The Relationship between Legibility, Age, and Gender:

Exploring Free-Form Handwriting

Julia Armenta

Department of Applied Linguistics and ESL

Georgia State University

21 March 2016

Abstract

This paper examines the potential relationship between free-form handwriting (i.e., unstructured ways of writing words, phrases, or sentences by hand) and two sociolinguistic variables: the writer's (1) age and (2) gender. While the accuracy of "gender guessing" or gender identification (e.g., Grieve, 2010; Al-Suwaiyan, 2010; Johnson, 2013) has been the focus in previous studies, this exploratory study looks at the potential correlation between numerically scaled legibility of handwritten samples and the observers' identification of the writer's gender and age based on survey data. This type of survey relates to studies of perceptions and attitudes, linguistic markers and identifications, and interlinguistic variation. Clearly, individuals may form judgments and prejudices based on factors such as accents, racial backgrounds, and socio-cultural orientations, but interpretations and judgments of writing samples (especially based on legibility) have not been sufficiently analyzed.

Conveniently sampled participants were given an informal survey in which they were asked to rate several handwriting samples for legibility and the age and gender of writer. All of the handwritten samples were sentence length and devoid of context. To identify if some patterns of responses were only related to English, two of the samples were written in Korean. In line with previous studies, a majority of participants correctly identified the gender (and age) of the samples with high percentages for English. The Korean samples, while still relatively accurate, were less precise, when compared with the English samples. Results of this study followed the expected trend with less legibly scaled samples having a higher percentage of being male

gendered and higher legibly scaled samples having an increased percentage of being female

gendered. Implications related to sociolinguistic perceptions, potential future research, and the

importance of awareness of stereotypes will be discussed during the presentation.

Introduction

Gender in Sociolinguistics is seen as a social variable that individuals use language to be a part of. As Elinor Ochs (cited in Meyerhoff, 2010) states, "Competent members of every community have been socialized […] and can without conscious control orchestrate messages to convey social meanings," in this context she is referring to speech, but it seems very possible that those core ideas could be transferred to written language specifically, handwriting. All children are socialized through gendered norms (although this does not have to correlate from biological sex to one's social identity through gender) in ways of communication. For example most girls will share secrets to create intimacy while boys tend to create competition for dominance. In terms of handwriting, many boys that have neat handwriting have often heard the criticism, "Your handwriting looks like a girls'," and it shows that even children have grouped themselves along gender lines based on handwriting characteristics. How do we know? How are children so quick to associating particular types and styles of handwritings to gender? Do these social norms of handwriting transcend through age or even types of script? Are we actually able to accurately identify a person's gender through only a small sample of writing? What else do we infer about a person just from their handwriting?

Handwriting and the Legibility Bias in Schools

Handwriting is very influential to how people are viewed and sometimes evaluated throughout schooling. The *legibility bias* states that legible handwriting is graded a letter grade higher than less legible handwriting in an academic setting (James 1929). These patterns and results were then replicated in a complicated experiment involving multiple essays of varying quality being copied for content by varying levels of legible handwriting. It did show; however, that while teachers are biased based on legibility, they were not biased due to the gender of the

writer (Greifender 2011). Gender is however an identifiable factor when looking at handwriting. In 2002, Burr came to the conclusion after several experiments that, "the ability […] to judge the gender of the author of a piece of handwriting was significantly better than chance and that this ability improved with practice"(Burr 2002). Burr's methods involved both prompting for gender and not prompting and the participants were only allowed to see the samples for a short period of time to avoid having the gendered decision based on the context of the sample. Participants also seemed to improve at identifying gender with more practice. More interesting though, Burr's results also included descriptions of the samples by the participants and found female samples included words such as, "neat, rounded, and consistent," while male samples included, "irregular, hurried, and scruffy" (Burr 2002). Descriptions even extended into personality traits such as, "confident, arrogant or passive" (Burr 2002).

Handwriting as a Social Identifier

Handwriting as a social identifier is not only common in the Western world but also has been the focus of studies in other cultures making use of written scripts. In Japanese the main interest has been on the individuality each person has in regards to handwriting which greatly depend on character class (Ueda 2009). This study by Ueda was very systematic as it measured distances of variance between multiple writers character samples. The idea of individuality through handwriting is fairly interesting as some handwritings are so clear that teachers are able to identify their own students through only their handwriting with enough exposure. The sense of individuality through handwriting seems very believable as even if two people have similar handwriting styles, it is likely with enough practice that participants could distinguish between the two (Burr 2002). Accuracy of gender identification has also been studied with regard to bilinguals of Urdu and English and showed that there were differences between what is

considered feminine and masculine between Urdu and English and therefore bilingual speakers used one set of social norms to assess the gender of English samples while another set of norms for Urdu samples(Hamid 1996). This stereotype of gendered handwriting extends past the Western World and is even different in terms of male and female handwriting in society. These differences in handwriting script styles is what this study was fueled by, to see if participants would be able to produce the same amount of accuracy in gender selection from English as to Korean without that exposure to the norms of gendered handwriting in Korean.

The focus of this exploratory research study is to draw connections and establish potential relationship between free-form handwriting (i.e., unstructured ways of writing words, phrases, or sentences by hand) and sociolinguistic variables especially the writer's (1) age and (2) gender. While the accuracy of "gender guessing" or gender identification (e.g., Grieve, 2010; Al-Suwaiyan, 2010; Johnson, 2013) has been examined in previous studies, this paper looks at the potential correlation between numerically scaled legibility of handwritten samples and the observers' identification of the writer's gender and age based on survey data. This type of survey relates to studies of perceptions and attitudes, linguistic markers and identifications, and interlinguistic variation. Some implications are provided in the concluding sections of this paper.

Methods

A thirteen question, electronic survey with writing samples in different languages were given to conveniently-sampled volunteer participants through social media. The first part of the experiment involved collecting handwriting samples. Between the balancing act of having a survey long enough to collect enough information and short enough to entice more participants; the survey consisted of four sentence length handwriting samples. This convenience sample was made available online and through the use of social media as a tool to reach as many participants

as possible in only four days (Friginal & Hardy 2014). Unlike Burr's study, the participants would have as long as they wanted to look at the samples so the samples used sentences that were of a "genderless context", or a sentence that could be written by either gender without having an indication of which gender actually wrote the sample. Two of the samples were in English and the other two were in Korean. All four samples had the same questions attached to them, "How legible would you consider this sample?" "How old is the person who wrote the sample?" and "What gender would you consider the person who wrote the sample?" (See Appendix 1 for full survey). The legibility question was given a scale numbered 1-5: 1 being very messy and 5 being very neat. The term "legible" was avoided in the scale itself due to the fact that all of the samples were by definition legible and the words "messy" and "neat" are easier to associate with.

While participants were given the choices, "Male, Female, and Other," with regard to their own gender, when selecting the gender for the samples only the choices "Male" and "Female" were given. This was to ensure that a clear choice was made and to prevent participants from selecting other as an "I don't know," answer. Also the study is not focused on forcing ideas upon male and female but more so to understand what is considered masculine or feminine in terms of handwriting.
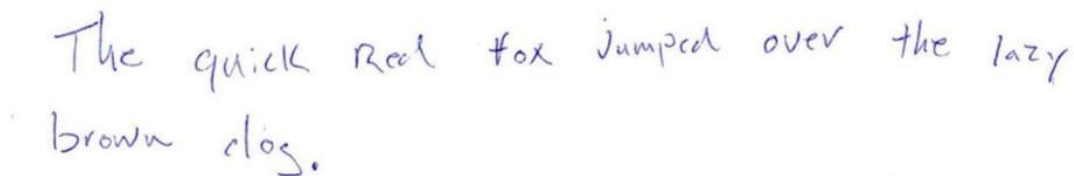
The quick Red fox jumped over the lazy brown dog.

Figure 1. Handwriting Sample A: English Male

Figure 1 was the first Handwriting Sample on the survey and was written by a 20 year-old male. The text was provided on paper where it was then scanned and turned into a digital image. This sample was chosen as it is legible and understandable but it is also clearly not a computer font.

The quick brown fox jumps over the lazy dog.

Figure 2. Handwriting Sample B: English Female

Figure 2 was the second sample in the survey and was written by a 19 year-old female. While originally the two English sentences were meant to be uniform in text, time dictated that the sentences remain as is. When the samples were requested, the sample writers were asked to write the sentence that contained all letters of the alphabet instead of providing them a sentence for them to copy. Since the samples remained genderless in context it suited the needs of the experiment. This sample was chosen because when compared to Sample A the scripts are distinguishable from one another.

간장 공장 공장장은 장 공장장이고
된장 공장 공장장은 강 공장장이다 .

Figure 3. Handwriting Sample C: Korean Female

Figure 3 was the third sample and the first of the Korean samples and was written by a 19 year-old female who is a native speaker of Korean. Participants were not given any prior knowledge that Korean would be a part of the survey (not in promotion, directions, or questions in the survey). This may have come as a shock to participants who potentially have never seen Korean before. Korean was chosen both out of convenience of obtaining the samples and because out of many scripts that are not Latin based, Korean is very different. It is structured in syllable blocks that give the language a very "organized" feel. Of course when handwritten the language does show many variations as with any script, but Korean seemed like a good "unfamiliar" script to use. The sentence is a Korean tongue twister that translates to English as: The manager of soy sauce factory is named Chang, and the manager of soy paste factory is named Kang.
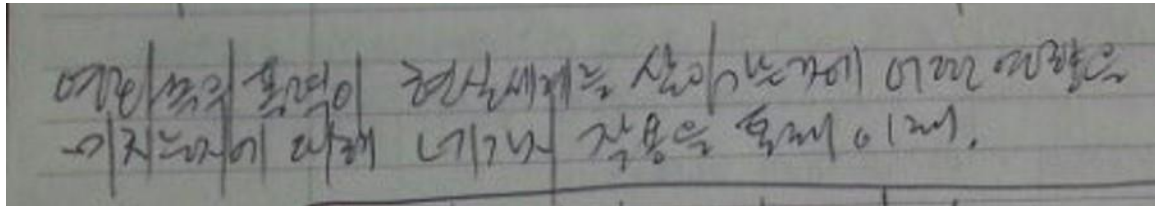

Figure 4. Handwriting Sample D: Korean Male

Figure 4 was the last sample in the survey written by a 46 year-old man who is a native speaker of Korean. This sample is the least cleanly presented as it was a picture taken from Christian Bible notes (not actual Bible quotes but comments on the readings). As stated earlier, Korean handwriting can end up straying from its block-like formation quite drastically and this sample shows that very clearly. This sample was chosen for its unorganized appearance and therefor, for the purposes of this experiment, it was suitable.

The participants were selected through convenience and the only people unqualified to be a participant would be the four individuals that gave their handwriting samples or those unable to read and comprehend the directions. While it was not asked on the survey, it may be noteworthy to know that people all over the world were participants. There were a handful of Germans, Americans, Koreans, and people from Singapore.

Table 1. Description of participants' age and gender

Gender	10-18	19-29	30-49	50+	Total
Male	6	30	8	12	**56**
Female	8	82	31	22	**143**
Other	0	1	0	0	**1**
Total	**14**	**113**	**39**	**34**	

As seen in Table 1, the vast majority of participants were between the ages of 19-29 and the Female group was largely over represented. Since the survey was promoted through social media and convenience it is not surprising that most of the participants were females of that age as it was a personal social media page and most of the people on it were females. While there is the large outlier of females and 19-29 year-olds, the rest of the sampling was fairly cohesive across the ages.

Results

In terms of accuracy, the results followed in the steps of previous research. Adding to the gender inference, age was also a point of interest as shown in the following tables.

Table 2. Accuracy in Gender & Age Selection

	A	B	C	D
Age	52.50%	67.50%	38%	21%
Gender	91%	95.50%	53%	76.50%

As seen in Table 2, participants were far more accurate when presented with the English handwriting samples (A and B) than the Korean sample (C and D). Across all of the samples age was far less accurate than gender. More than the findings of Burr who found that participants accurately selected the gender "significantly better than chance" (Burr 2002), the results in Table 2 show 95% for the English female and 67% of accurate age selecting. The results of accuracy seem very language specific, which is unsurprising. While it was not a part of the survey, most of the participants would have little to no exposure to the Korean language and may have felt hesitant about assigning a gender to an unfamiliar script. It is noteworthy to mention, that while the participants were mostly female, both men and woman participants were similarly accurate in their decisions for the English samples in terms of percentages.

The Korean samples showed higher accuracy percentages (minimally) for male participants. Participants were also divided into two other groups: liberal arts majors or non-liberal arts majors. Surprisingly, those with math, science, and business backgrounds were about 20% more accurate in identifying both age and gender on the English samples and 10% more accurate on Korean samples. It is not clear if this result shows anything truly noteworthy, but seeing as participants who were Linguistic majors or majors of a foreign language were less

successful as a group at identifying features from foreign and domestic script is something

interesting if nothing else.

The main focus of the survey was less on accuracy and more on the legibility selection in

relation to the gender. The anticipated results were that 1-2 on the legibility scale would be

connected to the male gender while 4-5 would be female, 3 would be in between the two.

Table 3. Relationship between Legibility Scale and Gender
Selection

	A	B	C	D
1-2:Male	54	0	13	109
1-2:Female	3	0	12	24
3: Male	106	1	30	26
3:Female	11	14	25	13
4-5:Male	22	8	51	18
4-5:Female	4	177	69	10

Table 3 shows the exact number of participants who chose a particular set of options. For Sample

A, there were 54 participants who chose either a 1 or 2 on the legibility scale along with

assigning that Sample as Male; only 3 participants assigned a 1 or 2 to female. These results

were fairly conclusive on all samples except for Sample C, the Korean female. The results, while

showing a general trend, are very closely split between genders. At first it appears to be a Korean

script problem, but Sample D, the Korean male, has no such split. While Sample D was not as

clear in its correlation as A or B, it did have over half the participants in the 1-2: Male category.

It does seem less clear cut as originally predicted as the trend shows that the participants were

more inclined to give Male's a 4-5 than giving Females a 1-2.

In regards to a 3 on the legibility scale, participants tended to give Male's more 3's than the Females in terms of the English Samples. The Korean samples are both fairly close in terms of participant's gender selection.

Summary and Conclusions

As exploratory data from this research shows, participants were relatively accurate in deciding gender when presented with English samples but much less accurate with Korean samples. Deciding age was much less conclusive but it seems like the participants were able to correctly select the age for the English samples with some accuracy. This cannot be said for the Korean samples however, as accuracy for both age and gender dropped significantly in comparison to the English samples. This could be due to a sheer lack of samples and comparable data but there is at least reason to believe that inferring gender from handwriting samples is less accurate in foreign scripts than a Latin based script.

There also does appear to be a correlation between a legibility scale and the gendered choice. The correlation results followed a predicted outcome. Less legible handwriting was associated more with males while neater handwriting was associated more with females and the "in between" handwriting tended to lean closer to male for English samples and split down the middle for Korean samples. It did show that participants were either more comfortable with judging English (or at the very least a Latin script) than with Korean. In particular the female Korean sample seemed almost unidentifiable for participants as it was ranked in all legibility numbers and across both genders. This could be due to that particular handwriting being hard to

identify (as the Korean male did have at least have a majority), but it does leave unanswered questions as to why it was so evenly split across age and gender.

Overall the English side of the study provided very clear results that reaffirmed many previous research and opened a door in a correlation between legibility and gender. While the foreign language aspect of the study was less conclusive, with some expansion, it could possibly be more informative. It could be said that this study is merely the first in many more to fully explore all the ideas and notions regarding handwriting identification between genders.

Limitations and Future Directions

This study was far from a perfect one at this stage. It could be easily be pointed out that there has been no distinction between participants that could read Korean and those that could not. This condition could have been fixed with one more question on the survey asking if participants could read and write Korean. It also would have been interesting to know the native language of participants along with all the languages they know up to an elementary level. Also, it would be interesting to make the legibility scale slightly larger, 1-7, with 1 being "un-readable or non-legible" and 7 being "computer-font like or extremely legible". It is important to have a middle number though, because seeing where the "average legible" samples fall in terms of gender lines is particularly interesting.

As if to say that there is a gender "neutral" style of writing that is not neat enough to be female but not messy enough to be male (if one abides by the correlation discussed in this paper). It could also be interesting to see how closely related participants find the sample handwriting to their own; however, this would have to be narrowed down to only the languages the participants could read and write. The study could also improve from more samples with more varying age

ranges. The study was lacking in expandable data due to the absence of more than 2 females and 2 males. Perhaps samples of all age ranges and genders would bring more conclusive results.

The primary limitation would, of course, be time. Participants were only gathered for four days and the survey was promoted through a personal social media page and convenience samples. If allowed more time to better construct and present the survey, perhaps results would have shown better conclusions. Sample sentences should also be made uniform if seeking a genderless context as to further alienate the samples from seeming random. They also need to stand alone with good visual detail and not look like they were photographed samples such as Sample D was. This unprofessional sample made the survey feel thrown together and less cohesive because it was not uniform in presentation to the others.

For future research, this study would need to include an identification of how legible the participants consider their own handwriting and their native language along with others. It would be beneficial to add more languages that used Latin script and those that do not, (German, Dutch, Spanish, Mandarin Chinese, Japanese, and Arabic). Of course, it would need to utilize findings from Hamid and understand the different expectations of what is considered neat and messy across different scripts in order to understand the stereotype, this could be done through a short interview with native speakers. This interview would also give several descriptive words about what is considered male handwriting and female handwriting in their culture and script. The participants might also create and use their own words to describe the handwriting to show another relationship between legibility, gender, and description words as too enhance on Burr's study. The study could then compare words from the interviewed participant and see if there are similarities between the descriptions for male and female.

Another interesting venue to peruse would be sexual orientation of the participant that provided handwriting samples, either explicitly prompted or discovered through descriptive words like, "gay" or "homosexual". If asked explicitly this idea could be its own research study and already leaving out the "other" category for identifying gender seems gray in terms of ethical practice. Perhaps for future research this would be avoided all together by using terms like masculine and feminine to describe the samples. In this way those who do not self-identify with male or female may feel more represented through this study.

While content has also been heavily veered away from in the samples, it would be interesting to see how heavily content actually effects the selection of gender or even age. If the same participant (a male for example) were to write three samples, a genderless context, a male context ("I was fixing the oil leak in my car the other day and found that the coolant was low,"), and a female context ("I was out shopping the other day and we decided to stop to get our hair done,"); it would be interesting to see the difference in gender selection based on the samples context. Perhaps in the descriptions of the samples written by males with a female context there would be comments of "awkward" or "off putting" due to the strange mixture of masculine handwriting and female content. The contexts would need to be more subtle than, "My name is Sarah," so perhaps they would need to rely on stereotypical norms for males and females in the United States.

This one study seems to be going in two different directions. One direction explores all of the English handwritings with gender, sexuality, and context. The other explores the foreign scripts and their different social norms of "masculine" and "feminine" and how Westerners view those differences from their own filter. Perhaps this study is better off involving many different surveys and many different participant groups as to better organize results and present

conclusions. With this in mind it seems as though this study was a good first step in a series of

more rigorous studies with larger scopes and greater numbers of participants.

References

Alsuwaiyan, L. (2015). Understanding patterns of written scripts. Unpublished manuscript.

Georgia State University.

Burr, V. (2002). Judging Gender From Samples Of Adult Handwriting: Accuracy And Use Of

Cues. The Journal of Social Psychology, 142(6), 691-700.

Friginal, E., & Hardy, J. (2014). Corpus-based sociolinguistics: A guide for students. New York:

Routledge.

Greifeneder, R., Zelt, S., Seele, T., Bottenberg, K., & Alt, A. (2011). Towards a better

understanding of the legibility bias in performance assessments: The case of gender-

based inferences. British Journal of Educational Psychology, 82, 361-374.

Grieve, J. (2010). Markers of social factors in free-form handwriting: Implications for

variationist studies. Journal of International Communication, 3(1), 67-81.

Hamid, S., & Loewenthal, K. (1996). Inferring Gender from Handwriting in Urdu and English.

      The Journal of Social Psychology, 136(6), 778-782.

James, H. W. (1929). The effect of handwriting upon grading. The English Journal, 16, 180–185.

Johnson, J. (2013). Dominance and social standards. Lingua, 34, 23-44.

Meyerhoff, M. (2010). Indexing Gender. In The Routledge sociolinguistics reader (pp. 483-497).

      London: Routledge.

Ueda, K., Matsuo, K., & Schwid, B. (2009). Experimental Analysis Of Individuality Of Japanese

      Handwriting. International Journal of Pattern Recognition and Artificial Intelligence,

      23(5), 869-885.

Appendix

Appendix 1: Survey

Handwriting Analysis

Please read and answer all questions. Make sure to read the Handwriting Samples carefully.

* Required

How old are you? *

o   ○   10-18

o   ○   19-29

o   ○   30-49

o   ○   50+

What is your biological gender? *

o   ○   Male

o   ○   Female

o   ○   Other

What was/is your major in college? *
i.e. business, education, music, arts, N/A

Handwriting Sample A

The quick Red fox jumped over the lazy brown dog.

How legible would you rate this sample? *
Handwriting Sample A

	1	2	3	4	5	
Very messy	○	○	○	○	○	Very neat

How old is the person who wrote this sample? *
Handwriting Sample A

o   ○   5-15

o   ○   16-23

    o ◯   24-45

    o ◯   46+

What gender is the person who wrote this sample? *
Handwriting Sample A

    o ◯   Male

    o ◯   Female

Handwriting Sample B

The quick brown fox jumps over the lazy dog.

How legible is this sample? *
Handwriting Sample B

	1	2	3	4	5	
Very messy	◯	◯	◯	◯	◯	Very neat

How old is the person who wrote this sample? *
Handwriting Sample B

    o ◯   5-15

    o ◯   16-23

    o ◯   24-45

    o ◯   46+

What gender is the person who wrote this sample? *
Handwriting Sample B

    o ◯   Male

    o ◯   Female

Handwriting Sample C

How legible would you consider this sample? *
Handwriting Sample C

	1	2	3	4	5	
Very Messy	○	○	○	○	○	Very Neat

How old is the person who wrote this sample? *
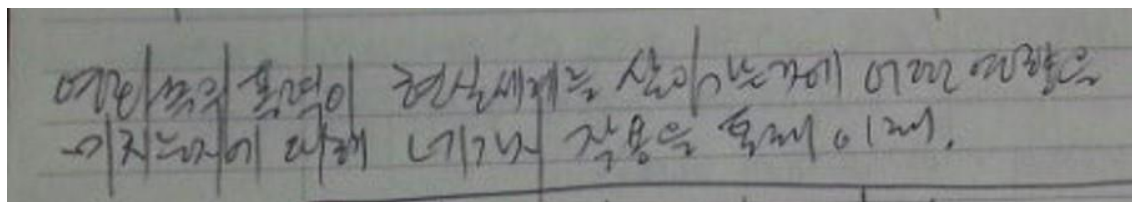Handwriting Sample C

- ○ 5-15
- ○ 16-23
- ○ 24-45
- ○ 46+

What gender is the person who wrote this sample? *
Handwriting Sample C

- ○ Male
- ○ Female

Handwriting Sample D

How legible is this sample? *
Handwriting Sample D

	1	2	3	4	5	
Very Messy	○	○	○	○	○	Very Neat

How old is the person who wrote this sample? *
Handwriting Sample D

- ○ 5-15

- ○ ○ 16-23
- ○ ○ 24-45
- ○ ○ 46+

What gender is the person who wrote this sample? *
Handwriting Sample D

- ○ ○ Male
- ○ ○ Female