FACTORS INVOLVED IN THE ASSESSMENT OF STUTTERING IN A FOREIGN LANGUAGE

Marjan Cosyns \textsuperscript{a}, Jóhanna T. Einarsdóttir \textsuperscript{b}, John Van Borsel \textsuperscript{a}

\textsuperscript{a} Ghent University, Ghent, Belgium
\textsuperscript{b} University of Iceland, Reykjavík, Iceland
Due to increased mobility and globalization, speech-language pathologists nowadays are more and more confronted with clients who speak a language different from their own mother tongue.

The assessment and treatment of persons who speak a foreign language poses particular challenges.
Two central questions with regard to the assessment of stuttering are:

1. How well are clinicians able to make reliable and valid judgments about the presence of stuttering in languages other than their own?
2. Which are the determining factors in such judgments?
Thus far, only a few studies looked into the assessment of stuttering in a foreign language.

Most of these focused on the importance of familiarity with the language to be judged.

Different languages and different conclusions
Introduction

* Humphrey (2004)
  * English/Spanish
  * familiarity makes no significant difference

* Van Borsel & Britto Pereira (2005)
  * Portuguese/Dutch
  * language familiarity does, at least to some extent, influence stuttering judgment
Einarsdóttir & Ingham (2009)

- US English/Icelandic
- Stuttering judgments do not necessarily rely on disfluency types that might assume familiarity with the speaker’s language, but also that language familiarity cannot be completely excluded as a basis for identifying stuttering.
Lee, Robb, Ormond, and Blomgren (2014)

- English/Spanish
- A language familiarity effect for judgment of finer characteristics of the disorder.
Introduction

* Hoffman, Wilson, Copley, Hewat, & Lim (2014)
  * Australian English/Mandarin
  * a language familiarity effect
In all of the above studies investigating the role of language familiarity:

- language familiarity conceived of as a binary phenomenon, i.e. participants were considered to be familiar with the language of the speaker(s) to be judged or not.

- more naturalistic to recognize the existence of various degrees of familiarity.
Other possible factors

- Closeness of the language to be judged to the mother tongue.
  - when the mother tongue is more remote from the foreign language judging stuttering is more difficult
  - there is especially a higher risk for false positive identification.
- Van Borsel, Leahy, and Britto Pereira (2008)
Stuttering severity
- identifying stuttering in a foreign language may not be so difficult when the client is a severe stutterer
- Studies?

Experience and training of the clinician
- clinicians with many years of experience who have seen many clients and hence are better at judging stuttering in their native language, may also be better at judging stuttering in a foreign language
- Studies?
Purpose

* To further explore the factors that play a role when assessing stuttering in a language that is not one’s native language.

* In contrast to previous studies we looked at several factors in an attempt to unravel a possible interplay and the relative importance of each factor.

* Factors thus considered were
  * experience of the clinician with assessing stuttering
  * stuttering severity of the client
  * familiarity of the clinician with the client’s language
  * closeness of the language of the client to that of the clinician.
Method
A listener experiment conducted during a meeting of the Stuttering Research and Education Network (STUREN) overall, participants had to judge stuttering severity in different languages.
Participants

* 19 individuals, 2 males and 17 females, ranging in age between 25 and 66 years (M = 40.2 years, SD = 13.7 years)

* from five different countries: Iceland, Sweden, Norway, Finland, Belgium

* Participants were divided into two groups:
  * experienced raters (i.e., at least 5 years working with stuttering clients)
  * non-experienced raters (i.e., students and individuals working with stuttering clients for less than 5 years).
Speech samples (1)

* videotaped spontaneous speech samples (n = 8)
* prepared by the participants of the STUREN meeting
* featuring persons who display mild to moderate stuttering or severe stuttering
* recorded in a therapy setting
* showing the clients’ face and shoulders against a not intrusive background
* various languages (Icelandic, Swedish, Norwegian, Dutch)
## Speech samples (2)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Language</th>
<th>Gender</th>
<th>Age (y;m)</th>
<th>Severity</th>
<th>Duration (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Icelandic</td>
<td>M</td>
<td>8;5</td>
<td>Mild to moderate</td>
<td>3,6</td>
</tr>
<tr>
<td>2</td>
<td>Norwegian</td>
<td>M</td>
<td>9;7</td>
<td>Mild to moderate</td>
<td>3,6</td>
</tr>
<tr>
<td>3</td>
<td>Swedish</td>
<td>M</td>
<td>9;6</td>
<td>Severe</td>
<td>6,7</td>
</tr>
<tr>
<td>4</td>
<td>Dutch</td>
<td>M</td>
<td>11;4</td>
<td>Severe</td>
<td>9,0</td>
</tr>
<tr>
<td>5</td>
<td>Icelandic</td>
<td>M</td>
<td>6;0</td>
<td>Severe</td>
<td>5,4</td>
</tr>
<tr>
<td>6</td>
<td>Norwegian</td>
<td>M</td>
<td>17;9</td>
<td>Severe</td>
<td>8,9</td>
</tr>
<tr>
<td>7</td>
<td>Dutch</td>
<td>M</td>
<td>8;11</td>
<td>Mild to moderate</td>
<td>4,0</td>
</tr>
<tr>
<td>8</td>
<td>Swedish</td>
<td>M</td>
<td>4;11</td>
<td>Mild to moderate</td>
<td>2,2</td>
</tr>
</tbody>
</table>
Speech samples (3)

* Stuttering severity of the samples

* calculated as the mean of scores on a 10.0 cm visual analogue scale given by the native experienced raters.
Procedure (1)

- Samples were presented using a laptop (DELL, Latitude E5520) connected to a projection screen and multimedia theater speakers (OZAKI).
- Participants were informed that they were going to judge stuttering severity in different languages without being told what the languages were.
- Each sample was presented twice in succession followed by an interval to complete an answering sheet.
Procedure (2)

* For each sample, participants first filled out a series of questions that assessed their familiarity with the language.

* They subsequently rated stuttering severity on a 10.0 cm visual analogue scale.

* Additionally, they were asked to score how easy they found it to rate stuttering severity, also on a 10.0 cm visual analogue scale.
Procedure (2a)

1. Is this a sample of yours? □ yes □ no
2. Is this sample in your mother tongue? □ yes □ no
3. Which language do you think this is? ...
4. How sure are you that you know which language it is?
   □ It is a pure guess. □ I think I could be □ I am rather sure. □ I am absolutely sure.
   correct.
5. Have you ever heard this language before?
   □ never □ seldom □ regularly □ often
6. Are you able to understand this language?

- [ ] not at all  
- [ ] a bit  
- [ ] reasonably  
- [ ] quite well

7. Are you able to speak this language?

- [ ] not at all  
- [ ] a bit  
- [ ] reasonably  
- [ ] quite well

8. Are you able to read this language?

- [ ] not at all  
- [ ] a bit  
- [ ] reasonably  
- [ ] quite well

9. Are you able to write this language?

- [ ] not at all  
- [ ] a bit  
- [ ] reasonably  
- [ ] quite well
10. How severe is the stuttering of this person according to you? (put a cross)

- no stuttering at all
- very severe stuttering

11. How easy did you find it to score the severity of this sample?

- very easy
- very difficult

12. Did you understand what was said in this sample?

- not at all
- a bit
- reasonably
- quite well
Outcome variables
- (1) accuracy with which the participants rated stuttering severity
- (2) reported ease of rating stuttering severity

Independent variables
- (1) experience of the raters
- (2) stuttering severity of the samples
- (3) familiarity with the languages shown
- (4) closeness of the languages to the own mother tongue.
Accuracy:

- The absolute value of the difference between the stuttering severity of the samples and the stuttering severity rating made by the participants.

- Thus, lower scores indicate that participants’ ratings were more in line with the scores given by the native experienced raters.
Ease:

- was based on the numerical value of the visual analogue scale.

- The higher the score, the more difficult participants experienced rating stuttering severity.
Experience:

- on the basis of the number of years working with stuttering clients
- Experienced: at least 5 years working with stuttering clients
- non-experienced: students and individuals working with stuttering clients for less than 5 years
Stuttering severity:

- according to the judgment of the participant who delivered the sample
- mild to moderate stuttering / severe stuttering
Familiarity:

- calculated from the language questions in the answering sheet

- language familiarity scores could range between 0 and 18, with higher scores representing more familiarity with a language
Closeness was established based on genetic classification:

- Score “0”: the language of a sample was in the participant’s mother tongue
- Score “1”: Closeness of languages from the same subgroup (such as Icelandic and Norwegian)
- Score “2”: languages came from different subgroups, but shared the same group (for instance Icelandic and Swedish)
- Score “3”: languages from different groups, but sharing the same branch (for instance Icelandic and Dutch)
- Score “4”: languages belonged to a different branch (for instance Icelandic and Finnish).
Figure 1. Genetic classification of languages (Huffman, 2014; van Bree, 1996)

West Sc. = West Scandinavian

East Sc. = East Scandinavian
Results

* Accuracy of rating stuttering severity
  * experienced raters
  * Non-experienced raters

* Ease of rating stuttering severity
Results

* Accuracy of rating stuttering severity: *experienced judges*

  * A significant difference in accuracy of rating stuttering severity between mild to moderate and severe samples (Mann-Whitney U test, $p = .004$).

  * Therefore, analysis of the relation between accuracy and familiarity and accuracy and closeness was separately conducted for mild to moderate and severe samples.
Boxplot of accuracy of stuttering severity scores given by the experienced raters in function of sample severity.
One statistically significant correlation emerged, i.e. accuracy positively correlated with closeness for mild to moderate samples only.

So, for mild to moderate samples, the following applies: the closer the language to the mother tongue, the more accurate an experienced rater will be.
## Results

<table>
<thead>
<tr>
<th></th>
<th>Familiarity</th>
<th>Closeness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy for mild to moderate samples</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_s$</td>
<td>-0.113</td>
<td>0.334</td>
</tr>
<tr>
<td>$p$</td>
<td>.488</td>
<td>.035*</td>
</tr>
<tr>
<td><strong>Accuracy for severe samples</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_s$</td>
<td>-0.103</td>
<td>0.111</td>
</tr>
<tr>
<td>$p$</td>
<td>.527</td>
<td>.494</td>
</tr>
</tbody>
</table>
Results

* **Accuracy of rating stuttering severity: non-experienced judges**

  * As in experienced raters, the difference in accuracy of rating stuttering severity between mild to moderate and severe samples was statistically significant (Mann-Whitney U test, $p = .001$)

  * Analogously to previous analyses, mild to moderate and severe samples were analyzed separately.
Boxplot of accuracy of stuttering severity scores given by the non-experienced raters in function of sample severity.
However, accuracy did not significantly correlate neither with familiarity nor with closeness.

<table>
<thead>
<tr>
<th></th>
<th>Familiarity</th>
<th>Closeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy for mild to moderate samples</td>
<td>$r_s$ = 0.107</td>
<td>$p = 0.535$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$r_s = 0.013$</td>
</tr>
<tr>
<td>Accuracy for severe samples</td>
<td>$r_s$ = 0.077</td>
<td>$p = 0.657$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$r_s = -0.104$</td>
</tr>
</tbody>
</table>
Results

- Ease of rating stuttering severity
  - results were the same for experienced and non-experienced raters
  - rating stuttering severity of severe samples was found to be as difficult as rating mild to moderate samples (Mann-Whitney U test, p = .135 for experienced raters and p = .367 for non-experienced raters).
  - Hence no distinction was made between mild to moderate and severe samples
Results

* All correlations were statistically significant.

* Thus, the more familiar with/closer to the mother tongue, the easier participants experienced rating stuttering severity.
## Results

<table>
<thead>
<tr>
<th></th>
<th>Familiarity</th>
<th>Closeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease for experienced raters</td>
<td>$r_s$ 0.318</td>
<td>0.229</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.004**</td>
<td>0.042*</td>
</tr>
<tr>
<td>Ease for non-experienced raters</td>
<td>$r_s$ 0.445</td>
<td>0.536</td>
</tr>
<tr>
<td></td>
<td>$p$ &lt;.001**</td>
<td>&lt;.001**</td>
</tr>
</tbody>
</table>
Rating stuttering severity in a foreign language

1. Accuracy is determined by severity of the stuttering. In severe stuttering rating more in line with that of native speakers, and this irrespective of the experience of the rater.
2. In experienced raters: the closer the language to the mother tongue, the more accurate an experienced rater will be in rating samples with mild to moderate stuttering.
3. No such pattern is seen in non-experienced raters.

Summing up (1)
4. Ease of rating stuttering severity does not depend on stuttering severity but depends on familiarity and closeness to the mother tongue. The more familiar with/closer to the mother tongue, the easier rating stuttering severity is experienced.
Discussion (1)

* Studies on the assessment stuttering in a foreign language almost all focused on language familiarity

* However, language familiarity is not the only factor involved.

* The picture is more complex

* Also severity of the stuttering, closeness to the mother tongue and experience of the rater play a role
The participants in the present study had to rate stuttering severity and they knew that all the samples they were going to rate were of people who stutter.

Therefore, the findings of the present study should not be generalized to situations where the presence and frequency of particular symptoms of stuttering in a foreign language has to be assessed or to situations where it is has to be determined whether or not a person speaking a foreign language presents with disfluencies that represent stuttering.

To what extent the patterns outlined above hold for these situations remains to be investigated.