

# Open Multilingual WordNet Extensions

Luís Morgado da Costa, Francis Bond, et al.

## Overview:

- *Chengyu* (4 character idioms)
- Pronouns, Determiners & Quantifiers
- Classifiers



# Chengyu (成语)

## *Chengyu* (?)

- prototypically four-character phrases;
- non or only partially compositional;
- to be treated as constituents;
- all / multiple wn parts-of-speech;
- derived meaning (historical lore or classical lit.);
- many good (Chinese) paper resources;
- no free computer tractable resources;

**Candidates: +59 000**  
**Matched (NTUMC): +1200**

**In Wordnet:**  
**156 (Types), 203 (Tokens)**

# Chengyu in NTUMC



大概一个星期以前，就是上星期二，我发现在一个窗台上画了一些跳舞的滑稽小人，跟那张纸上的一模一样，是粉笔画的。

Dàgài yīgè xīngqī yǐqián, jiùshì shàng xīngqīèr, wǒ fāxiàn zài yīgè chuāngtái shàng huàle yīxiē tiàowǔ de huájī xiǎo rén, gēn nà zhāng zhǐ shàng de **yīmóyíàng**, shì fěnbǐ huà de.

About a week ago — it was the Tuesday of last week — I found on one of the window-sills a number of funny little dancing figures drawn in chalk, **identical** to the ones on the paper.

(lit. “one mold one shape”)

*In The Adventure of The Dancing Men*

# More Examples



畫蛇添足 (lit. “paint snake add feet”) overdo /superfluous

紙上談兵 (lit. “paper on talk about soldiers”) to engage in idle theories

九牛一毛 (lit. “nine cows one hair strand) a drop in a bucket

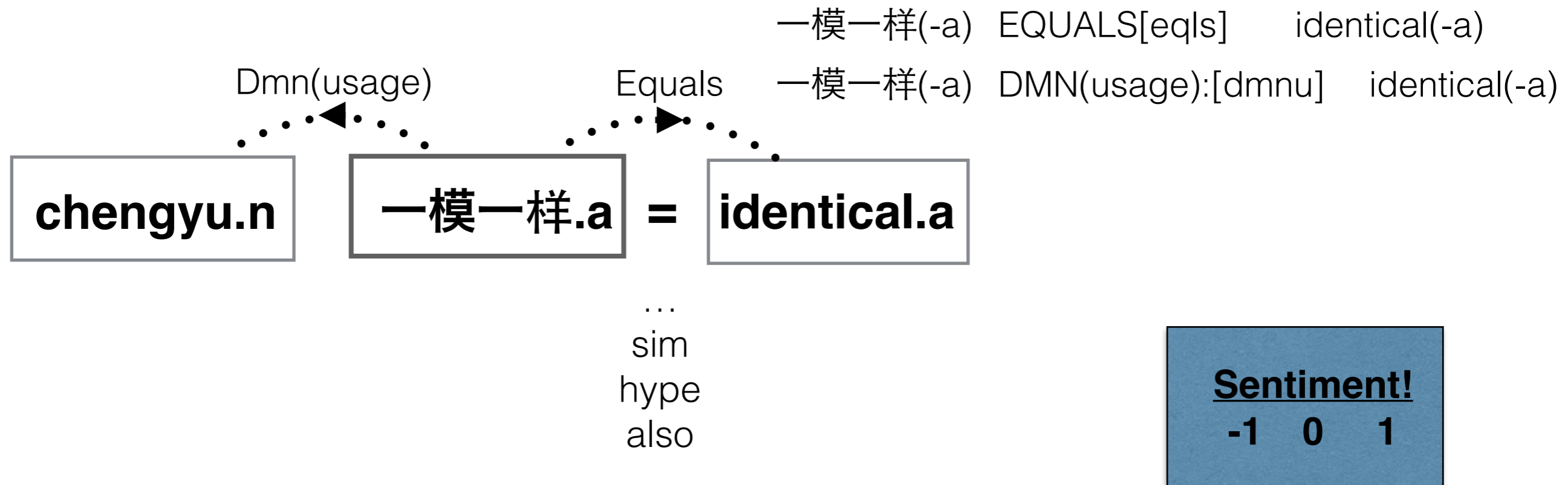
笨鳥先飛 (lit. “dumb bird early fly”) those with less ability have to try harder

如魚得水 (lit. “as fish get water”) like a fish in water

# Chengyu in the WN

## New relation: Equals[eqls]

*yī mó yī yàng*  
**e.g.** 一模一样 exactly the same (lit: one mold one shape)  
= identical (02068946-a) 'being the exact same one'



# Pronouns, Determiners & Quantifiers



Features: (11 categories, 41 tokens)

## Usage

Personal  
Possessive  
Demonstrative  
Interrogative  
General

## Person

1st  
1st (inclsiv.)  
1st (exclsv.)  
2nd  
3rd

## Proximity

Proximal  
Distal  
Medial  
Remote

## Gender Speech

Women's Speech  
Men's Speech

## Formality

Formal  
Informal

## Head (Pron.)

Entity  
Person  
Thing  
Location  
Kind  
Time  
Cause  
Manner  
Duration

## Gender

Neuter  
Feminine  
Masculine

## Type

Assertive Elective  
Negative  
Other  
Reciprocal  
Universal  
Interrogative

## Politeness

Polite

## Other

Reflexive

## Pronouns:

**Types: 140**

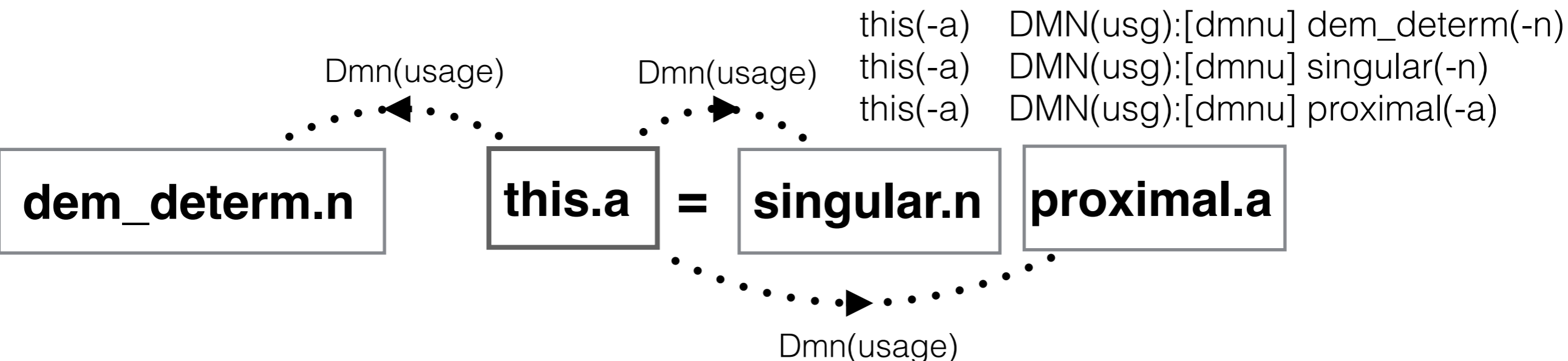
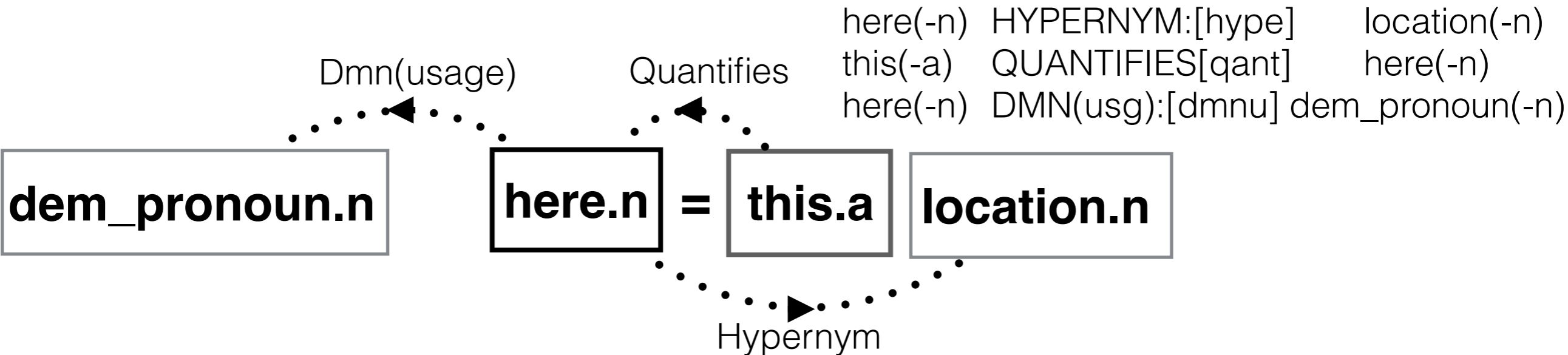
**Tokens: 333** English (96) Japanese (129)  
Chinese (64) Indonesian (40)

# Pronouns, Determiners & Quantifiers



**New relation:** Quantifies[qant] / Quantifier[hasq]

**e.g.** here = 'this + location'

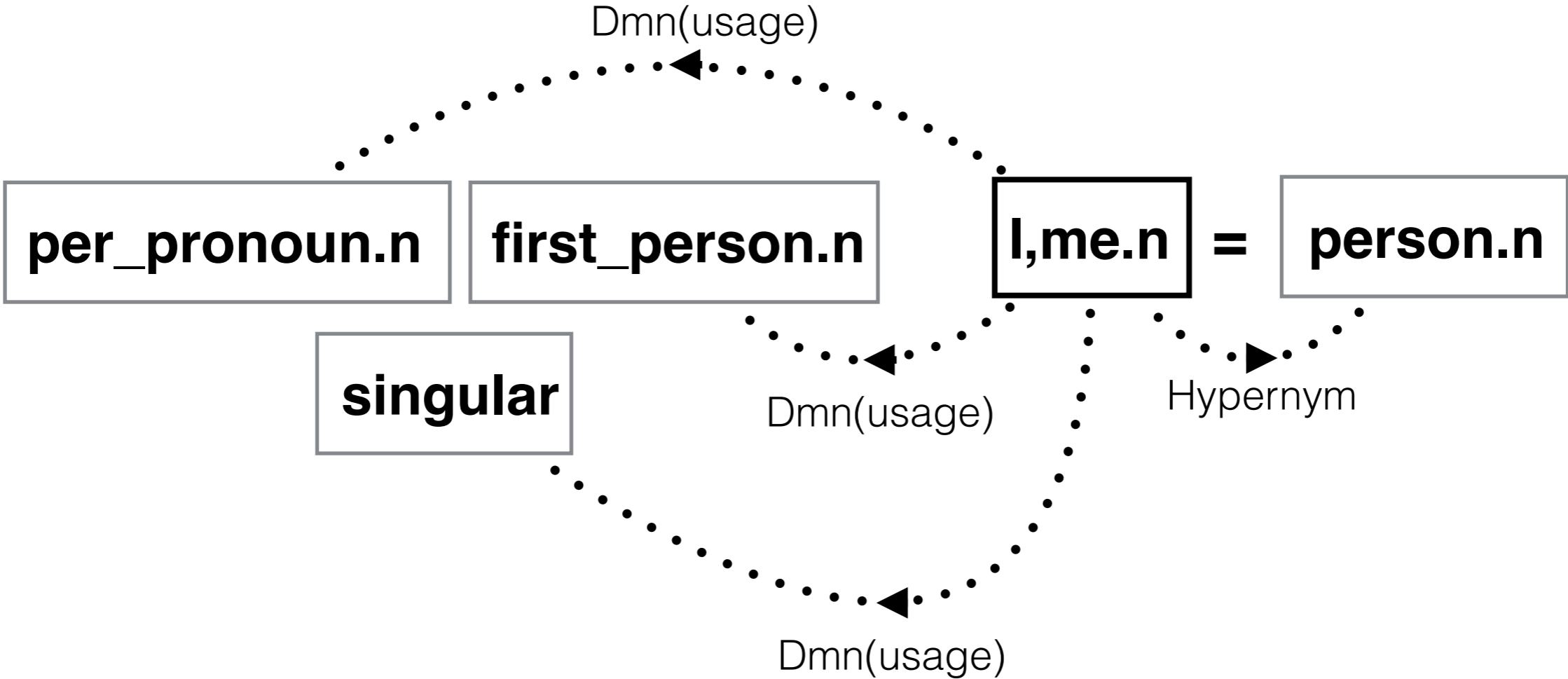


# Pronouns, Determiners & Quantifiers



e.g. 'I,me'

I,me(-n) HYPERNYM:[hype] person(-n)  
I,me(-n) DMN(usage):[dmnu] per\_pronoun(-n)  
I,me(-n) DMN(usage):[dmnu] first\_person(-n)  
I,me(-n) DMN(usage):[dmnu] singular(-n)





# Classifiers

- 3 out of 4 languages in NTUMC make frequent use of Classifiers;
- Most frequent: Chinese ( $\approx 200$ ); Japanese ( $\approx 80$ ); Indonesian (?)
- Important for generation/machine translation;
- No free computer tractable resources;

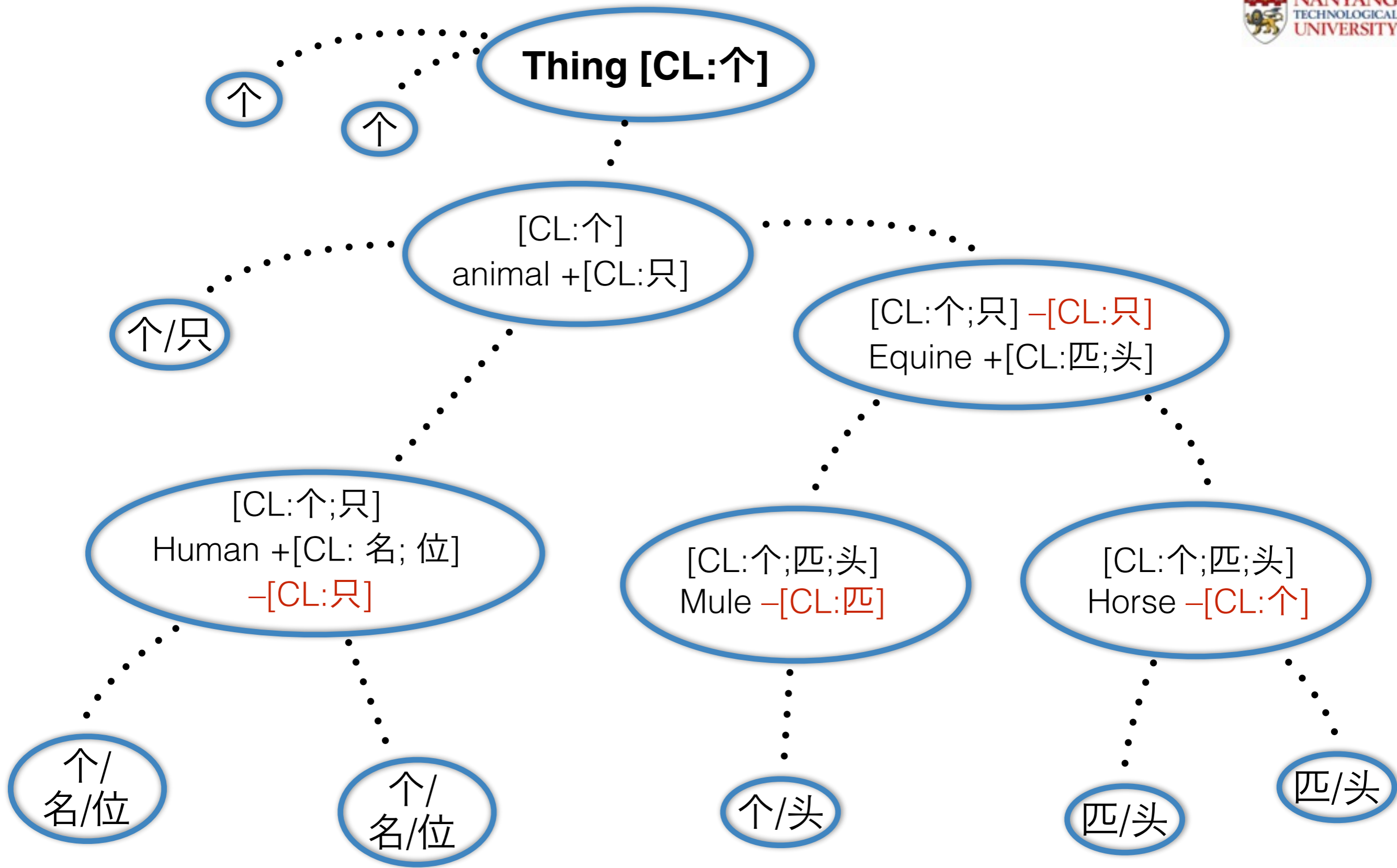
Our experimental analysis makes use of propagation rules to project classifiers over the WN hierarchy:

## **Propagation Rules:**

Japanese ( $> 120$ )

Chinese ( $> 250$ )

# Projecting Classifiers over WN



# Projecting Classifiers over WN

02374451-n • 'solid-hoofed herbivorous quadruped domesticated since prehistoric times';

Albanian	<i>kalë</i>
Arabic	فَرَس , جَوَاد , جِصَان
Bulgarian	<i>KOH</i>
Catalan	<i>cavall</i>
Chinese (simplified)	马 [CL: 匹; 只]
Danish	<i>hest, prubest, byphest</i>
Greek	<i>ἵππος</i>
English	<i>horse</i> 103 , <i>Equus caballus</i>
Basque	<i>zaldi</i>
Finnish	<i>Equus caballus</i> , <i>hevonen</i>
French	<i>cheval</i> , <i>equus caballus</i>
Galician	<i>Equus caballus</i> , <i>cabalo</i>
Indonesian	<i>kuda</i>
Italian	<i>Equus caballus</i> , <i>cavallo</i>
Japanese	馬 , 雄馬 , 牡馬 , 馬匹 , ウマ [CL: 匹; 個; 頭]
Nynorsk	<i>hest</i>

We're far from perfect!

**We don't really have conclusions yet...  
We want suggestions and discussion!  
Thank You**

**Luís Morgado da Costa**  
**[luis.passos.morgado@gmail.com](mailto:luis.passos.morgado@gmail.com)**

