Age of Arrival vs. Time of Arrival: A Reconsideration of Johnson and Newport (1989) and Johnson (1992)

Vernon Percival, Francis Howerd and Benjamin Hill.

Manx Plaque Institut für Mundhygieneforschung
Douglas, Isle of Man
Great Britain

Abstract

This paper examines claims by Johnson and Newport (1989) and Johnson (1992) concerning critical period effects in second language acquisition. We report on a replication of the original Johnson and Newport experiment using a different population of learners, both children and adults, the results of which can only be partially reconciled with Johnson and Newport's findings. Instead, we posit the existence of a crucial variable not considered by these two researchers which we believe has more general value in accounting for our results, and perhaps for those of Johnson and Newport as well.

In an oft-cited paper Johnson and Newport (1989) have claimed that native-like fluency cannot be acquired by second language learners unless they have taken up residence in their new host country by the age of 7. If they arrive after that age, performance declines; if they arrive after puberty, it declines sharply. On the basis of their findings, Johnson and Newport (henceforth JN) state that there is a critical period for second language acquisition that ends at the age of 7. This claim is re-iterated in Johnson (1992), a replication of Johnson and Newport with the same subjects one year later responding to the same test stimuli, though this time in written rather than spoken form.

Johnson and Newport's subjects were well-educated Korean and Chinese immigrants who had been resident in the US for between 3 and 26 years. Half the group had arrived in the US between the ages of 3 and 15, and half between 17 and 39. They can be broken down as follows:
The results of the experiment are presented in Table 2 below, where the figures given represent the percentages of correct judgements for each group:

<table>
<thead>
<tr>
<th>AGE OF ARRIVAL</th>
<th>3-15*</th>
<th>17-39**</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>7-10</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>11-15</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>23-26</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1: The distribution of early (≤15) and late (≥17) arrivals in terms of the number of years they lived in the US.
* "For the most part freshmen or sophomore undergraduates"
** "Primarily professors, research associates, and graduate students, (with) 2-12 years formal instruction in their native country"
From Johnson & Newport (1989)

As will be apparent from Table 2, only those Ss arriving in America between the ages of 3 and 7 performed at native speaker levels of fluency, and those arriving after the age of 17 did least well. The differences between the native speakers and the 3-7 age group are not significant, while there are significant differences between these two groups and the remaining groups. There was a strong negative correlation (−0.77) between the variables Age...
posttest session, a number stated that they had had enough damn fool questions just trying to get into the country, and were sometimes unable to distinguish test items from announcements broadcast over the PA system. In many cases, it turned out that they were indeed better at judging the grammaticality of these broadcast announcements than the test items themselves.

**Results**

The results of Experiment I are summed up in Table 3 below:

<table>
<thead>
<tr>
<th>NATIVE SPEAKERS</th>
<th>CHILDREN</th>
<th>THIRD YEARS</th>
</tr>
</thead>
<tbody>
<tr>
<td>92%</td>
<td>50%</td>
<td>42%</td>
</tr>
</tbody>
</table>

Table 3. Performance on JN grammar test by Belgian learners of English

We see there what appears to be corroborating evidence for JN’s position – Third Years tested at Time I show the non-native fluency characteristic of the late starter. They are easily outperformed by Children (< .01). However, it must be admitted that the latter group themselves only scored an unimpressive 50%, showing that they have as yet only mastered half of the English language to native levels of fluency. However, in interpreting their results, we must not forget that some of the children in this group were older than 7, the age which JN show forms the upper limit for achieving native-like fluency. Further analysis will perhaps reveal that these older children were responsible for the performance deficit compared to Native Speakers, who scored an almost perfect 92%. At this stage in the proceedings, we must conclude that our findings tend to confirm JN’s in broad outline.

The performance of the adult learner group led us to hypothesise that students who had yet to go to the States would prove to be even less fluent than the Third Years. In order to see if there was any effect for proficiency among adult learners who had not been to America, we extended our testing to include students in the First and Second Years of their courses at RPAB; to measure the effects of exposure on proficiency we also took up the option of testing the Third Years again at the end of their year in America (Expt. II below).

**EXPERIMENT II**

In addition to the Third Years at Time II, 200 learners of English at the RPAB, were also
tested for the purposes of this experiment. All had begun learning English at secondary school after the age of 12. 100 of the Ss were in their first year of English studies at the Academy, and 100 were in their second year. Each group consisted of 50 male and 50 female students. None had at this stage even the slightest desire to go to the US, though one was present at the opening of EuroDisneyland. Like their Third Year peers, all had begun learning English after the age of 12. The average age of the First Year students was being 20.5 and that of Second Year students 21.5. We shall call these two groups First and Second Years respectively.

The results are presented in Table 4:

<table>
<thead>
<tr>
<th>FIRST YEARS</th>
<th>SECOND YEARS</th>
<th>THIRD YEARS**</th>
<th>THIRD YEARS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n=27)</td>
<td>(n=27)*</td>
<td>Time I (n=100)</td>
<td>Time II (n=100)</td>
</tr>
</tbody>
</table>

Table 4: Performance on IN grammar test by First and Second Years, and Third Years at beginning and end of American year.

* See text for explanation of this low N-size.

** Data from Expt I

Of the original 200 Ss in these two groups, 54 subjects (27%; 27 men and 27 women) managed to complete the test. Among those that failed to complete it there were two noteworthy tendencies: a) some Ss fell into a torpor, and b) some walked out after between 30 and 100 sentences had been presented. Upon being asked for an explanation in a post-hoc appraisal of the testing procedure, several of this subgroup claimed that their 'intelligence was being insulted and they had better things to do'. We have no idea what this can mean, but an extra percentage point was awarded to those that gave this response in English. It is possible that these students found the test too difficult and were fabricating an excuse so as not to lose face. Surprisingly, on checking their answers as far as they had gone, we found that they had make no error.

We wish now we had tested them individually so as to be able to offer them on-site counselling and encouragement.

In contradistinction to the results of Expt. I, a perusal of Table 4 shows that our hypothesis that pre-Third Year adults would show lower levels of fluency than Third Years could not be supported. First and Second Years, who had yet to even think about setting foot on American soil, also performed significantly better than Third Years at Time I. Remarkably, both these groups also performed significantly better than the pre-pubescent Children, and as just as well as the Native Speaker controls from Expt. I. Third Years at Time II, after one year in the States,
show a dramatic improvement over their performance at Time I. These results certainly cast doubt on JN’s interpretation of their data. Not only do adult learners seem to be able to improve their performance to native speaker levels after a mere year of exposure in the US, but there was now evidence that adult learners could outperform children, and that to perform to native levels of fluency, it was not even necessary to have arrived in the US at all. What price the notion of a critical period for second language acquisition now?

Discussion

How can we account for these seemingly contradictory findings? Careful reflection suggests that there is a key variable, unfortunately ignored by JN, which we think has considerable explanatory force. This variable is Time of Arrival. Let us now consider the performance of the Third Years at Time I when this variable is taken into account (Table 5):

<table>
<thead>
<tr>
<th>Time of Arrival</th>
<th>0300-0700*</th>
<th>0800-1000*</th>
<th>1100-1500*</th>
<th>1700-2355*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n=34)</td>
<td>(n=34)</td>
<td>(n=15)</td>
<td>(n=17)</td>
<td></td>
</tr>
<tr>
<td>15%</td>
<td>37%</td>
<td>90%</td>
<td>68%</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Score on test by group according to Time of Arrival, Third Years, Time II

*These times are local (i.e. EST) times in New York. To calculate Central European Time, please add 6 hours (but only 5 when summer time is in operation). Actual flying time to New York is 6 hours. Members of our Frequent Flyer program receive one and a half times the actual mileage when they fly our Silver Arrow class. Testing did not take place until (and only if) subjects had cleared US Customs and Immigration. The average time for this was 2 hours.

Only those students arriving in America in the middle of the day and early afternoon achieve levels of fluency comparable to First and Second Year groups back home. Those students arriving at other local times performed much less well, with the early morning group performing worst of all. Bear in mind that Ss arriving at these hours would themselves have been present at Brussels airport at times of considerable incipient matitutinality. These results show quite clearly that it is Time of Arrival and not Age of Arrival that is the crucial factor in grammar test performance. Using Valdispert’s standard cor-
rection tables, we were able to show a correlation between Score on Test and Circadian Rhythm of .88 (p<.01). (See also Percival 1982, 1983.)

Is it possible that these results might have predicted from a second, theoretical variable, Estimated Time of Arrival (ETA)? Our research shows that there is a significant positive correlation between Projected Circadian Rhythms (calculated on the basis of published ETAs) and Score on Test (r =.62, p<.05), but it is not nearly as strong a predictor as actual Time of Arrival. This seems to be due to the fact that neither predicted take-off times nor predicted arrival times offer anything more than the roughest approximation to reality. In fact even those students arriving at times which should have found them at the peak of their circadian rhythms, had been subjected to the usual delays involving departure and disembarkation procedures, or had experienced the in-flight hell of a prolonged visit to the smoking section aft while waiting for a lavatory door to unfold in their favour.

Conclusions
Given JN’s unimpeachable findings about the acquisition of native fluency on a basis of grammaticality judgement test (where the authors assume that a ‘yes’ or ‘no’ awarded to a sentence, whether written or spoken) is inevitably targeted by the subject at just the grammatical structure the authors were interested in), we must admit by these results to be puzzled be. They seem to support the now discredited ‘older is better’ position (cf. Snow and Hoefnagel-Höhle 1978). Adults in our sample consistently achieve higher levels of fluency than children. Specifically, these results show that adult learners who have yet to arrive in the US initially outperform nominally more proficient adult learners who have already arrive actually in America (and therefore have a Length of Residence value of > 0, as well as an actual Age of Arrival), though this advantage does not last.

What do these results told us about the critical period and second language acquisition? Let us return to JN’s three conclusions above:

1. “… if one is immersed in a second language before the age of 7, one is able to achieve native fluency in the language; however, immersion even soon after that age results in a decrement in ultimate performance” (p 78, emphasis added)

2. “Success in learning is almost entirely predicted by the age at which it begins” (p 81)

3. “For adults, later age of acquisition determines that one will not become native or near-native in a language; however, there are large individual variations in ultimate ability in the language, within the lowered range of performance” (p 81)
In the light of our data, which, like Johnson and Newport's, are derived from the performance of learners asked to judge the grammaticality of a large number of sentences, it will be necessary to re-cast two of JN's three statements listed at the beginning of this article. Statement 1 must be rewritten as follows:

1'. "... if one is immersed in a second language after the age of 12, one is able to achieve native fluency in the language; however, testing even soon before that age reveals a decrement in performance, and testing even earlier leads to tantrums and stains on the carpet"

Statement 2 need not be rewritten, however, except perhaps in a fancier font:

2'. "Success in learning is almost entirely predicted by the age at which it begins."

Statement 3 needs reconsideration (and, like the original, not a little parsing):

3'. "For adults, later age of acquisition does not determine that one will or will not become native or near-native in a language. Adult learners who do not venture abroad or who recover from jet lag are the most likely ultimately to achieve native levels of performance"

Perhaps we could add an additional statement of our own:

4. "Ask 276 silly questions – get a silly answer"3

NOTES

*We would like to thank the first author's former colleagues at the Royal Pedagogical Academy, Brussels, for providing subjects and assisting in the carrying-out of the experimentation reported here. An earlier version of this paper appeared in The Newsletter of the European Second Language Association in 1993. A slightly less early version was rejected by Language Learning after heated debate.

1 Hawkins, Towell, & Bazergui (1993: 190) describe JN as a particularly representative study showing that learners older than seven "fail to achieve full native-like proficiency, even after long exposure to the L2". Long (1990: 271), in his invaluable survey of the age literature, labels JN as 'The least ambiguous evidence to date of maturational constraints operating in the morpho-syntactic domain ...' Strömqvist and Day (1993: 136) claim that 'Johnson and Newport (1989) found substantial support for the critical period hypothesis'.

2 Since the conclusions in Johnson (1992) confirm those of Johnson and Newport, we will only refer to the later paper where it is necessary to distinguish between the two. The reader may assume that the present article refers equally well to both of them.
3 Their paper does raise the question as to whether there may be a critical period in the acquisition of parts of a language, given certain pairings of L1 and L2. For instance, it is conceivable that poor performance on some of JN's items could reflect the fact that where Korean, Chinese and English differ substantially (e.g. in respect of the existence of an article system), successful acquisition may only take place before a certain age.

REFERENCES


Appendix

EXAMPLES OF TEST MATERIALS TAKEN FROM JN (HOW MANY COULD YOU GET RIGHT AFTER 8-10 YEARS OF EXPOSURE IN AN ACQUISITION-POOR CLASSROOM ENVIRONMENT IN BELGIUM?)

Grammatical judgement test, spoken sentences
(n=276, 140 ungrammatical)

1. PAST TENSE
a) Yesterday the hunter shot a deer
b) Yesterday the hunter shoots a deer
c) A bat flew into our attic last night
d) A bat flewed into our attic last night

2. PLURAL
a) The farmer bought two pigs at the market
b) The farmer bought two pig at the market
c) A shoe salesman sees many feet throughout the day
d) A shoe salesman sees many fooots throughout the day

3. PRESENT PROGRESSIVE
a) The little boy is speaking to a policeman
b) The little boy is speak to a policeman

4. DETERMINERS
a) Tom is reading a book in the bathtub
b) Tom is reading book in the bathtub
c) The boys are going to the zoo this Saturday
d) A boys are going to the zoo this Saturday

5. PRONOUNS
a) Susan is making some cookies for us
b) Susan is making some cookies for we

6. PARTICLE MOVEMENT
a) The man climbed up the ladder carefully
b) The man climbed the ladder up carefully

7. SUBCATHERISATION
a) The man allows his son to watch TV
b) The man allows his son watch TV

8. AUXILIARIES
a) Leonard should have written a letter to his mother
b) Leonard should has written a letter to his mother

9. YES/NO QUESTIONS
a) Has the king been served his dinner?
b) Has been the king served his dinner?

10. WH-QUESTIONS
a) When will Sam fix his car?
b) When Sam will fix his car?

11. WORD ORDER
a) The woman paints
b) Paints the woman

12. 3RD PERSON SING.
No examples given. Make your own up.