Multimodal distributional semantic models and conceptual representations in sensory deprived subjects

Giovanni Cassani and Alessandro Lopopolo
giovanni.cassani@uantwerpen.be, alopopolo@vu.nl

VRIJE UNIVERSITEIT AMSTERDAM
Universiteit Antwerpen

1. Introduction

Multimodal distributional semantics (MDS) aims to encode information extracted from large collections of texts, images, sounds and olfactory chemicals (so far, at least) in **vector spaces**.
This encoding has been shown to capture interesting semantic information. Here we compare MDS models with descriptions of 70 common words provided by congenitally blind (CB) and sighted subjects (S).

2. Datasets

BLIND A parallel collection of semantic feature norms collected for 70 concepts from a group of congenitally blind Italian subjects and a group of sighted subjects [1]. We manually annotated features in generally perceptual, visual, and acoustic.

ImageNet A large collection of labeled images organized along the WordNet hierarchy.

SoundFX A large dataset of labeled real world sounds referring to different concepts.

4. Research questions

1. Does **SoundFX** space compare differently between CB and S acoustic description of concepts?
2. Does **ImageNet** space fare better for visual description of concepts provided by S rather than by CB?

We expect that a MDS model shouldn't behave differently wrt perceptual features provided by CB and S, if the perceptual modality on which it is grounded is available to both populations. Vice versa, ImageNet space should fare worse for approximating concept relatedness in CB subjects, since they cannot rely on this source of perceptual information.

6. Conclusions

Although tentative and provisional conclusions can be drawn given the small sample sizes:

- no between-group differences emerge at the general perceptual level
- visual information explains sighted data better, as expected by availability of perceptual information,
- acoustic information is a better predictor in blind subjects in spite of being available to both groups, suggesting that a **compensatory strategy** might be in play.

A. Acknowledgements

We thank Marco Baroni for his insightful advice.
The first author is funded by a BOF/TOP grant (ID 29072) of the Research Council of the University of Antwerp; the second author is funded by NWO Spinoza Prize grant.
This poster was designed on Overleaf, and is based on the template developed by Brian Amberg.

3. Semantic representations

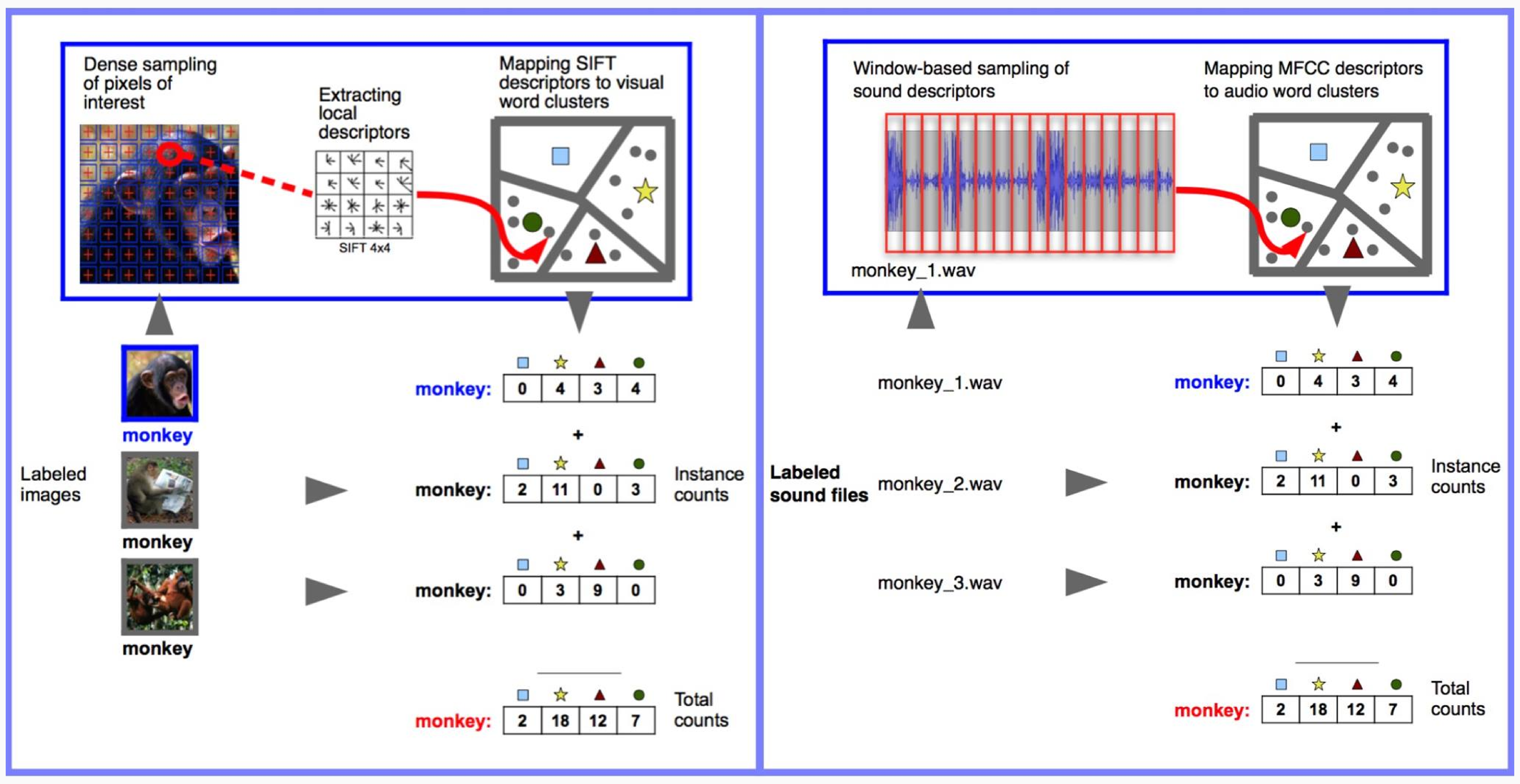

Figure 1: Multimodal distributional semantics representation of visual information (left panel, [2]) and acoustic information (right panel, [3])

Group	Target	Feature:Category	Production	Perceptual	Visual	Acoustic
blind	apple	round:ppe	12	1	1	0
blind	dog	to_bark:eve	5	1	0	1
sighted	ship	transport:isa	15	0	0	0
sighted	cat	soft:ppe	1	1	0	0
...	...	...	...	...	...	...

Table 1: Feature norms from the BLIND dataset. The production frequency of a given norm for a given target was taken as co-occurrence count. The three rightmost columns were added for the present study and each norm was manually tagged by the authors

5. Results

Feature sets	Target concepts	Norms $Blind \cup Sighted$	Spearman's Rho $Blind \sim Sighted$
All	70	1584	0.847
Perceptual	62	187	0.701
Visual	49	92	0.717
Acoustic	18	42	0.449

Table 2: PRELIMINARY EXPERIMENT: Between-groups correlations when different sets of feature norms are taken into account

Group	Norms	ImageNet space	SoundFX space
Sighted	perceptual	0.298	0.008
	visual	0.284	.
	acoustic	.	0.077
Blind	perceptual	0.282	0.223
	visual	0.245	.
	acoustic	.	-0.023

Table 3: EXPERIMENT 1 - Correlations (Spearman's ρ) between modality-specific MDS spaces, and modality-general and modality-specific norms-derived semantic spaces in sighted and congenitally blind subjects

Norms	Group	ImageNet space	SoundFX space
Visual	blind	0.245	.
	sighted	0.284	.
Acoustic	blind	.	0.115
	sighted	.	0.071

Table 4: EXPERIMENT 2 - Correlations (Spearman's ρ) between modality-specific MDS spaces and modality-specific norms-derived semantic spaces across blind and sighted subjects

B. References

[1] Alessandro Lenci, Marco Baroni, Giulia Cazzolli, and Giovanna Marotta. Blind: a set of semantic feature norms from the congenitally blind. *Behavior Research Methods*, 45(4):1218–1233, 2013.

[2] Elia Bruni, Nam Khanh Tran, and Marco Baroni. Multimodal distributional semantics. *Journal of Artificial Intelligence Research*, 49(1):1–47, 2014.

[3] Alessandro Lopopolo and Emiel van Miltenburg. Sound-based distributional models. In *Proceedings of the 11th International Conference on Computational Semantics*, pages 70–75, London, UK, 2015. Association for Computational Linguistics.