Institute of Phonetic Sciences,
University of Amsterdam,
Proceedings 25 (2003), 43-52.

COMPARING UTTERANCE STRUCTURES OF DEAF AND HEARING INFANTS[1]

Florien J. van Beinum[2] and Lillian Doppen[3]

Abstract

Both deaf and hearing infants produce many speech-like sounds in their first years of life. This paper concentrates on the question as to how far deaf or severely hearing-impaired children differ from their hearing peers with respect to their utterance structures. Utterances of five hearing and five deaf Dutch children from 10.5 to 17.5 months of age were analyzed for a number of speech characteristics, like phonation and articulation type, number of syllables, utterance structure, place of articulation and preferred combinations of vowel-like and consonant-like elements. Results show that although deaf children produce many multi-syllabic utterances, alternations of CV movements are scarce compared to hearing children.

1 Introduction

Babbling in its canonical form is generally considered to be a crucial phase in the development towards the production of full-fledged speech. In normally developing hearing infants the onset of babbling is about seven months of age and can be understood as a result of a nicely hierarchical course of mastering phonatory and articulatory skills (Koopmans-van Beinum & Van der Stelt, 1986). Structures of the babbling utterances are primarily based on biomechanical or sensorimotor principals: frames of open-close alternations of mandibular movements. Subsequently and under the influence of the environmental language, consonantal and vocalic content is imposed resulting in utterance structures in agreement with the specific mother language, revealing a clear continuity between the pre-lexical and lexical stages. The predominant role of frames has been claimed to be obvious in a number of aspects of babbling and early use of words (MacNeilage & Davis, 1990), independent of the environmental language. Three of these aspects concern favored consonant-vowel combinations: a) central vowels co-occur with labial consonants, b) front vowels co-occur with alveolar consonants, and c) back vowels co-occur with velar consonants. The availability of a Dutch database of deaf and hearing infants (Koopmans-van Beinum, Clement & Van den Dikkenberg-Pot, 2001) provided us with the possibility to investigate the influence of spoken language input and auditory feedback on assumed preferences in consonant-vowel combinations.

---

[1] This paper is a revised and extended version of the paper presented at the 15th International Congress of Phonetic Sciences, Barcelona, 3-9 August 2003 (Koopmans-van Beinum and Doppen, 2003).
[2] formerly: Koopmans-van Beinum
[3] Nijmegen University

2 Methods

The total Dutch database of deaf and hearing infants was described extensively in previous publications (Koopmans-van Beinum et al., 2001). We'll summarize here the information necessary to understand the present paper.

2.1 Participants

For the present study we used speech material of ten mother-infant pairs who participated in a longitudinal study. Five infants were profoundly hearing impaired (D), the five other infants had normal hearing (H). All infants (all boys) had hearing parents. No clear health problems were found in a health screening right after birth (Apgar score), nor later on as tested by the Denver Developmental Screening Test and the Bayley Developmental Scales, applied at 12 and 18 months. The hearing infants (H) were matched with the deaf infants (D) along the following criteria: sex, birth order, duration of pregnancy, age of the mother, social-economic status of the parents (defined in terms of education), dialect of the parents (defined in terms of residence and regional origin) and degree of dialect use. All infants were the second or third child in the family.

The D infants had an average hearing loss of over 90 dB in the best ear, which was established by Auditory Brainstem Response audiometry (ABR) in the first months of life. The loss was confirmed by pure-tone audiometric tests at a later age. All D infants participated in early intervention programs. They all used hearing aids, although not all of the children used them regularly during the first year. In all cases the cause of deafness was genetically based. One D child was raised with Total Communication (TC), two with Dutch Sign Language (NGT) / TC, and two mainly by the Oral Method.

Table 1. Overview of the auditory characteristics of the hearing-impaired children (D). The various language methods concern Oral Method, Total Communication (TC) and Dutch Sign Language (NGT).

Participant	Hearing loss (dB) best ear	Loss with Hearing aids (dB)	Hearing-Aids from age (m)	Language method	Start of recording (months)
D-1	97	55	2.0	Oral	2.5
D-2	93	55	3.5	TC	5.5
D-3	110	65	4.5	Oral	5.5
D-4	> 120	not tested	no	NGT/TC	2.5
D-5	> 120	not tested	6.5	NGT/TC	3.5

2.2 Speech material

From the audio-recordings, made in interactive situations by the parents themselves at their homes, we selected per month 50 non-vegetative comfort utterances, starting at the age of 10.5 months until 17.5 months. This resulted in a data set of 4000 utterances (8 x 50 x 10) in total. Starting our present study at the age of 10.5 months implied for all children a rather mature speech production instrument, at least a speech production instrument that anatomically and physiologically is capable to make canonical syllables. As published before (Koopmans-van Beinum et al., 2001), onset of canonical babbling in the five hearing infants was between 5.5 and 7.5 months, and even one of the deaf infants (D2 with a loss of 55 dB) had started to babble at 7.5 months of age.

2.3 Data analysis

All selected utterances were digitized and subsequently coded auditorily for a number of characteristics. Use was made of the speech analysis program PRAAT (Boersma & Weenink, 1996) in order to combine the audible sound form with an oscillographic display and to make coding decisions easier and more reliable thanks to possible acoustic analyses like pitch detection and spectrography. In this way each utterance was analyzed for the following characteristics:

- *phonation type*: classifying phonation into one of five possible types, e.g., no phonation (0), simple uninterrupted (1) or interrupted (2) phonation, variegated uninterrupted (3) or variegated interrupted (4) phonation.
- *articulation type*: classifying articulation into one of three possible types, e.g., no articulation movement (0), one articulation movement (1), two or more articu-lation movements during two- or more-syllabic utterances (2).
- *number of syllables*: ranging from 0 to ≥5, the criteria for a syllable or syllable-like element being a minimal rhythmic unit containing a vowel-like phase with or without a preceding or following consonant-like closing phase. These criteria are broader than those for a canonical syllable (Oller, 1986) and although we are well aware of the problems that are inherent in our working definition, we decided to use this in order to include pre-canonical babbling as well.
- *structure of the utterance*: indicating co-occurrences of vowel-like (V) and consonant-like (C) elements within each utterance. To make our results manageable we distinguished twelve classes: V = single vowel-like sound; VV…= two or more vowel-like sounds ('more' is indicated by …); C…– = voiceless consonant-like sound(s); C…+ = voiced consonant-like sound(s); V…C…= one or more vowel-like sound(s) followed by one or more consonant-like sound(s); C…V…= one or more consonant-like sound(s) followed by one or more vowel-like sound(s); V…C…V…= one or more vowel-like sound(s) followed by one or more consonant-like sound(s) followed by one or more vowel-like sound(s); C…V…C…= one or more consonant-like sound(s) followed by one or more vowel-like sound(s) followed by one or more consonant-like sound(s); and finally four types of babbling series: utterances beginning with a consonant-like element and ending with a vowel-like element; the same series but ending with a consonant-like element; beginning with a vowel-like element and ending with a consonant-like element; the same series but ending with a vowel-like element. The distinction in open and closed utterances was considered to be probably essential for the Dutch language.
- *place of articulation per syllable*: classifying vowel-like elements into one of three possible categories (front, central, or back), and classifying consonant-like elements into one of three possible categories (labial, alveolar, or velar). Palatal elements were left aside, uvular and pharyngeal elements were grouped with velars. Per child we classified each syllable that met the requirements of a CV or VC structure. Since it is absolutely not clear in advance whether syllable boundaries in babbling series have to be thought after CV- or after VC-, we decided to make a double analysis: one for all occurring CV syllables and one for

all occurring VC syllables. If order of consonant and vowel has no influence on the preferred combinations, both our analyses would give identical results.

2.4 Reliability

Primarily a single transcriber analyzed all utterances, after an intensive training period. Two other transcribers analyzed parts of the material. Transcription reliability was examined in several ways. First of all intra-transcriber reliability was controlled by reanalyzing randomly chosen parts of the speech material. Subsequently a second transcriber analyzed ten percent of the material as well, whereas a third transcriber analyzed all speech material from 12.5 to 17.5 months of age for phonation type, articulation type, and number of syllables. Agreement tested so far was rather high (over 80%), which was considered to be sufficient.

3 Results

In order to investigate the influence of spoken language input and auditory feedback on assumed preferences in consonant-vowel combinations as mentioned in the Introduction, we will give the results of the data analysis for the hearing-impaired (D) and the normally hearing (H) group by means of frequency counts and chi-square tests in a number of steps (sections 3.1 to 3.4). For each of the children we counted frequencies over the whole period together (8 x 50 = 400 utterances per child).

3.1 Phonation and Articulation Type

As a first step frequency counts of the five phonation types and the three articulation types were processed per group (D and H) over the total 8-months period. In Table 2 for both groups frequencies of occurrence (in %) of each of the 15 combinations (five phonation types and three articulation types) have been displayed. Significant differences are underlined (single line for p < 0.05, double line for p < .001).

Table 2. Frequency of occurrence (in %) for each combination of phonation and articulation type for the D (n=5) and the H (n=5) infants over the eight months together. Significant differences are underlined (single line for p < 0.05, double line for p < .001).

	Articulation					
	NoArt		OneArt		TwoArt	
Phonation	**D**	**H**	**D**	**H**	**D**	**H**
NoPhon	--	--	6.12	1.05	0.50	0.30
UnIntPhon	30.74	32.01	8.83	18.46	0.90	3.85
IntPhon	18.90	9.70	2.60	3.65	1.15	3.10
VarUnIntPhon	5.21	5.45	1.35	4.30	1.60	4.45
VarIntPhon	17.80	4.40	2.00	3.10	2.26	6.15

A number of striking differences between the D and the H infants are:

- D infants produce far more utterances with interrupted phonation without articulation than H infants do. These simple repetitive phonation movements might be seen as substitutions for the more complex repetitive babbling movements.
- D infants produce more utterances with one articulation movement but without phonation than H infants do. The tactile sensation may play a role here.
- H infants produce more utterances with uninterrupted phonation combined with one articulation movement than D infants do.
- H infants produce more utterances with two or more articulations movements like in canonical babbling than D infants do.

3.2 Number of syllables

The analysis of phonation and articulation movements as described above suggests different utterance structures for the two groups of children, although both seem to make multi-syllabic utterances. Table 3 presents frequencies of occurrence (in %) of syllables per utterance, as well as the absolute numbers of syllables for the two groups. All differences are highly significant ($p < .001$). Specific observations are:

- D infants outnumber H infants considerably for the total number of syllables over all utterances (4318 vs 3323). In view of the results on phonation and articulation as mentioned above, it is quite plausible that the high number of utterances with interrupted phonation without articulation in the D group is cause of this difference.
- D infants produce considerably more utterances with a zero syllable count than H infants do. A zero syllable count has been assigned to those utterances that consist in an articulation movement without phonation. As mentioned above the tactile sensation might play a role here for the D infants.
- D infants produce more utterances with three or more syllables than H infants do. Again the cause might be in the high number of utterances with interrupted phonation without articulation.
- H infants produce more utterances with one or two syllables than D infants do. The considerably higher number of utterances with phonation plus articulation movements in the H group (see Table 2) is probably the cause of this difference.

Table 3. Frequency of occurrence (in %) of number of syllables within an utterance for the D (n=5) and the H (n=5) infants over the eight months together. All differences are highly significant ($p < .001$).

Number of syllables per utterance	D	H
0 syllables	6.77	1.50
1 syllable	43.88	57.27
2 syllables	20.11	27.06
3 syllables	11.88	8.05
4 syllables	6.87	2.90
5 or more syllables	10.48	1.90
total number of syllables	4318	3323

3.3 Utterance structure

Since the two groups differ considerably in the number of syllables in their utterances, it is not unlikely that they will differ in the complexity of their utterances as well. We thus counted frequencies of occurrence of the most common structures for the two groups, with the results as given in Table 4.

Table 4. Frequency of occurrence (in %) of types of utterance structures for the D (n=5) and the H (n=5) infants over the eight months together. Significant differences are underlined (single line for $p < 0.05$, double line for $p < .001$).

Structure of the utterance	D	H
V	35.35	37.4
VV…	37.21	14.15
C…–	6.51	1.10
C…+	4.26	3.30
C…V…	3.96	12.95
V…C…	4.36	3.95
V…C…V…	3.30	11.10
C…V…C…	0.60	2.90
C…V…C…V…(C…V…)	1.30	6.70
C…V…C…V…C…(V…C…)	0.20	0.70
V…C…V…C…(V…C…)	1.10	2.60
V…C…V…C…V…(C…V…)	1.80	3.10

Here again some striking differences arise, mainly in line with the results of the preceding sections:

- D infants and H infants produce a similar percentage of utterances consisting in a single vowel-like sound.
- D infants produce, percentage-wise, considerably more utterances with series of vowel-like sounds than H infants do.
- D infants produce, percentage-wise, considerably more utterances consisting in a single voiceless consonant-like sound than H infants do.
- H infants outnumber D infants, percentage-wise, considerably in all structures with consonant-vowel alternations, except utterances with single VC structure.

3.4 Place of articulation

Each syllable that met the requirements of a CV or VC structure was classified in one of the nine combinations of front, central, or back (for vowel-like elements) with alveolar, labial, or velar (for consonant-like elements). Frequency counts of all possible combinations (in % per group) gave the results as shown in Table 5 (CV structures) and Table 6 (VC structures). As could be foreseen already from the results in the preceding sections, the total number of CV and VC occurrences for the D children is considerably smaller than for the H children (CV for D = 368, for H = 1164; VC for D = 349, for H = 891). The H children, percentage-wise, outnumber the D children in the frequencies of most of the vowel-consonant combinations, both for the CV (Table 5) and for the VC combinations (Table 6). No essential differences are found between both tables, so we'll confine our comments to the CV data as given in Table 5. Differences still to be mentioned are:

- D children outnumber, percentage-wise, H children as for the use of labial consonants in combination with central vowels.
- D children outnumber, percentage-wise, H children as for the use of velar consonants in combination with central vowels.

The visual and tactile sensations, respectively, may play an important role here.

Table 5 shows for the D group a significant preference to alveolar-central and to labial-central combinations. The difference in preference between these two combinations is not significant (z=.58). In the H group too alveolar-central combinations are preferred, but here they are significantly more favourite than the labial-central combinations (z=14.05). Besides, the alveolar-central combinations are significantly more favourite in the H group than in the D group (z=-2.62).

Table 5. Total absolute number of CV syllables for D and for H, and frequency of occurrence (in %) of alveolar, labial, and velar consonants in combination with front, central, and back vowels in CV syllables for the D (n=5) and the H (n=5) infants over the eight months together. Percentages for D and for H add up to 100.

CV (total D = 368) (total H = 1164)	**Consonants**					
	Alveolar		Labial		Velar	
Vowels	**D**	**H**	**D**	**H**	**D**	**H**
Front	2.45	8.59	2.17	2.49	3.81	1.80
Central	37.60	45.36	33.24	18.32	13.07	6.70
Back	2.72	5.67	4.63	7.64	0.27	2.40

Table 6. Total absolute number of VC syllables for D and for H, and frequency of occurrence (in %) of alveolar, labial, and velar consonants in combination with front, central, and back vowels in VC syllables for the DS (n=5) and the H (n=5) infants over the eight months together. Percentages for D and for H add up to 100.

VC (total D = 349) (total H = 891)	**Consonants**					
	Alveolar		Labial		Velar	
Vowels	**D**	**H**	**D**	**H**	**D**	**H**
Front	2.29	9.76	1.14	0.78	3.15	1.34
Central	31.80	47.25	29.79	16.27	19.77	7.07
Back	6.87	8.19	2.57	6.06	2.57	4.25

Table 7. Observed-to-expected ratios of occurrences of alveolar, labial, and velar consonants in combination with front, central, and back vowels, for hearing impaired infants (D), for hearing infants (H), and for Dutch adults (DUTCH). Values on the diagonal (in grey) should be clearly above-chance (> 1.0) for the infant groups.

D	Alveolar	Labial	Velar
Front	0.68	0.64	2.64
Central	1.05	0.99	0.91
Back	0.83	1.53	0.19

H	Alveolar	Labial	Velar
Front	1.11	0.67	1.27
Central	1.07	0.90	0.87
Back	0.59	1.71	1.37

DUTCH	Alveolar	Labial	Velar
Front	1.39	0.56	1.07
Central	0.91	1.15	0.79
Back	0.65	1.36	1.09

In Table 7 we present the observed-to-expected ratios of occurrences of alveolar, labial, and velar consonants in combination with front, central, and back vowels, calculated for hearing impaired infants (D), for hearing infants (H), and for Dutch adults (DUTCH). These latter results are derived from adult frequencies of occurrence of vowels and consonants known from the Dutch CELEX database, based on a million of Dutch words. The expected values in each cell would be 1.0 if the distribution of consonant-vowel combinations is based on chance. However, if the hypothesis concerning the predominant role of frames of prefered consonant-vowel combinations is valid, the values of the cells on the diagonal should be clearly above-chance (i.e. >1.0) in the infants groups.

When inspecting the frequencies in Tables 5 and 6 for the H and the D infants, and the observed-to-expected ratios in Table 7 for D and H, it becomes clear that the claims about favored consonant-vowel combinations (the diagonal in the tables) as made by MacNeilage & Davis (1999), cannot be supported by our results.

In Table 8 we compared the data presented above on our H and D infants and the Dutch adults, with data on American-English learning children in the same age as our H and D children, as given by Davis and MacNeilage (1995). Separately for vowels and consonants we calculated percentages of front, central, and back vowels, as well as percentages of alveolar, labial, and velar consonants in order to see where main differences come from.

Table 8. Frequencies of occurrence of front, central, and back vowels (VOWELS), and of alveolar, labial, and velar consonants (CONSONANTS) calculated in percentages for American-English learning children, for Dutch H children, for Dutch D children, and for Dutch adults.

	in %	Amer child.	Dutch H	Dutch D	Dutch Adults
VOWELS	Front	69	13	8	32
	Central	28	70	84	44
	Back	3	16	8	24
CONSONANTS	Alveolar	78	60	43	65
	Labial	17	28	40	23
	Velar	5	11	17	12

As for the consonants the D children produce a much higher percentage of labials at the expense of alveolar consonants than the others do. This preference may be explained by the greater visual component of labial consonants.

More striking, however, are the differences in vowels, especially those between the American-English and the Dutch hearing (H) children. In our Dutch data we found a very high number of central vowels (for H = 70% and for D = 84%), different from the American-English data (28%). Back vowels are rather few (for H = 16% and for D = 8%); front vowels in our dataset are lowest in frequency (for H = 13% and for D = 8%), whereas American-English children prefer front vowels (69%). This might have resulted in a preferred combination of central vowels especially with alveolar consonants in the Dutch children, since alveolar consonants (for H = 60% and for D = 43%) outnumber labial and velar consonants largely, especially in the H group (labial 28% and velar 11%; for the D group 40% and 17%, respectively).

4 Discussion

When asking ourselves what might be the cause of the deviations in consonant-vowel preferences, several possibilities come to the fore. One cause may be found in the Dutch mother language. As can be seen in Table 8, the adult frequency data, clustered in the same way as we did for the D and H infants, irrespective of other consonants, provided as frequencies for vowels: 32% front, 44% central, and 24% back, and as frequencies for consonants: 65% alveolar, 23% labial, and 12% velar. The Dutch infants show a very high frequency of central vowels, contrary to the American-English children. However, central vowels are the most frequent group for Dutch adults as well. So an early influence of the Dutch mother tongue could be present here. But since the frequency data of the D children are quite comparable to those of the H children in spite of the lack of auditory input in the D children, influence of the mother language then could be only in the perception of the transcribers instead of in the production of the children. This would mean that the reliability and comparability of the transcription is crucial in cross-linguistic studies.

Another explanation might be found in the selection of the syllables used in the present study. Our working definition of a syllable was: a minimal rhythmic unit containing a vowel-like phase with or without a preceding or following consonant-like closing phase. As said before (section 2.3) these criteria are broader than those for a canonicle syllable as given by Oller (1986). So because of this definition more pseudo-resonant vowels may have been included, transcribed as central, than in the study of Davis and MacNeilage (1995).

5 Conclusion

The results of our study all point into the same direction: although D children between 10.5 and 17.5 months of age produce many multi-syllabic utterances, the structure of these utterances is quite different from those produced by the H children. Alternations of CV movements in D children are rather scarce compared to what H children do. As stated before (Koopmans-van Beinum et al., 2001) coordination of articulation and phonation movements seems to require auditory feedback in order to provide the possibility to produce all variations in utterance structure as needed to achieve full-fledged adult speech in the end. The development of the utterance structures for each of the ten children individually during the period mentioned above will be next subject of our study. That also may give an answer why the Dutch D group and the Dutch H group are more alike in their CV preferences, than the Dutch H group and the hearing American-English children are.

References

Boersma, P. & Weenink, D. (1996). *Praat, a system for doing phonetics by computer.* Institute of Phonetic Sciences, University of Amsterdam, report 132.
[see also http://www.fon.hum.uva.nl/praat/ for the present version].

Davis, B.L. & MacNeilage, P.F. (1995). “The articulatory basis of babbling”, *Journal of Speech and Hearing Research* **38**, pp. 1199–1211.

Koopmans-van Beinum, F.J., Clement, C.J. & Van den Dikkenberg-Pot, I. (2001). “Babbling and the lack of auditory speech perception: a matter of coordination?” *Developmental Science* **4**, pp. 61-70.

Koopmans-van Beinum, F.J. & Doppen, L. (2003). “Development in utterance structures of deaf and hearing infants”. In: M.J. Solé, D. Recasens & J. Romero (Eds), *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona: Universitat Autonoma de Barcelona, pp. 1033-1036.

Koopmans-van Beinum, F.J. & Van der Stelt, J.M. (1986). “Early stages in the development of speech movements”. In: B. Lindblom & R. Zetterström (Eds), *Precursors of Early Speech*, Wenner Gren. Int. Symp. Series 44, New York: Stockton Press, pp. 37-50.

MacNeilage, P.F. & Davis, B.L. (1990). “Acquisition of speech production: Frames then content”. In: M. Jeannerod (Ed.), *Attention and performance XIII: Motor representation and control*, Hillsdale, NJ: Lawrence Erlbaum, pp. 453-475.

Oller, D.K. (1986). “Metaphonology and infant vocalizations”. In: B. Lindblom & R. Zetterström (Eds.), *Precursors of Early Speech*, Wenner-Gren Int. Symp. Series 44, New York: Stockton Press, pp. 21-35.