

## **A Comparison of Two Studies Examining Age Effects on L2 Learning**

**Tess Forest**  
**Nov. 27<sup>th</sup>, 2014**

### Abstract:

The first study, Guijarro-Fuentes and Geeslin's 2003 paper, *Age Related Factors in Copula Choice in Steady State L2 Spanish Grammars*, examined the effect of age on L2 learners eventual command of Spanish copula verbs, *ser* and *estar*. They asked whether any factor related to age (time of acquisition, age at time of study, or time speaking Spanish) had an effect on participants' eventual use of these verbs. (Citations 3)

The second paper, *Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language* also examines age effects on L2 learning. Johnson and Newport's 1989 study asked what effects age of initial exposure to English had on Korean learner's abilities to eventually arrive at a native-like command of their L2. (Citations 2012)

## **i. Quality of the Writing**

The first study, Guijarro-Fuentes and Geeslin (2003), is written in a long-winded and extravagant way, lacking clarity and conciseness, and making it hard for the reader to keep track of the main point. The authors used jargon when simpler words would have gotten the point across more effectively. They spent long sections of the paper explaining things with little obvious relevance to their research, and would have been more effective if they had been more selective in choosing what to include. The paper also lacked organization. For example, it starts out with a seven-page introduction, which serves as more of a literature review than a concise introduction to the particular study. All the expected parts of a research report are included, but the weight given to each section isn't organizationally effective; while the introduction is seven pages long, the conclusion is one paragraph, making it unclear how the studies they discuss in the introduction relate to their study and not providing few connections to further research in the conclusion. The organizational issues in the paper made it hard to stay captivated in what the authors were trying to convey, and created a paper that lacked a clear story. Despite a broad introduction to the subject, and a funneling towards explaining their study, the authors failed to effectively return to the real world applications of their research or convey why their work was interesting.

Johnson and Newport's study, in contrast was generally concisely worded, and conveyed their point clearly. Like the first paper, it was longer than it needed to be, including excessive background in L2 acquisition in an introduction and detailed methods section, which would have been better as an appendix, as to not distract readers from the main question of the study. Although what they did mention was relevant, and clearly linked to their own study, there was an excess of information presented to readers in the introduction, methods, and conclusion sections, which detracted from the main focus of the paper. Despite individual sections that should have been more selective in the material presented, the paper was concise, and overall well structured. Organizationally, the paper made sense. It included all sections one would expect to find in a research paper, and they all made mention of how the topic discussed in their study related to the work they were discussing at the time. The paper's story was funnel-shaped; providing lots of information at the beginning, narrowing it to a particular study, and then zooming back out into the broader implications of their work.

Take, for instance this section of Guijarro-Fuentes and Geeslin's paper:

It will be recalled that the first research question asks whether or not the non-native participants in the study have achieved native-like copula choice. In order to assess whether or not this was the case, the average rates of use for each copula for the native and non-native groups were tabulated and compared statistically, using an  $\chi^2$  test. It will be recalled that each non-native group will be compared to its respective target group. A statistically significant difference in the rate of use for the

two groups (i.e. a correlation between group and copula choice) will constitute evidence that the non-native groups do not use copulas in a native-like manner.

It is wordy and could be rewritten more effectively, as shown below:

Our first research question asked if non-native participants reach a native-like ability in copula choice. To test this, we compared the average rates of use for each copula using an  $X^2$  test. This was done comparing native and non-native groups, making sure that each non-native group was paired for comparison with its respective target/native group. A statistically significant difference in the rate of use for the two groups will suggest that non-native speakers do not use copulas in a native-like manner.

By deleting the instances of “it will be recalled that,” I got rid of unnecessary words, streamlining sentences. This helps readers to see the main point, and not get confused by extra words used in the original excerpt. Changing the verb tense from a combination of past, present, and future to a uniform use of past tense makes this passage less confusing for readers by giving them one less thing to keep track of while reading. Instead of trying to figure out what acts were going on when, they can focus on the content of the sentences. Similarly, by substituting jargon in the original with more approachable language, the readers can focus on understanding the content of the passage instead of the words being used.

## **ii. Quality of the Arguments**

Johnson and Newport’s argument was structured well. At times, terms were not as clearly defined as they could have been. For example, it was ambiguous whether “second language learning” refers to third, or additional, language learning as well. Although they define “second language learning” as “undertaken only after a native language is already acquired,” they could mean an additional language, after a native and second language. The effect that speaking an additional language could have on participants’ performance could be influential. The authors occasionally make assumptions that the reader has background in psychology, and don’t define terms like “plasticity,” which readers without a psychology background might not understand. Additionally, when discussing past work, they make the assumption that congenitally deaf people not exposed to sign language until age 12, are receiving *no* linguistic input until then. Although this might be the case for some, others may have developed homesign systems, or non-standard systems of visual communication with their family members (bottom of p.64). The paper didn’t make errors in logical reasoning, and walked readers through most steps of their logic quite clearly. Although the authors did not directly manipulate variables, or randomly assign participants to condition, the causal link drawn between age of exposure to English and eventual fluency was clear, particularly because the authors ruled out so many alternatives through statistical tests for their significance.

Like the first study, Guijarro-Fuentes and Geeslin's paper also makes assumptions that the reader has some background in their topic, particularly in their introduction. They fail to define all the key terms that need to be clear for full understanding of the study. For example, on page 92 they mention the "variable dependence on existence." They then explain what the variable does, but fail to clearly explain exactly what it is. Throughout the paper, the authors attempt to define terms, but aren't effective in doing so, which could be attributed to a vague writing style, instead of a failure to remember to include definitions. Another major unstated assumption is that participants' judgments on grammaticality ratings in the lab are reflective of their actual behavior when speaking. It is possible that when speaking, the context of their surroundings would change the likelihood that they would use one copula over the other. The paper has one major flaw in reasoning. In their results, they claim the non-significant difference between the American L1 and American L2 speakers "is likely due to too much variation in both groups." Although coming up with an explanation for their finding is helpful to readers, they neglect the possibility that it is simply because there isn't an effect. Like the Johnson and Newport paper, the authors attempted to explain potentially confounding variables through statistical tests, thereby avoiding most faulty causation claims.

As mentioned, the Guijarro-Fuentes and Geeslin's study made the unstated assumption that both L1 and L2 speakers' choice of which copula verb to use on a test in the lab would be representative of that choice in the real world. One way to improve this study would have been to address the reason for why they believe this is the case. They could have said that even if this is the case, they expect that L1 and L2 speakers of Spanish will perform equally differently in the lab than in the real world, or could have addressed the concern by mentioning that there is previous work that supports the idea that this will not have an effect on one group's results but not the other's. Secondly, the paper could have been improved by having clearer definitions of key terms, especially if they used simple wording to define the terms they wanted to use. Although this is particularly important for the introduction, the entire paper could benefit from an elaboration on the definitions of target words. In doing so, the authors would make it easier for people to keep track of their results, by having well-defined terms to link the results to. It would also make their analysis more straightforward for readers to understand.

### **iii. Design of the Study**

As discussed, the Guijarro-Fuentes and Geeslin paper operationalized "command of their L2" as how well L2 learners could perform on a task which asked them to judge the grammaticality of sentences that used copula verbs (*ser* and *estar*). Although motivated by the idea that mastering the use of these verbs is a late appearance in L1 Spanish Learners, their decision neglected to study L2 learners mastering of pronunciation or morphological markers. One important factor that the paper failed to address in its discussion was that English, Portuguese, and Spanish speakers come from varying cultural backgrounds, and could see different traits as being constant vs. changing. A Portuguese and Spanish speaker might have

more similar views about the constancy of characteristics than the English speaker, which could influence the L2 learner's correct copula use in Spanish. Because the study was correlational in nature, the number of experimental conditions was limited to the counterbalance order of the questions, which the experimenters randomized. As mentioned (pg. 94), the data collection varied based on the testing location (Spain vs. the United States). Participants in the US knew they were recruited using academic networks, and in Spain they were recruited using social ones. This could have been a confounding variable, because the L2 speakers might have been less comfortable by feeling like they were taking a test.

Like the first paper, the Johnson and Newport study operationalized "full command of the L2" as how well subjects performed on a test. Although theirs tested syntactic and morphological structures, it still didn't capture whether or not the participants had a full-fledged ability to produce of the phonology or phonetics of their L2. Pronunciation could have been vastly different between early and late-arriving groups, which this operationalization fails to capture. Similarly, this study mirrored the first in its correlational design, and didn't use experimental conditions. One potential confound was the nature of the input that late vs. early arrivals could be getting. If the early arrival group was systematically learning from more informal settings, the type of exposure could account for the variation in performance instead of the age of arrival. Otherwise, major factors were acknowledged, such as the possibility that age of arrival was the determining factor in ultimate control of an L2, instead of age of learning English.

Interestingly, both studies could have benefitted from a change in design to include some aspect of the sound system of the L2 in their definition of "command of a language." They could have done this by having participants judge the pronunciation of recorded speakers, or by having them read sentences out loud, and then having a native speaker code them for correct or incorrect pronunciation. Doing so would have either shown that their results held for all aspects of L2 learning, or shown that once other domains were included, the results were not significantly different between age groups. The Guijarro-Fuentes and Geeslin paper also would have been improved if the researchers had done additional work to determine if the L1 of participants had an effect on how they rated the constancy of certain traits, to ensure those rating aren't a confounding variable.

#### **iv. Consideration of the Alternatives**

Thanks to the lengthy introduction, the first paper did a good job of addressing the reasons for designing the study the way they did. By including a mini-literature review as their introduction, they laid out their reasons for choosing the research design they did. Particularly, lots of attention was given to why they chose the particular syntactic task they did to study competency in their experiment. Overall, the choice of the task was logical, because syntax is learned later than other aspects of a language. But, it could be that South American Spanish and European Spanish don't always use the two copula verbs identically. If L2 learners of Spanish learned it

from a South American speaker of Spanish, their results could be skewed compared to the native Spanish speakers because they were familiar with a different dialect. When examining their results, overall the authors did a decent job of addressing concerns within their results, particularly when discussing the differences between English and Portuguese learners of Spanish. They addressed many of the other possibilities, and effectively explained how they ruled out alternatives to arrive at their conclusion.

Similarly, Johnson and Newport's paper addressed their reasons for doing their research the way they did because of their extensive review of the previous research in the field. By discussing this work, the authors covered many of the possible ways to address their question. They discuss the two studies they know of that examined child-adult differences in L2 performance, and provide a clear link between the short comings they saw there and their study design. This led to one strong aspect of this study's design: the attention they paid to potentially confounding factors, like maturational, motivational, or self-esteem related effects on L2 learning. Additionally, the authors' decision to measure multiple aspects of grammaticality was a well-motivated one, as it ensured that speakers' general command of the L2 was tested. That said, the test still only measured their understanding of morphological and syntactic understanding in their L2, not any phonological or pronunciation task. Johnson and Newport first present their results and then discuss their reasons for attributing certain meaning to their findings. They present multiple potential reasons their numbers could have turned out the way they did, and by doing so, arrive at the most plausible explanation of their results.

That said, the Johnson and Newport paper didn't address all the possible explanations for their results. For example, the authors present a figure (pg. 80) that graphs the relationship between age of arrival and test score for the late-arriving group. They claim there is little correlation because after a critical period your ability to learn a language isn't guaranteed, and your individual abilities become more relevant. But, the authors don't present alternative explanations. For example, it is possible that the late-arrivals received a non-standard treatment from the people they met once moving to the United States, whereas the treatment for people who moved to the United States before a critical age was more uniform. People arriving later, for instance, could have been treated differently based on their job; maybe employers worked harder to teach those who needed English for their profession than those who didn't. This would mean that the graph's randomness had nothing to do with personal differences, but external factors surrounding the late-arrivers. The authors discuss age of exposure to English in school as a main factor to be ruled out in favor of age of Arrival in America, but another possible explanation is that people who left Korea or China later in life had more exposure to formal instruction in their L1. This could have affected participants' ability to learn a new grammar once they arrived in the United States.

## **v. Quality of the Results**

The first paper (Guijarro-Fuentes and Geeslin, 2003) did not have a major impact on existing theory. Surprisingly, it produced somewhat different results than earlier work in the field, but the differences reported between the two non-native groups prevented the study from impacting existing practices. People were not interested in shifting their mindset based on a study that didn't have consistent results, or enough further research on the reason behind these results to explain them fully. The results were not well articulated; the authors' lack of clear wording and inability to account fully for the results made hindered readers ability to understand the study or its results fully. Although Guijarro-Fuentes and Geeslin attempted to make their results maximally generalizable by including multiple languages in their non-native group, their mixed results led to a less generalizable result. Besides the contradictory results, the study was not interesting to the field, because of nearly identical work carried out earlier, and reported in a much clearer way.

The Johnson and Newport (1989) paper, on the other hand, was quite influential, and had a major impact on the mindset of the time. Prior to their research, very few people had studied the pattern of language learning in adults compared to children, and their work produced clearly articulated results. This articulation, both in general writing style and their reporting of results, makes the paper easy for readers to follow despite its length. Although their results were more generalizable than the first papers because of their success in explaining two non-native language groups, the paper mentions homogeneity in their testing population. This homogenation means that many of the test subjects came from similar walks of life, limiting the generalizability of this study to people involved in academia. Because the study examined something that very few people had looked at previously, they were interesting to many people. Additionally, because the results supported a critical period hypothesis this work fit in neatly to an existing theory, which helped to engulf it into the literature.

The first papers research question could have been improved by asking a question that had not already been answered. For example, it could have asked, "what effect does similarity between an L1 and L2 have on eventual command of the L2?" In doing this, they would have been able to address the question of why Portuguese and English participants differed so much in their results. Even if researching this question led to the discovery that there is no effect, their first study's results would be stronger if this research had been done. Alternatively, they could have examined the difference in eventual command of an L2 learned primarily in the classroom vs. learned outside of the classroom, and whether this had a larger impact on children vs. adults. If they had done their study with this edition, they would have been able to attempt a non-direct replication of past studies as well as add to the literature. Alternatively, the researchers could have asked a causal question instead, and manipulated a fictional grammar that children or adults learned in the lab to

address these questions in a way that would have truly allowed them to claim causation.

## **General Conclusions**

Overall, the papers differed the most in the quality of their writing, the quality of their design, and the clarity of their results. Throughout the first paper, Guijarro-Fuentes and Geeslin used long-winded and overly elaborate sentences to convey their ideas. The paper did not do a good job of picking the introductory research well, and included anecdotal information where it detracted from the main focus. Conversely, the Johnson and Newport paper, although equally long, had more relevant examples, and clearer explanations for how the previous research they presented was relevant. In their results and discussion, they avoided discussing things not crucial to the readers understanding of their findings. The quality of their design differed as well; the first paper only tested syntactic competence, the second paper tested morphological understanding as well. Although both studies could have benefitted by including additional domains, the Guijarro-Fuentes and Geeslin paper was more affected. Finally, the clarity of their results also differed. Where Guijarro-Fuentes and Geeslin's study had mismatched results between groups, and lacked a clear explanation as to why, Johnson and Newport's paper presented uniform and significant results between groups.

These differences also impacted the effect the two studies had on the field. The quality of Guijarro-Fuentes and Geeslin's writing made their paper hard for readers to tackle, presumably deterring them from wanting to finish reading it. Their use of jargon also prevented people from clearly understanding the point they were trying to make, and inhibited readers from fully understanding their motivation for, or design of, the study, as well as their results. If readers cannot understand a paper, as in this case, they are unlikely to explain it to others, or discuss the findings as a scientific community. The difference in designs also affected the scientific impact the two papers had. Johnson and Newport's incorporation of morphological competency made their results more wide reaching. This, in turn, meant that more people could use their paper as background research in their own field, be it syntax or morphology. Finally, the fact that Guijarro-Fuentes and Geeslin's results were not consistent with one another, made people more likely to discredit the study as a whole. Johnson and Newport's clear results made it easy for readers to remember them when doing future work in the field. Additionally, their clear results led to further research in the field, whereas Guijarro-Fuentes and Geeslin's led to the need for replication of their own study. In combination, these three factors led to Johnson and Newport's paper having a much bigger impact on the field.



## **Appendix**

### **Paper One:**

Guijarro-Fuentes and Geeslin's, 2003. *Age Related Factors in Copula Choice in Steady State L2 Spanish Grammars*.

First Link, which downloads automatically as a PDF:

[http://scholar.google.com/scholar?q=Age+related+Factors+In+Copula+Choice+In+Steady+State+L2&hl=en&as\\_sdt=0&as\\_vis=1&oi=scholart&sa=X&ei=M2d3VPHEE9bnoASBuoGQAw&ved=0CBoQgQMwAA](http://scholar.google.com/scholar?q=Age+related+Factors+In+Copula+Choice+In+Steady+State+L2&hl=en&as_sdt=0&as_vis=1&oi=scholart&sa=X&ei=M2d3VPHEE9bnoASBuoGQAw&ved=0CBoQgQMwAA)

### **Paper Two:**

Johnson and Newport, 1989. *Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language*.

[http://ac.els-cdn.com/0010028589900030/1-s2.0-0010028589900030-main.pdf?\\_tid=229a453c-746c-11e4-a9b9-00000aab0f26&acdnat=1416897031\\_565e410ee9ee4f703e2777689d2f1653](http://ac.els-cdn.com/0010028589900030/1-s2.0-0010028589900030-main.pdf?_tid=229a453c-746c-11e4-a9b9-00000aab0f26&acdnat=1416897031_565e410ee9ee4f703e2777689d2f1653)