Principles for a Quantitative Speech Learning Model
Geoffrey Stewart Morrison, University of Alberta

L1 learning

- Infants
  - begin tabula rasa
  - exposed to speech sounds produced by adults
  - sum of multidimensional distributions of acoustic properties
  - posit categories based on sum of distributions
    - valleys between peaks: boundaries
  - associate categories with differences in meaning
  - formation proceeds category labelling
  - production categories based on perception categories

L2 learning

- L2 learners
  - begin with an L1 perception system
  - exposed to speech sounds produced by L2 speakers
  - sum of multidimensional distributions of acoustic properties
  - posit categories based on sum of distributions
    - same mechanisms as for L1 learning
    - no dimension-specific stipulations
      - cf. Bohm’s Desensitization Hypothesis: desensitisation to spectral differences
      - cf. Escudero & Boersma (2004): no categories on the duration dimension
  - slow track
    - number of instances of L2 sounds heard is small compared to past exposure to L1 sounds
    - perceptual inertia
    - distributions of L2 sounds gradually added to distributions of L1 sounds
    - phenomena specified qualitatively in Flege’s SLM emerge from first principles of distribution-based learning
      - “new” L2 categories have properties which match those of native speakers of the L2
        - e.g., prototypical values for new /l/ category
      - “similar” diaphone categories develop properties which are a mixture of the L1 and L2 category properties
        - e.g., increase in duration range for /e/ diaphone
  - fast track
    - instances of L2 sounds are assimilated to L1 sounds
    - sounds assimilated to an L1 sound but further from the L1 prototype are more noticeable as deviant members of the L1 category (Best’s PAM)
    - multidimensional category-goodness assimilation
      - example:
        - multidimensional category-goodness assimilation of English /i/ and /e/ to Spanish /i/:
          - Spanish /i/-/e/ boundary is upper F1 boundary for assimilation of higher-F1 instances of English vowels to Spanish /i/:
          - no Spanish vowel category, of similar or lower F1, is longer than Spanish /i/:
            - therefore no upper boundary on assimilation of longer-duration instances of English vowels to Spanish /i/:
          - instances of long English /i/ are farther from the Spanish /i/-prototype than any instances of English /i/-/l/-English /i/-assimilated to Spanish /i/:
          - instances of long English /i/ are very far from the Spanish /i/-prototype and are therefore immediately noticeable
          - listeners notice the duration difference and (consciously) use Spanish /i/ and /l/ to differentiate English /i/ and /l/ in perception and production

Future Research

- additional data collection
  - development of computational implementation of model
  - inclusion of multiple vowels
    - Spanish /i/-/e/-/o/:
    - English /i/-/l/-/o/:
    - inclusion of multiple acoustic dimensions:
      - duration
      - F1
      - F2
      - vowel inherent spectral change

Research supported by:
Natural Sciences and Humanities Research Council of Canada
Conseil de recherches en sciences humaines du Canada