Bilingual representation and learning in a connectionist framework

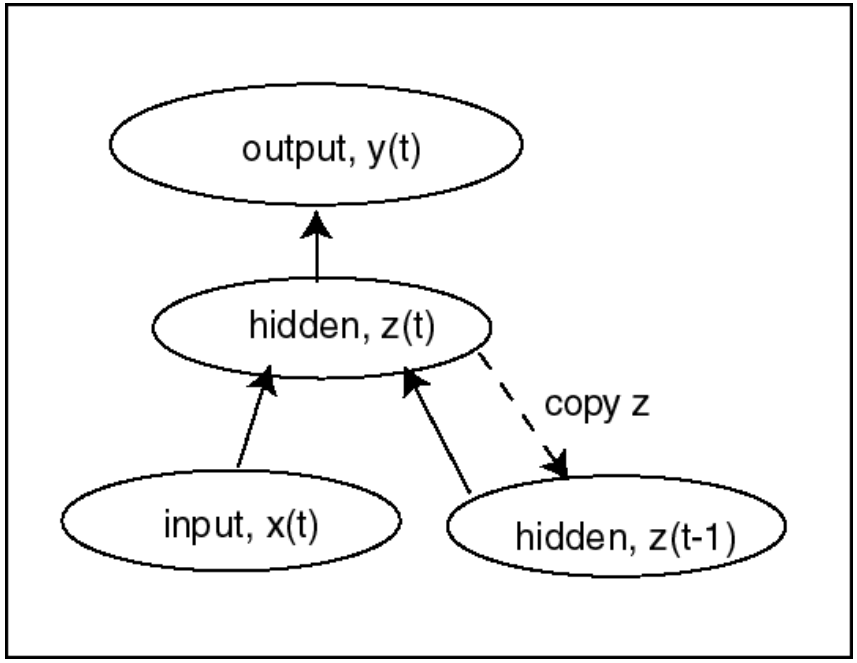
CS784: Language Acquisition

Arindam Biswas
Indradeep Saha

Introduction

- At least half of the world's population is bilingual
- Questions
  1. 2 separate lexicons, one for each language, or 1 large 'bilingual' lexicon?
  2. Are there separate conceptual and lexical levels?
     - conceptual level shared by both languages
     - lexical representations specific to each language
- Connectionist frameworks

Elman Network



- The Elman network commonly is a two-layer network with feedback from the first-layer output to the first layer input. This recurrent connection allows the Elman network to both detect and generate time-varying patterns.

Simulation Environment

- Two micro languages with 12 words each
  - Alpha:
    - Subject Nouns: BOY, GIRL, MAN, WOMAN
    - Verbs: LIFTS, TOUCHES, SEES, PUSHES
    - Object Nouns: TOY, BALL, BOOK, PEN
  - Beta: consisting of 1
    - Subject Nouns: GARÇON, FILLE, HOMME, FEMME
    - Verbs: SOULEVE, TOUCHE, VOIT, POUSSE
    - Object Nouns: JOUET, BALLON, LIVRE, STYLO
- The words “BOY” “GIRL” carry no semantic information. They can be anything.

Simulation Environment contd …

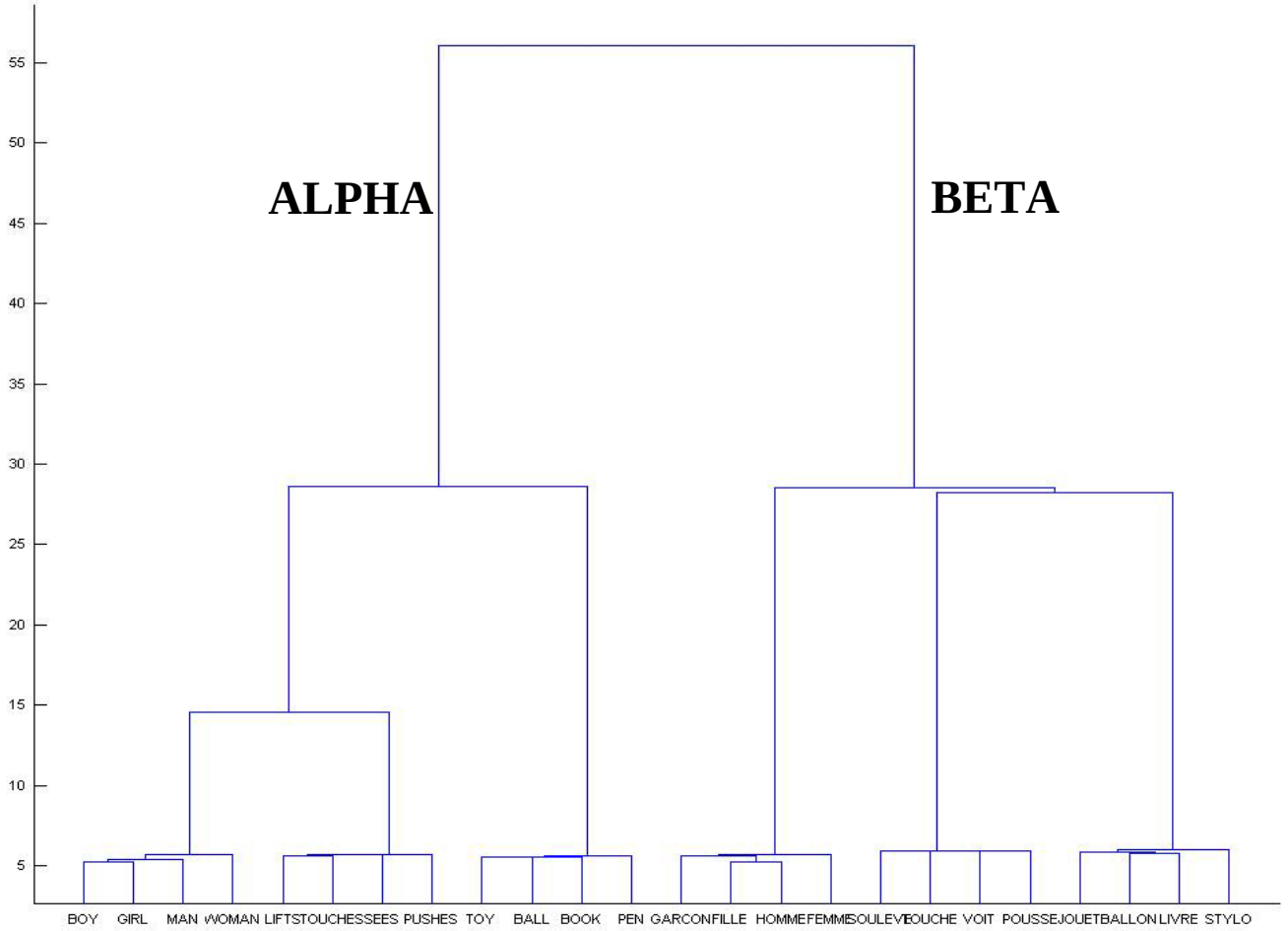
- Sentence from each language have the form Subject-Verb-Object (SVO)
- A language generator used to simulate the bilingual environment
  - Fixed probability of switching from one language to another(0.001)
  - No switching in the middle of a sentence

  E.g.
  - BOY LIFTS TOY MAN SEES PEN MAN TOUCHES BOOK GIRL PUSHES BALL WOMAN TOUCHES TOY BOY PUSHES BOOK …….

Methodology

- 24-32-24 Elman network with a bias node
- For each word of the sentence network's task is to predict the following word
- Words coded to binary form
  - BOY = 100000000000000000000000,
  - GIRL = 010000000000000000000000,
  - MAN = 001000000000000000000000,
- Iteration 90,000
- Agglomerative hierarchical cluster analysis using a Euclidean distance metric and Ward's method to determine linkage.

Language and grammatical clustering



Bilingual Simple Recurrent Network (BSRN)

The SRN network exhibits:

- Progressive development of hidden-unit representations that cluster according to grammatical forms (subject, verb, object) and languages, even though there are no explicit markers on input distinguishing the languages or their grammatical forms
- Inter-lingual interference effects
- Considerable resistance to lesioning
- Significant disruption of internal organization that can be produced, on rare occasions, by lesioning a very small number of nodes.

Initial models

- Kolers(1963) first to try to explain bilingualism
  - took information processing approach
- Common storage model: Interdependence Hypothesis
- Separate storage model: Independent Hypothesis

Initial models

- Independence Hypothesis
  - Balanced bilinguals should react as monolinguals in both languages
  - Exptl. Observations:
    - Response different in their 2 languages
    - Fail to transfer from one language to another
- Interdependence Hypothesis
  - Exptl. Observations:
    - Inter-lingual behavior is not different from intra-lingual behavior

DCT

- Dual coding theory (Paivio, 1971)
- 2 symbolic systems
  - Verbal
  - Non verbal - imagery
- Both representations interconnected (representation in one system can affect other) but capable of functioning independently(additive effect on free recall)
- Representational processing
- Referential processing
- Associative processing -connections between linguistic units and between images in each system.

Bilingual Dual coding Theory

- Paivio & Desrochers (1980)
- 2 verbal representational systems, one representation in the imagery system.
- 3 systems autonomous but interconnected at the referential level
- Mix of both independence and interdependence hypotheses

Exptl. Verification of BDCT

- French-English balanced bilinguals
- Shown pictures, French words and English words
- Name pictures
- Translate words
- Just copy down words
- Result: 3:2:1 (47%: 31%: 18%)

Exptl. Verification of BDCT

- Sung & Padilla (1991)
- Unbalanced Korean-English(Korean dominant)
- Result: 2:2:1 (43%: 44%: 22.5%)
- Why??
- Another experiment among statistically balanced Korean-English and Spanish-English bilingual groups by Sung & Padilla.
- Inference: Influence of orthographies not the reason.

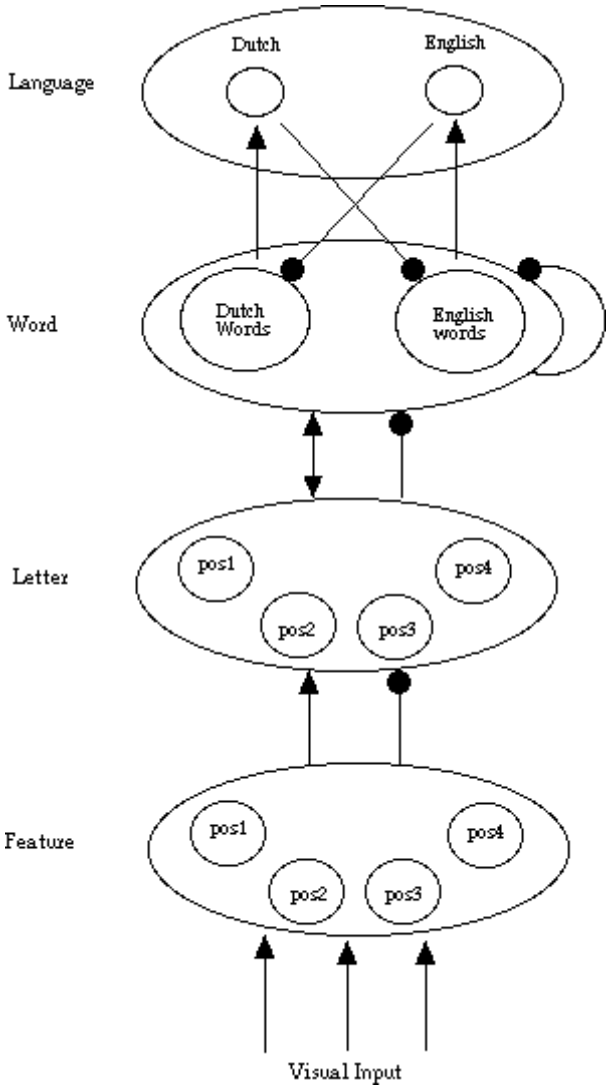
Representation

- Homographs
- Non-cognate
  - MAIN and FIN mean hand and end in French
- Cognate
  - TRAIN means the same thing in French
- Non cognate homographs have a different meaning, hence a different frequency of occurrence
- Higher frequency word recognized more in the corresponding language context.
- Since they show the same frequency response as Singles suggests that the lexical representations are therefore independent.
- Presentation of a non-cognate homograph in one language context does not facilitate later recognition of the word form in the other language context
- Conclusion: bilingual has independent representations for a word and its translation equivalent at the lexical level, but a common representation at the semantic level

Caveats

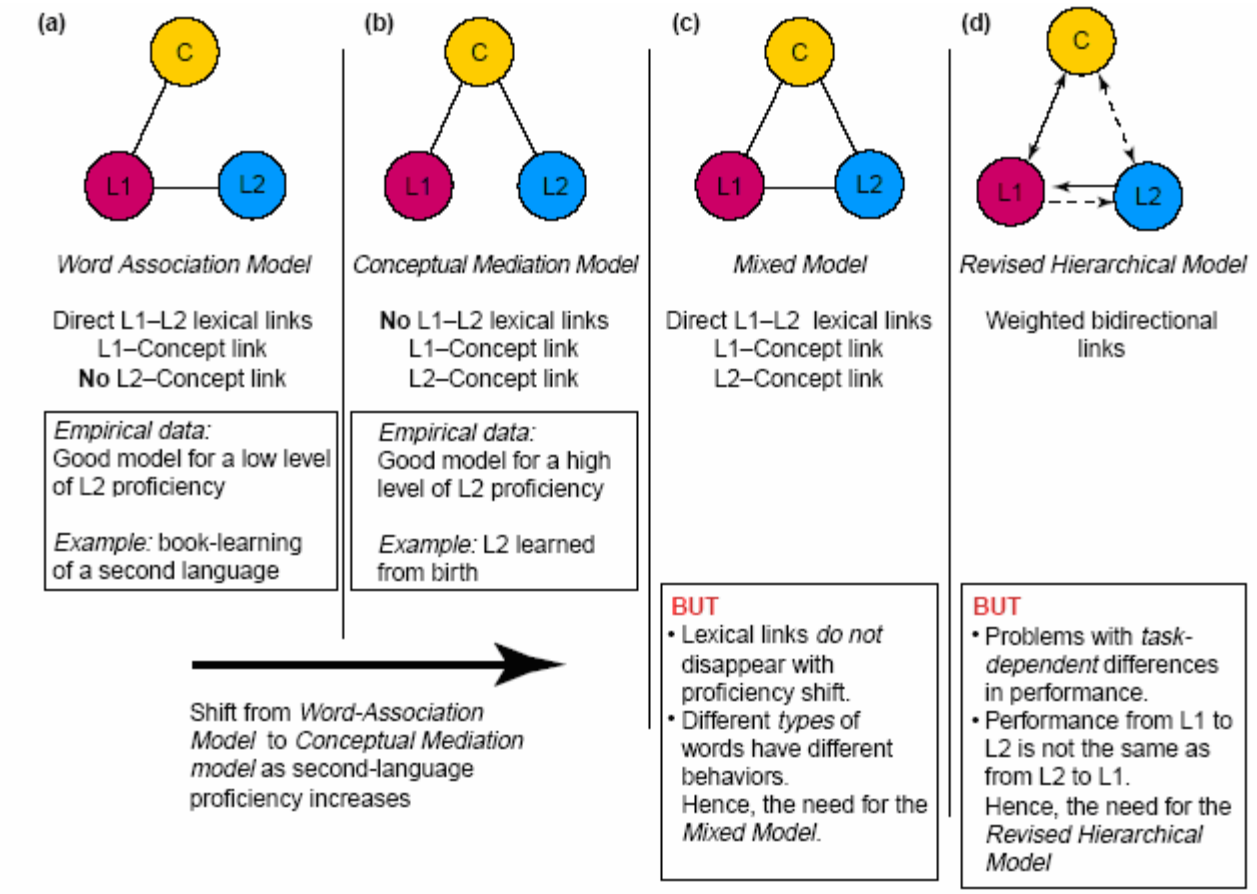
- Under some circumstances, between language similarity effects are found
- Non-cognate homographs were recognised more slowly than matched cognate homographs
- Cognate homographs in a bilingual's weaker language were recognised more quickly than Singles of matched frequency, as if the stronger language were helping the weaker language on words they had in common
- Bilingual subjects recognised words with orthographic patterns specific to that language more quickly than words with orthographic patterns common to both languages

Bilingual Interactive Activation



- Affects particular features at each letter position
- Activated letters excite words of both languages at its position while all other words are inhibited
- All words inhibit each other irrespective to the language they belong to
- Activated words of a language sends activation to its language node and these nodes send inhibitory feedback to all word nodes in the other lang.

Hierarchical Models



Hierarchical Models

Arguments supporting two lexicons:

- Absence of **Priming effects** between orthographically dissimilar translation-equivalents
- **Release from proactive interference:** A release from proactive interference is observed by changing the language between two lists to be memorized.

Counter Arguments:

- RPI is observed in monolinguals also

Conclusion

- Most of the theories favor the need of a conceptual and lexical separation.
- Almost all the researchers agrees on the fact of having a single conceptual memory.
- It seems from the different experiment results there is no two different lexicons for the two languages in the bilinguals rather there is a common lexicon with two different sub lexica for two languages.
- The early results of BSRN seems to be very much promising as it is able to explain to some extent different cases of bilingual aphasia.

Bibliography

- Robert M. French. *A Simple Recurrent Network Model of Bilingual Memory*
- Jacquet, M. & French, R. M. (2002). *The BIA++: Extending the BIA+ to a dynamical distributed connectionist framework. Bilingualism, 5(3), 202-205.*
- Michael S.C. & Thomas Kim Plunkett. *Representing the bilingual's two lexicons.*
- Hyekyung Sung & Amado M. Padilla. *Bilingual Memory*: *Representation in Korean-English and Spanish-English Bilinguals*
- Robert M. French and Maud Jacquet. TRENDS in Cognitive Sciences Vol.8 No.2, February 2004. *Understanding bilingual memory: models and data*
- David W. Green. *Regulating the bilingual lexico-semantic system*
- Jing Liang. *Bilingual Language Learning in Connectionist Networks*
- Michael S. C. Thomas. *Bilingualism and the Single route / Dual route debate*
- Michael S. C. Thomas. *Distributed representations and the bilingual lexicon: One store or two?*
- Walter J. B. van Heuven and Ton Dijkstra, *Orthographic Neighborhood Effects in Bilingual Word Recognition*
- *http://www.semioticon.com/virtuals/talks/tomioka.htm*