# Marco Antonio Flores Coronado

## Education

**Bachelor in Hispanic Language and Literature**, National Autonomous University of Mexico
**Master in Science** (Computer Modeling and Scientific Calculation) Autonomous University of Morelos State

## Research Experience:

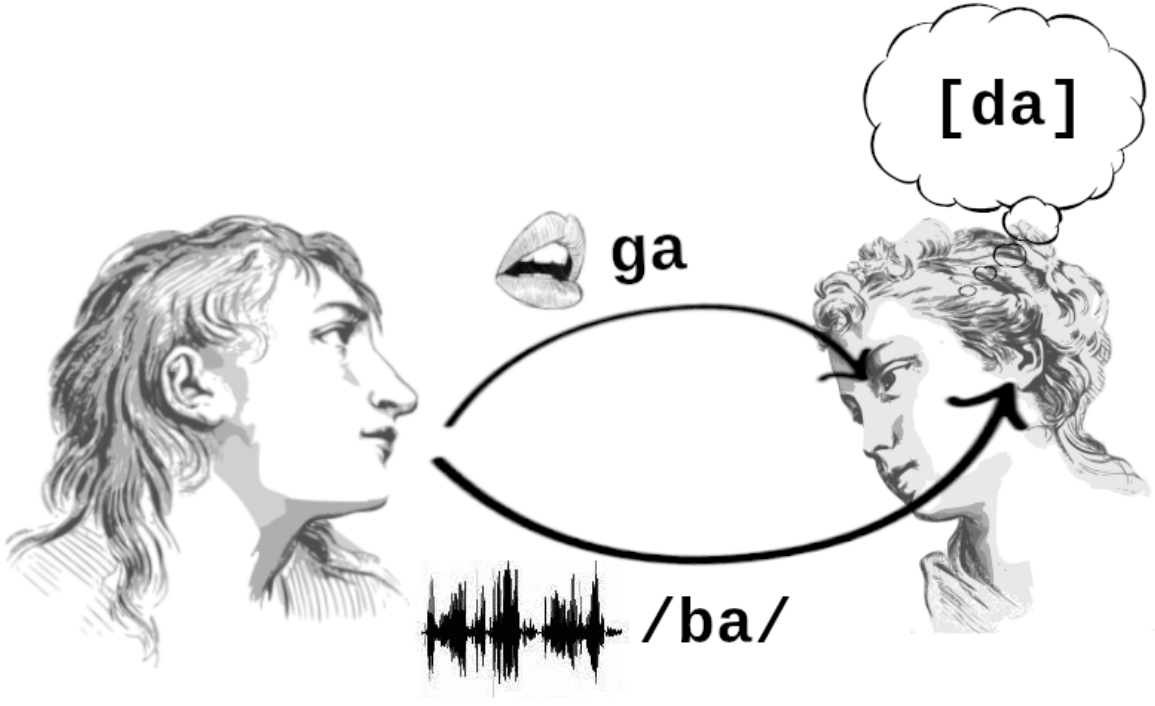
**Cognitive Robotics Lab** (2019-), Center for Science Investigation, UAEM
**Cognitive and Language Development Lab** (2018-),Psychology School, UNAM
**Psycholinguistics Lab** (2016-),Psychology School, UNAM

## Research Interests:

Multisensory integration -language-
Embodied cognition -meaning-
Language Processing
Cognitive Robotics
Language Aquisition -word and meaning- (typical and atypical)
Lexical networks -semantic and grammar interaction-

# Modelling of the McGurk effect, a multisensory integration illusion

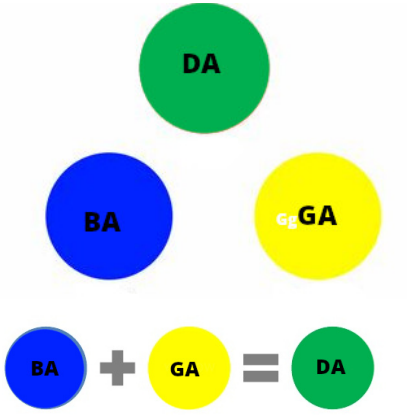
**visual** /*ga*/
**audition** /*ba*/
**perception** [*da*]

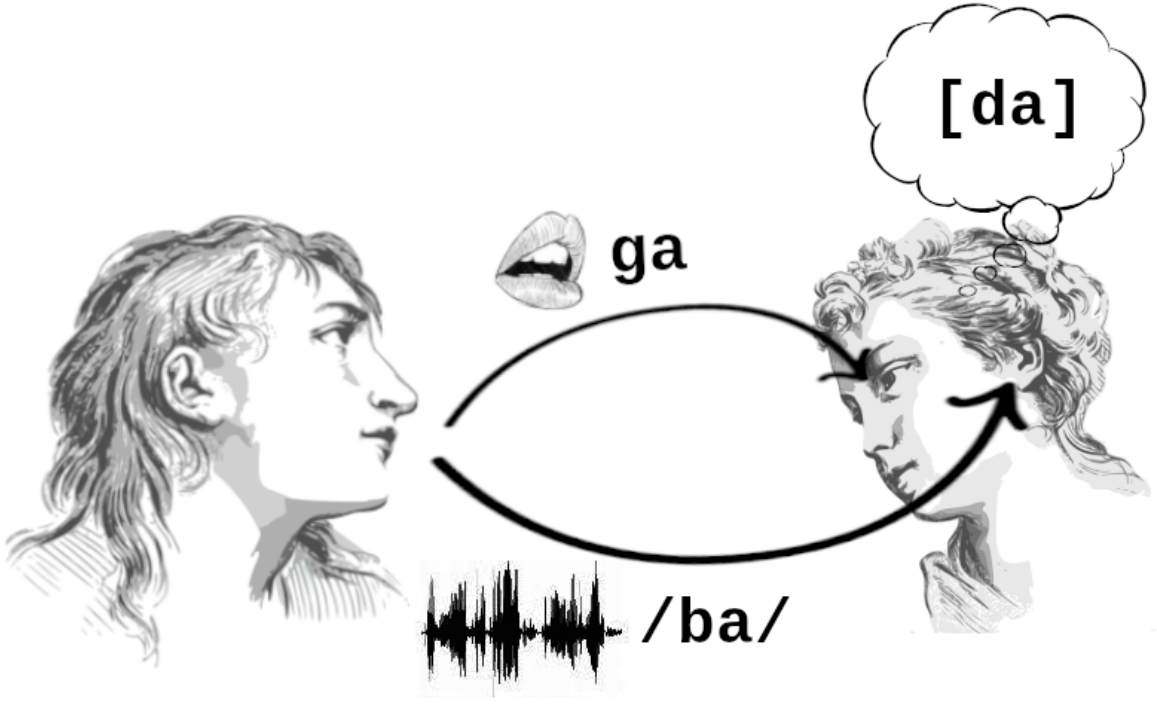



[Mcgurk and Macdonald, 1976, Van Engen et al., 2019, Mitchel et al., 2014]

# Modelling of the McGurk effect, a multisensory integration illusion

**visual** /*ga*/
**audition** /*ba*/
**perception** [*da*]

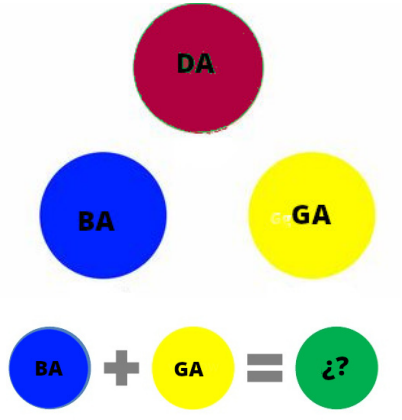


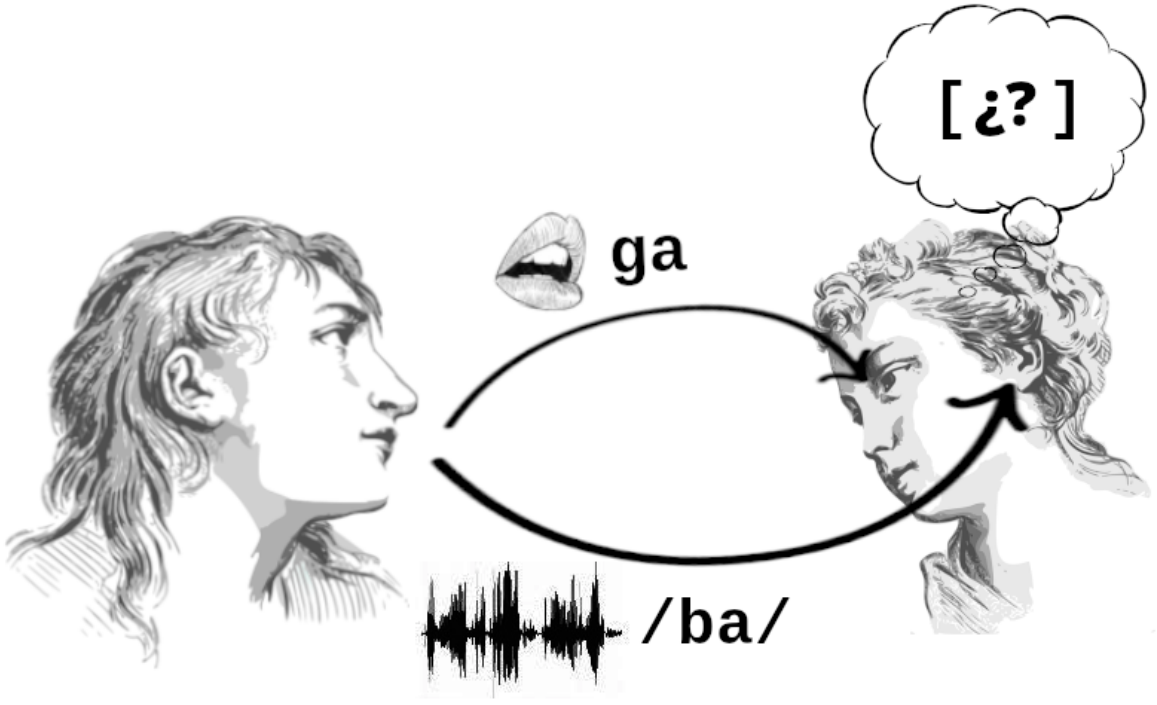




[Mcgurk and Macdonald, 1976, Van Engen et al., 2019, Mitchel et al., 2014]

# Modelling of the McGurk effect, a multisensory integration illusion

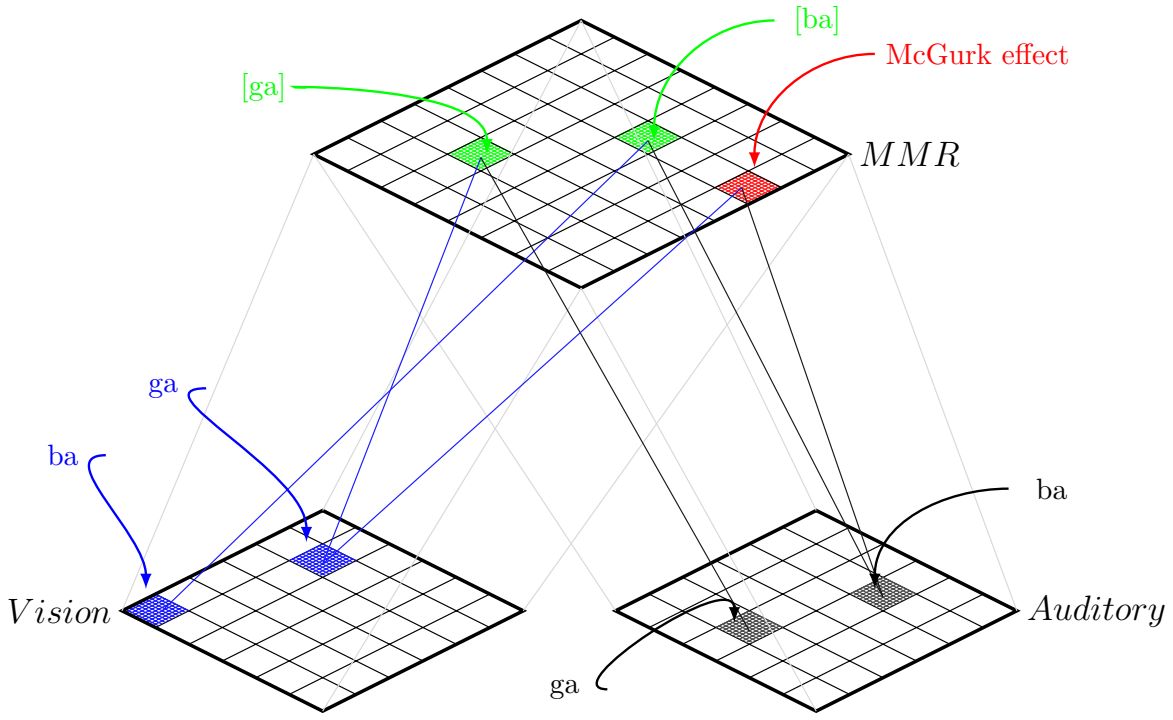
**visual** /*ga*/
**audition** /*ba*/
**perception** [*da*]




[Mcgurk and Macdonald, 1976, Van Engen et al., 2019, Mitchel et al., 2014]

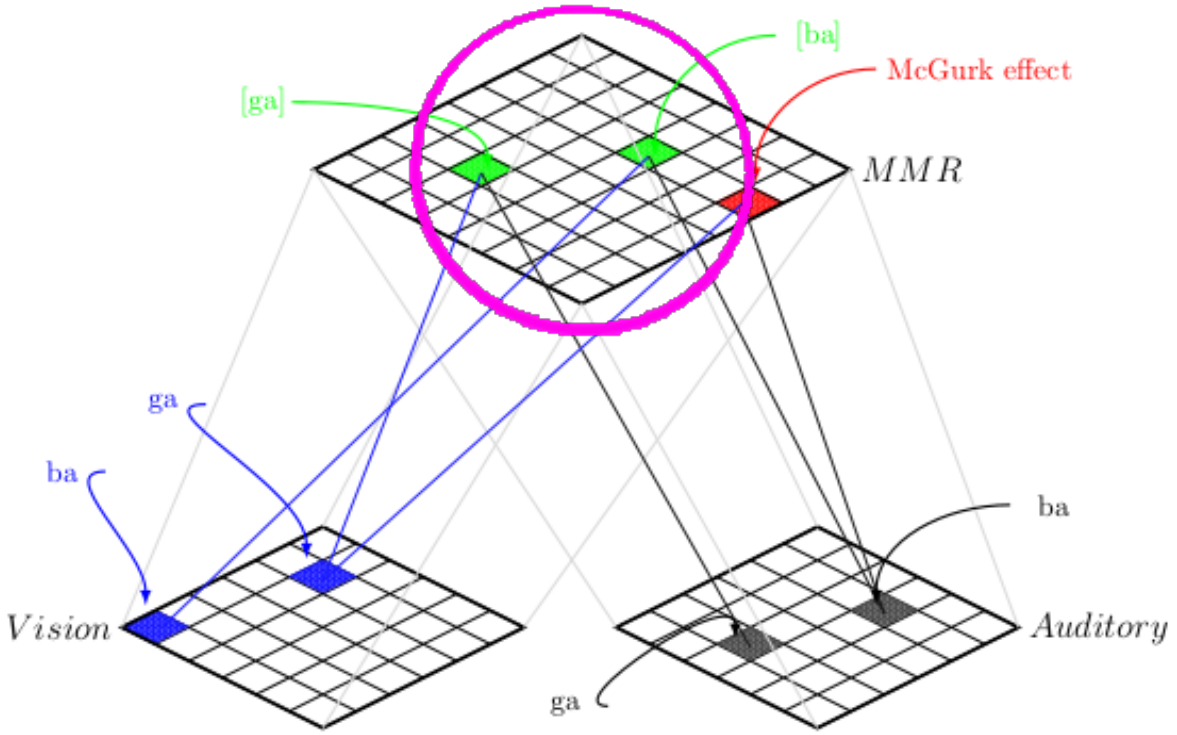
# Self-Organized Internal Model Architecture (SOIMA)




Note. Computer Achitecture. Multysensory integration happens in the Multimodal Representation Map (MMR).

[Escobar-Juárez et al., 2016, Morse and Cangelosi, 2017]

# Self-Organized Internal Model Architecture (SOIMA)




Note. Computer Achitecture. Multysensory integration happens in the Multimodal Representation Map (MMR).

[Escobar-Juárez et al., 2016, Morse and Cangelosi, 2017]

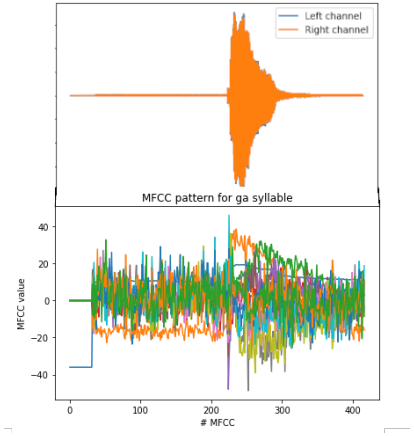
# Feature extraction/model training

## Vision

*Oriented Histograms of Regional Optic Flow*

## Audition

Mel-frequency cepstral coefficients




## Model validation

Psychopy -online-
Lipreading experiment

N= 36, n= 20


[Basu Mallick et al., 2015, Viola and Jones, 2001, Kazemi and Sullivan, 2014, Liu et al., 2016, Gold et al., 2011, Hoffman and Gelman, 2014]

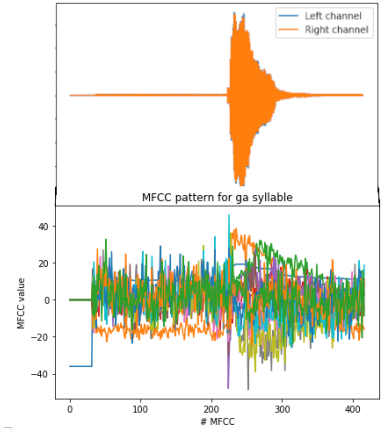
# Feature extraction/model training

## Vision

*Oriented Histograms of Regional Optic Flow*

## Audition

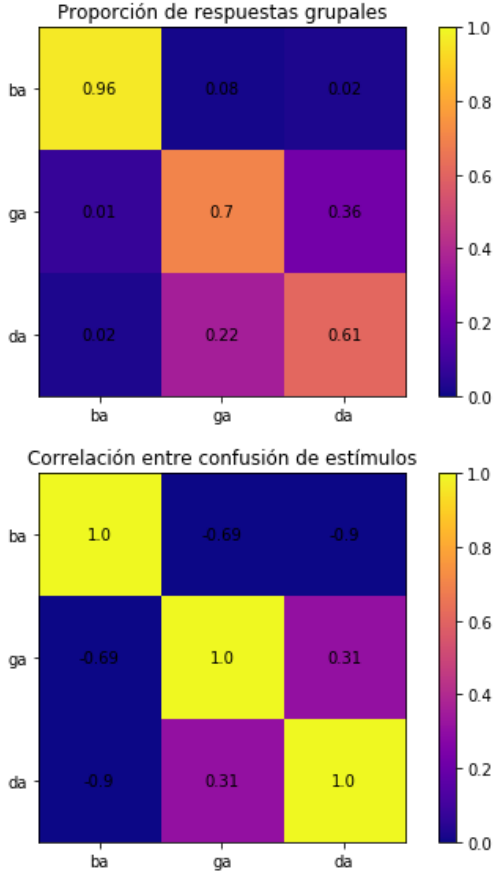
Mel-frequency cepstral coefficients




## Model validation

Psychopy -online-
Lipreading experiment

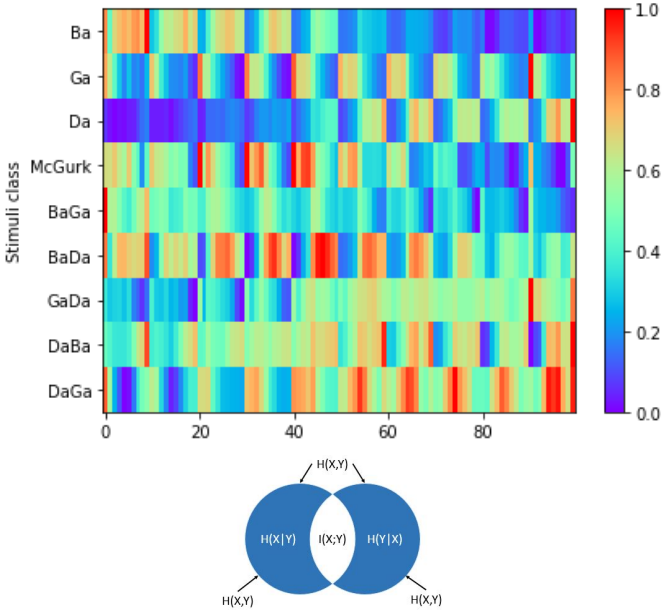
N= 36, n= 20



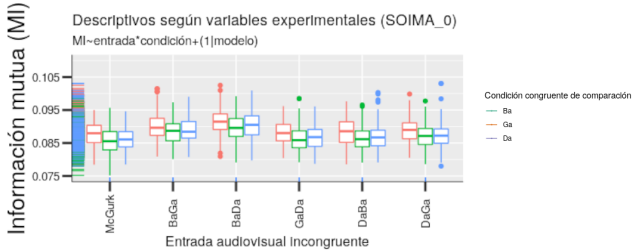

[Basu Mallick et al., 2015, Viola and Jones, 2001, Kazemi and Sullivan, 2014, Liu et al., 2016, Gold et al., 2011, Hoffman and Gelman, 2014]

# Results

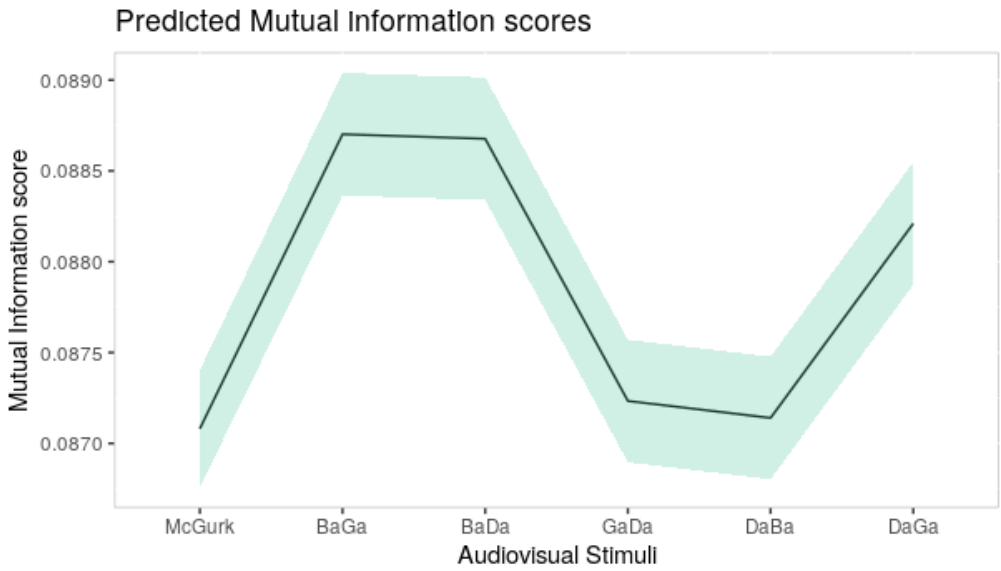
Congruent and Incongruent stimuli activation




Mutual Information: incongruent - congruent stimuli




Linear Mixed Effects analysis




# Future Research

Which is the relevance of multisensory integration in meaning and syntax emergence?

Which are the neural correlates of affordance based meaning? (i.e., tools Vs. food)

How do multisensory integration explain non-referential meaning?

How can we model language development/processing?

[Barsalou et al., 2003, Barsalou et al., 2018, Kuhnke et al., 2020, Twomey and Cangelosi, 2020]

# Future Research

Which is the relevance of multisensory integration in meaning and syntax emergence?

Which are the neural correlates of affordance based meaning? (i.e., tools Vs. food)

How do multisensory integration explain non-referential meaning?

How can we model language development/processing?

[Barsalou et al., 2003, Barsalou et al., 2018, Kuhnke et al., 2020, Twomey and Cangelosi, 2020]

# Future Research

Which is the relevance of multisensory integration in meaning and syntax emergence?

Which are the neural correlates of affordance based meaning? (i.e., tools Vs. food)

How do multisensory integration explain non-referential meaning?

How can we model language development/processing?

[Barsalou et al., 2003, Barsalou et al., 2018, Kuhnke et al., 2020, Twomey and Cangelosi, 2020]

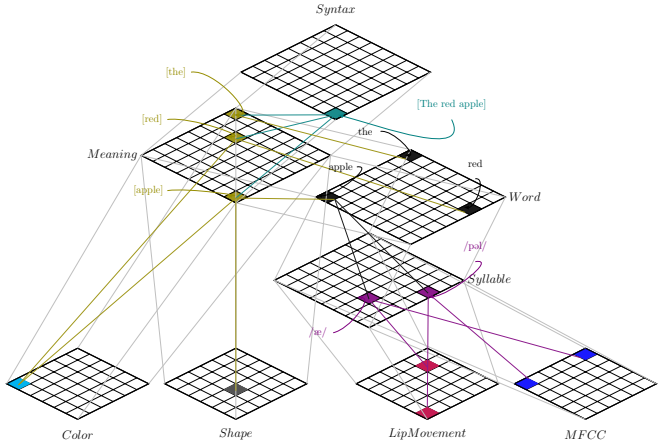
# Future Research

Which is the relevance of multisensory integration in meaning and syntax emergence?

Which are the neural correlates of affordance based meaning? (i.e., tools Vs. food)

How do multisensory integration explain non-referential meaning?

How can we model language development/processing?




[Barsalou et al., 2003, Barsalou et al., 2018, Kuhnke et al., 2020, Twomey and Cangelosi, 2020]

Laboratorio de
Robótica
Cognitiva

Centro de Investigación en Ciencias

CONACYT
Consejo Nacional de Ciencia y Tecnología

**Scholarship No. 1004634**

## Thank you for your attention!

*www.roboticacognitiva.mx*

*robotica.cognitiva@uaem.mx*

*Laboratorio de Robótica Cognitiva*

# SOIMA (entrenamiento SOMs)

Activación de los mapas modales en ecuación **??**

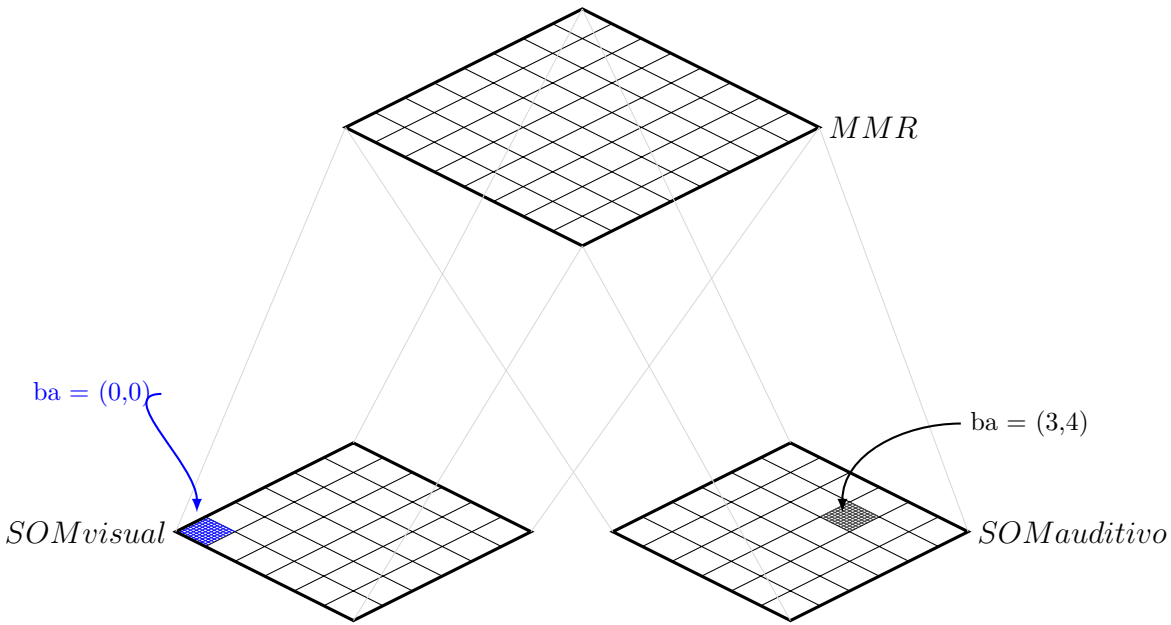
$$Act_j = \sqrt{\sum (v_i - w_j)^2} \quad (1)$$

Vecindario de activación en ecuación **??**

$$h_j = \lambda \sqrt{\sum (Dim_1, \ldots, Dim_n)^2} \quad (2)$$

Incremento pesos SOM en ecuación **??**

$$\Delta w_{ij} = (\alpha\lambda) h_j (v_i - w_j) \quad (3)$$



# SOIMA (entrenamiento SOMs)

Activación del MMR

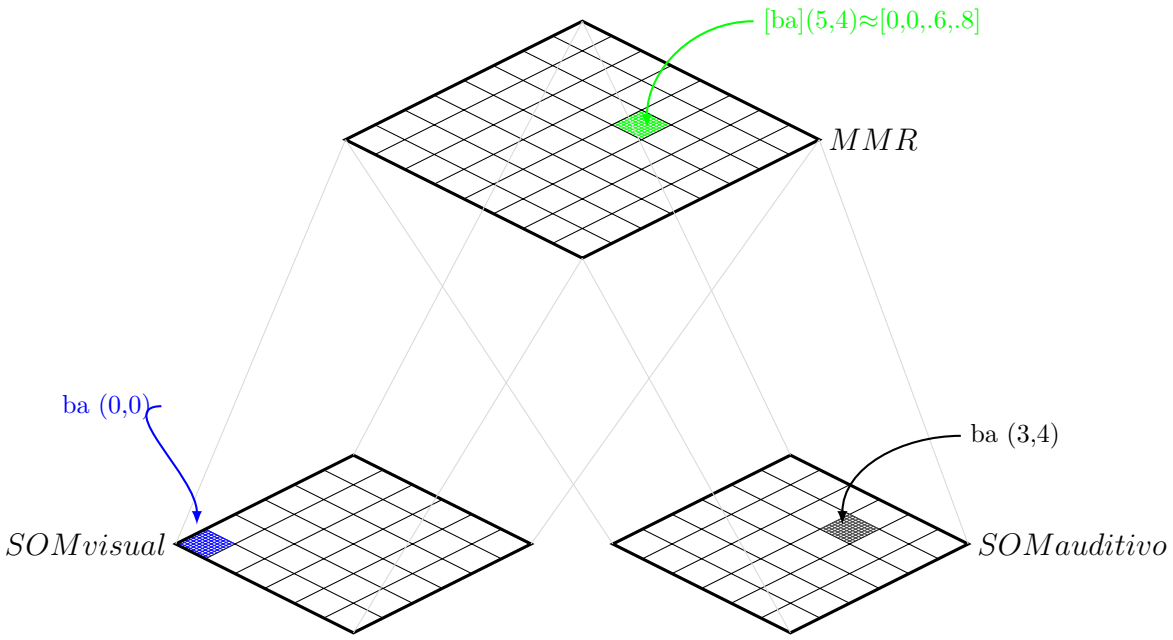
$$Act_j = \sqrt{\sum (v_i - w_j)^2} \quad (4)$$

Vecindario de activación del MMR

$$h_j = \lambda \sqrt{\sum (Dim_1, \ldots, Dim_n)^2} \quad (5)$$

Incremento pesos MMR

$$\Delta w_{ij} = (\alpha\lambda) h_j (v_i - w_j) \quad (6)$$




El proceso de entrenamiento es el mismo en todos los SOMs

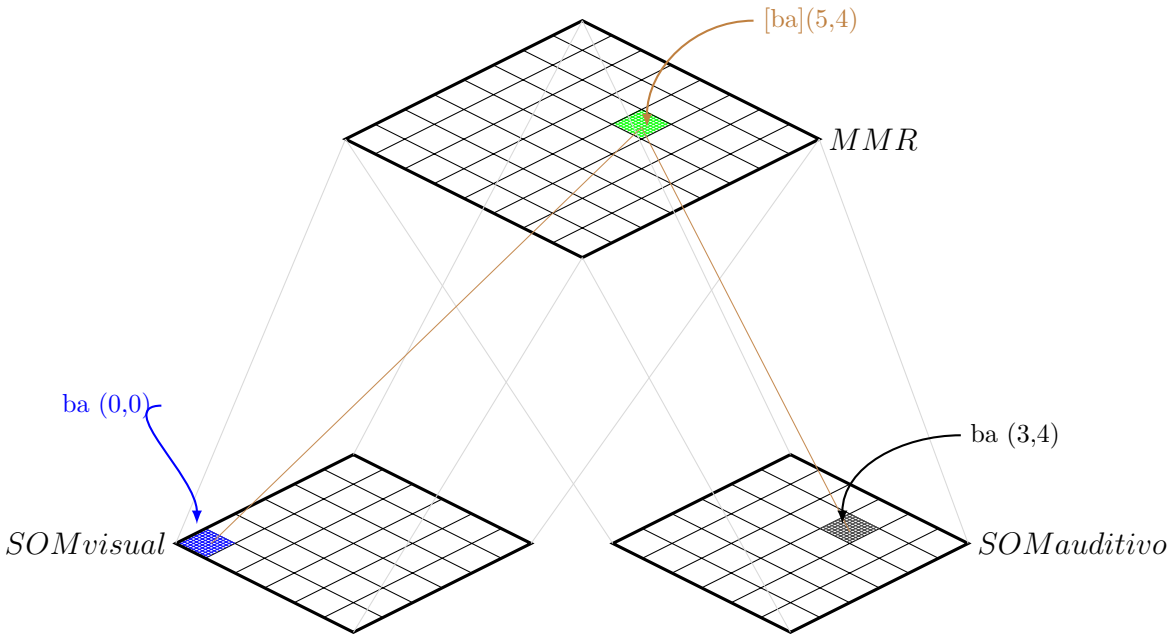
# SOIMA (aprendizaje multimodal)

Distancia entre entrada y BMU en ecuación **??**

$$D_{BMU} = \sqrt{\sum (e - w_{bmu})^2} \quad (7)$$

Incremento pesos asociación multimodal en ecuación **??**

$$\Delta w_{kl} = \alpha_h (D_k)(D_l) \quad (8)$$



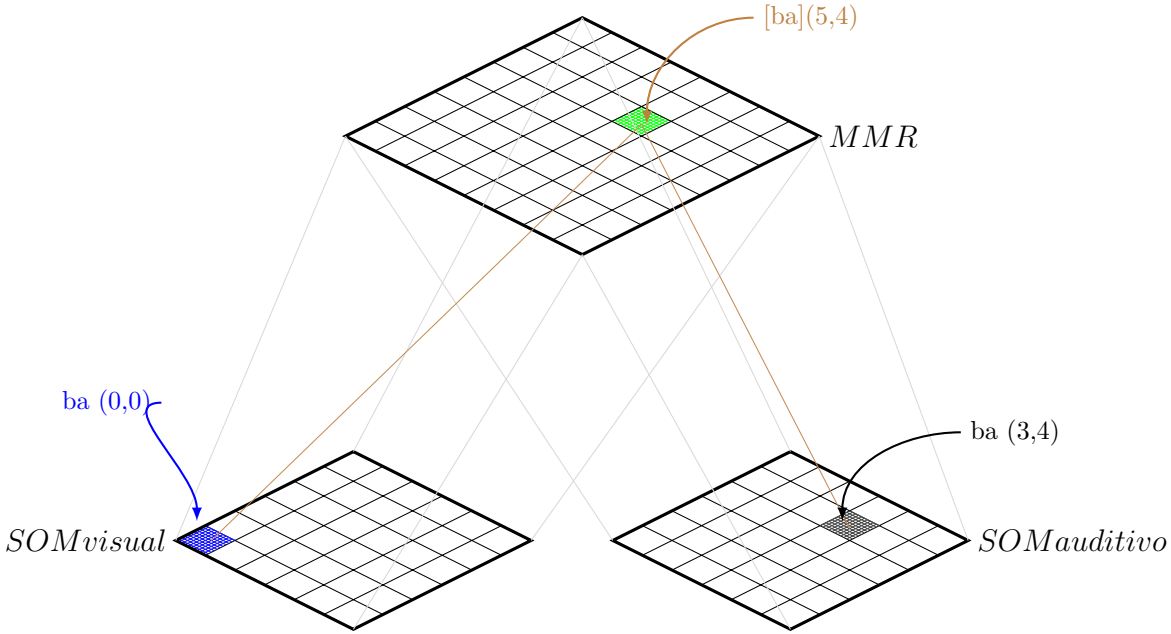
# SOIMA (procesamiento del error predictivo)

Predicción de activación multimodal en ecuación **??**

$$PredAct_j = (\zeta w_{j|BMU_A}) + (\zeta w_{j|BMU_V}) \quad (9)$$

Activación multimodal corregida en ecuación **??**

$$CorrAct_j = (1 - \eta Act_j) + (\eta PredAct_j) \quad (10)$$



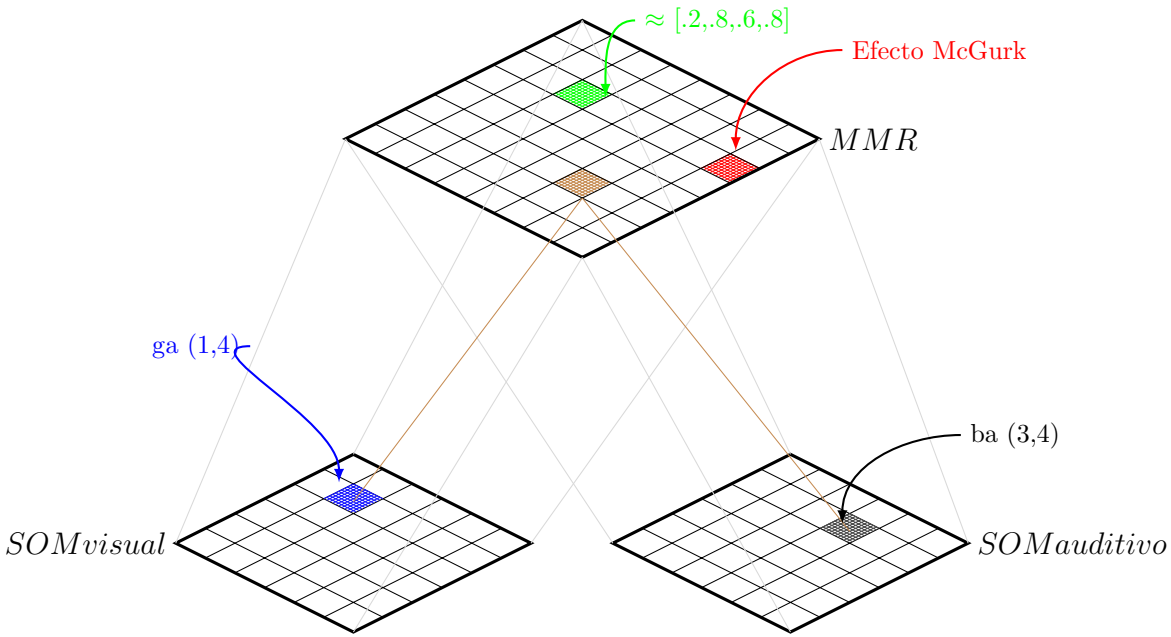
# SOIMA (procesamiento del error predictivo)

Predicción de activación multimodal en ecuación **??**

$$PredAct_j = (\zeta w_{j|BMU_A}) + (\zeta w_{j|BMU_V}) \quad (9)$$

Activación multimodal corregida en ecuación **??**

$$CorrAct_j = (1 - \eta Act_j) + (\eta PredAct_j) \quad (10)$$



# References

Barsalou, L. W., Dutriaux, L., and Scheepers, C. (2018). Moving beyond the distinction between concrete and abstract concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1752).

Barsalou, L. W., Simmons, W. K., Barbey, A. K., and Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7(2):84–91.

Basu Mallick, D., F. Magnotti, J., and S. Beauchamp, M. (2015). Variability and stability in the McGurk effect: contributions of participants, stimuli, time, and response type. *Psychonomic Bulletin  Review*, 22(5):1299–1307.

Escobar-Juárez, E., Schillaci, G., Hermosillo-Valadez, J., and Lara-Guzmán, B. (2016). A Self-Organized Internal Models Architecture for Coding Sensory–Motor Schemes. *Frontiers*

*in Robotics and AI*, 3.

Gold, B., Morgan, N., and Ellis, D. (2011). *Speech and Audio Signal Processing*. John Wiley  Sons, Inc., Hoboken, NJ, USA, second edi edition.

Hoffman, M. D. and Gelman, A. (2014). The No-U-Turn Sampler: Adaptively Setting Path Lengthsin Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15.

Kazemi, V. and Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1867–1874. IEEE.

Kuhnke, P., Kiefer, M., and Hartwigsen, G. (2020). Task-Dependent Recruitment of Modality-Specific and Multimodal Regions during Conceptual Processing. *Cerebral cortex (New York, N.Y. : 1991)*, 30(7):3938–3959.

Liu, X., Cheung, Y.-m., and Tang, Y. Y. (2016). Lip event detection using oriented histograms of regional optical flow and low rank affinity pursuit. *Computer Vision and Image Understanding*, 148:153–163.

Mcgurk, H. and Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588):746–748.

Mitchel, A. D., Christiansen, M. H., and Weiss, D. J. (2014). Multimodal integration in statistical learning: evidence from the McGurk illusion. *Frontiers in Psychology*, 5.

Morse, A. F. and Cangelosi, A. (2017). Why Are There Developmental Stages in Language Learning? A Developmental Robotics Model of Language Development. *Cognitive Science*, 41:32–51.

Twomey, K. E. and Cangelosi, A. (2020). Heads, shoulders, knees and toes: What developmental robotics can tell us about

language acquisition. In *Current Perspectives on Child Language Acquisition*, pages 39–64. John Benjamins.

Van Engen, K. J., Dey, A., Sommers, M., and Peelle, J. (2019). Audiovisual Speech perception: Moving beyond McGurk.

Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–511–I–518. IEEE Comput. Soc.