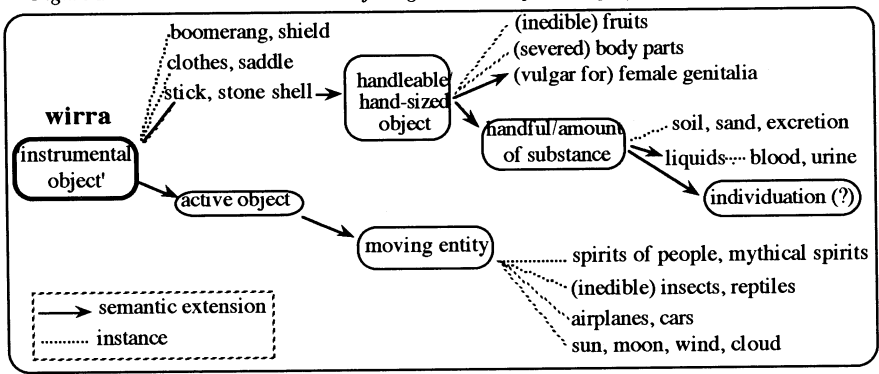
Table 1 summarizes the properties of general classifiers in some typologically diverse languages. Diyari (Austin, 1981) and Kala Lagau Ya (Bani, 1987), both Australian languages, give a clear and simple picture of complement function. They each have a two-gender system; in Diyari one pronoun is restricted to females, and the other *nu* is used for males, non-sex specific animate terms, and inanimates; in other words, for all possible referents outside the *female* category. Kala Lagau Ya's system has the inverse structure: the pronoun *na* for all possible referents outside the *male* category. These pronouns, despite their complement function, still retain the core semantic properties *male* and *female* respectively, e.g. although *nu* can refer to a wide variety of entities, male beings are privileged exemplars in the category. Dyirbal (Dixon, 1982) is another Australian language with a semantically more complex gender system. The complement category *bala* has no overall coherent structure, unlike the other three categories in the system, which have semantically coherent radial structure (Lakoff, 1986). But *bala* does have privileged members, which are inanimate. German (Zubin & Köpcke, 1986), characteristic of the gender systems of Indo-European, has highly complex semantic structure in its gender categories, and within specific semantic domains any of the categories may have complement function. For example, the complement gender category in the *beverage* domain is *masculine*, within the *ship* domain is *feminine*, and within the *building* domain is *neuter*. But there is no one complement category at a more global level of conceptual structure.<sup>8</sup>

Panare (Carlson & Payne, 1989) is a Cariban language with a well-developed set of genitive (possessed) classifiers for nouns which themselves cannot be inflected for possession. The classifiers divide the world up into 20 or more functional categories, most of which are quite specific (e.g. *ko* for weapons or *uyung* for artificial light sources). When referents not falling into these specific categories are possessed, they take the *complement classifier iyu*, which in addition is productive for novel referents, such as plastic toys.

Navajo (Young & Morgan, 1980), an Athapaskan language with verb-stem classifiers characteristic of this family, has two general categories: the (shape) category for 3D objects is also used for a variety of entities not admissible to other categories, i.e. serves complement function, and the 1D flexible category (see Table 1c).

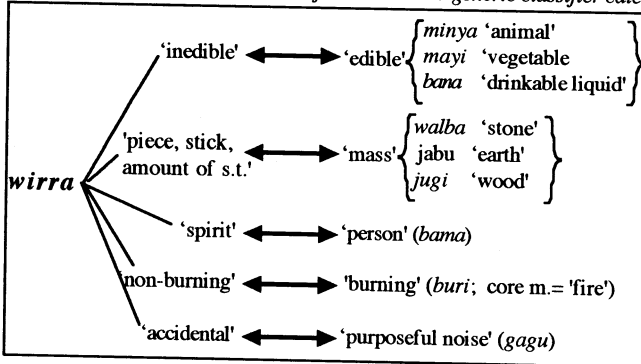
Coming closer to numeral classifier systems, Yidiny has a larger inventory of superordinate noun classifier categories, all of which have highly coherent semantic structure,<sup>9</sup> with the exception of *wirra*, which has complex radial structure (Fig. 2)

Figure 2: Radial semantic structure of the generic classifier category *wirra* in Yidiny



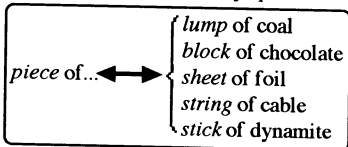
inferable from Dixon's (1982) description, and near-global complement function (Fig. 3). Fig 3 shows that *wirra* is used just in case the semantic domain of one of the other categories is not applicable. For example, *minya* is used for edible animals, and *wirra* for all animals considered inedible. The radial structure in Figure 2 suggests, however, that handsized instrumental objects are privileged exemplars in the category.

Figure 3: Complement relations of *wirra* to other generic classifier categories



**Default function** (Table 1b). The individuation construction in English (Fig 4) gives a clear illustration. Speakers can freely alternate between the shape-neutral general classifier "piece" and the shape-specific classifiers such as "sheet," depending on the amount of information required at that point. Erbaugh's (1986) pear-story data show that in spontaneous narrative production Mandarin speakers will substitute the general classifier *ge* for more specific classifiers even though they may insist on the specific ones in an elicitation task. Recent work by Ahrens (1993) shows that when Mandarin speakers' attention is drawn to the issue of classifier

Figure 4: default function of "piece"



choice by an elicitation task they more easily substitute *ge* for peripheral members than for central members of specific categories, with the same prototype effects that we will show for Japanese general classifiers below, i.e. *ge* does show some tendency toward *complement* function. Finally, Chichewa reveals (Corbett & Mtenje, 1987) default categories in the choice of verb agreement when nominals from different noun classes are conjoined. Human nouns take class-1 plural agreement, even if neither of the nouns is in class 1; and conjoined inanimates take class-6 plural agreement, even if neither noun is in class 6. The plurals of classes 1 and 6 thus serve as default categories for human and inanimate plurals when conjoining blocks the grammatical pairing of singular and plural noun classes.

**Unspecified referent function** (Table 1c) is clearly illustrated by the *neuter* gender category in German. Examples (1a,b) necessarily refer to entities which are at least implicitly lexicalized with nouns of *masc*- and *fem*-gender, respectively, e.g. a spoon (*masc*) and a fork (*fem*). But (1c) can refer either to a specifically *neuter* item (e.g. a

- 1a. Gibst Du mir **den**?    b. Gibst Du mir **die**?    c. Gibst Du mir **das**?  
 give you me **it(masc)**    give you me **it(fem)**    give you me **it(neut)**  
 'would you give it to me?'

knife) or to any item, the properties of which the speaker is unsure of, e.g. something the addressee is concealing in his hand. This function is further evident in the use of *neuter* pronouns with "weather" and other "environment" predicates, and in the *neut*-gender of non-taxonomic superordinate nouns such as *Zeug* 'stuff' and *Ding* 'thing' (Zubin & Köpcke, 1986).

In Navajo the predicate classifier for 1D flexible objects has a number of radial extensions, and is also used for objects whose properties are unknown. And in Japanese, the numeral classifier *tsu*, whose complement and default functions will be discussed in the next sections, is also used for objects of unknown properties.

**General Classifiers in Japanese.**<sup>10</sup> Against this background picture of different functions for general classifiers the situation in Japanese is particularly instructive, since there are really two classifier inventories (exemplified in Table 2), one

**Table 2:** the two classifier sets

	classifiers co-occurring with:	
	native numerals (1-10 only)	Sino-J numerals (full set)
specific classifiers	tsubu 'grain' kabu 'root veg' mune 'building' eda 'branch'	hon '1D' mai '2D' ko '3D' dai 'vehicle' ki 'airplane' ken 'house'
general classifier	<i>tsu</i>	<i>ko</i>

syntactically tied to a limited set of native Japanese numerals from one to ten, the other to a full set of Sino-Japanese numerals,<sup>11</sup> each of which has a "general classifier." *tsu* with the native numerals, and *ko* with the Sino-Japanese ones, allowing a parallel study of the two. Table 3 summarizes the results of a simple listing task (cf. Rosch, 1973), in which 11 native speakers were asked to give 5 nouns they could use with each of the classifiers *ko* and *tsu*, without time constraint. With few exceptions, all nouns listed were either outside of specific

classifier categories, or at most were peripheral members of these.<sup>12</sup> This demonstrates the basic complement function of *tsu* and *ko*: they apply primarily to semantic territory outside the semantic domains of specific classifiers.

**Table 3:** Nouns given for *tsu* and *ko* in a free listing task. Full data is in Table 7.

	Nouns Listed for <i>Ko</i>	Nouns Listed for <i>Tsu</i>
Central Members of Specific Classifier Categories	isu (chair) - kyaku	sanrinsya (tricycle) - dai osara (dish) - mai naironbukuro (plastic bag) - mai hon (book) - satsu isu (chair) - kyaku
Peripheral Members of Specific Classifier Categories	(none)	mado (window) - mai kasettoteepu (cassette tape) - hon ronbun (thesis) - hon
Not Members of Any specific Category	everything else (22 nouns)	everything else (34 nouns)

A more precise measure of the complement status of *tsu* and *ko* is given by the results of a rating task with 102 native speakers, summarized in Table 4. They were given

short generic sentences each containing a classifier construction with one of the general classifiers *ko* or *tsu*, as in (2)

2. mukoo-ni mikan-ga san-ko aru 'There are three oranges over there'  
 over:there-LOC orange-nom 3-NC exist

Speakers were asked to rate the acceptability of the sentence on a 5-point scale. *Tsu* and *ko* were paired with nouns which are members of specific classifier categories--these are listed across the top of the table. Nouns were chosen so that some would be central, others peripheral members of their respective classifier categories.<sup>13</sup> The results show that *ko* and *tsu* are more acceptable with peripheral members of other categories than with central ones, as indicated by the positive values of the difference scores in the right-hand column. It is also crucial to note that this increase in acceptability occurred for each of the five specific classifier categories tested.

**Table 4.** Central and peripheral member nouns of specific classifier categories - *tsu* and *ko* are more acceptable with peripheral than with central members. Rating scale 1 - 5 (5=fully acceptable)

noun combined with ...	specific classifier category to which noun belongs						Overall Average	Diff. of Average
	hon	mai	dai	ken	ki			
<b>tsu</b>	central	2.33	1.38	2.24	3.27	2.08	2.26	+1.06
	peripheral	2.87	3.46	3.50	3.83	2.95	3.32	
<b>ko</b>	central	1.60	1.24	1.69	1.62	1.59	1.55	+0.68
	peripheral	1.68	2.31	2.84	2.21	2.13	2.23	

These results suggest that the complement function of *ko* and *tsu* begins to exert itself as the graded structure of specific categories fades out at their peripheries. Note also that the difference score for *tsu* is greater than for *ko*, and that the absolute scores for *tsu* are over all about one rating point higher than for *ko*, suggesting that *tsu* fills this complement function more robustly than *ko* does.

The relatively low absolute scores for *tsu* and *ko* in this test might appear to be a problem, unless we keep in mind that we are dealing with their tendency to extend into the semantic territory of the specific classifiers. *Tsu* and *ko* are *less unacceptable* with the peripheral members of these categories. A pilot test in which some of these nouns were combined with other *specific* classifiers (e.g. *denwa* 'telephone call' with *mai*) yielded scores approaching 1.0, i.e. absolute unacceptability.<sup>14</sup>

Further evidence for the complement function of *tsu* and *ko* is provided by distorting the form of everyday objects. Our reasoning is that stereotypic exemplars of such objects are an important component of semantic representations, and criterial for the applicability of classifiers. When these objects are distorted in a way that makes them non-stereotypic for the noun, then speakers will tend to abandon the classifier with which the noun is associated, and turn to one of the general classifiers. Table 5 shows that as we spatially distorted several different objects in ways corresponding to the everyday use or variations of these objects, the 11 speakers we tested shifted from the relevant specific classifier to either *tsu* or *ko*, depending on which choice we gave them. The flat cutting board was distorted by making it thicker and thicker until it became more of a 3 dimensional cutting block. The flat napkin was crumpled so that it was no longer flat. The candle was changed into a candle stub, and the piece of chalk was made shorter and then into a stub. Note that all distortions used are actual-

**Table 5:** Distortions of *mai* (2D) and *hon*(1D) type objects. Subjects tended to use *tsu* or *ko* for the distortions.

Lexical item	Percept shown to the subjects	no. of subjects choosing each classifier			
		<i>mai</i> vs <i>tsu</i>		<i>mai</i> vs <i>ko</i>	
<b>cutting board</b>	1. flat and thin	7	4	10	1
	2. thicker	7	4	10	1
	3. still thicker	2	9	3	8
	4. still thicker	1	10	2	9
	5. blocky	1	10	0	11
		<i>mai</i> vs <i>tsu</i>		<i>mai</i> vs <i>ko</i>	
<b>napkin</b>	1. unfolded flat.	11	0	11	0
	2. crumpled	0	11	0	11
		<i>hon</i> vs <i>tsu</i>		<i>hon</i> vs <i>ko</i>	
<b>candle</b>	1. full length	11	0	11	0
	2. stub	2	9	2	9
		<i>hon</i> vs <i>tsu</i>		<i>hon</i> vs <i>ko</i>	
<b>chalk</b>	1. full length	11	0	11	0
	2. shorter	9	2	10	1
	3. stub	0	11	2	9

ly encountered in everyday life. All the percepts tested in this image test were distorted from 1- and 2-D objects into 3-D ones.

To further examine the role of dimensionality we have begun a study of distortion from 3 into 2 dimensions, as represented by examples (3) and (4). When the car and the can are undistorted, they take their usual classifier, *dai* and *ko* respectively. But when they have been squashed flat, at least some speakers prefer *tsu*. Note that they do not prefer *mai*, the classifier for flat objects, since this would clash with the semantic representation of the nouns *car* and *can*. Speakers tend not to shift to *ko* as a complement classifier in these cases because *ko* retains a core-semantic represen-

tation for 3D objects, to be discussed below.

*undistorted objects*

- 3a. kuruma-ga san-dai aru  
car-Sub 3-NC exist  
'There are three cars.'
- 4a. biiru-no kan-ga san-ko aru  
beer-Gen can-Sub 3-NC exist  
'There are three beer cans.'

*distorted objects*

- 3b. pesyanko no kuruma-ga mit-tsu aru  
squashed-flat GEN car-Sub 3-NC exist  
'There are three squashed-flat cars.'
- 4b. pesyanko no biiru-no kan-ga mit-tsu aru  
squashed-flat GEN beer-Gencan-Sub 3-NC exist  
'There are three squashed-flat beer cans.'

**Default function of *tsu* and *ko*.** The evidence provided so far suggests that *tsu* and *ko* complement the semantic categories of a range of specific classifiers. Further evidence also suggests that they have a tendency toward **default** function. First, in the free listing task summarized in table 3 five nouns were volunteered for *tsu* which are central members of specific classifier categories. Since the experimental task focuses the speaker's awareness on classifier choice, the results should be more conservative than spontaneous conversation. Indeed an ongoing observational study of spontaneous conversation (Shimojo, 1993c) shows that *ko* is also extended to nouns that are central members of specific classifier categories, as in (5).

- 5a. (sooseeji) ni-ko-zutsu hoshii? '(Do we) want two (sausages) each?'  
(sausage) 2-NC-each want [specific CL = *hon*]
- 5b. watashi ee-wa ni-ko mottekita 'I brought 2 English-Japanese (dictionaries).'  
I English-Japanese 2-NC brought [specific CL = *satsu*]

Second, table 6 shows a shift away from specific classifiers and toward *tsu* and *ko* as more focus is placed on the numeral itself, a pragmatic substitution of a general for a specific classifier when attention is shifted away from the nature of the referent.

**Table 6:** default function - *Tsu* and *ko* are more acceptable when more focus is on the numeral. Numeral focussed with an adverb. Rating scale 1 - 5 (5=fully acceptable)

Noun	Numeral+ <i>tsu</i>	<i>exactly</i>	Numeral+ <i>ko</i>	<i>exactly</i>
		+Numeral+ <i>tsu</i>		+Numeral+ <i>ko</i>
<i>enpitsu</i> (pencil)	2.43	3.30	1.56	1.95
<i>manaita</i> (cutting brd)	3.53	4.26	2.71	3.15
<i>terebe</i> (TV set)	3.39	4.00	2.67	2.90
<i>biru</i> (building)	3.84	4.70	2.35	3.40
<i>roketto</i> (rocket)	3.10	4.05	2.48	3.40
<b>Overall Average</b>	<b>3.26</b>	<b>4.06</b>	<b>2.35</b>	<b>2.96</b>

**Internal structure of *ko* and *tsu*.** One of the problematic issues surrounding general classifiers is whether they have core semantic structure. If a category accepts the "rejects" of all other categories in the system, then how can it itself have an internal structure? Our evidence suggests that *tsu* indeed seems not to have any core semantic representation,<sup>15</sup> but that *ko* does, namely for 3D hand-sized manipulable solid objects, and that the core semantics retained by *ko* limits its exploitability in complement and default function. In other words we think that *tsu* is fully developed as a general classifier, while *ko* is on its way, with its core semantics stretched out of shape to be sure, but still not freely extendible to any (inanimate) noun regardless of its semantic properties. The most direct evidence (Table 7) comes from the Roschian free listing

**Table 7** Free listing task: Nouns listed for *ko* and *tsu* by 21 subjects. No. of subjects listing each noun precedes the noun.

---Nouns Listed for <i>ko</i> ---	-----Nouns Listed for <i>tsu</i> -----	
7 <i>keshigomu</i> 'pencil eraser'	3 <i>kaban</i> 'bag'	1 <i>kutsu</i> 'shoe'
5 <i>ringo</i> 'apple'	2 <i>isu</i> 'chair'	1 <i>baketsu</i> 'pail'
4 <i>tamago</i> 'egg'	2 <i>mado</i> 'window'	1 <i>suutsukeesu</i> 'suitcase'
4 <i>ishikoro</i> 'small stone'	2 <i>kumo</i> 'cloud'	1 <i>kasettoteepu</i> 'cassette tape'
4 <i>booru</i> 'ball'	1 <i>mezamashidokee</i>	1 <i>ohajiki</i> 'marbles'
4 <i>mikan</i> 'orange'	'alarm clock'	1 <i>ningyoo</i> 'doll'
4 <i>koppu</i> 'drinking glass'	1 <i>booru</i> 'bowl'	1 <i>ronbun</i> 'thesis'
2 <i>manjyuu</i> 'bean-jam bun' <sup>16</sup>	1 <i>wagon</i> 'dinner wagon'	1 <i>sekai</i> 'world'
2 <i>kyandii</i> 'candy'	1 <i>manjyuu</i> 'bean-jam bun' <sup>16</sup>	1 <i>syokugyoo</i> 'occupation'
2 <i>jjagaimo</i> 'baking potato'	1 <i>ishikoro</i> 'small stone'	1 <i>hooahoo</i> 'method'
2 <i>tomato</i> 'tomato'	1 <i>hoshi</i> 'star'	1 <i>uso</i> 'lie'
2 <i>denkyuu</i> 'light bulb'	1 <i>honoo</i> 'flame'	1 <i>zenka</i> 'criminal record'
1 <i>saikoro</i> 'dice'	1 <i>gakubuchi</i> 'picture frame'	1 <i>yume</i> 'dream'
1 <i>isu</i> 'chair'	1 <i>hangaa</i> 'hanger'	1 <i>yotee</i> 'schedule'
1 <i>syootokeeki</i> 'pc of cake'	1 <i>osara</i> 'dish'	1 <i>koosu</i> 'course of lectures'
1 <i>omochi</i> 'rice cake'*	1 <i>sanrinsya</i> 'tricycle'	1 <i>yoyaku</i> 'appointment'
1 <i>roorupan</i> 'bread roll'	1 <i>naironbukuro</i> 'plastic bag'	1 <i>machigai</i> 'mistake'
1 <i>tamanegi</i> 'onion'	1 <i>tsukue</i> 'desk'	1 <i>himitsu</i> 'secret'
1 <i>ningyoo</i> 'doll'	1 <i>udedokei</i> 'watch'	1 <i>iken</i> 'opinion'
1 <i>hudebako</i> 'pencil case'	1 <i>hachiue</i> 'potted plant'	1 <i>hanashi</i> 'story'
1 <i>denchi</i> 'dry battery'	1 <i>hon</i> 'book'	
1 <i>hashioki</i> 'chopstick rest'	1 <i>nabe</i> 'sauce pan'	
1 <i>tsukue</i> 'desk'	1 <i>machi</i> 'town'	
type-token Ratio = .43	type-token Ratio = .89	

task which was summarized earlier in Table 3. Note that the nouns listed in Table 7 for *ko* are, with two exceptions, limited to solid hand-size manipulable 3D objects<sup>17</sup>. *Tsu*, on the other hand, is spread out across a number of dimensions: concrete to abstract, small to very large, hollow or solid, 1-,2-, or 3- dimensional; there doesn't seem to be any semantic focus at all.

Second, *ko* has a short list of 7 nouns listed four times or more, i.e. by 1/5 or more of the subjects, suggesting a cluster of prototypical exemplars in the semantic core of the category. The nouns listed for *tsu* come pretty close to having no concentration at all, suggesting the absence of privileged exemplars. This observation is quantified by the type-token ratio for *tsu* of .89, which is close to 1 and double the ratio for *ko*.

Third, in performing the task, subjects were able to list nouns for *ko* much more quickly than for *tsu*. When listing nouns for *tsu*, subjects needed more prompting to find 5 nouns, and several were unable to complete the task. These observations suggest that subjects were more likely to have privileged exemplars readily available for *ko* than for *tsu*.

Finally, while desk and chair are non-conforming exemplars for the summary description of *ko*, they are highly manipulable. In Japan students move their desks and chairs around every day so that they can clean the classroom. And there is a difference among chairs: easily moveable straight chairs are more readily classifiable with *ko* than are living-room easy chairs.

Corroborating evidence comes from a rating study of nouns, not in the domain of other specific classifiers (Table 8). Stereotypic exemplars that are hand-size at the top of the table get the highest ratings, button and thumb-tack in the middle which are typically handled with the fingertips rather than the hand, and are flat in shape, get somewhat lower ratings. Desk and chest of drawers, which are clearly out of the hand-size range, get markedly lower ratings with *ko*. Now table 9 shows what happens when *ko* loses its grip: *tsu* becomes the preferred classifier for desk and chest at the bottom of the table. But note that *ko* may still be used.

**Table 8:** Core Semantics of *ko*: Effect of object size on ratings (1-5)

Noun	Average score
mikan 'orange'	4.47
saikoro 'dice'	
ishikoro 'small stone'	
suika 'melon'	
botan 'button'	4.24
gabyoo 'thumb tack'	
tsukue 'desk'	2.77
tansu 'chest of drawers'	

**Table 9:** Core Semantics of *ko*: Forced Choice of Classifiers

Noun	no of subjects (out of 11) choosing <i>tsu</i> and <i>ko</i>
saikoro (dice)	<b>ko</b> (11) <b>tsu</b> (0)
ishikoro (small stone)	<b>ko</b> (10) <b>tsu</b> (1)
botan (button)	<b>ko</b> (9) <b>tsu</b> (2)
mikan (orange)	<b>ko</b> (8) <b>tsu</b> (3)
gabyoo (thumb tack)	<b>ko</b> (7) <b>tsu</b> (4)
suika (melon)	<b>ko</b> (6) <b>tsu</b> (5)
tsukue (desk)	<b>tsu</b> (7) <b>ko</b> (3) <b>other</b> (1)
tansu (chest of drawers)	<b>tsu</b> (5) <b>ko</b> (3) <b>other</b> (3)

Table 10, with further results from the deformation experiment discussed earlier, reinforces the same point. Subjects were shown four sizes of the same statue in

Table 10: Statue size variations

statue size	No. of subjects	
	tsu	ko
a. handsized	4	7
b. tablesized	8	3
c. bodysized	8	3
d. monumental	10	1

appropriate context: a) held in the hand; b) on a coffee table; c) freestanding on the floor; and d) on the plaza outside a large public building. As the statue increased out of the hand-size range, subjects increasingly shifted from *ko* to *tsu*, suggesting that *ko* has a semantic core which restricts its applicability as a general classifier, while *tsu* does not.

A final piece of evidence suggesting that *ko* has core semantics restricting its applicability as a complement classifier, but that *tsu* does not, comes from the so-called  $\emptyset$ -classifier construction, a syntactic context in which bare numerals are combined with nouns without an intervening classifier, as exemplified in number 6,

- 6a. manjyuu-o jyuugo-ko tabeta. '(I) ate 15 bean-jam buns.'  
 bean=jam=bun-OB<sub>1</sub> 15-CL ate
- 6b.sono heya-ni-wa isu-ga jyuunana-ko/Øaru. 'There are 17 chairs in the room.'  
 the room-LOC-TOP chair-NOM 17-CL exist
- 6c.sono hon-ni-wa hanashi-ga jyuuni-Ø aru. 'There are 12 stories in the book.'  
 the book-LOC-TOP story-NOM 12-CL exist

and in table 11. Note that since the numerals are over 10, a Sino-Japanese classifier must be chosen, hence *ko* rather than *tsu*. In (6a) the classifier is obligatory, in (6b) it is optional, and in (6c) no classifier is acceptable. These differences are largely controlled by lexical selection, as depicted in table 11. Note the semantic grouping: column A contains nouns fitting the semantic core of *ko*; column B has nouns that lack one or more core semantic properties of *ko*: solidity in the case of *bag* and *dish*; hand-size and manipulability in the case of *window*, and apparently all properties in

Table 11: The  $\emptyset$ -classifier construction. *Tsu* is excluded by the selection of a Sino-Japanese numeral. Nouns in (A) meet the summary description of *ko*; in (B) they meet only some parts, and in (C) they don't at all.

Sino-Japanese numeral 'over 10' occurs with...					
A: Only With <i>Ko</i>		B: With Either <i>Ko</i> or $\emptyset$ Cl		C: Only With $\emptyset$ Classifier	
<i>mezamashi</i>	'alarm clock'	<i>kaban</i>	'bag'	<i>honoo</i>	'flame'
<i>dokee</i>		<i>isu</i>	'chair'	<i>gakubuchi</i>	'picture frame'
<i>manjyuu</i>	'bean-jam bun'	<i>booru</i>	'bowl'	<i>machi</i>	'town'
<i>ishikoro</i>	'small stone'	<i>wagon</i>	'dinner cart'	<i>kumo</i>	'cloud'
<i>udedokee</i>	'watch'	<i>hoshi</i>	'star'	<i>ronbun</i>	'thesis'
<i>kasetoteepu</i>	'cassette tape'	<i>hangaa</i>	'hanger'	<i>sekai</i>	'world'
<i>ohajiki</i>	'marbles'	<i>tsukue</i>	'desk'	<i>syokugyoo</i>	'occupation'
<i>ningyoo</i>	'doll'	<i>hachiue</i>	'potted plant'	<i>hoofoo</i>	'method'
		<i>nabe</i>	'sauce pan'	<i>uso</i>	'lie'
		<i>kutsu</i>	'shoe'	<i>zenka</i>	'criminal record'
		<i>baketsu</i>	'pail'	<i>yume</i>	'dream'
		<i>suutsukeesu</i>	'suitcase'	<i>yotee</i>	'schedule'
		<i>koosu</i>	'course'	<i>yoyaku</i>	'appointment'
		<i>machigai</i>	'mistake'	<i>himitsu</i>	'secret'
				<i>iken</i>	'opinion'
				<i>hanashi</i>	'story'

the case of the noun "mistake."F Finally, column C has nouns that completely violate the semantic core of *ko*, with few exceptions. This distribution clearly shows the semantic restrictedness of *ko*, provided, of course, that we allow a semantic explanation for the existence of the  $\emptyset$ -classifier construction. Note that if a native numeral and hence *tsu* is used in the sentences in (3), the classifier is always obligatory, i.e. there is no  $\emptyset$ -classifier option. In fact, the nouns that we used in table 11 for illustrating the restriction on *ko* are precisely the nouns given by our subjects in the free listing task for *tsu* (Table 7)! So lexical selection in the  $\emptyset$ -classifier construction once again suggests that the use of *ko* is restricted by a summary description, while the use of *tsu* is not.

Several conclusions follow (invarying degrees) from these studies of *ko* and *tsu*:

1. There is such a thing as a semantic system in the Saussurean sense, in which a sign may be chosen by speakers not only for its compatibility with the message being communicated, but also because it is the least incompatible. *Tsu* demonstrates this in a dramatic way, as does *wirra* from the Yidiny classifier system.
2. Numeral classification, as exemplified in Japanese, sits squarely on the semantic boundary between grammar and lexicon, the native numeral classifiers (with *tsu*) on one side, and the Sino-Japanese classifiers (with *ko*) on the other.
3. So-called "general" categories can be broken down into more precise semantic and pragmatic functions in a way that is theoretically insightful. We have been able to delineate three distinct functions: *complement*, *default*, and *unspecified referent*.
4. Complement function has a competitive relationship with internal semantic structure; hence a fully complementary category such as *tsu* has no summary description for purely systemic reasons. This competitive relationship can lead to intractable data. Only by sorting out the influences of core meaning and complement function are we able to make any sense out of the behavior of the classifier *ko*.
5. Classifiers may serve complement function at different levels of generality. Mandarin *ge* has no restriction: it complements (and defaults) across all semantic domains. Our work on Japanese suggests that *tsu* is absolutely limited to non-animate concepts, and *ko* to concrete inanimates, in agreement with Matsumoto (1993). In German, all three genders serve complement function within highly specific semantic domains.

## Footnotes

\* Department of Linguistics, Baldy Hall 685, Buffalo 14260; linzubin@ubvms, v050tb58@ubvms (Shimojo)

1. See Corbett (1991) for one attempt to distinguish semantically from formally motivated systems.
2. Downing (1984), in a ground-breaking study of the Japanese classifier system, says that *ko* and *tsu* are general classifiers for inanimates. Matsumoto (1993) says that *tsu* is "a classifier for inanimate entities in general," and that *ko* is "a classifier for concrete objects in general." But neither of them explicitly differentiates distinct semantic/pragmatic functions associated with these classifiers.
3. Core meaning, roughly defined, is a semantic representation that captures the *privileged* exemplars in a category, but falls far short of the category's full extension, and thus cannot be considered a summary representation (Smith & Medin, 1981).
4. But all three genders serve complement function within specific semantic domains.
5. expressed by verb stem alternations, primarily of verbs of location and motion.
6. a closed class of generic nouns distinguished by their obligatory poss. marking.

7. a closed class of generic nouns occurring in N+N and in relative constructions, and used anaphorically.
8. Traditional analyses of German claim that *neuter* is the complement gender. *Neuter* does in fact serve *unspecified referent* function (see Table 1c).
9. E.g. *minya* classifies edible flesh, animals from which it is derived, and meat dishes
10. The data for this section come from Shimojo (1993a,b)
11. *Tsubu*, *kabu* and *mune* tend, to varying degrees, to shift to the Sino-Japanese system with numbers over 4/5.
12. The exceptions (central members of other categories) are given in Table 3. Non-central membership in a classifier category can be determined by comparing (a) the judgement of native speakers that they can use that classifier with the noun with (b) the dictionary definition (Kindaichi et al., 1981) for the classifier, which lists typical members. Downing (1984) also provides information on graded membership of some classifier categories based on a Roschian listing task.
13. Specific nouns used were (central members in boldface):  
*hon*: *ya* (arrow), *enpitsu* (pencil), *ki* (tree), *hitto* (bb hit), *denwa* (tel call) and *terebi bangumi* (TV program)  
*mai*: *kami* (paper), *manaita* (cutting board), *zabuton* (cushion)  
*dai*: *kuruma* (car), *jitensya* (bicycle), *rajio* (radio), *terebi* (TV set), *reezooko* (refrigerator)  
*ken*: *ie* (house), *apaato* (apt bldg), *biru* (building)  
*ki*: *hikooki* (airplane), *herikoputaa* (helicopter), *roketto* (rocket)
14. The *dai* and *ki* categories are exceptions, since they share some members on the basis of their specific semantic overlap (Denny, 1979).
15. beyond a restriction to inanimates, as Matsumoto (1993) points out. He suggests that *tsu* is the general classifier for inanimates, *hiki* for animals, and *ri/nin* for humans.
16. Bean-jam bun' is a traditional food which is stuffed with sweetened bean paste. It is usually round and its diameter is about 2 inches. The size and shape of typical 'rice cake' is about the same as 'bean-jam bun'.
17. Tanihara et al (1990) say: 'concrete, small, solid, independent.'

## References

- Ahrens, K. 1993. Classifier neutralization in Mandarin Chinese. Mms, Dept of Linguistics, UC San Diego.
- Austin, P. 1981. *A grammar of Diyari*. Cambridge: Cambridge UP.
- Bani, E. 1987. Gurka a ipika: masculine and feminine grammatical gender in Kala Lagau Ya. *Australian Journal of Linguistics* 7, 189-201.
- Bybee, J. 1985. *Morphology: a study in the relation between form and meaning*. Amsterdam: Benjamins
- Carlson, R. and D. Payne. 1989. Genitive classifiers. *Proceedings from the fourth meeting of the Pacific Linguistics Conference*. Eugene: U Oregon, pp. 87-119
- Corbett, GG. 1991 *Gender*. Cambridge: Cambridge UP.
- Corbett, GG. and AD. Mtenje. 1987. Gender agreement in Chichewa. *Studies in African Linguistics* 18, 1: 1-38.
- Denny, P. 1979. Semantic analysis of selected Japanese numeral classifiers for units. *Linguistics* 17, 317-335.
- Dixon, RMW. 1982. Nominal classification. Part D of *Where Have All the Adjectives Gone? and Other Essays in Semantics and Syntax*. Berlin: Mouton, pp. 157-233.

- Downing, P. 1984. *Japanese numeral classifiers: a syntactic, semantic, and functional profile*. University of California at Berkeley diss.
- Erbaugh, M. 1986. Taking stock: the development of Chinese noun classifiers historically and in young children. In C. Craig (Ed.), *Noun Classes and Categorization*. Amsterdam: Benjamins.
- Gilliéron, J. 1918. *Généalogie des mots qui désignent l'abeille*. Paris: Champion.
- Jacobson, R. 1936. Beitrag zur allgemeinen Kasuslehre. *Travaux du Cercle Linguistique de Prague* 6, 240-288. Repr. in EP Hamp, FW Householder & R Austerlitz (Eds), *Readings in Linguistics II*. Chicago: Chicago UP, 1966, pp. 51-89
- Kindaichi, K. et al. 1981. *Nihon kokugo daijiten*. Tokyo: Syoogakkan.
- Lakoff, G. 1986. Classifiers as a reflection of mind. In C. Craig, *Noun classes and categorization*. Amsterdam: Benjamins, pp. 13-51
- Matsumoto, Y. 1993. Japanese numeral classifiers: a study on semantic categories and lexical organization *Linguistics*, in press.
- Rosch, E. 1973. On the internal structure of perceptual and semantic categories. In TE. Moore (Ed.), *Cognitive development and the acquisition of language*. NY: Academic Press.
- Shimojo, M. 1993a. *Complement and default numeral classifier functions of tsu and ko in Japanese: a psycholinguistic study*. MA essay, Department of Linguistics, SUNY Buffalo.
- Shimojo, M. 1993b. Complement and default numeral classifier functions of *tsu* and *ko* in Japanese. In *ESCOL 9*. Ithaca: Cornell U Linguistics Dept.
- Shimojo, M. 1993c. Historical and conversational aspects of the general classifiers *ko* and *tsu* in Japanese. In preparation.
- Smith, EE. and DL. Medin. 1981. *Categories and concepts*. Cambridge: Harvard UP.
- Tanihara, K., JC. Yen, and D. Lee. 1990. A study on Japanese classifiers: *men, mai, hon, ko, and tsu*. *Mathematical Linguistics* 17, pp. 209-226.
- Young, RW. and W. Morgan. 1980 *The Navajo language : a grammar and colloquial dictionary*. Albuquerque : U New Mexico P.
- Zubin, D. and K-M. Köpcke. 1986. Gender and folk taxonomy. In C. Craig (Ed.). *Noun Classes and Categorization*. Amsterdam: Benjamins, pp. 139-180.