# The development of segment inventories

Claartje Levelt & Marc van Oostendorp
Leiden University (LUCL) & Meertens Instituut, Amsterdam

TIN-dag, Feb. 3, 2007

## Summary

In this talk, we argue in favour of a traditional, Jakobsonian view of segment acquisition (acquisition of features) based on data from the CLPF database and using a constraint-based framework.

## 1 The issue

(1) a. { p, t, k } →
b. { p, t, k, b, d, g } →
c. { p, t, k, b, d, g, f, s, x, v, z, ɣ }

(2) a. Feature-based: children learn feature by feature; frequent and infrequent sounds are acquired around the same time if they are in the same natural class
b. Segment-based: children learn more frequent sounds first; natural class behaviour is unexpected or epiphenomenal
c. Word-based: children initially use all kinds of segments, provided they occur in frequent words; sounds spread through the lexicon

(3) Restrictions

a. We did not (yet) consider the target sounds
b. We disregarded the glides /j/ and /w/, since it is unclear whether to regard them as part of the consonantal system
c. Similarly, we disregarded /ʃ/, /ʔ/ and /h/, of which the place in the segment inventory is unclear
d. We ordered the remaining segments on Guttman scales

(4) The data: Child 2

| Ch | Dh | b | d | t | h | p | m | s | z | n | f | w | k | X | l | x | R | j | | S | v |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 529 | 3 | 7 | 5 | 4 | 2 | | | | | | | | | | | | | | | |
| 2 | 540 | 3 | 6 | 7 | 2 | | 2 | | | | | | | | | | | | | | |
| 2 | 554 | 4 | 8 | 4 | 2 | | | | 5 | | | | | | | | | | | | |
| 2 | 567 | 5 | 15 | 7 | | 2 | | 3 | | | | | | | | | | 2 | | | |
| 2 | 588 | 10 | 15 | 13 | | 3 | 5 | 3 | 5 | | | | | 2 | | | | | | | |
| 2 | 602 | 5 | | 9 | 2 | 12 | 2 | 4 | | | | | | | | | | | | | |
| 2 | 615 | 5 | 5 | 12 | | 23 | 2 | 4 | | 4 | 6 | 3 | | | | | | | | | |
| 2 | 629 | 19 | 10 | 11 | | 9 | | 3 | | 2 | 5 | 2 | | | | | | | | | |
| 2 | 643 | 65 | 19 | 94 | 26 | 62 | 6 | 27 | 4 | 18 | 66 | 6 | 4 | 3 | | | 2 | | | | 2 |
| 2 | 715 | 14 | 15 | 28 | 21 | 35 | 13 | 23 | 3 | 9 | 22 | 7 | 3 | | | | | | | | |
| 2 | 766 | 34 | 16 | 32 | 10 | 32 | 12 | 12 | | 10 | 20 | 8 | | | 2 | | | | | | |
| 2 | 785 | 6 | 8 | 26 | 24 | 8 | 8 | 12 | | 10 | 22 | 14 | 2 | | | | | | | 4 | |
| 2 | 817 | 28 | 10 | 12 | 12 | 8 | 16 | 20 | | 4 | 12 | 8 | 18 | 4 | 4 | 2 | 2 | | | | |
| 2 | 830 | 84 | 30 | 40 | 52 | 24 | 22 | 83 | 2 | 18 | 34 | 32 | 52 | 18 | 34 | | | 8 | | 2 | 2 |

| Ch | Dh | s | t | p | m | n | k | r | R | X | f | l | N | S |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 529 | 6 | 1 | 3 | | | | | | | | | | |
| 2 | 540 | 3 | 1 | 4 | 1 | 2 | | | | | | | 1 | |
| 2 | 554 | 6 | 5 | 2 | 2 | 4 | 1 | | | | | | | |
| 2 | 567 | 7 | 6 | | 2 | 7 | | | | | | | | |
| 2 | 588 | 10 | 4 | 4 | 5 | 9 | | | | | | | | |
| 2 | 602 | 8 | 6 | 1 | 1 | 4 | | | | | | 1 | | |
| 2 | 615 | 11 | 17 | 1 | 2 | 7 | | | | | | | | |
| 2 | 629 | 6 | 16 | 2 | 1 | 11 | | | | | 1 | | | |
| 2 | 643 | 48 | 55 | 16 | 9 | 62 | 20 | 11 | 13 | 8 | 4 | 2 | 1 | |
| 2 | 715 | 19 | 19 | 4 | 6 | 20 | 12 | 9 | 7 | 4 | | | 1 | |
| 2 | 766 | 28 | 24 | 6 | 4 | 24 | 8 | 14 | | | 6 | | | |
| 2 | 785 | 20 | 12 | | 2 | 30 | 10 | 2 | 4 | 6 | 4 | | | |
| 2 | 817 | 38 | 18 | | 2 | 20 | 8 | 10 | 2 | 2 | 2 | 2 | 2 | |
| 2 | 830 | 82 | 62 | 27 | 12 | 104 | 24 | 20 | 12 | 15 | 14 | 8 | 4 | 4 |

(5) Frequency (from the Joost van de Weijer Corpus)

| Onsets | Offsets |
|---|---|
| j = 10,6 | n = 10,3 |
| m = 10,6 | t = 10,1 |
| d = 9,6 | r = 6,9 |
| h = 7,2 | m = 6,1 |
| n = 6,9 | s = 5,3 |
| z = 5,6 | k = 2,9 |
| b = 5,4 | x = 2,9 |
| w = 4,4 | p = 2,3 |
| k = 4 | l = 2,2 |
| x = 3,6 | nt = 1 |
| v = 2,1 | j = 0,5 |
| l = 2 | f = 0,4 |
| p = 1,4 | xt = 0,4 |
| t = 1,3 | st = 0,5 |

(6) Frequency would predict the following order:
j, m > d > h > n > z > b > w > k > x > v > l> p > t

(7) The word-based approach

a. Target onset /l/ is [h] or [s] up until 2;2.27. In subsequent recordings target words starting with /l/ are produced with onset [l] (100% correct): *leeuw, lift, lezen, lepel, lettertjes, lopen, lekker, luier, laarzen* etc.

b. Target onset [m]. Is [m] only in three fossilized forms: *mamma, mij, meer*, [b] or [p] otherwise up until 1;11.20. In subsequent recording all target words starting with /m/ are produced with onset [m] (100% correct): *mag, mee, mooi, mannetje*.

c. Target onset /f/ ([v]). Is produced [s] or [z] up until 1;8.10. In subsequent recording we find [f] onsets for all target words starting with /v/, like *vallen, vis, vogel*.

d. Target onset /x/ is produced [s] and later [f] up until 2;1.25. In subsequent recording we find 100% [x] productions for target words with onset /x/: *grote, gegeten, ga, gek, glijbaan*.

## 2 A feature-based analysis

(8) An important reason why a feature-based analysis seems to fail, is that we find gaps: natural classes are not always learned as a whole.

(9) Example: Beers (1996)

a. inventory 1: { p, m, t, n, j }
acquired features: [consonantal], [sonorant], [labial] [coronal]
problem: how do we distinguish /j/ and /n/?

b. inventory 2: { p, m, t, n, j, k }
acquired feature: [dorsal]
problem: no [ŋ] in inventory

c. inventory 3: { p, m, t, n, j, k, s, x, h }
acquired feature: [continuant]
problem: no [f] in inventory

(10) Two types of fcc (Itô, Mester and Padgett 1994):

a. *[F,G]: No segment has both F and G

b. [F]⊃[G]: If a segment has F, it also has G

(11) Only the following constraints seem necessary:

a. **General:** *[nasal,velar], *[velar,voice], *[continuant,voice], [continuant]⊃[coronal], *[continuant,velar]

b. **Onset:** [continuant]⊃[labial], [nasal]⊃[labial], [labial]⊃[nasal]

c. **Coda:** [velar]⊃[continuant]

(12) Example: Child 2 / Onset

| | **Features** | **Constraints** | **Predicted inventory** | **Day** |
|---|---|---|---|---|
| 1. | [voice]<br>[labial], [coronal] | - | { b, p, t, d } | 529 |
| 2. | [nasal] | i. [nasal]⊃[labial] | { b, p, t, d, m } | 540 |
| 3. | [continuant] | ii. [continuant]⊃[coronal] | { b, p, t, d, m, s, z } | 554 |
| 4. | - | Revoke i.<br>Revoke ii.<br>(Assuming w=v) | { b, p, t, d, m, n, s, z, f, v } | 615 |
| 5. | [velar] | iii. *[voice,velar] | { b, p, t, d, m, n, s, z, f, v, k, x } | 643 |
| 6. | [lateral] | - | { b, p, t, d, m, n, s, z, f, v, k, x, l } | 766 |
| 7. | [rhotic] | - | { b, p, t, d, m, n, s, z, f, v, k, x, l, r } | 817 |

(13) Example: Child 2 / Coda

| | **Features** | **Constraints** | **Predicted inventory** | **Day** |
|---|---|---|---|---|
| 1. | [labial],[coronal]<br>[continuant] | a. *[continuant, Place] | { p, t, s } | 529 |
| 2. | [nasal] | - | { p, t, s, n, m } | 540 |
| 3. | [velar]<br>[rhotic] | Revoke a.<br>b. *[nasal,velar] | { p, t, s, n, m, x, f, k, r } | 643 |
| 4. | [lateral] | Revoke b. | { p, t, s, n, m, x, f, k, r, ŋ, l } | 817 |

(14) Example: Child 10 / Onset

| | **Features** | **Constraints** | **Predicted inventory** | **Day** |
|---|---|---|---|---|
| 1. | [labial]<br>[continuant] | a. *[continuant,Place] | { p, s } | 777 |
| 2. | [velar], [coronal]<br>[nasal] | Revoke a. | { p, s, k, t, f, x, n, m, ŋ } | 915 |
| 3. | [lateral] | - | { p, s, k, t, f, x, n, m, ŋ, l } | 1065 |

(15) Example: Child 4 / Onset

| | **Features** | **Constraints** | **Predicted inventory** | **Day** |
|---|---|---|---|---|
| 1. | [labial], [coronal], [velar]<br>[continuant]<br>[nasal] | a. [velar]⊃[continuant] | { p, t, f, s, x, m, n } | 497 |
| 2. | [rhotic] | | { p, t, f, s, x, m, n, r } | 590 |
| 3. | - | Revoke a.<br>b. *[nasal,velar] | { p, t, f, s, x, m, n, r, k } | 643 |
| 4. | - | Revoke b. | { p, t, f, s, x, m, n, r, k, ŋ } | 703 |

(16) Example: Child 7 / Onset

| | **Features** | **Constraints** | **Predicted inventory** | **Day** |
|---|---|---|---|---|
| 1. | [coronal] | - | { t } | 392 |
| 2. | [labial]<br>[nasal]<br><br>[continuant] | <br>a. [nasal]⊃[labial]<br>b. [labial]⊃[nasal] | { t, m, s } | 429 |
| 3. | - | Revoke b. | { t, m, s, p, f } | 460 |
| 4. | [velar] | Revoke a.<br>c. (✘inexpressible constraint against /k/) | { t, m, n, s, p, f, ŋ, x } | 524 |
| 5. | [lateral] | | { t, m, n, s, p, f, ŋ, x, k, l } | 537 |

(17) Example: Child 8 / Onset

| | **Features** | **Constraints** | **Predicted inventory** | **Day** |
|---|---|---|---|---|
| 1. | [labial], [velar] | - | { p, k } | 517 |
| 2. | [continuant] | a. *[continuant,Place] | { p, k, s } | 572 |
| 3. | [coronal]<br>[lateral] | - | { p, k, s, t, l } | 590 |
| 4. | [nasal] | b. [nasal]⊃[labial] - | { p, k, s, t, l, m } | 608 |
| 5. | - | Revoke a. | { p, k, s, t, l, m, f, x } | 636 |
| 6. | - | Revoke b. | { p, k, s, t, l, m, f, x, n, ŋ } | 649 |

(18) Example: Child 9 / Onset

| | **Features** | **Constraints** | **Predicted inventory** | **Day** |
|---|---|---|---|---|
| 1. | [voice]<br>[labial], [coronal] | - | { p, b, t, d } | 569 |
| 2. | [velar]<br>[nasal] | i. *[velar,voice]<br>ii. [nasal]⊃[labial] | { p, b, t, d, k, m } | 583 |
| 3. | | Revoke ii.<br>iii. *[nasal,velar] | { p, b, t, d, k, m, n } | 639 |
| 4. | [continuant] | iv. *[continuant,velar]<br>vii. *[continuant,voice] | { p, b, t, d, k, m, s, f } | 691 |
| 5. | [lateral] | - | { p, b, t, d, k, m, s, f, l } | 741 |
| 6. | - | Revoke iv. | { p, b, t, d, k, m, s, f, l, x } | 846 |

# 3 Discussion / conclusion

- A feature-based analysis of acquisition of segment inventories seems feasible, if supplemented with a restrictive theory of fcc
- However, we still need to find out what determines the order in which features are acquired
- Variation might still be due to relative input frequency
- We also need to consider the relevance of the target words

(Slides: `http://www.vanoostendorp.nl/pdf/tin2007.pdf`)