# Unseen features.
# Collecting semantic data from congenital blind subjects

**Alessandro Lenci♣, Marco Baroni♠, Giovanna Marotta♣**

♣University of Pisa (Italy)
♠University of Trento (Italy)
alessandro.lenci@ling.unipi.it, marco.baroni@unitn.it, gmarotta@ling.unipi.it

Congenital blind subjects are able to learn how to use color terms and other types of vision-related words in a way that is de facto undistinguishable from sighted people. It has actually been proposed that language provides a rich source of information that blind subjects can exploit to acquire aspects of word meaning that are related to visual experience, such as the color of fruits or animals. Despite this, whether and how sensory deprivation affects the structure of semantic representations is still an open question. In this talk, we present a new, freely available collection of feature norms produced by congenital blind subjects and normal sighted people. Subjects were asked to produce semantic features describing the meaning of concrete and abstract nouns and verbs. Data were collected from Italian subjects, translated into English, and categorized with respect to their semantic type (e.g. hypernym, meronym, physical property, etc.). First analyses of the feature norms highlight important differences between blind and sighted subjects, for instance for the role of color and other visual features in the produced semantic descriptions. This resource can provide new evidence on the role of perceptual experience in shaping concepts, as well as on its interplay with information extracted from linguistic data. The norms will also be used to carry out computational experiments with distributional semantic models to simulate blind and sighted semantic spaces.



*Proceedings of the EACL 2012 Workshop on Computational Models of Language Acquisition and Loss*, page 32,
Avignon, France, April 24 2012. 