DEFINING CHINESE IDEOPHONES: A FAMILY OF CONSTRUCTIONS

Nanzan University, 2019/8/3
International Workshop on Mimetics
(Ideophones, Expressives) III: Crucibles of Mimetics

Thomas Van Hoey 司馬智
National Taiwan University

Roadmap

Introduction

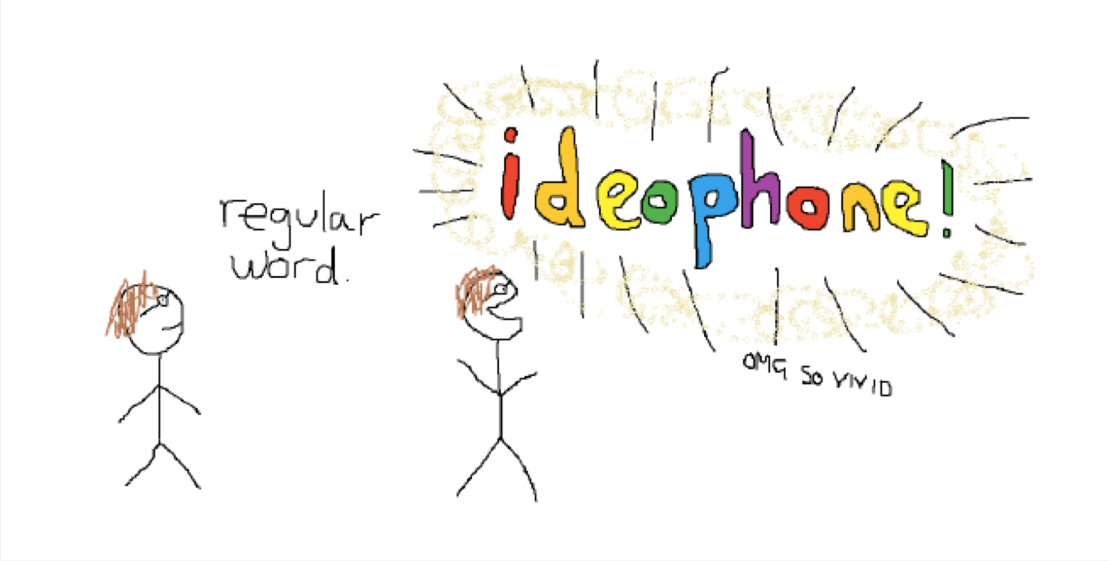
Goal and previous research
- Japanese example
- Crosslinguistic approach (Canonical Typology)
- First exclusion of non-ideophones
- Research questions

Data: CHIDEOD

Methodology
- Phonological form, written form, sensory modality
- Multiple Correspondence Analysis

Analysis and results

Discussion and conclusions



(Lockwood 2017)

Chinese ideophones

霧　茫茫

wù　máng~máng

mist　hazy.IDEO

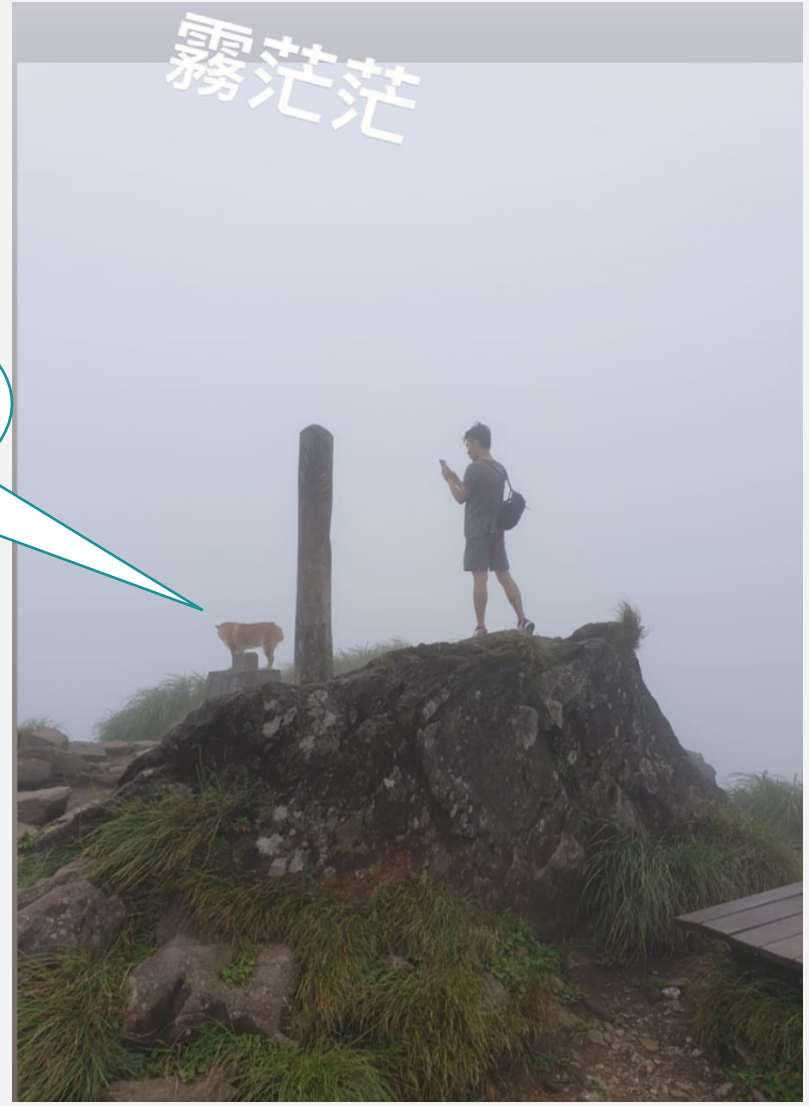
“It’s foggy.”

狗　吠　汪汪

gǒu fèi　wāng~wāng

dog bark　woofwoof.IDEO

“The dog is barking woofwoof.”



Instagram: @nickprometheus31

Chinese ideophones

這　篇　文章　　讀起來　凌亂。

zhè piān wénzhāng dú-qǐ.lái líng~luàn

DEM CL paper read-MID in.a.mess.IDEO

“This paper reads sloppily.”

摸起來　滑滑的，　冷冷的。

mō-qǐ.lái huá~huá=de léng~léng=de

touch-MID slippery.IDEO=LNK cold.IDEO=LNK

“It feels slippery and very cold.”

Ree Lin’s pet snake Seysey

Introduction

From these four examples, it can be readily seen that Chinese ideophones cover many sensory modalities, and that their constructions are not equal.

Moreover, the question needs to be asked if they are all equally ideophonic.

That is the main topic of this presentation.

PREVIOUS RESEARCH AND GOAL

Research on Chinese ideophones

Chinese research on Chinese ideophones is mostly concentrated on onomatopoeia (ideophones that depict SOUND).

Zhao Aiwu 赵爱武 (2008)
Qiu Di 邱迪 (2018)

However, there are some broader discussions of ideophones

- phonology — Mok (2001); Thompson (2018)
- Beijing dialect — Meng (2012)
- Southern Sinitic — Wu (2014)
- Cantonese (vs. Dagaare) — Bodomo (2006)
- Mandarin (vs. Japanese) — You (2015)
- Japanese (vs. Mandarin) — Lu (2006)
- reduplication in Old Chinese — Sun (1999)
- Middle Chinese — Van Hoey (2015)

Research on Chinese ideophones

Most of these studies define their terms ‘neatly’:

- sound similarity for onomatopoeia
- “vivid expressions related to manner, color, sound, smell, action, state, intensity, height, length, weight, size, shape and other physical attributes” for ideophones (Mok 2001:3-7)

But they barely touch upon the gradedness / representativeness of certain ideophones for the ideophone category.

In other words, the notion of **prototypicality** is not discussed.

The goal: Japanese prototypes

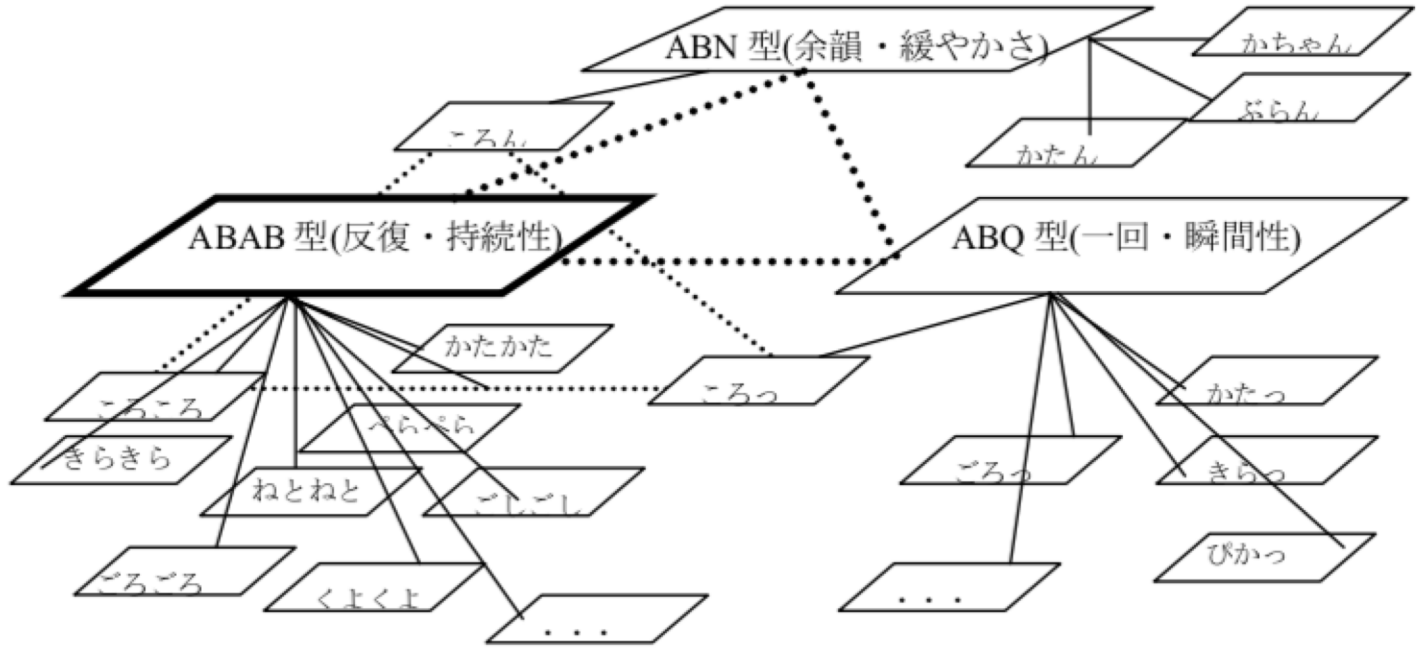


図 4.5　擬音語・擬態語のパターン・ネットワーク（立体図）

直線：事例化リンク　　点線：パターン同士のリンク

表4.10　基本語彙における擬音語・擬態語の分

	ABAB型	ABQ型	ABN型
語彙数	156	4	10
%	68%	2%	4%

分布（総語彙数229）

ABri型	AQBri型	ANBri型	その他
13	20	3	23
6%	9%	0.4%	10%

Lu (2006:97 ff.)

The goal: Japanese prototypes

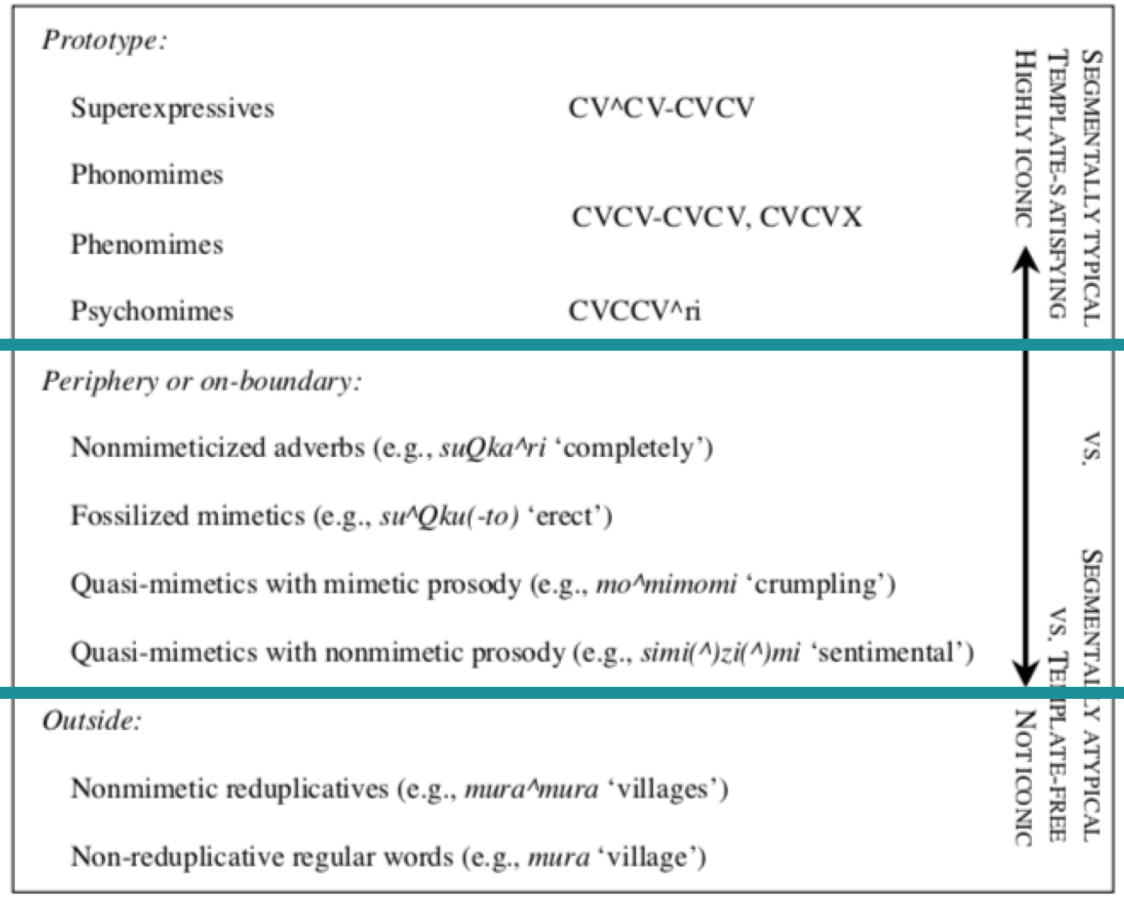


Figure 4.2. The internal structure of the prototype category of Japanese mimetics

Table 4.2. The coverage of mimetic morphophonological template

a. Mimetics satisfying a template	1,643 (99.46%)
Monomoraic root-based	214 (12.95%)
CVQ(^)	50 (3.03%)
CV(^)N(^)	29 (1.76%)
CViQ^	14 (.85%)
CV(^)V(^)	21 (1.27%)
CV(^)V-CVV	46 (2.78%)
CV(^)N-CVN	45 (2.72%)
CV^i-CVi	9 (.54%)
Bimoraic root-based	1062 (64.29%)
CVCVQ^	213 (12.89%)
CVCV(^)N(^)	101 (6.11%)
CVCV^ri	130 (7.87%)
CVCCV^ri	134 (8.11%)
CV(^)CV-CVCV, etc.	484 (29.30%)
Derivatives	332 (20.10%)
Fossilized templates	35 (2.12%)
b. Mimetics satisfying no template	9 (.54%)

Akita (2009)

Japanese quasi-mimetics

(10) Ideophones:

a. Demo gozyuu-en-kurai [Quot-adv **bababababat-te**] oti-tyau-node
but 50-yen-about IDPH-QUOT fall-end.up-because
[V **bikkuri-si**]-ta-koto-ga aru-n-desu-kedo-ne. (032, affirmative-declarative)
IDPH-do-PST-NML-NOM be-NML-COP.POL-but-SFP
'But as [the payphone] consumed about ¥50 *very quickly*, [I] *was astonished*.'

b. Un, wasyoku-da-kedo [N **kotekote-zya**]-naku-tte. (080, negative)
yeah Japanese.food-COP-but IDPH-COP.TOP-NEG-and
'Yeah, [I want to have] a Japanese cuisine but not an overly Japanese one.'

(11) Quasi-ideophones:

a. Yoosuruni, "[Bare-adv **wazawaza**] uke-ru-no?" -tte-i-u
in.short IDPH take-NPST-QUES -QUOT-say-NPST
sekai-dat-ta-n-da-kedo... (047) (interrogative)
world-COP-PST-NML-COP-but
'In short, it was like "Do you *bother to* take [the entrance exam]?"'

b. Toriaezu, [Quot-adv **zut-to**] massugu it-tyat-te-kudasai. (004 imperative)
for.now IDPH-QUOT straight go-end.up-CONJ-POL.IMP
'For now, please go straight *on*.'

Quasi-ideophones don't have the same preferences for sentence types / sentence restrictions as 'real ideophones' (Akita 2017)
→adverbial usage

But before, Akita (2009) also talked about **quasi-mimetics** like

nu(^)kenu(^)ke ヌケヌケ 'impudent' (< 抜く nuku 'come off')

which are deverbal (in this case)

Ideophones are prototypes, with fuzzy boundaries

Childs (1994:181) already mentioned that ”It is thus best to think of ideophones as a prototype category with a core of good members.”

The previous two examples of Lu and Akita’s prototypical constructions show that there are a number of words, or groups of words, that will have to be excluded from ‘the ideophone’-proper.

This is at the language-particular level, but on the cross-linguistic level there there have also been some advances.

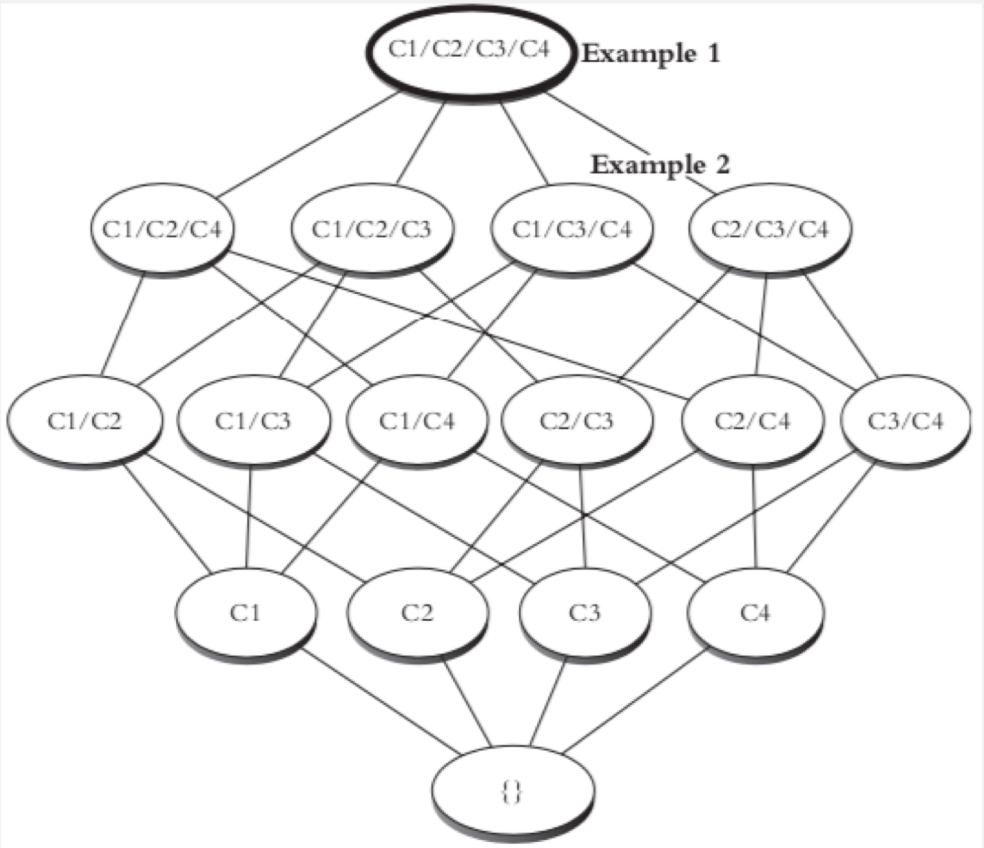
Enter Canonical Typology.

Canonical Typology

"**Data comparability** is a fundamental problem within typology and it has far-reaching implications for our ability to define natural languages.
[…]
Canonical Typology seeks to avoid the tendency to use linguistic terms with vague and shifting definitions by placing emphasis on the criteria used to associate particular linguistic phenomena with cross-linguistic categories.

It therefore demands greater detail and rigour in terms of description, because it requires the typologist to be clear about the basis on which a phenomenon might be considered an instance of a particular **concept**." (Brown & Chumakina 2013)

For a given concept criteria are proposed, until the canonical is theorized (but it may not occur in real language!)



Canonical Typology <-> Prototype

Canonical typology is thus a philosophical exercise based on data (see Brown, Chumakina & Corbett 2013 a.o.).

“The canonical ideal at which the “indisputable”, “best” or “clearest” instances of the phenomenon are found is not to be confused with the prototype, which may be the most visible and frequent instance of the given domain (Corbett 2005).

An analogy for canonicity is the system of cardinal vowels, which display the maximum possible in their degrees of frontness and closeness, within the space where specific vowels can be populated (Baerman & Corbett 2012).

On the other hand, **an analogy for prototype is Venus, which is the most visible planet, but which does not hold special logical status among other planets** (Corbett 2010).

In reality, the canonical ideal, which is a maximal logical standard, may often be rare or even non-existent. Nevertheless, its role is essential in canonical methodology because it unambiguously sets a logically maximal endpoint from which examples of the phenomenon can be calibrated within the base. For a more detailed summary of Canonical Typology, see Brown et al. (2013).” (Kwon 2017)

Canonical Typology: phonesthemes

Kwon & Round (2015) have treated **phonesthemes** (like *gl-* in *glitter, glimmer, glisten*) using this framework.

Table 1 Criteria for the canonicity of phonaesthemes

Recurrence	1	The phonaestheme occurs in many lexical stems > in few
	2	It occurs in many parts of speech > in few
Form and meaning	3	It is strongly image-iconic > weakly > not image-iconic
	4	Its form is paired with only one meaning > with many
	5	Its meaning is paired with only one form > with many
Distribution	6	It combines only with non-recurring residues > also with recurring residues > is able to occur alone
Transparency	7	It combines agglutinatively with residue > non-agglutinatively

My problem with this is that they identify these criteria but then don't take the last step of populating the different groups they found (e.g. Criterion 1 + 2).

I want to know how many of this and that, so that we can get at a converged picture of the phenomenon that rests upon canonical typology *and prototypicality*.

Ideophones and Canonical Typology

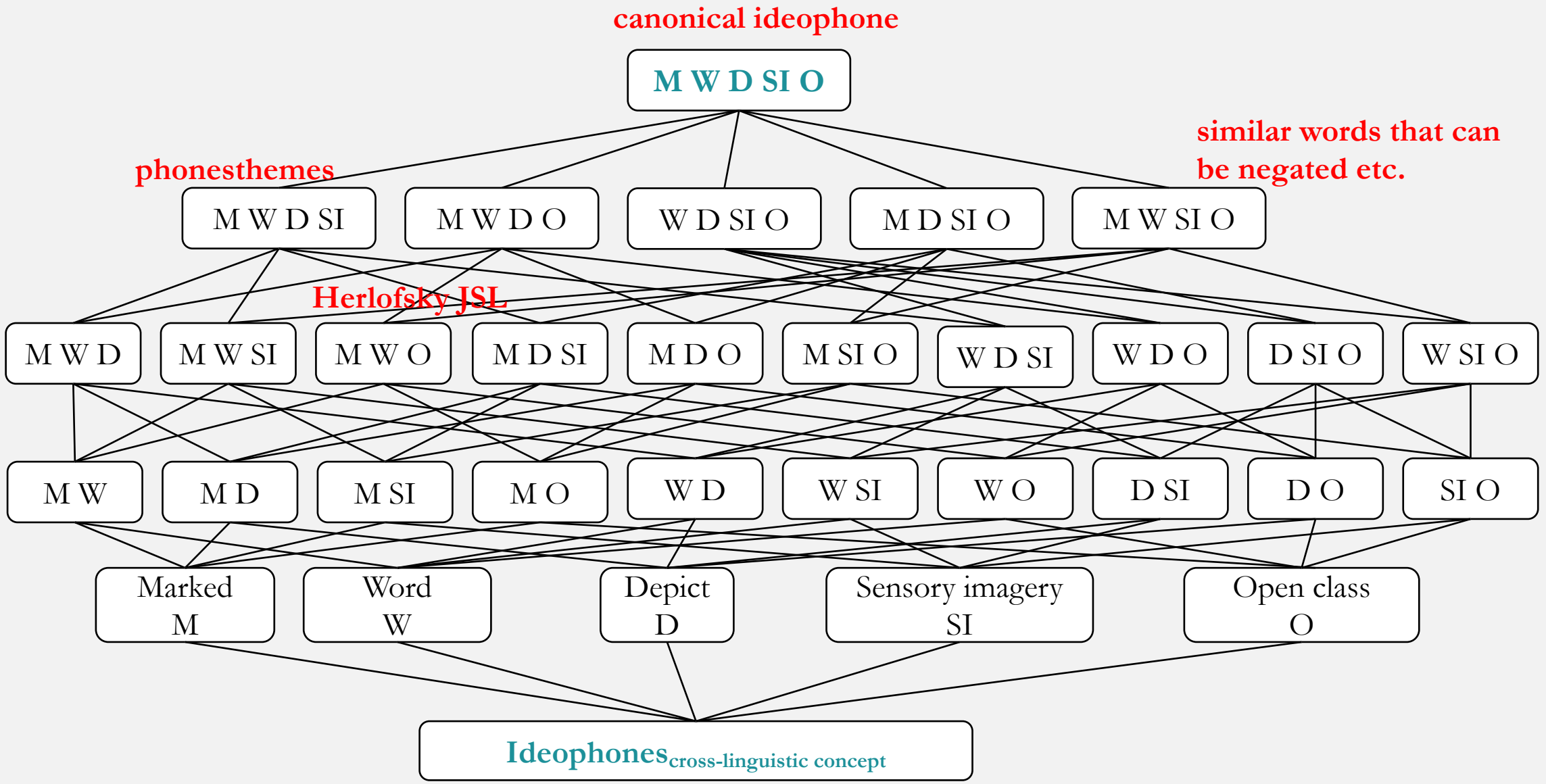
Why am I talking about canonical typology?

Akita & Pardeshi's (2019) proceedings from the Mimetics conference held at NINJAL in 2016 has a chapter by Dingemanse's, which discusses a revision of his definition for ideophones, as a comparative concept.

Canonical ideophones display all of these criterions:

i. ideophones are **marked**, i.e. they have structural properties that make them stand out from other words
ii. they are **words**, i.e., conventionalized lexical items that can be listed and defined
iii. they **depict**, i.e., they represent scenes by means of structural resemblances between aspects of form and meaning
iv. their meanings lie in the broad domain of **sensory imagery**, which covers perceptions of the external world as well as inner sensations and feelings
v. ideophones form an **open lexical class**, i.e., a set of lexical items open to new additions

My own research into Chinese ideophones hopes thus to benefit from insights of this canonical definition as well as prototype theoretical insights.


canonical ideophone
M W D SI O
similar words that can be negated etc.
phonesthemes
M W D SI
M W D O
W D SI O
M D SI O
M W SI O
Herlofsky JSL
M W D
M W SI
M W O
M D SI
M D O
M SI O
W D SI
W D O
D SI O
W SI O
M W
M D
M SI
M O
W D
W SI
W O
D SI
D O
SI O
Marked M
Word W
Depict D
Sensory imagery SI
Open class O
Ideophones cross-linguistic concept

Words that are *not* Chinese ideophones

- **Normal repetitions** (cf. Dingemanse 2015)
  duìduìduì 對對對 ‘yeah…’ *shìshìshì*是是是 ‘yes…’ *láilái*+ 來來+ ‘come…’

- **Non-qualificative words** (so referential words) (cf. Newman 1968)
  rìrì 日日 ‘every day’、*rénrén* 人人 ‘everybody’
  but also all these plants and animals e.g. *tángláng* 螳螂 ‘mantis’、*zhāngláng* 蟑螂 ‘cockroach’ which may have an ideophonic basis (that has been lost)

- **Unmarked prosaic words that mean “semantic imagery”** (Winter 2019)
  hóng 紅 ‘red’
  red, but also words like *sticky, cheesy, gooey*

So all of the canonical criteria are important, but may be manifested differently.

Research questions

What is the internal structure of the ideophone category in Chinese?

- What types of constructions occur most frequently?
- Correlations between the phonology, orthography and semantic category?

$$\left[\frac{phonological\ form}{written\ form} \mid meaning\right]$$

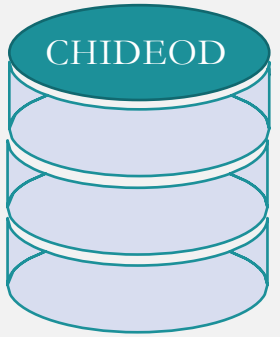
(漢字的形音義, the form, sound and meaning of Chinese characters cf. Hsieh 2006)

In order to address these problems, we make use of our new database, CHIDEOD.

DATA CHIDEOD

The Chinese Ideophone Database

Data: CHIDEOD



To answer the data reusability problem and provide better analyses we have been developing the Chinese Ideophone Database, or CHIDEOD.

The database repository can be accessed through https://osf.io/kpwgf/ and is open source. We have made an app (written in R shiny) that lets you easily subset the the many formal and semantic variables we have provided for each entry.

This was first presented this year at ILL 12, held at Lund University, Sweden (Van Hoey & Thompson 2019).

CHIDEOD: why and where

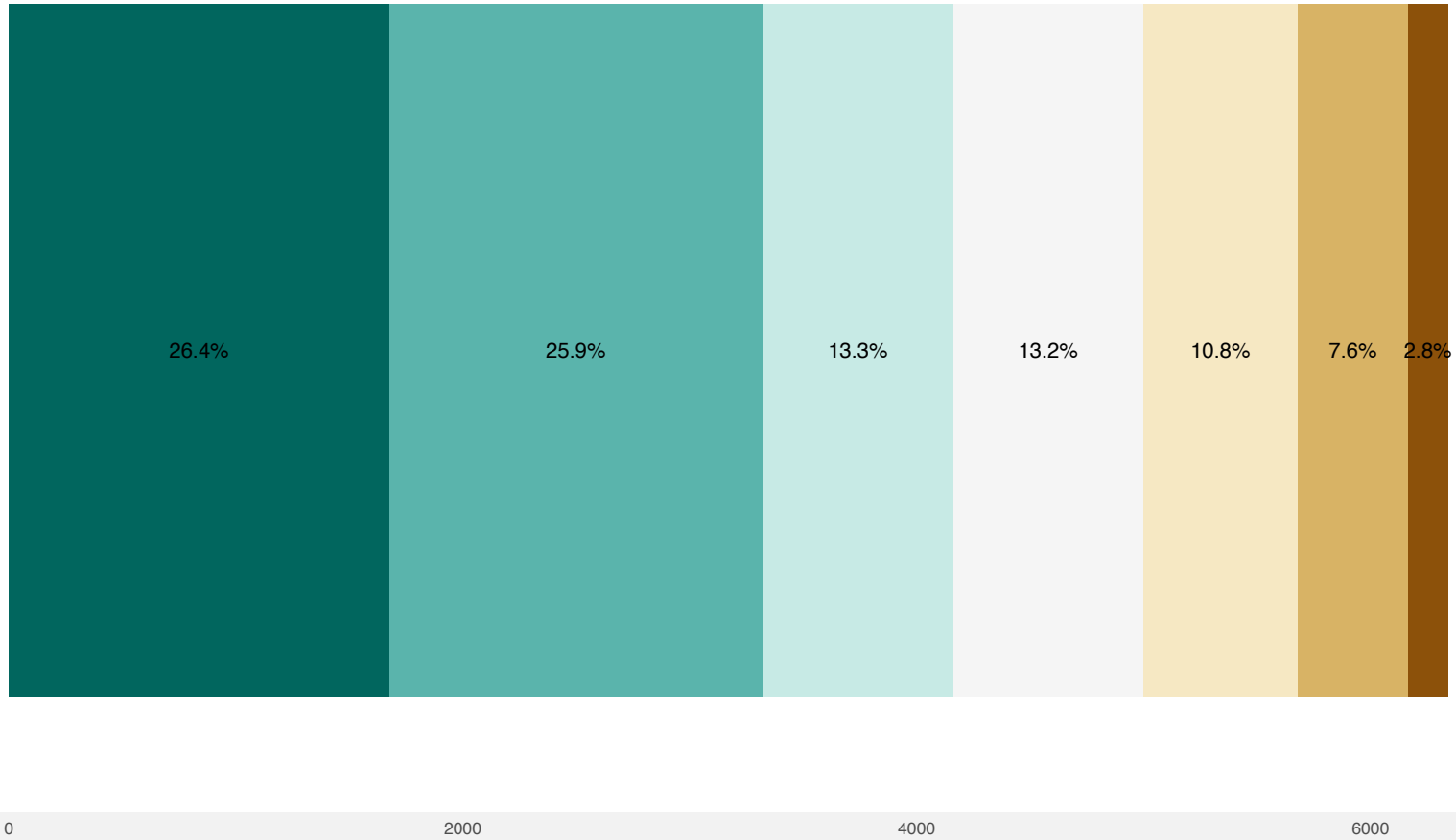
- digitization of data
- centralization of data
- exploration
  - semantics
  - phonology
  - ortography
  - historical
- expandable research resource rather than the finalized tool
- type frequencies
  (not yet token frequencies)

Open source project available at OSF
https://osf.io/kpwgf/

Available as online app
https://simazhi.shinyapps.io/Chineseideophone/

Also available as an R package
(see osf website)

CHIDEOD: basic stats



	Dictionaries	Dedicated studies
Mandarin	Gong (1991) Wang (1987)	Li (2007)
	Hanyu Da Cidian (2006)	
Premodern Chinese	Kroll (2015)	Van Hoey (2015) Van Hoey (2016)

variables in CHIDEOD

Formal variables

- phonology
  - pinyintone, pinyinnum, pinyinnone
  - Middle Chinese (MC)
  - Old Chinese (OC)
- orthography
  - traditional
  - simplified
  - T1-T4
  - S1-S4
  - S1-S4.charfreq
  - S1-S4.famfreq
  - morphology
  - S1-S4.sem
  - S1-S4.semfreq
  - S1-S4.semfam
  - radical support
  - S1-S4.phon
  - S1-S4.phonfreq
  - S1-S4.phonfam

Semantic variables

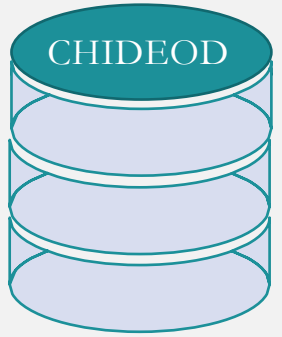
- word level
  - Kroll dictionary
  - Handian (zdic)
  - Hanyu Da Cidian
  - sensory modality
- character level
- below-character level

Other variables

- variants
- note
- datasource

Abbreviations:
S = simplified,
T = traditional,
sem = semantic radical,
phon = phonetic radical,
freq = (token) frequency,
fam = (type) family frequency

Data



As data we will use a slightly updated version of **CHIDEOD** since ILL 12.

This means that we look at **type frequencies**, rather than token frequencies.

The analysis will thus focus on the **linguistic system**, rather than linguistic usage — which is an important part of the meaning and use of ideophones.

It also means that diachronic data and synchronic data is a bit lumped together (disclaimers).

An **analogy** would be to describe the PAST TENSE of English.

Construction wise: you would find the **high type frequency construction VERB+ED**, as well as a number of **strong verb constructions**.

To really get the full picture, you would then want to look at the **token frequencies**, after this stage.

METHODOLOGY

Three variables

In CHIDEOD, interactions between

- morphology
- whether or not there is a form of reduplication present in the graphical form
- what sensory domain

Based on identifying a

- **BASE**
- **REDUPLICANT**
- **extra elements**

(cf. Sun 1999)

BASE

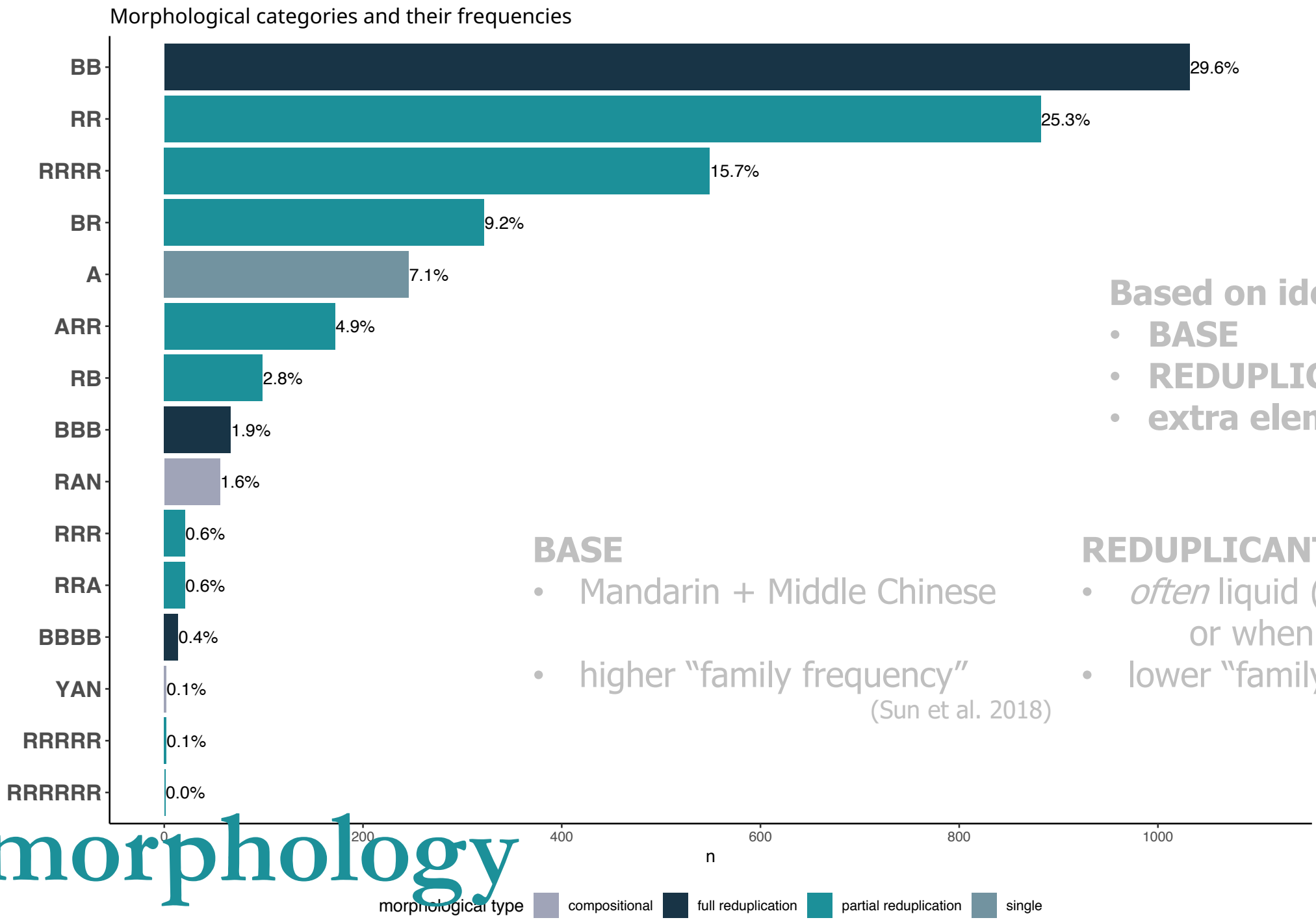
- Mandarin + Middle Chinese
- higher "family frequency"

(Sun et al. 2018)

REDUPLICANT

- *often* liquid (or MC reflex) or when base unclear
- lower "family frequency"

morphology



Based on identifying a

- **BASE**
- **REDUPLICANT**
- **extra elements**

(cf. Sun 1999)

BASE

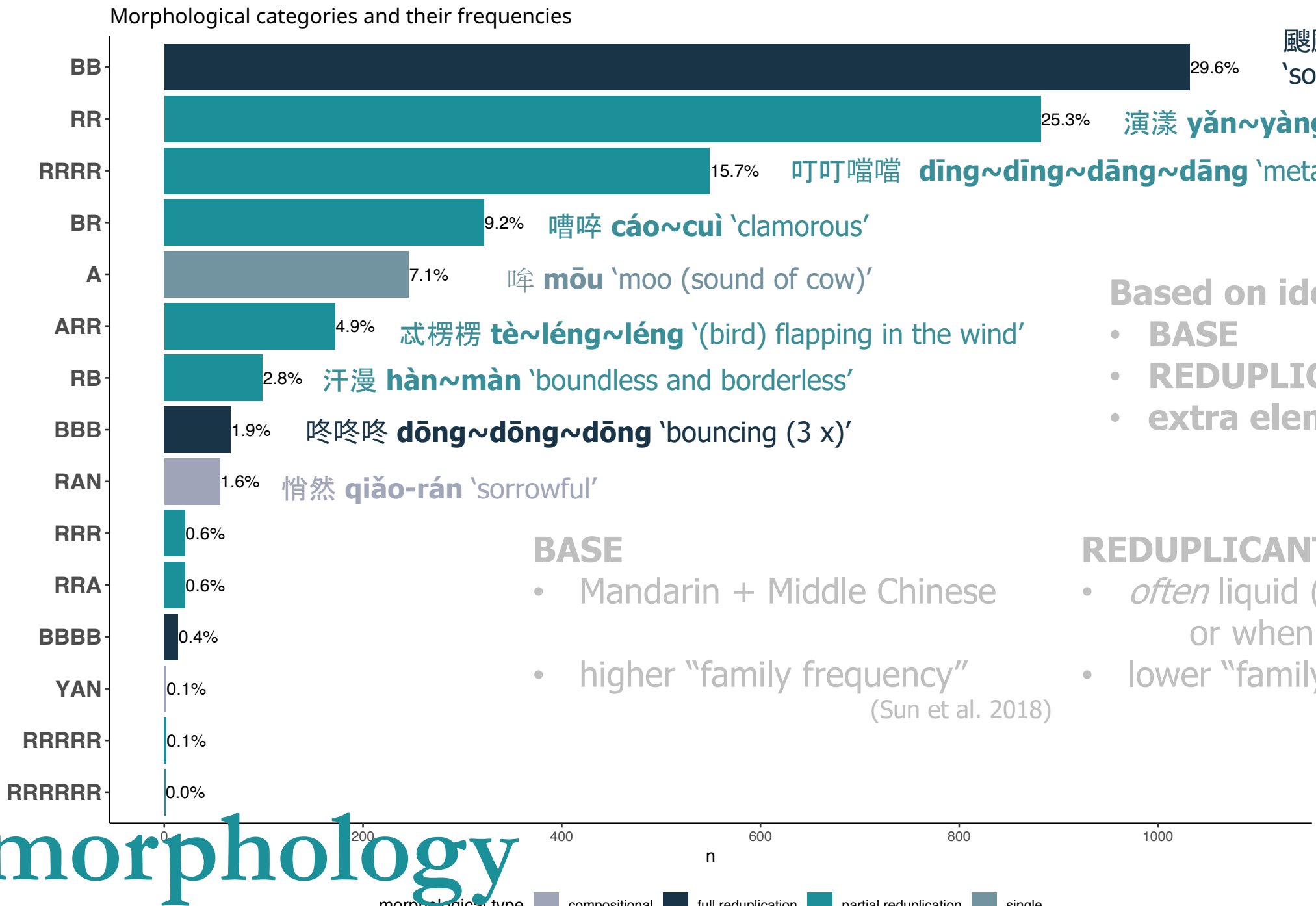
- Mandarin + Middle Chinese
- higher "family frequency"

(Sun et al. 2018)

REDUPLICANT

- *often* liquid (or MC reflex) or when base unclear
- lower "family frequency"

morphology



颼颼 **sōu~ sōu** 'sound of wind'

演漾 **yǎn~yàng** 'waves crashing'

叮叮噹噹 **dīng~dīng~dāng~dāng** 'metal clincking'

嘈啐 **cáo~cuì** 'clamorous'

哞 **mōu** 'moo (sound of cow)'

忒楞楞 **tè~léng~léng** '(bird) flapping in the wind'

汗漫 **hàn~màn** 'boundless and borderless'

咚咚咚 **dōng~dōng~dōng** 'bouncing (3 x)'

悄然 **qiǎo-rán** 'sorrowful'

Based on identifying a

- **BASE**
- **REDUPLICANT**
- **extra elements**

(cf. Sun 1999)

BASE

- Mandarin + Middle Chinese
- higher "family frequency"

(Sun et al. 2018)

REDUPLICANT

- *often* liquid (or MC reflex) or when base unclear
- lower "family frequency"

morphology

Morphophonology ~ **radical support**

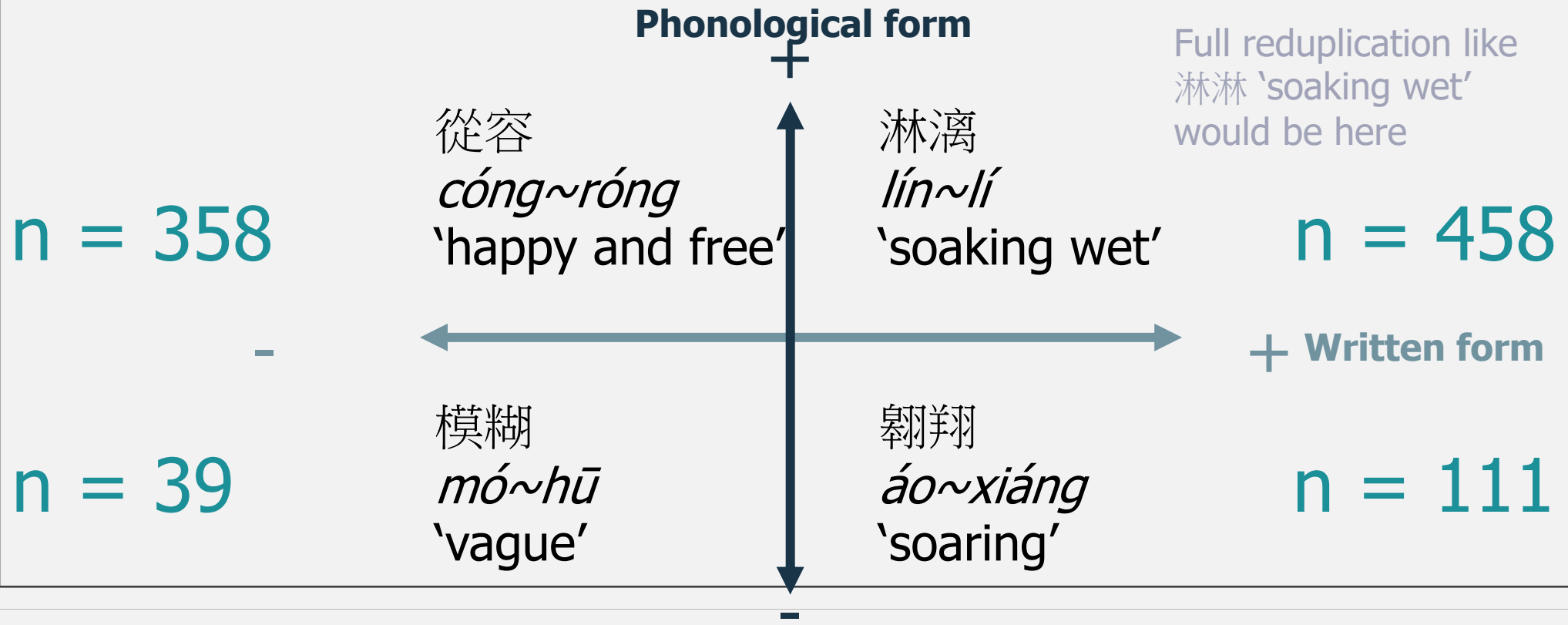
- **Phonological form** — Morphophonological motivation / markedness
- **Written form** — Orthographic motivation / markedness

Subset: datasource == Kroll (because best analyzed)

Phonology ~ radical support

- **Phonological form** — Morphophonological motivation / markedness
- **Written form** — Orthographic motivation / markedness

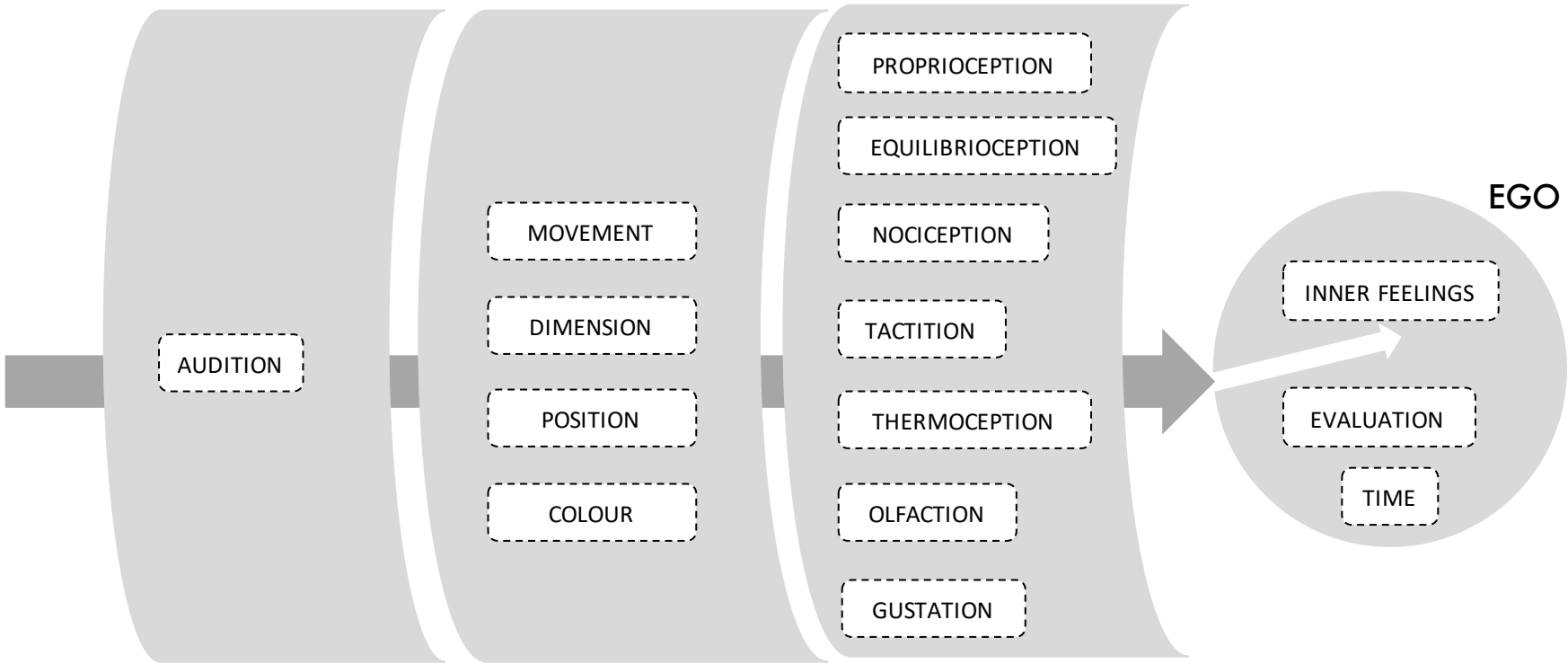
Subset: datasource == Kroll (because best analyzed)



Full reduplication like 淋淋 'soaking wet' would be here

$\chi 2 = 15.99$
$p = 6.369e\text{-}05$

Sensory domains



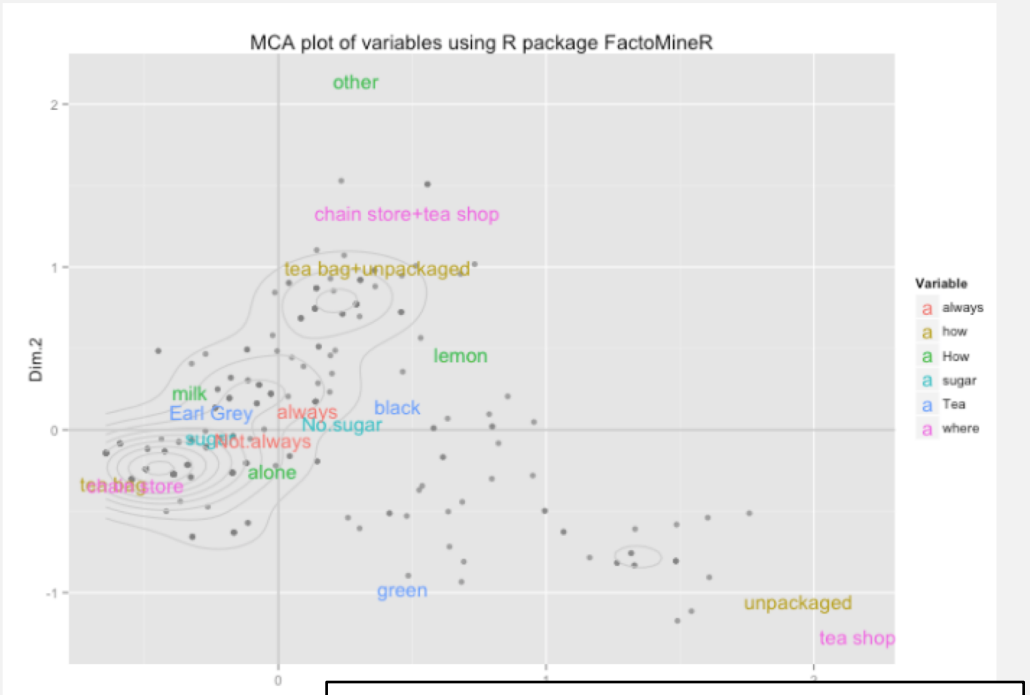
Van Hoey (2016, in prep.);
McLean (2019)

Multiple Correspondence Analysis

Multiple Correspondence Analysis (Glynn 2014; but also Levshina 2015; Greenacre 2007; Lê et al. 2008) or **MCA** is a statistical technique to identify the **associations between different variable qualitative categories**.

1. What kind of tea do you drink (black, green, flavored)
2. How do you drink it (alone, w/milk, w/lemon, other)
3. What kind of presentation do you buy (tea bags, loose tea, both)
4. Do you add sugar (yes, no)
5. Where do you buy it (supermarket, shops, both)
6. Do you always drink tea (always, not always)

```
##         Tea   How     how    sugar       where     always
## 1     black alone tea bag    sugar chain store Not.always
## 2     black  milk tea bag No.sugar chain store Not.always
## 3 Earl Grey alone tea bag No.sugar chain store Not.always
## 4 Earl Grey alone tea bag    sugar chain store Not.always
## 5 Earl Grey alone tea bag No.sugar chain store     always
## 6 Earl Grey alone tea bag No.sugar chain store Not.always
```



A point is a row in the table.
The labels represent **barycenters** (centroids)

http://www.gastonsanchez.com/visually-enforced/how-to/2012/10/13/MCA-in-R/

ANALYSIS & RESULTS

Coding

Analysis performed with R (R Core Team 2018)

- tidyverse (Wickham 2017)
- FactoMineR (Lê, Josse & Husson 2008)
- factoextra (Kassambra & Mundt 2017)

Data preparation

Based on Glynn (2014), be aware of the following four things when doing correspondence analyses: *fishing for results*, *over-simplifying the model*, *over-complicating the model*, avoiding *sparseness*

So this is a bit of a give-and-take, when deciding what the values of the categories are.

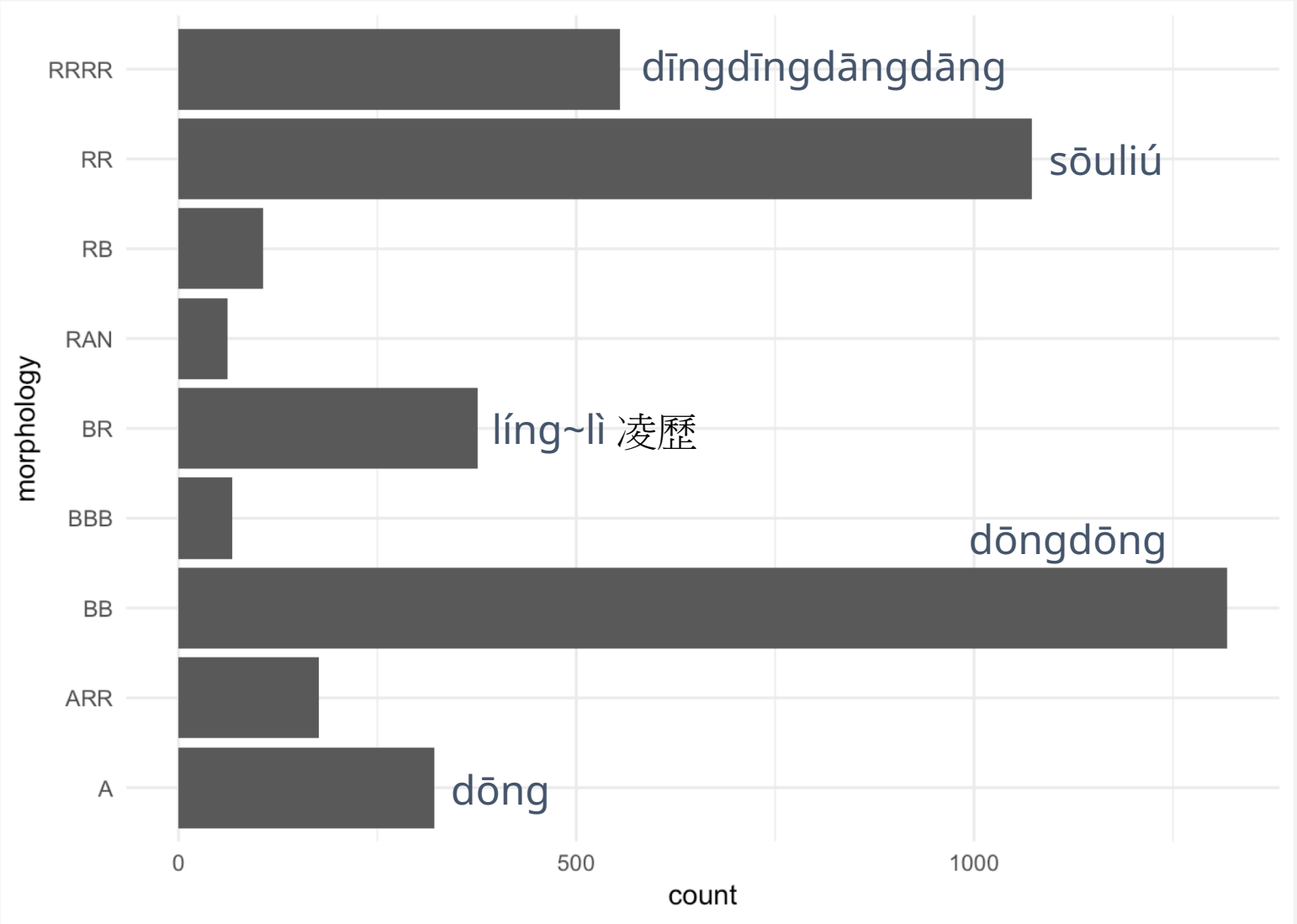
- multisensorial (e.g. x = SOUND+VISUAL) was split into two times the same word (x = SOUND + x = VISUAL)
- COGNITIVE_STATE, FEELING, INNER_FEELING, INNERFEELINGS were recoded as INNERSTATE
- WETNESS, TACTILE were recoded as TEXTURE
- SMELL (n = 8), TASTE (n = 2) were recoded as NASAL (Glynn 2014: at least n > 10)
- $n_{morphology} < 50$ was left out

Morphology

The major construction is full reduplication (BB).

But BR, RR, RRRR, and A occur often as well.

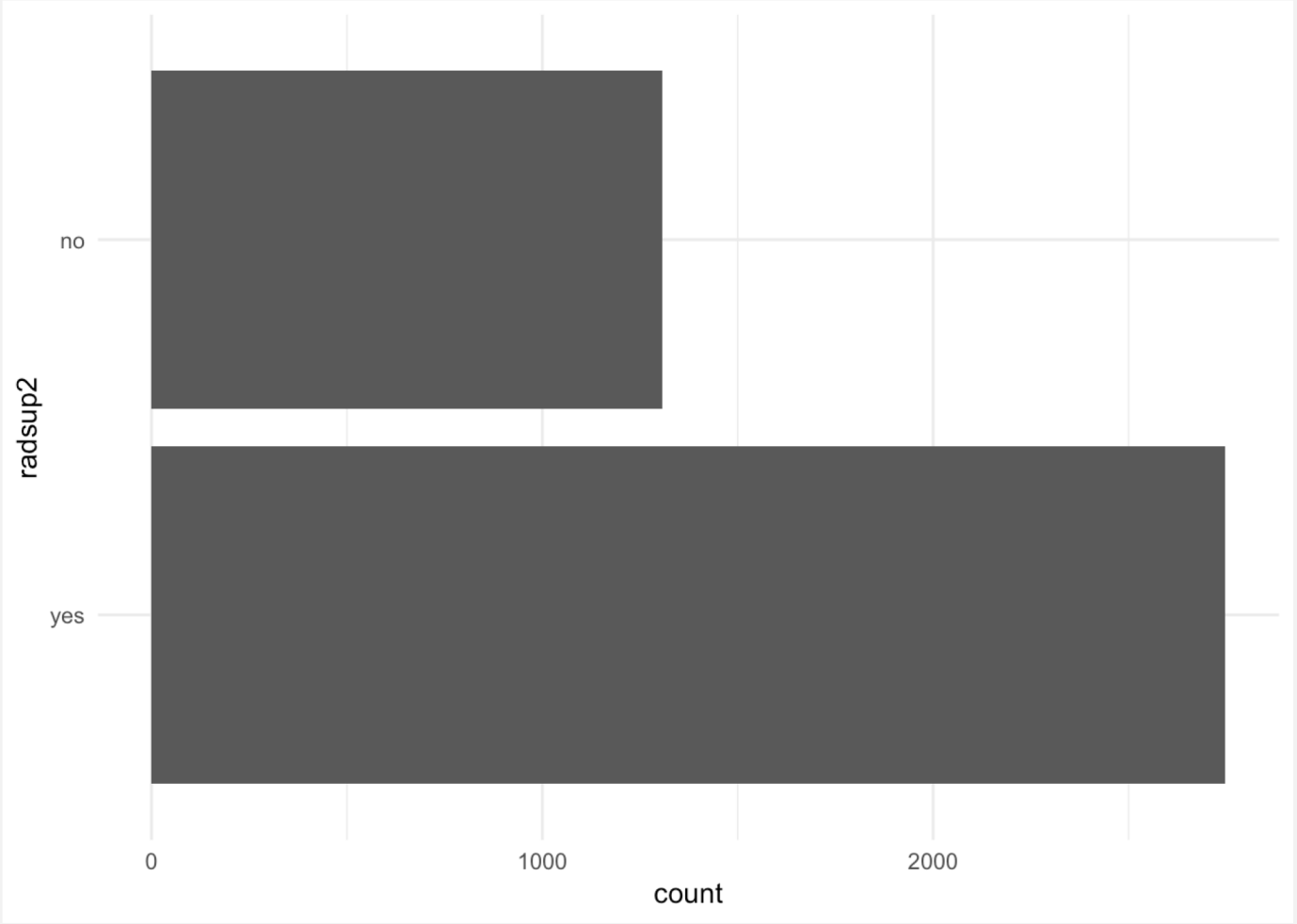
(We try to provide accurate constructions but for many the base is simply unknown, so RR is very high.)



Radical support

This was also recoded into a binary parameter.

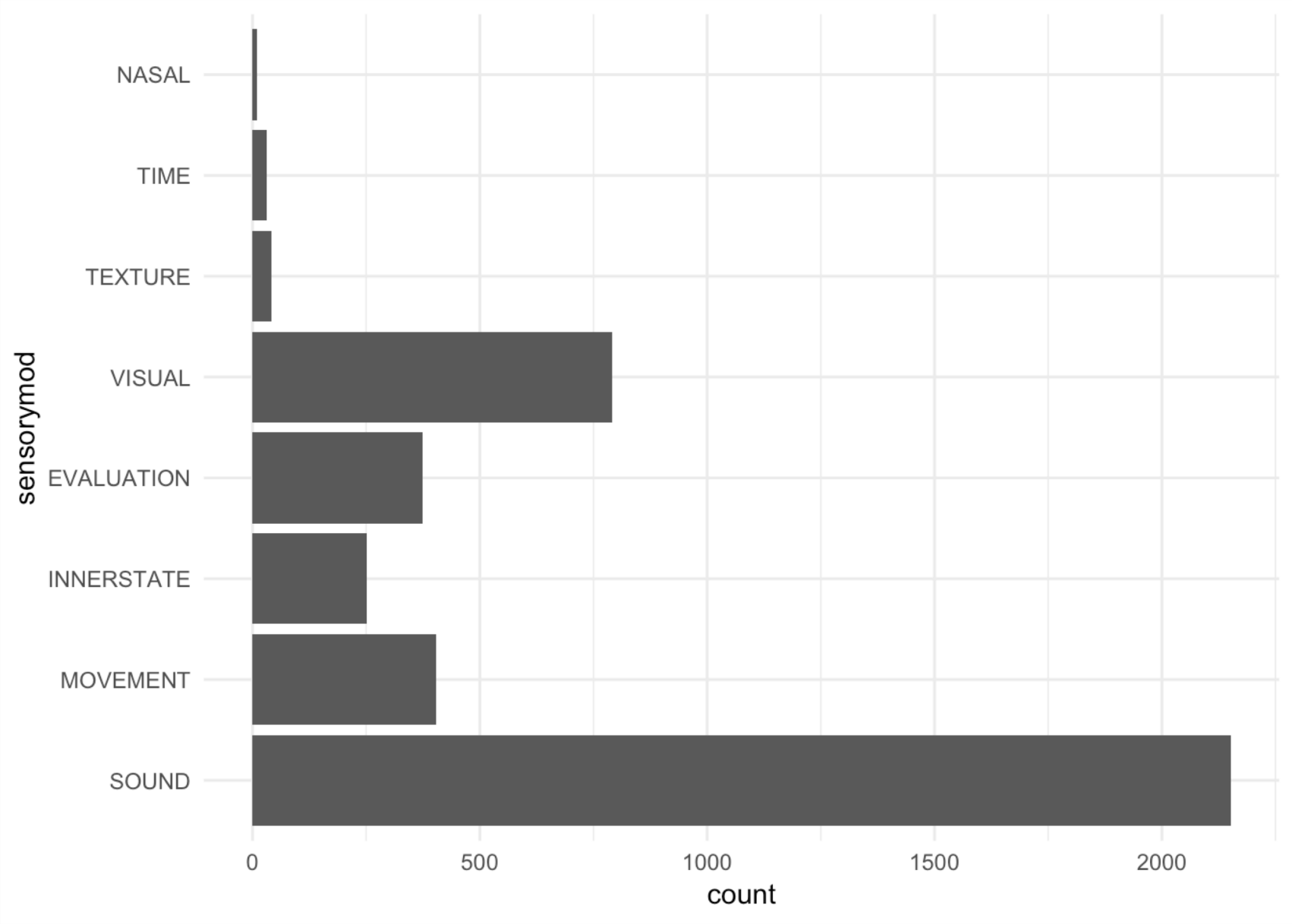
The data here thus suggests that the big majority of items in CHIDEOD has a form of radical support, where the radical 部首 somehow indicates the markedness.



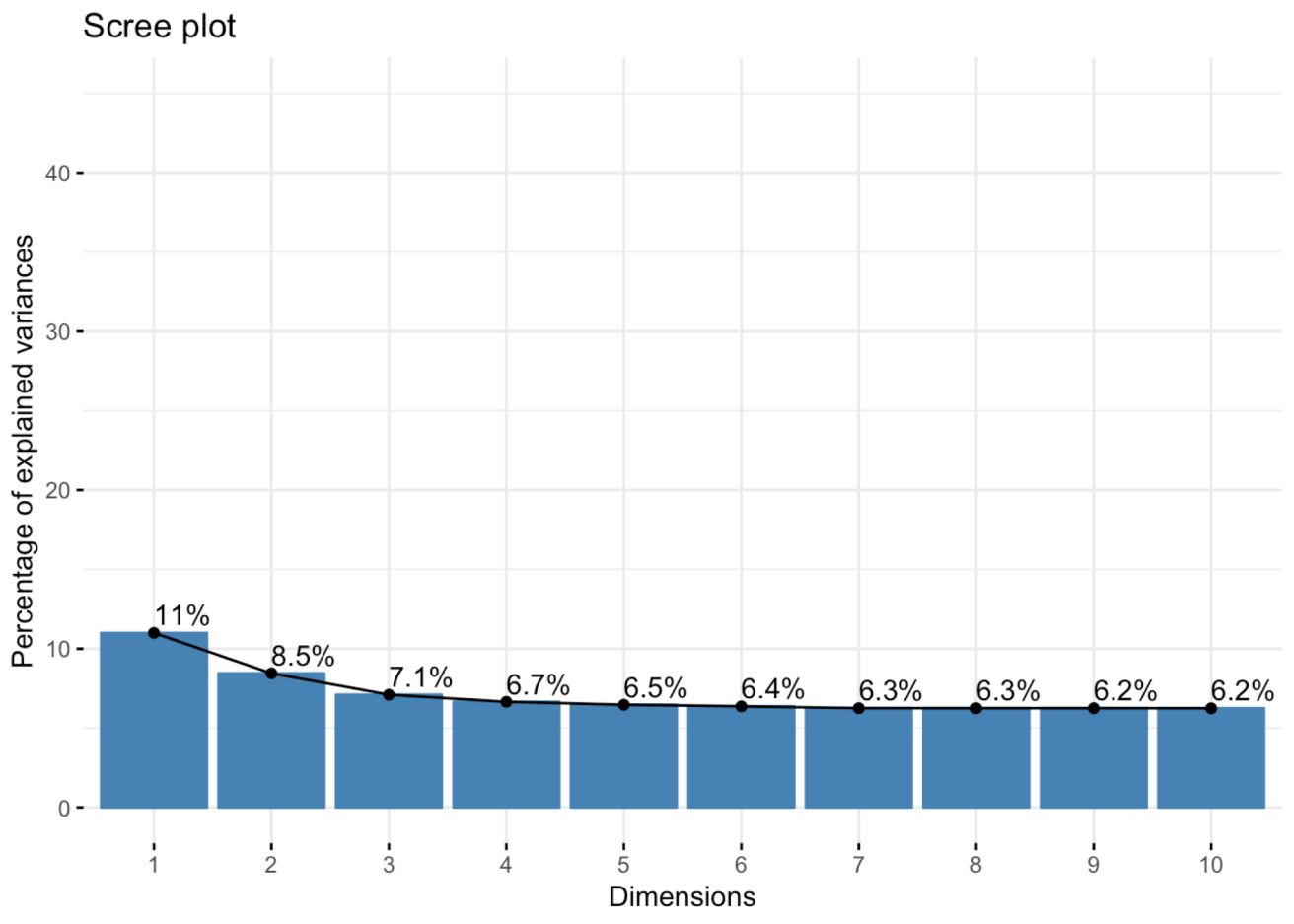
Sensory domain

Good indication of which categories occur often: SOUND, VISION-MOVEMENT, EVALUATION-INNERSTATE

But not: NASAL, TIME, TEXTURE



Coverage of the data



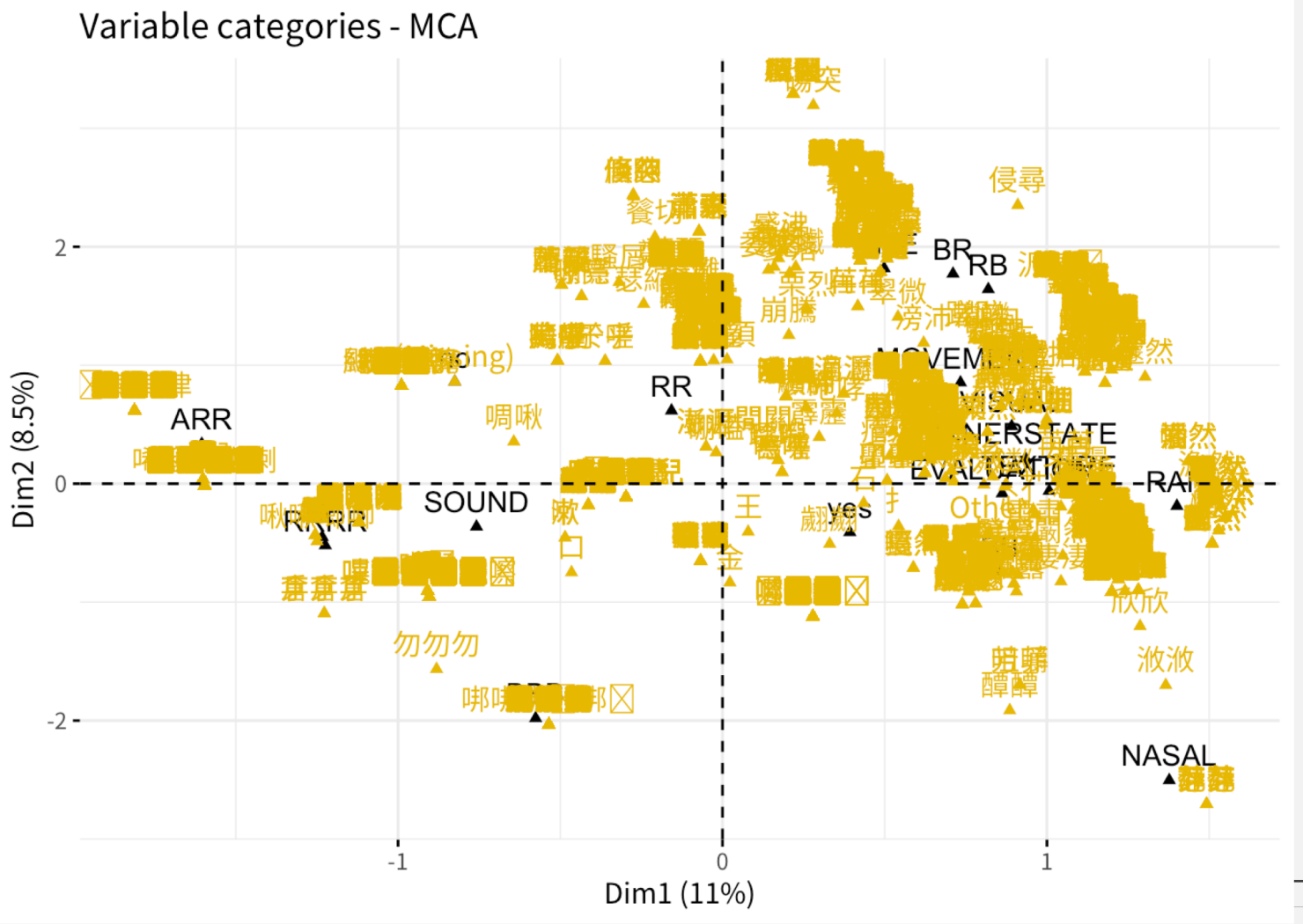
If this was simple Correspondence Analysis (CA), these values could just be trusted.

But for Multiple Correspondence Analysis (MCA) they are judged too harshly.

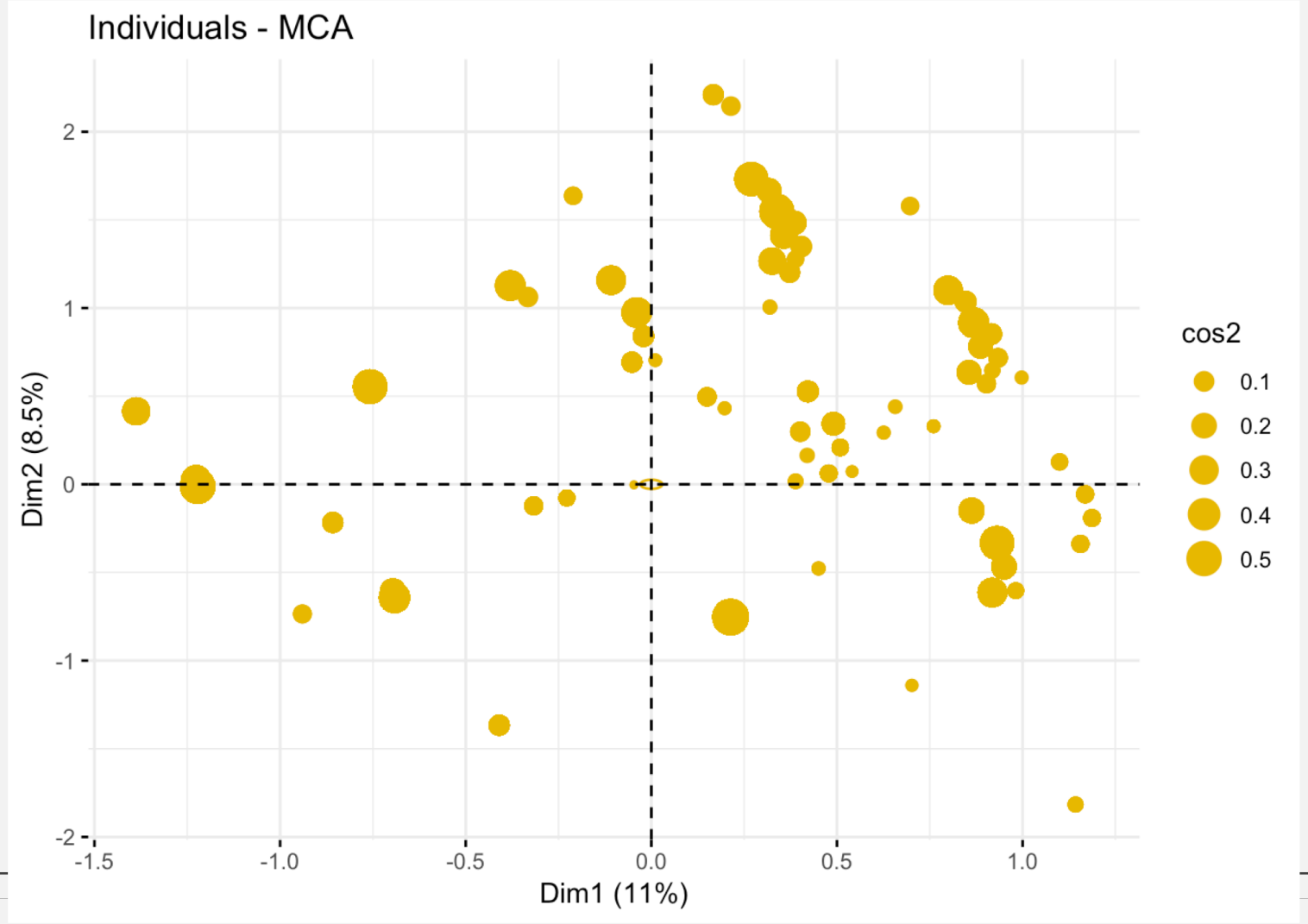
So using Greenacre's correction it yields:

Dim 1: 63.1 %
Dim 2: 13.6 %
Tot: 76.7 % > 75 (so is okay)

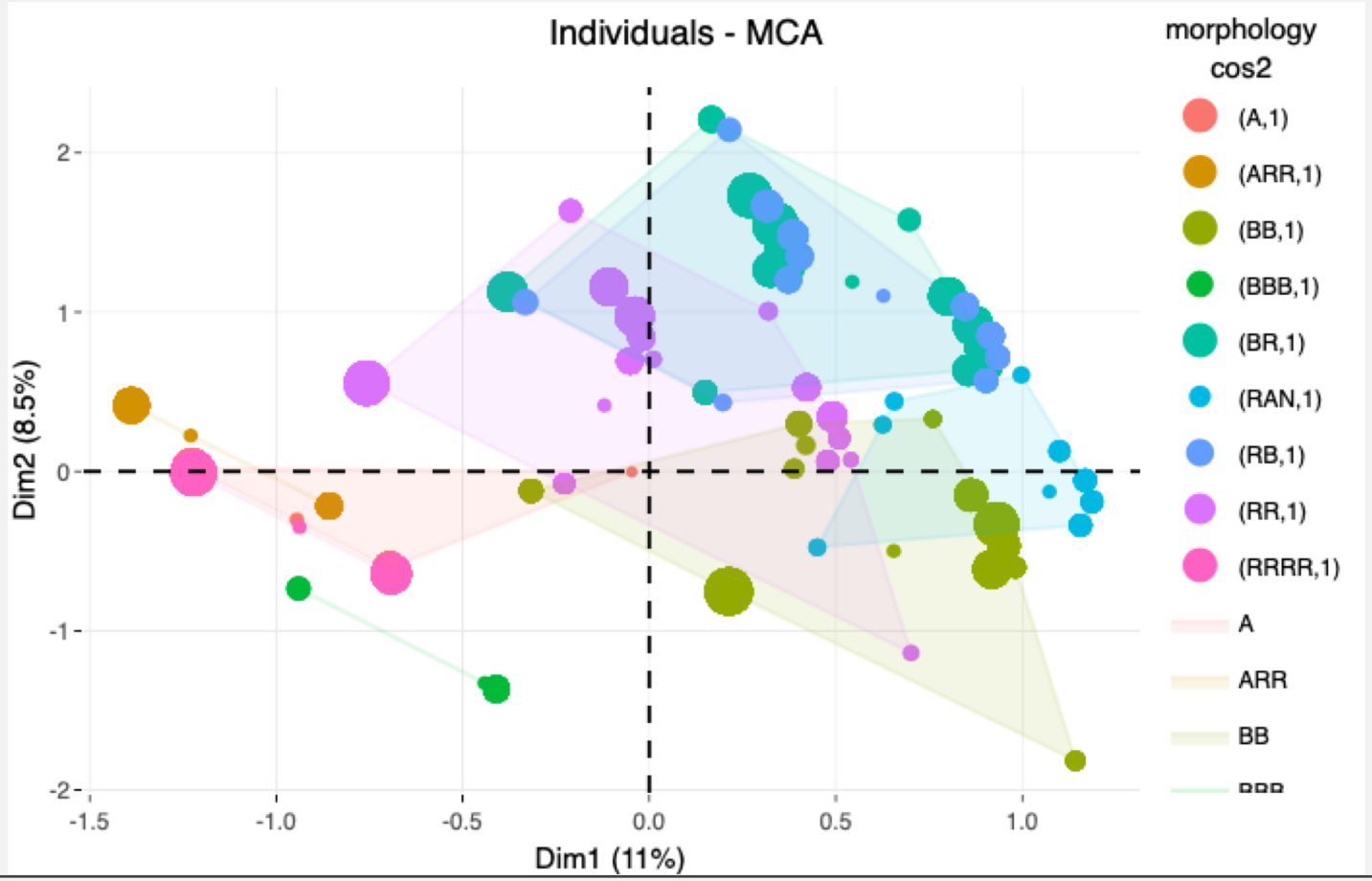
MCA of variable categories



MCA of variable categories

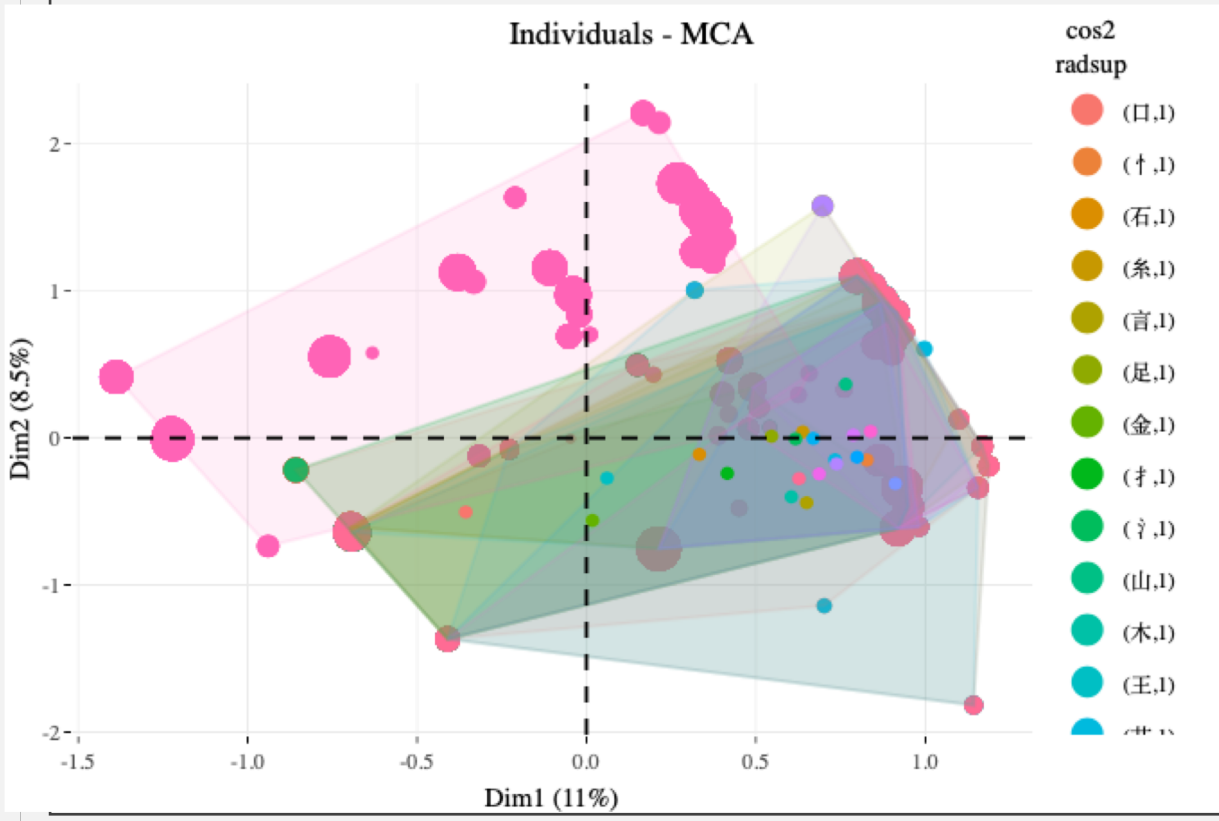


Morphology

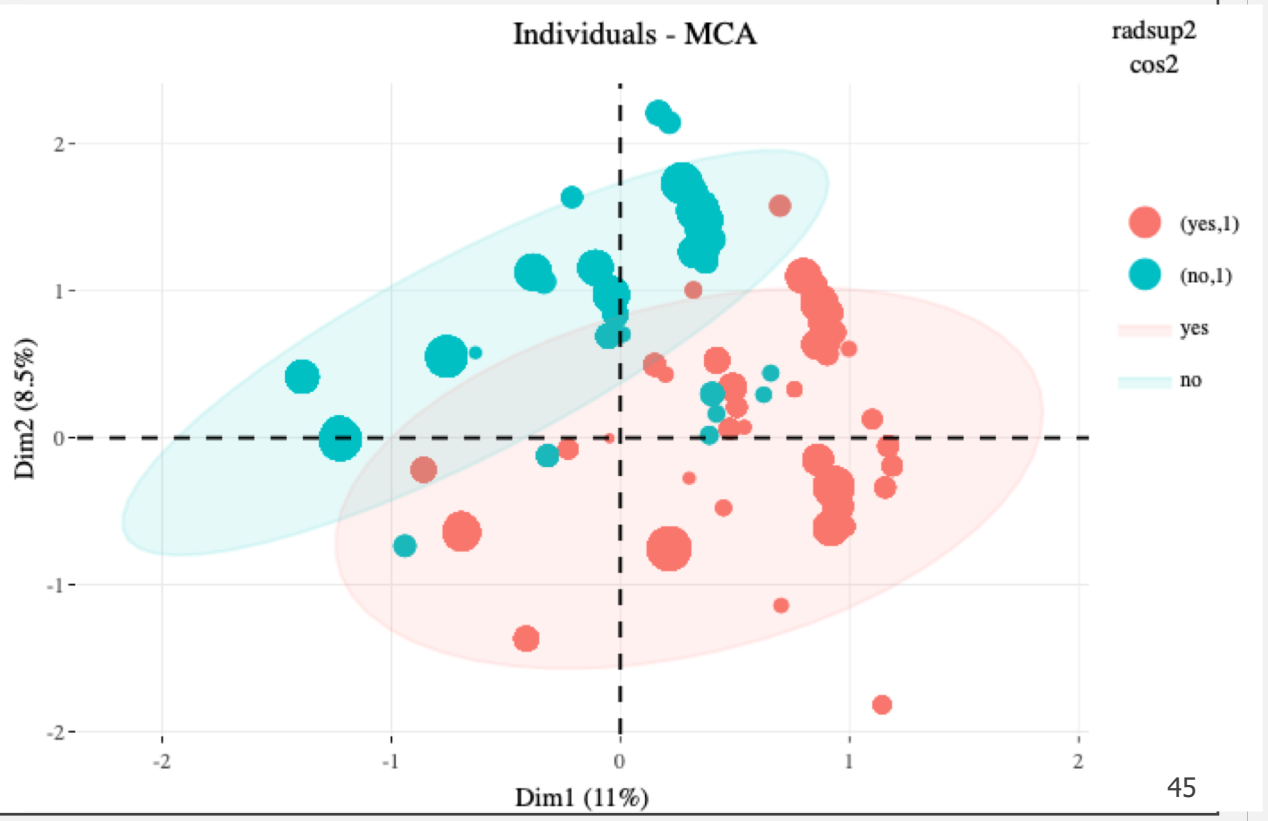

Individuals - MCA
morphology
cos2
(A,1)
(ARR,1)
(BB,1)
(BBB,1)
(BR,1)
(RAN,1)
(RB,1)
(RR,1)
(RRRR,1)
A
ARR
BB
Dim2 (8.5%)
Dim1 (11%)
2
1
0
-1
-2
-1.5
-1.0
-0.5
0.0
0.5
1.0

Radical support

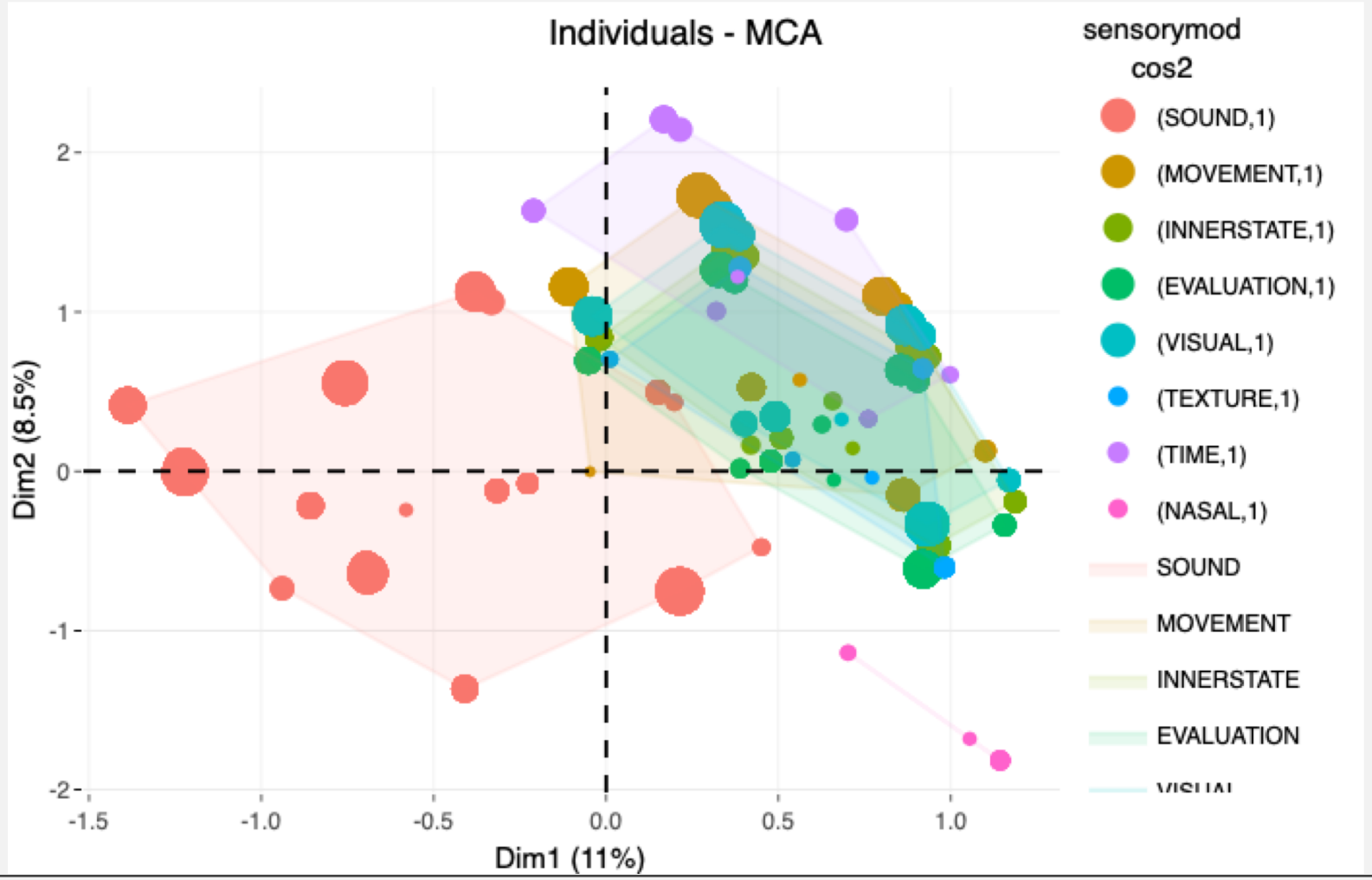
Real radicals



Binary



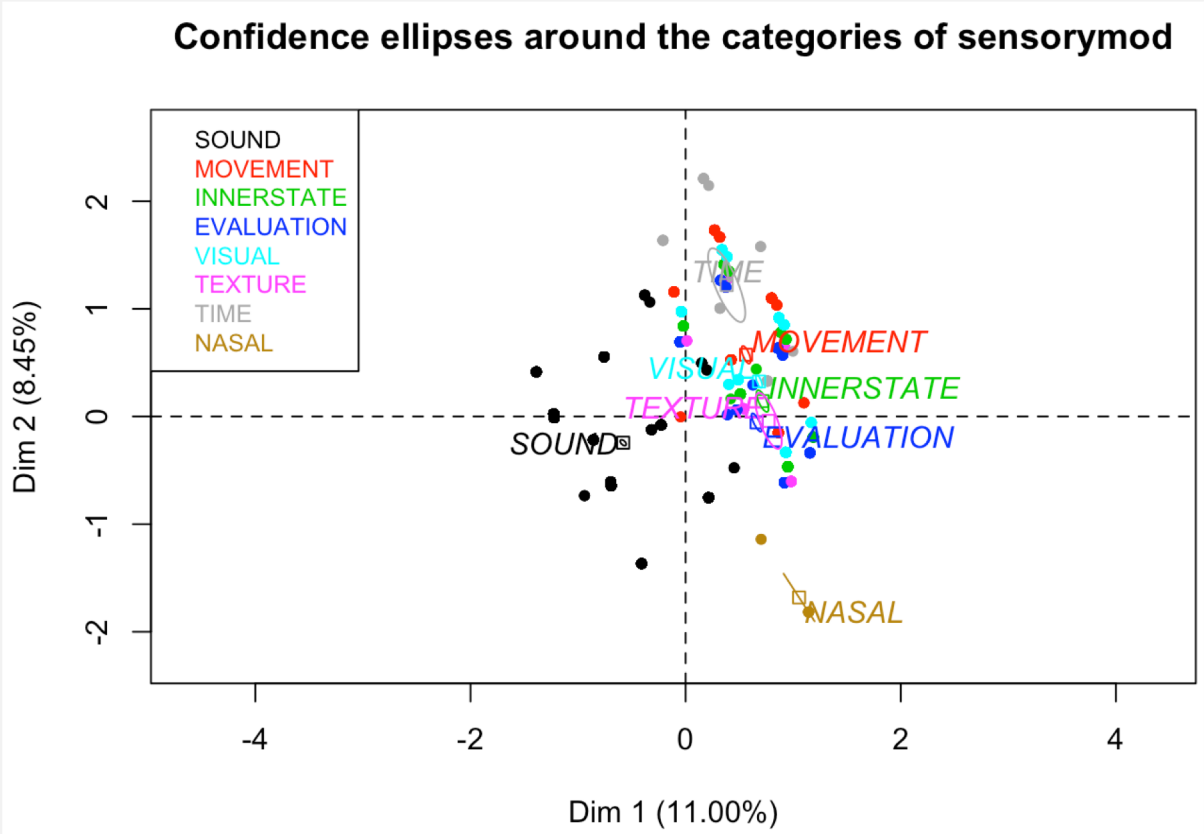
Sensory modality

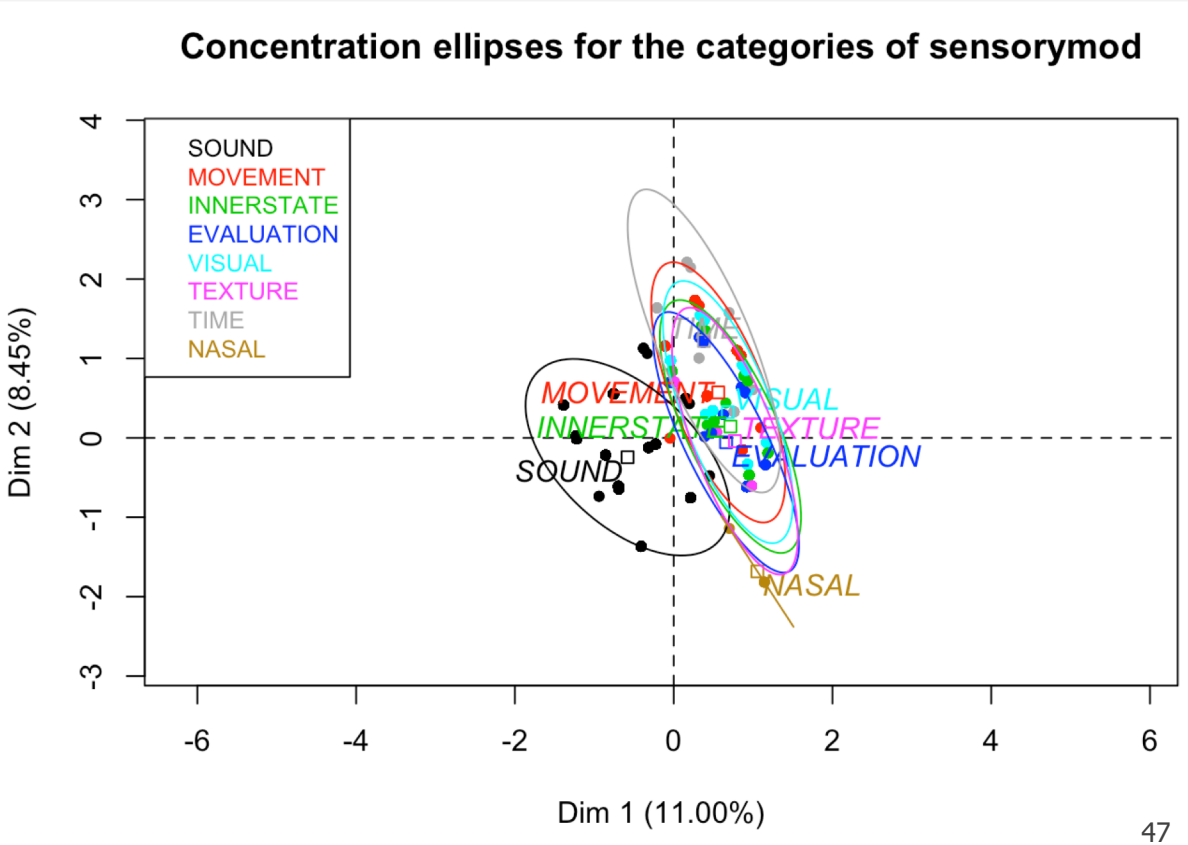


Sensory modality

Prototypical core

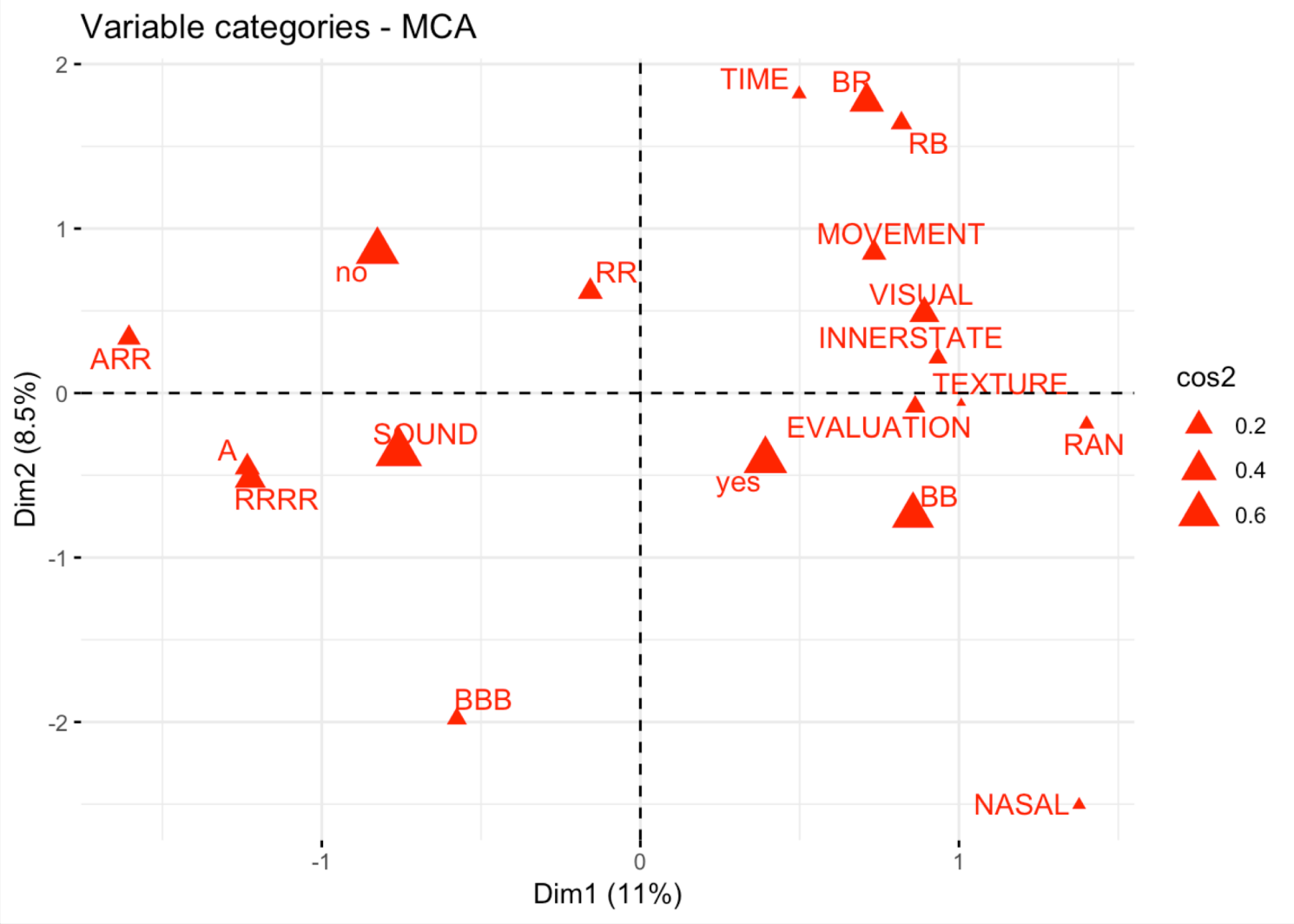
Fuzziness at edges





Cf. Levshina (2015)

Summary of MCA



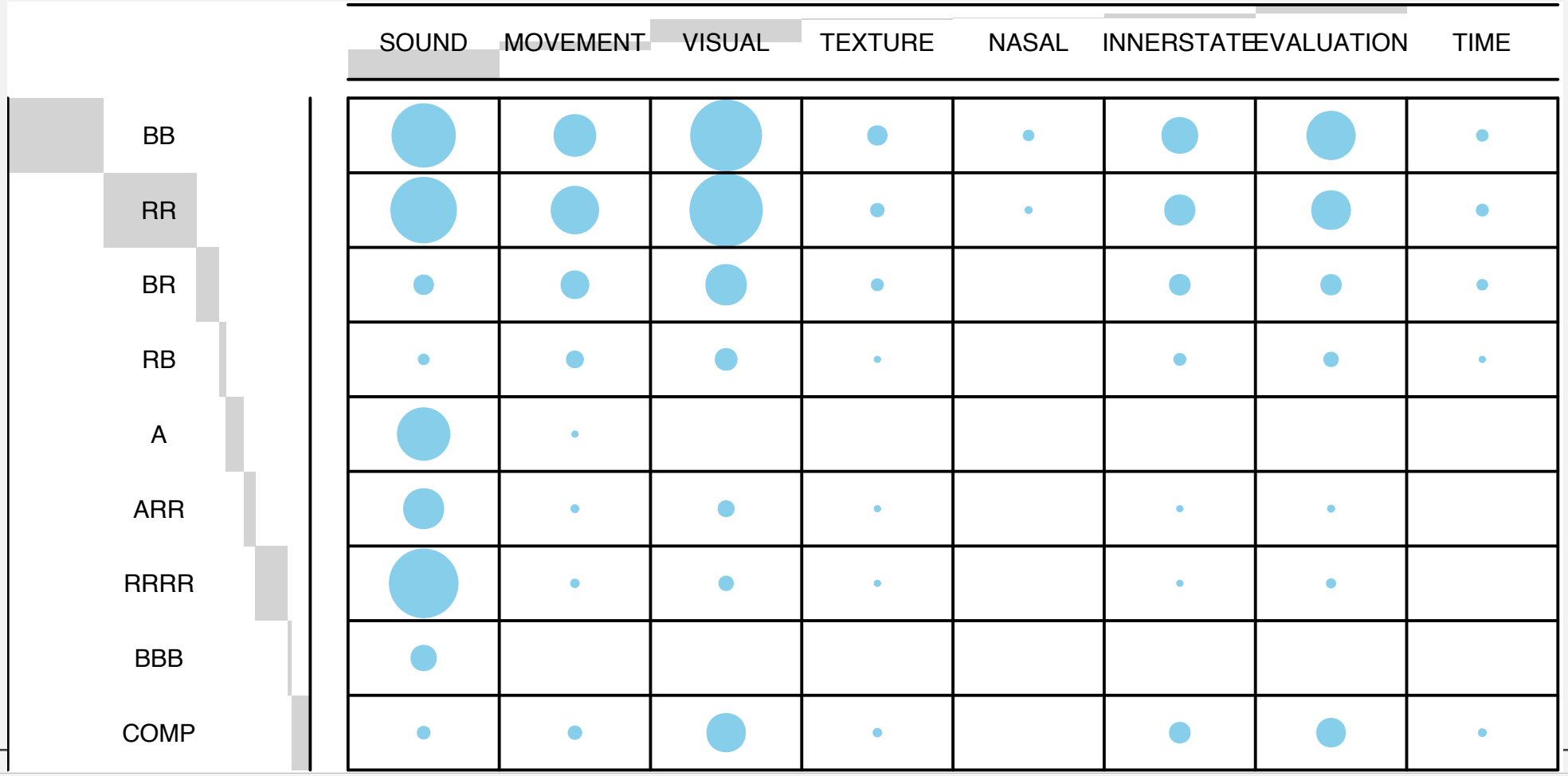
The barycenters of the factors / variables are shown as how they relate to one another.

Remember: based on the coverage (eigen values), the horizontal distance is the most important way of interpreting.

The goodness of fit of this plot is also confirmed statistically, through logistic regression (cf. Levshina 2015):

- Morphology C = 0.66 quite weak
- Radical support C = 0.99 outstanding
- Sensory modality C = 0.88 excellent

What if only CA with morphology and sensory domain?



DISCUSSION AND CONCLUSIONS

Goal vs. Result of MCA

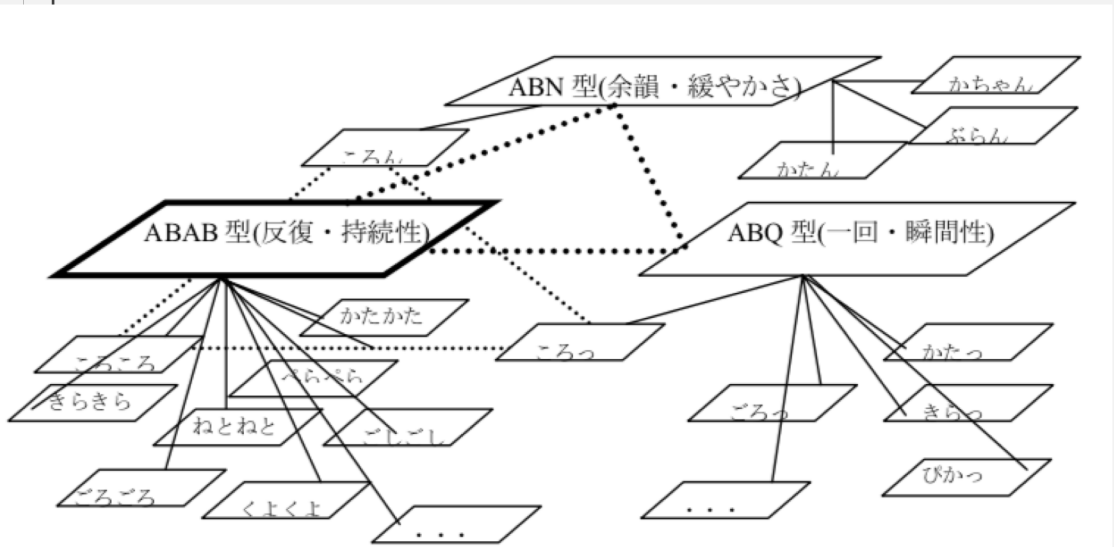
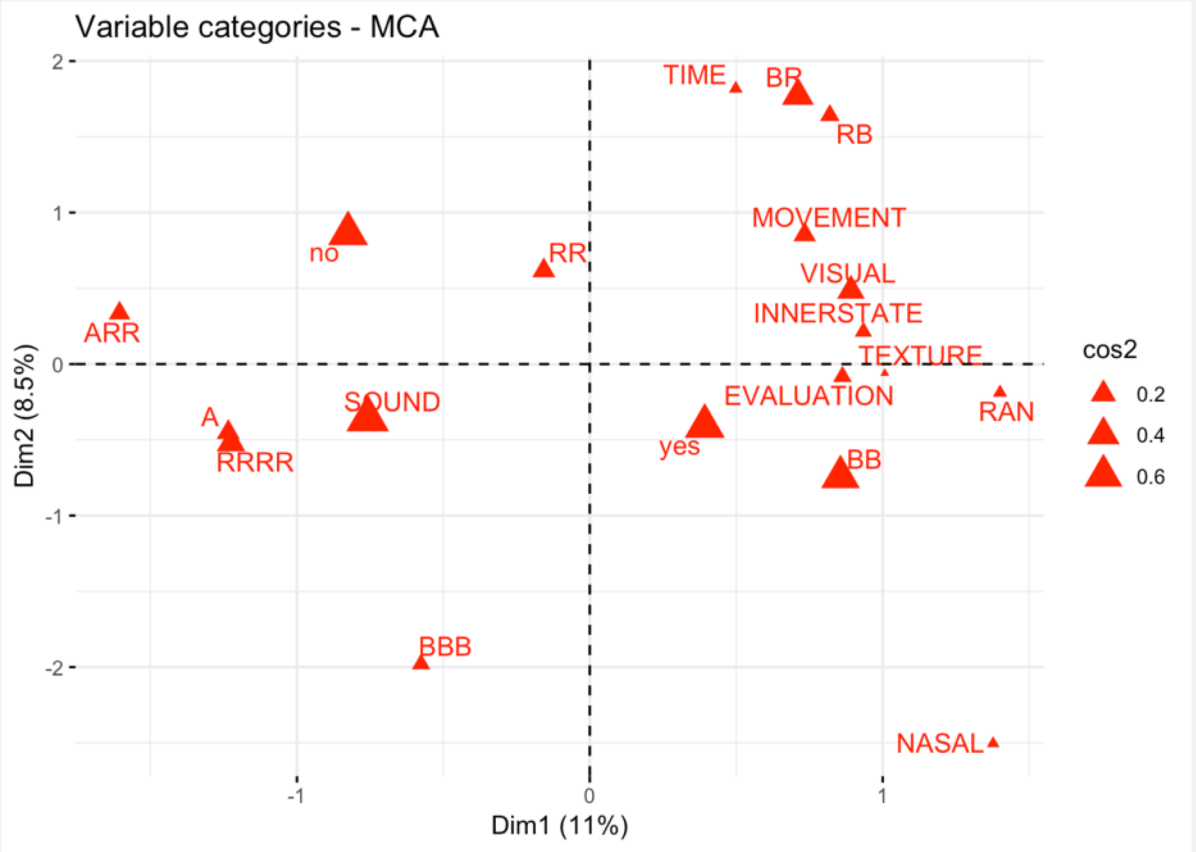


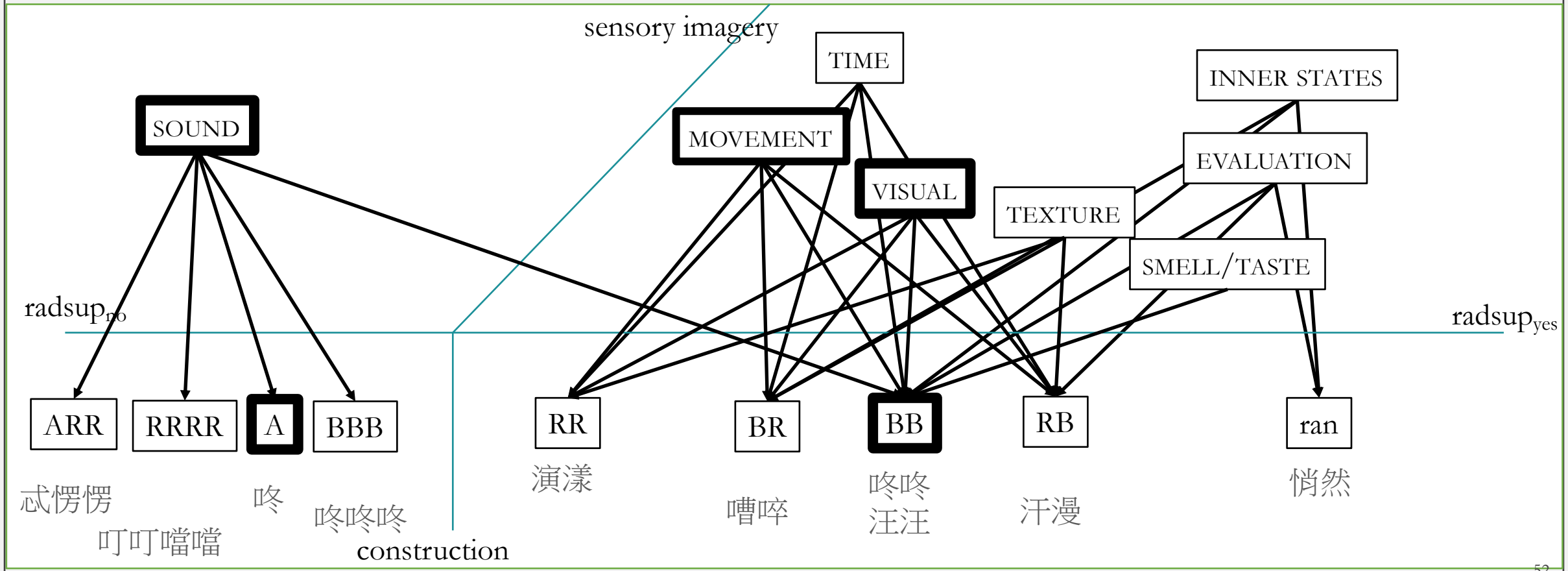
図 4.5　擬音語・擬態語のパターン・ネットワーク（立体図）

直線：事例化リンク　　点線：パターン同士のリンク



Let's transform this into a diagram that aims to capture the most important relations.

The prototypes for Chinese



Dual prototypical cores

There seem to be **two prototypical cores (dyadic prototype structure)**:

- a set of constructions focused around **SOUND**
- the rest **VISUAL, MOVEMENT** but also TEXTURE, TASTE/SMELL, INNER STATES, EVALUATION, TIME

This difference between SOUND and the rest may also be the reason why Chinese research traditionally only focused on onomatopoeia (SOUND ideophones). Yet, it is clear that there is also enough overlap to justify including other sensory modalities in a more functional framework of ideophones.

Quite a large number of constructions, but they do correlate with the two cores.

Future research

- Is there a historical evolution from one core to the other? Did one of the cores become more important / less important?
- If possible: inter-rater agreement (Fleiss's kappa) of sensory domains
- But first a familiarity rating is needed (working on it)
- Take into account the groups that were excluded in the beginning – to see how effective this analysis is.

Is this the final story?

No.

A comprehensive view of ideophones (in any language) needs to look at **token frequencies** and usage as well.

Maybe we can look at the interaction between the current variables + POS / construction
(in the same way Akita argued about quasi-mimetics in the beginning).

However, if you want to define Chinese ideophones in the *langue,* in the *system,* then this is a convincing way.

Is this the final story?

In other words, by the final paper, I hope to take these results and discuss them properly in depth.

However, I hope you can see how I approached the definitional question:

1. **Look at previous examples, e.g. Japanese (in this presentation) + cross-linguistic approaches, e.g. Canonical Typology etc.**
2. **See how these criteria are to be applied in Chinese.**
3. **Include — exclude groups (the basis of CHIDEOD).**
4. **Analyze what we have.**
5. **Interpret the results.**

References

Akita, Kimi. 2016. A multimedia encyclopedia of Japanese mimetics: A frame-semantic approach to L2 sound-symbolic words. In Kaori Kabata & Kiyoko Toratani (eds.), *Cognitive-functional approaches to the study of Japanese as a second language* (Studies on Language Acquisition volume 46), 139–167. Boston: De Gruyter Mouton.

Akita, Kimi. 2017. The linguistic integration of Japanese ideophones and its typological implications. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 62(02). 314–334. doi:10.1017/cnj.2017.6.

Akita, Kimi & Prashant Pardeshi (eds.). 2019. *Ideophones, mimetics and expressives* (Iconicity in Language and Literature, ILL 16). Amsterdam : Philadelphia: John Benjamins Pub. Co.

Bodomo, Adams. 2005. A corpus of Cantonese ideophones. *University of Hong Kong* 59.

Brown, Dunstan & Marina Chumakina. 2013. What there might be and what there is: An introduction to Canonical Typology. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 1–19. First edition. Oxford: Oxford University Press.

Brown, Dunstan, Marina Chumakina & Greville G. Corbett (eds.). 2013. *Canonical morphology and syntax*. First edition. Oxford: Oxford University Press.

Childs, G. Tucker. 1994. African ideophones. In Leanne Hinton, Johanna Nichols & John J. Ohala (eds.), *Sound symbolism*, 178–204. Cambridge [England]: Cambridge University Press.

Corbett, Greville G. 2007. Canonical Typology, suppletion, and possible words. *Language* 83(1). 8–42. doi:10.1353/lan.2007.0006.

Dingemanse, Mark. 2011. *The meaning and use of ideophones in Siwu*. Nijmegen: Radboud University Nijmegen dissertation.

Dingemanse, Mark. 2012. Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass* 6(10). 654–672.

Dingemanse, Mark. 2015. Ideophones and reduplication: Depiction, description, and the interpretation of repeated talk in discourse. *Studies in Language* 39(4). 946–970. doi:10.1075/sl.39.4.05din.

Dingemanse, Mark. 2019. "Ideophone" as a comparative concept. In Kimi Akita & Prashant Pardeshi (eds.), *Ideophones, mimetics and expressives* (Iconicity in Language and Literature, ILL 16), 13–33. Amsterdam : Philadelphia: John Benjamins Pub. Co.

Glynn, Dylan. 2014. Correspondence analysis: Exploring data and identifying patterns. In Dylan Glynn & Justyna A. Robinson (eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy* (Human Cognitive Processing volume 43), 307–341. Amsterdam ; Philadelphia: John Benjamins Publishing Company.

Gōng, Liángyù 龚良玉 (ed.). 1991. *Xiàngshēngcí cídiǎn 象声词词典 [Dictionary of onomatopoeia]*. Guìyáng: Guìzhōu jiàoyù chūbǎnshè.

Herlofsky, William J. 2019. Re-vision: Are there ideophones in signed languages? Oral presentation presented at the ILL 12 [International Symposium on Iconicity in Language and Literature], Lund: Lund University. https://konferens.ht.lu.se/en/ill-12/.

References

Kassambra, Alboukadel & Fabian Mundt. 2017. *factoextra: Extract and visualize the results of multivariate data analyses (v. 1.0.5.999)*. http://www.sthda.com/english/rpkgs/factoextra.

Kroll, Paul W. 2015. *A student's dictionary of Classical and Medieval Chinese* (Handbook of Oriental Studies: Section 4 China 30). Leiden: Brill.

Kwon, Nahyun. 2017. Total reduplication in Japanese ideophones: An exercise in Localized Canonical Typology. *Glossa: a journal of general linguistics* 2(1). 40. doi:10.5334/gjgl.267.

Kwon, Nahyun & Erich R. Round. 2015. Phonaesthemes in morphological theory. *Morphology* 25(1). 1–27. doi:10.1007/s11525-014-9250-z.

Lê, Sébastien, Julie Josse & François Husson. 2008. FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software* 25(1). 1–18.

Levshina, Natalia. 2015. *How to do linguistics with R: data exploration and statistical analysis*. Amsterdam ; Philadelphia: John Benjamins Publishing Company.

Lǐ, Jìng'ér 李镜儿. 2007. *Xiàndài Hànyǔ níshēngcí yánjiǔ 现代汉语拟声词研究 (Onomatopoeics in Modern Chinese)*. Shànghǎi: Xuélín chūbǎnshè.

Lockwood, Gwilym. 2017. *Talking sense: The behavioural and neural correlates of sound symbolism*. Nijmegen: Radboud University Nijmegen PhD dissertation.

Lu Chiarung 呂佳蓉. 2006. *Giongo, gitaigo no hiyuteki kakuchō no shosō: ninchi gengogaku to ruikeiron no kanten kara 擬音語・擬態語の比喩的拡張の諸相——認知言語学と類型論の観点から [Figurative extensions of mimetics: A Cognitive Linguistic and typological study]*. Kyōto: Kyōto University PhD dissertation.

McLean, Bonnie. 2019. One form, many meanings: Iconicity in phonological and semantic development. Canberra: The Australian National University Honours thesis.

Meng, Chenxi. 2012. *A description of ideophonic words in Mandarin Chinese*. Leiden: Leiden University Research Master in Linguistics.

Mok, Waiching Enid. 2001. *Chinese sound symbolism: A phonological perspective*. Hawai'i: University of Hawai'i PhD dissertation.

Newman, Paul. 1968. Ideophones from a syntactic point of view. *Journal of West African Languages* 5. 107–117.

NINJAL. 2016. Lago Word Profiler for the Balanced Corpus of Contemporary Written Japanese (NINJAL-LWP for BCCWJ). Database. http://nlb.ninjal.ac.jp/search/ (25 March, 2019).

Nuckolls, Janis B. & Tod D. Swanson. 2019. Quechua Real Words: An audiovisual ANTI-dictioanry of expressive Quechua ideophones. http://quechuarealwords-dev.byu.edu/index.php (12 March, 2019).

Nuckolls, Janis B., Tod D. Swanson, Diana Shelton, Alexander Rice & Sarah Hatton. 2017. Lexicography in-your-face: The active semantics of Pastaza Quichua ideophones. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 62(02). 154–172. doi:10.1017/cnj.2017.9.

Qiu, Di 邱迪. 2018. Xiandai Hanyu zhong de "gezhong + nishengci" jiegou yanjiu 现代汉语中的"各种+拟声词"结构研究 [The construction "ALL KINDS OF + onomatopoeia" in Modern Chinese]. *Liaoning jiaoyu xingzheng xueyuan xuebao 辽宁教育行政学院学报* 4. 69–72.

Sun, Jingtao. 1999. *Reduplication in Old Chinese*. Vancouver: University of British Columbia PhD dissertation.

References

Thompson, Arthur Lewis. 2018. Are tones in the expressive lexicon iconic? Evidence from three Chinese languages. *PLOS ONE* 13(12). e0204270. doi:10.1371/journal.pone.0204270.

Van Hoey, Thomas. 2015. *Ideophones in Middle Chinese: A typological study of a Tang dynasty poetic corpus*. Leuven: KU Leuven Master thesis.

Van Hoey, Thomas. 2016. Ideophones in Old Chinese: the case of the Shijing 詩經. *ISACG 9 [International Symposium on Ancient Chinese Grammar]*. Berlin: Humboldt University.

Van Hoey, Thomas & Arthur Lewis Thompson. under review. The Chinese Ideophone Database (CHIDEOD). *International Journal of Corpus Linguistics*.

Van Hoey, Thomas & Arthur Lewis Thompson. 2019. Bridging phonology, meaning, and written form across time: Introducing a database of Chinese literary ideophones. Oral presentation presented at the ILL 12 [International Symposium on Iconicity in Language and Literature], Lund: Lund University. https://konferens.ht.lu.se/en/ill-12/.

Wáng, Wànrén 王万仁 (ed.). 1987. *Xiàngshēngcí lì shì 象声词例释 [Examples and explanations of onomatopoeia]*. Nánníng: Guǎngxī jiàoyù chūbǎnshè : Guǎngxī xīnhuá shūdiàn fāxíng.

Wickham, Hadley. 2017. *tidyverse: Easily install and load the "Tidyverse" (v. 1.2.1)*. https://CRAN.R-project.org/package=tidyverse.

Winter, Bodo. 2019. *Sensory Linguistics: Language, perception and metaphor* (Converging Evidence in Language and Communication Research 20). Amsterdam: John Benjamins Publishing Company. doi:10.1075/celcr.20.

Wu, Mengqi. 2014. *The structure of ideophones in Southern Sinitic*. Hong Kong: University of Hong Kong Master thesis.

Xu Zhongshu 徐中舒 & Hanyu da zidian bianji weiyuanhui 汉语大字典编辑委员会. 1995. *Hanyu da zidian 漢語大字典 (Great Compendium of Chinese Characters)*. 3 vols. Wuhan: Sichuan cishu chubanshe.

You, Weilun 游韋倫. 2015. *Nichū ryō gengo ni okeru giongo no imi to imi kakuchō: Furēmu imi-ron no kanten kara no apurōchi 日中両言語における擬音語の意味と意味拡張 ―フレーム意味論の観 点からのアプローチ― [The semantics and semantic extensions of onomatopoeia in Japanese and Chinese: A Frame Semantics approach]*. Kobe: Kobe University PhD dissertation.

Zhao, Aiwu 赵爱武. 2008. Jin 20 nian Hanyu xiangshengci yanjiu zongshu 近20年汉语象声词研究综述 (Chinese onomatopoeia study in recent 20 years: summary). *Wuhan daxue xuebao 武汉大学学报 (renwen kexue ban 人文科学版)* 61(2). 180–185.

ご清聴ありがとうございました

MORE ABOUT CHIDEOD

Similar databases: BCCWJ's word profiler

Balanced Corpus of Contemporary Written Japanese (NINJAL 2016) has a word profiler (LWP)

The word profiler looks up words in the BCCWJ and provides sketch grammar-like statistics.

The goal of CHIDEOD is to collect all TYPES,

which later could be used in a corpus study

Similar databases: MEJaM

The Multimedia Encylcopedia of Japanese Mimetics (Akita 2016)

bururu
[collocation]
verb
adjective
nouns

Body movement / 体の動き

ぶるぶる
［コロケーション］
動詞：震える / shiver
形容詞：寒い / cold
名詞：手 / hand、唇 / lip、身体 / body、眼 / eye
［Google画像］

buru~buru

CHIDEOD might eventually evolve into a Multimedia CHIDEOD

Which would include pictures and video clips to illustrate the depictive nature.

However, given the diachronic and synchronic scope, this may not be realizable for all items.

Similar databases: Quechua Real Words

Audiovisual ANTI-dictionary of expressive Quechua ideophones

(Nuckolls 2017; Nuckolls & Swanson 2019)

The goal is to study the multimodal interaction between Quechua ideophones and gesture through video clips

polan

Subsets of data provided in CHIDEOD can aid in researching how multimodality(gesture) interacts with Chinese ideophones.