Sensori-motor constraints and the organization of sound patterns

Lucie Ménard
Laboratoire de phonétique
Université du Québec à Montréal
Center for Research on Language, Mind, and Brain
Institut des sciences cognitives
www.phonetique.uqam.ca
and:

- Louis-Jean Boë (GIPSA, Grenoble)
- Jean-Luc Schwartz (GIPSA, Grenoble)
- Pierre Poirier (ISC, UQAM)
- Collaborators from the Laboratoire de phonétique (UQAM), in alphabetical order:
  Jérôme Aubin, Annie Brasseur, Serge Drouin, Caroline Émond, Marilyn Giroux, Annie Leclerc, Marilène C. Rousseau, Mélanie Thibeault, Corinne Toupin
- the subjects.....
Sensori-motor constraints: related to the speaker’s production system and listener’s perception system.

Languages usually do not use all possible sounds that can be produced and perceived by humans, but rather use sounds related to sensori-motor constraints.
Understanding those physical constraints can shed light on:

- Sound changes in diachrony
- Infant’s speech development
The issue of the role played by sensori-motor constraints in speech became more important with the emergence of “embodied cognition”.

It has been claimed that the abstract representations (=mental representations, for some researchers) should be the focus of linguistic studies.

But since the brain interacts with the physical world (and sometimes develops with sensori-motor experience), representations and their implementation in the body are both related.
The notion of “articulatory ease” or “naturalness” has been taken into account in generative grammar (phonology) through the notion of “markedness”.

After SPE, an unmarked feature was one considered more “natural” phonetically, generally favored in languages of the world.

Other phonological models (feature geometry, natural phonology,...) have integrated mechanisms to take into account biomechanic links between features.
Two central questions:

- Could knowledge of the articulatory processes involved in speech production and vocal tract anatomy explain sound patterns in languages of the world and in speech development?

- Could knowledge of auditory mechanisms involved in human speech perception explain sound patterns in languages of the world and in speech development?
And related questions:

- Does the evolution of sound categories require the evolution of abstract representations...

- ...or could the evolution of sound categories result from the evolution of the vocal tract and perceptual system?

- Could vocal tract constraints influence the nature of abstract representations of sound?
Vowels
The UPSID database

- UPSID (UCLA Phonological Segment Inventory Database) surveys sound inventories of 317 languages in the world (updated to more than 400). (Maddieson, 1984; Maddieson, 1991)

The UPSID database

- In this database, vowels and consonants are represented by
  
i) their symbol (quality, identity)

ii) their localization in a « system », representing articulatory and acoustic dimensions

Schwartz et al. (1997a, 236)
Sound systems

- This form corresponds to articulation and acoustics.

- The vocal tract offers more possibilities than those used by languages.
Number of vowels per language

Languages tend to reduce the number of vowels in their inventories.

Vallée (1994)
Most frequent vowels

- 90% of languages have /i u a/ in their inventories
- The most frequent vowels are peripheral vowels, high vowels and internal vowels

Vallée (1994)
### System organization

- When the number of vowels is greater than 3, a constraint of structural organization is respected.

---

#### Decreasing order of frequency

<table>
<thead>
<tr>
<th>Number of Vowels</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>88</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
</tr>
</tbody>
</table>

---

Vallée (1994)
The role of production constraints

- The most frequent vowels are those representing the greatest contrast.
**The role of production constraints**

- The tendency to align peripheral vowels in the system along “straight lines” would come from a tendency to use maximal available controls (Schwartz *et al.*, 2007; Ménard *et al.*, 2008)
The search for auditory constraints

- Already present in Troubetzkoy

- In 1972, two milestone papers were published:


- Those papers postulate that universal trends in languages result from sensori-motor constraints
The search for auditory constraints

- Liljencrants & Lindblom (1972):
  Dispersion theory (DT):
  Sound systems are composed of units organized in order to respect the “maximal perceptual contrast” constraint (later, “sufficient contrast”). This criteria explains why peripheral forms like /i u a/ and /i e a o u/ are so frequent.
The search for auditory constraints

- Within the DT, criteria are global or relational

- Another theory, Stevens (1972) and the Quantal Theory (QT):
  There are regions in the vocal tract for which articulatory-acoustic relationships are quantal, in the sense that a large articulatory movement is related to a small acoustic change, and, conversely, a small articulatory displacement yields a rapid change from one acoustic state to the other. Those quantal relationships are crucial in shaping languages sound inventories.
The search for auditory constraints

- Stevens (1972) and the Quantal Theory (QT):
  - Basis for categorical perception
The search for auditory constraints

➢ Zones of articulatory-acoustic stability would correspond to sound categories. This would allow a speaker to produce a sound with less precision, while the acoustic result remains unchanged

➢ Principle of economy of effort

➢ Those quantal relationships are also found between:
  - neural commands and muscle contractions
  - acoustic parameters and perceptual products
The search for auditory constraints

➢ But, both the DT and the QT have weaknesses and can not account for all universal trends of the world’s languages.

➢ An improved version of the DT, the Dispersion-Focalization Theory (DFT) has been proposed (Schwartz et al., 1997. The Dispersion-Focalization Theory of vowel systems, *Journal of Phonetics*, 25, 255-286)

➢ In the DFT, a more local auditory constraint is proposed: languages would primarily select vowels for which adjacent spectral peaks (formants) are close to each other. This is called “focalization”.

- Background
- Questions
- **Vowels**
  - Trends
  - Sensori-motor
- Consonants
  - Trends
  - Sensori-motor
- Syllables
  - Trends
  - Sensori-motor
- Conclusion
The search for auditory constraints

➢ Focalization:

Preferred vowels are those easier to perceive by the human ear, more auditorily salient (Schwartz et al., 2005)
Consonants
Consonants

Most favored places of articulation for consonants in UPSID:

- Alveolar: 15.3%
- Labial: 14.3%
- Velar: 12.6%

As for vowels, favored places of articulation are those related to the greatest contrast.
Most favored consonants

- In UPSID, stop consonants (for which the vocal tract is completely closed at one point) are more frequent than fricative consonants (for which the vocal tract is not completely closed, so that noise is generated.) Ex.: /p t k b d g/ are more frequent than /f s v z/

- Voiced consonants (those requiring vocal fold vibration) are disfavored compared to unvoiced consonants (those for which vocal folds are not vibrating). Ex.: /b d g v z/ are less frequent than /p t k f s/
The role of aerodynamic constraints

- Producing a fricative requires much more articulatory precision than producing a stop to generate air flow turbulence and noise.

Producing a stop /t/       Producing a fricative /s/
The role of aerodynamic constraints

- Maintaining voicing requires fine motor control and is more difficult:

Producing /k/  Producing /g/
The role of aerodynamic constraints

- Maintaining voicing in a fricative /z/ is even more complex:

Producing a voiced fricative

- Latest sounds to be mastered by children
Production constraints applied to sound change in diachrony

- Anatomical and aerodynamic properties of the vocal tract also explain sound changes

- The complex aerodynamic constraints involved in voicing was likely responsible for the loss of /g/ or its change into /k/ from Proto-Bantu to Duala and Ngom (Ohala, 1981: 199)
Perceptual constraints applied to sound change in diachrony

- It is known that vowels following unvoiced consonants (ex.: /p t k f s/) are higher in pitch than vowels following voiced consonants (ex.: /b d g v z/)

- Listeners have been shown to use this cue in the vowel to identify the quality of the preceding consonant (experimental evidence)

- This phenomenon played a role in the emergence of tonal languages like Chinese (Ohala, 1981; Hombert et al., 1979)
**Summary 1 (vowels and consonants)**

- Languages do not use all possible sounds in their inventories, some are preferred

- Recurrent sound patterns found in the world’s languages can be explained by physical constraints related to the speaker’s vocal tract shape and motor control and to the listener’s perceptual mechanisms

- Those constraints can be verified experimentally in synchrony

- Those constraints likely played a role in initiating sound changes diachronically
Syllables
Number of syllables per word

- Another version of the language database, ULSID (UCLA Lexical and Syllabic Inventory Database, Ian Maddieson) surveys 32 languages (more than 150000 syllable entries) (here, 14 languages)

- Languages tend to limit the number of syllables per word (Rousset, 2008: 89):
Favored syllable structure

- Languages only use a subset of all possible syllable structures (Rousset, 2008: 111):
  C = consonant; V = vowel

Proportion of syllable structures in ULSID

- Consonant clusters are more frequent in onsets (like « true ») than in codas (like « farm »)
The labial-coronal effect

- One universal trend observed in the database is the so called “Labial-coronal effect”: Languages prefer sequences of syllables of the form CVCV or CVC in which the first consonant is a labial, and the second consonant is a coronal, instead of the reverse pattern (2.5 times more frequent)

Ex.: /pati/ is more frequent than /tapi/
    /pat/ is more frequent than /tap/

- This LC effect would be present in the children’s first words inventory (MacNeilage and Davis, 2000; MacNeilage et al., 1999)
The role played by production constraints
- The predominance of the CV syllable can be explained by constraints related to motor control

- According to the frame-then-content Theory (MacNeilage, 1990; MacNeilage and Davis, 1998), the open-close oscillatory cycle of the mandible is the basic pattern of speech
The role played by production constraints

- The mandible movement is the first one to be acquired by babies (canonical babbling, around 7 months)

- The most frequent vowels and consonants produced at that stage: bilabial stops and open-mid vowels (pure frames) (Ex.: « baba », « meme »)

Kent, 1997: 145
The role played by production constraints

- The jaw movement is the first one to be acquired by babies (canonical babbling, around 7 months)

AVI examples

8 m.-o.  12 m.-o.
The role played by production constraints

- A property of the jaw cycle (Redford, 1999): the opening phase is longer than the closing phase

- This pattern would explain the recurrence of the CV syllable, the preference for consonant clusters in onsets rather than codas
The role played by production constraints

- Principles of speech motor control can explain the labial-coronal effect (Sato et al., 2007; Rochet-Capellan and Schwartz, 2007)

- /\p/ and /\t/ are realized by different articulators (lips and tongue)
The role played by production constraints

- Experimental evidence: Rochet-Capellan and Schwartz (2007)
  Method: speakers had to produce « pata » and « tapa » at an increased speech rate.
  The majority of trials yielded transformations of sequences to « pta » (Labial-Coronal)

- In a /pta/ sequence, the tongue-palate constriction associated to /t/ can be maximally anticipated into the /p/, while this is not possible in /tpa/. The labial-coronal sequence thus represents an optimally phased sequence
The role played by production constraints

- It has been proven that the “labial-coronal” structure is indeed more “natural” for listeners to perceive (Sato et al., 2007)

- Method: a listener hears sequences of /sep/, at a high speech rate
  - at the end of the trial, speakers report having switched to /pse/, for which lips-tongue-jaw phasing is maximal

- This pattern is also found in the babies’ first word inventories (MacNeilage et al., 1999)

- Similar motor constraints to those found in finger tapping (Kelso et al., 1995)
Summary 2 (syllables)

- Languages combine sounds in a way that is optimal in articulatory terms and perceptual saliency

- Similar constraints related to motor control and perceptual ease can explain preferred syllabic patterns in the world’s languages

- The alternating open-close jaw cycle, presumably at the origin of speech, is a key component of articulatory and perceptual organization
Universal tendencies of languages can be explained in light of the speakers’ and listeners’ sensori-motor constraints.

Motor control properties and auditory properties shape sound systems.

Substance is now seen to be related to form, in that forms emerge from the substance (perceptuo-motor processes), as in the Perception for Action Control Theory (PACT) (Schwartz et al., 2007).
Other issues not addressed in this talk:

- social pressure

- other changes coming from external factors

- concepts of auto-organization (Davis and MacNeilage, 2000)

- computational modelling of the emergence of speech (Oudeyer, 2006, 2005; de Boer, 2001; Steels, 2003)
-the multisensory nature of the speech perception mechanisms

- Complementarity of visual and auditory cues
  (Robert-Ribès et al., 1998)
  
  Visual --------------------------→ Auditory
  rounding, height, place

- Blind speakers show (Lewis, 1975; Ménard et al., 2009):
  - enhanced auditory abilities
  - no imitation of labial movements in pre-babbling and less differentiated labial gestures
  - less labial dynamics in adults

- recurrence of /m/-/n/ contrast in languages
Perceptual constraints should be studied in their multimodal nature

Experimental studies of speech sound can shed light on the emergence of sound representations in languages
Thanks!