An Investigation of Infant Babbling

by M.C. Dinger and J.G. Blom

Contents:

1.0 Introduction.

1.1   The investigation concerning infant babbling was actuated by a previous
pilot-investigation carried out in the U.S.A. by professor B. Tervoort
between 1960 - 1963.
Some tapes with babbling noises of five- and ten-months old babies,
2 Dutch, 2 American and 1 Chinese baby (Chinatown, San Fransisco) were
played for a group of listeners (University students) in the Netherlands
and in the United States of America.
The listeners were requested to indicate:

     a) the recordings taken were of the babblings of non-American
        c.q. non-Dutch babies.
     b) no opinion.

	Dutch listeners	U.S. listeners
Dutch baby 5 months	b	b
Dutch baby 10 months	b	a
U.S. baby 5 months	b	b
U.S. baby 10 months	a	b
Chinese baby 12 months	a	a

The American listeners only heard that the 10 months old Dutch infant
was non-American, for the Dutch listeners the situation was the reverse;
both groups heard that the Chinese baby was non-American c.q. not a
Dutch infant.

In order to verify the results of the investigation done in the U.S.A.
professor Tervoort decided to repeat his experiment in the Netherlands.
He contacted professor H. Mol of the Institute of Phonetic Sciences in
Amsterdam. A mutual arrangement between members of both the Phonetics
Department and professor Tervoort's Institute of General Linguistics
resulted in the following division of labour: a literature study of
Infant Babbling would be undertaken by an assistant of professor
Tervoort's; the actual execution of the investigation by an assistant
of professor Mol.

2.0 Definition of the Problem.

Discussions on the organization of the experiment led to the following definition of the problem

2.1 Is it possible for a group of adult listeners to determine whether a baby belongs to their own language background or not on the strength of the babbling monologues produced by an infant?

2.2 If 2.1 is possible the following question is raised. From what age onward can a baby's babbling be identified?

2.3 It would be of interest to determine if listeners have a higher score for a baby of their own language background than for a baby from a different background.

2.4 To what degree knowledge of a different language system and of having children influences recognition could be sorted out with the aid of a questionnaire which testees would be requested to fill in after the experiment.

3.0 Set-up.

In order to gain some insight into the problem involved the investigation was set up as follows:

3.1 The choice of babies for the listening experiment was to be confined to two nationalities: American and Dutch infants. Babies would have to belong to families whose social backgrounds would be matched.

3.2 A longitudinal investigation was preferred. Listeners could then be confronted with the babbling noises of one and the same baby at different stages of its development.

3.3 Recordings would be made of the babbling of a number of American and of Dutch babies between 7 and 18 months of age. (It would have been without point to continue the recordings after 18 months as the infants were expected to produce their first one-word-sentences at

about that age). For this pilot-investigation a minimum of 3 American
and 3 Dutch babies were deemed desirable.

3.4   From the babbling material a tape would be made up, using fragments
of productions of each baby taken at various times. These would be
offered for judgment to 50 American and 50 Dutch women.

4.0   Data Collecting.

4.1   For more than a year babbling noises were recorded of 3 American and
4 Dutch babies at intervals of approximately 3 weeks. (It was im-
possible to find more suitable American babies in Amsterdam at the
time). The development of the babies was followed in the period ranging
from 7 to 18 months.

4.2   The babbling monologues were recorded with an UHER-4000 Report-S.

5.0   Data Processing.

5.1   From the babbling material a condensed record was made per baby. These
tape-recordings contained fragments which were unimpaired by:
      a) noises made by other people than the infant
      b) background noises
      c) periods of silence lasting more than 2 seconds.

5.2   As the tapes of only 2 American babies were complete (one of the
American families moved abroad unexpectedly) and as it was thought
best to keep the number of Dutch and American babies equal, the final
test version of the recordings was made up of the condensed tapes of
2 American and 2 Dutch babies (1 boy, 3 girls).

5.3   This test recording consisted of 20 fragments of babbling monologues;
each fragment lasting 90 seconds exactly. The tape contained five
recordings of each infant at the ages of circa 33, 43, 53, 61 and 70
weeks. The 20 fragments were fully randomized. Each fragment was
preceded by a serial number spoken in Dutch and in English. After each
fragment there was a 15 second pause in which the listener had to score.

6.0 Execution.

6.1 The listening experiment took place at the Institute of Phonetic Sciences, where a number of listening-booths are available for such purposes. Each booth has its own Revox G 36 with separate test-recording.

6.2 The group of listeners consisted of:

50 American women, some living in the Netherlands temporarily.

50 Dutch women, mostly students of speech-therapy and university students.

The testees were instructed to listen to the babbling noises and to indicate after each fragment whether the babbling was produced by a Dutch or an American baby by putting a cross under Dutch or American baby respectively after the pertaining item number. The scoring was done by way of a forced choice.

The testees did not know how many babies had been recorded. Neither did they know the ages of the infants. The listeners were asked to fill in a questionnaire after finishing their scoring.

7.0 Results and Discussion.

Statistical analysis of the babbling investigation shows the following.

7.1 The table at the end of this paper summarizes the statistical analysis.

The code for the Dutch babies is NL 1 and NL 2.

The code for the American babies is AM 1 and AM 2.

The Dutch listening group is coded NL.

The American listening group is coded AM.

The table shows the number of correct scores for the five points of time presented in the test-recording.

Significantly high scores are marked with an asterisk ($\overset{*}{}$), significantly low scores have been underlined.

The scores have been tested against the hypothesis that the babies' nationality is not recognized. Therefore:

$$Ho : p = \tfrac{1}{2}$$

The expected number of correct scores then becomes:

$$np = \frac{n}{2} \text{ in which } \underline{n} \text{ is the number of trials.}$$

The distribution of the number of correct scores follows a Binomial Distribution with $p = \frac{1}{2}$. This distribution can be approximated by a Normal Distribution with mean value $\mu = np = \frac{n}{2}$ and one standard deviation

$$\sigma = npq = \frac{1}{4} n \text{ in which } q = 1 - p$$

So that $z = \dfrac{m - \mu}{\sigma}$ in which m = number of correct scores.

The criterium of significance chosen is

$z > 2.05$ that is a 2% one-sided chance of being exceeded.

A clear trend can be found in a number of columns only, e.g. the total score of the NL group for both Dutch babies together NL 1 + 2. The last column especially, viz all listeners (NL + AM) for all infants (NL 1 + 2 + AM 1 + 2) clearly has a general direction. When infants reach the age of 53 weeks they are recognized a significant number of times. Beyond this age the number of correct scores increases regular- ly. At the age of 70 weeks the chance of recognition is 273/400 ≈ 68%. Chance of recognition is therefore significant albeit still small. After all, without listening a subject would have a correct score of 50%. Besides, the question arises whether the babbling monologues are still babbling sounds pure and simple or whether one-word sentences are occurring at this stage.

7.2   The great many times a very low score is produced cannot be explained. This would point to the hypothesis that $p = \frac{1}{2}$ is incorrect, or perhaps Dutch babies produce a striking "American" sound and American babies a striking "Dutch" one. This last theory is confuted by the fact that scores on NL 1 at 33 weeks of age are significantly low. The same goes for NL 2 and AM 2 at the age of 43 weeks.

7.3   If we compare the scores of the two listening groups we see that the Dutch group scores correctly more often than the English-speaking group.

Perhaps the composition of the Dutch listeners was of some importance.
They were, for 80%, students of speech therapy from the day course for
"Logopedie en Foniatrie" in Amsterdam, while the American-English
language group was not homogeneous in its composition. It was clear
too, that both Dutch and American listeners recognized Dutch babies
better than American infants. This is contrary to the idea that a
listener is supposedly able to hear that a baby belongs to a different
language background. The phenomenon might be explained by the fact
that the tape-recordings of the Dutch infants show a greater variety
of babbling noises when taken all together.

7.4   How great the importance is that can be attached to a knowledge of a
language system not one's own when judging fragments of babbling noises
is difficult to assess. The questionnaire filled in by the listeners
makes it clear that the Dutch group had a better passive and active
knowledge of English than the English-speaking group had of Dutch.

7.5   Another point which was checked was whether listeners with children did
better in recognizing babies than those without children of their own.
It became clear that having children of one's own did not add to the
recognition of a baby's language background.

7.6   The level of significance in recognition is only slight, as is shown
by the above experiment. This fact is supported by the reaction of both
groups of listeners. Their response to the question whether they had
any idea if they had recognized American and Dutch babies as such cor-
rectly showed a rather great amount of hesitations.

Baby	NL1	NL2	NL1+2	AM1	AM2	AM1+2	AM1+2 NL1+2	NL1	NL2	NL1+2	AM1	AM2	AM1+2	NL1+2 AM1+2	NL1	NL2	NL1+2	AM1	AM2	AM1+2	NL1+2 AM1+2
Listener	NL	NL	NL	NL	NL	NL	NL	AM	AM	AM	AM	AM	AM	AM	NL+AM	NL+AM	NL+AM	NL+AM	NL+AM	NL+AM	NL+AM
np	25	25	50	25	25	50	100	25	25	50	25	25	50	100	50	50	100	50	50	100	200
σ	3.5	3.5	5	3.5	3.5	5	7	3.5	3.5	5	3.5	3.5	5	7	5	5	7	5	5	7	10
Weeks																					
33	12	35*	48	24	31	55	103	16	20	36	19	32*	51	87	29	55	84	43	63*	106	190
43	55*	18	53	36*	18	56	109	24	16	40	23	11	34	74	59	34	93	61*	29	90	183
53	29	41*	70*	31	24	55	125*	34*	21	55	23	24	47	102	63*	62*	125*	54	48	102	227*
62	13	49*	62*	12	45*	57	119*	30	38*	68*	30	40*	70*	138*	43	87*	130*	42	85*	127*	257*
70	50*	31	81*	21	48*	69*	150*	48*	28	76*	15	32*	47	129*	98*	59	157*	36	80*	116*	273*
Total	140	174*	314*	126	166*	292*	606*	152*	123	275*	110	139	249	524	292*	297*	589*	236	305*	541*	1130*
np	125	125	250	125	125	250	500	125	125	250	125	125	250	500	250	250	500	250	250	500	1000
σ	7.9	7.9	11.2	7.9	7.9	11.2	15.8	7.9	7.9	11.2	7.9	7.9	11.2	15.8	11.2	11.2	15.8	11.2	11.2	15.8	22.5
	1	2	3 = 1+2	4	5	6 = 4+5	7 = 3+6	8	9	10 = 8+9	11	12	13 = 11+12	14 = 10+13	15 = 1+8	16 = 2+9	17 = 15+16	18 = 4+11	19 = 5+12	20 = 18+19	21 = 17+20

* Score significantly high

− Score significantly low