Is Swedish more beautiful than Danish? – A matched-guise investigation.

Charlotte Gooskens, Nanna H. Hilton & Anja Schüppert

1. Introduction

It has repeatedly been suggested that Swedes have less positive attitudes toward the Danish language, culture and people, than vice versa (Delsing & Lundin Åkesson 2005; Gooskens 2006). It is tempting to assume that the asymmetric attitudes found in previous investigations should be explained by imposed norms and social connotations. In the present paper we will show experimental evidence for the fact that inherent values also play a role in the aesthetic evaluations of the two languages.

2. Previous studies

In the literature, the three mainland Scandinavian languages, i.e. Danish, Swedish and Norwegian, are often mentioned as examples of so-called ‘Ausbausprachen’, that is, languages which are separate standard languages only for political and historical reasons. Linguistically they are so closely related that they may be considered as dialects of one language (Kloss 1967). However, the fact that there is no common Scandinavian standard language makes the communication between Scandinavians different from communication between speakers of, for example, Dutch dialects. In the Dutch language area, speakers of different dialects in general turn to Standard Dutch when they talk to each other. When speakers of the Scandinavian languages meet, they usually communicate in their own language.

Speakers of the Scandinavian languages are strongly encouraged by the Scandinavian authorities to use their own language rather than a lingua franca such as English when communicating with other Scandinavians. Among other incentives this was secured by the Sprakkonvention (‘language convention’), released by the Nordic Council of Ministers in 1981. It was implemented in 1987 and its goal was to ensure that citizens of the
Nordic countries are entitled to use their native language in written communication with authorities in other Nordic countries. In the past, a number of studies were carried out in order to get a precise picture of the actual level of understanding between speakers of the Scandinavian languages (e.g. Maurud 1976, Bø 1978, Börestam Uhlmann 1991, Delsing & Lundin Åkesson 2005, Schüppert & Gooskens 2011). The results of these investigations invariably show that the spoken communication between Swedes and Danes is especially problematic and that the intelligibility is asymmetric: Danes understand Swedish better than Swedes understand Danish.

One of the explanations that have been put forward to explain this asymmetry in intelligibility is a corresponding asymmetry in attitudes towards the neighboring languages. The existence of negative attitudes, or social stigmas, attached to languages is often seen as a potential obstruction for successful intergroup communication. It has repeatedly been suggested that the asymmetric intelligibility between Swedes and Danes can be traced back to less positive attitudes among Swedes towards the Danish language, culture and people (Delsing & Lundin Åkesson 2005; Gooskens 2006). But it has not been possible, so far, to establish whether negative attitudes are a result of poor intelligibility, or whether poor intelligibility is a result of negative attitudes, caused by some other factor.

Haugen (1966) elicited language attitudes in a survey on mutual intelligibility in Scandinavia, and found that 42% of the Danish participants thought that Swedish sounded more beautiful than their own language, while none of the Swedish participants thought that Danish sounded more beautiful than Swedish. Delsing & Lundin Åkesson (2005) tested the intelligibility of the three Scandinavian languages in a large number of places in the Nordic countries and also asked the participants to complete a detailed questionnaire. Among other things, the participants were asked two questions about their attitude towards the neighboring countries and languages: 1. ‘Do you think that language x sounds beautiful?’ 2. ‘Would you like to live in country x?’ The results showed that Danish subjects considered the Swedish language more beautiful than Swedish subjects considered the Danish language. However, the asymmetry of the attitude scores was reversed when the same subjects were asked if they would like to live in the neighboring country. Danes were less willing to move to Sweden than vice versa. Schüppert & Gooskens (2011) investigated auditory word recognition of the neighboring language among Danish and Swedish children and adolescents in a picture pointing task. In addition they also assessed the subjects’ attitudes towards the neighboring language by asking whether the language they had heard during the experiment sounded (1) less nice than, (2) as nice as, or (3) nicer than their native language. The children generally had neutral attitudes while the adults were more negative. 40% of the Danish adults and 84% of the Swedish adults chose ‘less nice than my native language’ and 25% of the Danish adults and 5% of the Swedish adults chose ‘nicer than my native language’. So, again, we see that the Danes are more positive
toward Swedish than Swedes are toward Danish; we also see that Danes are less positive toward their own language than the Swedes are.

A major methodological problem with the listening experiments in the investigations discussed above is that the Swedish and Danish speech fragments were recordings of different persons. Aesthetic evaluations of languages may be affected by individual speaker characteristics such as voice quality, mean pitch level and intonation. In order to avoid this methodological problem a larger number of speakers could be used for the listening experiment. In this way effects of variability between speakers could be averaged out (‘verbal guise’). The speakers should be homogenous in age and gender. Another way to neutralize the influence of voice characteristics on the aesthetic judgments is to use the ‘matched-guise’ technique. This technique was first developed for the investigations of language attitudes in the French-English bilingual setting in Quebec, Canada (Lambert, Hodgson, Gardner & Fillenbaum 1960).

A matched-guise test consists of lexically identical speech samples from a balanced bilingual speaker (i.e. a bilingual with equally high proficiency levels in both languages). The recordings of the bilingual are played interspersed with other recordings (distracters) to avoid listeners being aware of hearing the same speaker twice. Listeners are then asked to evaluate the speakers that they are hearing for different personality traits such as kindness, richness and beauty. Since the two varieties spoken by the bilingual are in fact produced by the same speaker, the results of the evaluations can be taken as indirect measures of language attitudes, language usage being the only feature between the two recordings that differs. Lambert et al. (1960) found that the participants’ judgments of personality traits of the bilingual speaker were strongly influenced by the language spoken. Both English and French-speaking participants rated the speakers more positively on status and solidarity traits when they spoke English, which is believed to reflect the English language’s higher status in Quebec. A matched-guise experiment also allows for direct questions about the languages, such as how beautiful a language sounds to the listener. In this case, too, different evaluations of the two guises must be based on characteristics of the language used by the bilingual speaker, rather than on speaker characteristics.

Schüppert et al. (2015) conducted a matched-guise experiment with recordings of a balanced bilingual speaker of Danish and Swedish. Groups of Danish and Swedish children between 7 and 16 years judged the Swedish and Danish recordings and four other languages on 5-point semantic differential scales indicating how normal, beautiful, smart, modern, kind and rich the speakers sounded to them. The results showed that the bilingual speaker was judged more positively when she spoke the listeners’ own language than when she spoke the neighboring language. Furthermore, the speaker was rated more positively by the Danes when she spoke Swedish than by the Swedes when she spoke Danish.
The investigation by Schüppert et al. (2015) thus confirms the results of previous Danish-Swedish attitude investigations. By applying the matched-guise technique we know for sure that the asymmetric attitudes are not caused by differences in voice quality. Since the listeners heard the same text in the two languages, the asymmetry could not be due to the content of the recording. The question that remains to be answered is where the aesthetic perceptions of linguistic features come from. Giles, Bourhis & Davies (1975) suggested two possible answers, termed the imposed-norm hypothesis and the inherent-value hypothesis. The imposed-norm hypothesis stresses the importance of non-linguistic factors such as social connotations and cultural norms. A language variety would be considered attractive when its speakers are socially privileged. This would explain why English listeners locate Received Pronunciation (RP or BBC English) at the top of the aesthetic hierarchy, regional English accents in the middle, and urban English accents at the bottom (e.g. Giles 1970; Trudgill & Giles 1978; Milroy & McClenaghan 1977). RP would be placed at the top because of cultural prestige, whereas regional accents are judged more positively than urban accents because the former are associated with a more attractive lifestyle and environmental setting. The imposed-norm hypothesis can only apply to the evaluations by persons having the relevant sociolinguistic knowledge. Laureys (1993) summarizes the Danish stereotype as ‘ambivalent, ironic and self-relativizing’ and the Swedish stereotype as ‘reserved and rational [.....] feel[ing] superior to the rest of the world’ (Laureys 1993: 15). Sweden has for a long time been referred to as the ‘big brother’ in Danish public opinion (Thorvaldson 2009, Sletten 2004). This label points to the fact that Sweden has previously been more influential than the other Scandinavian countries. Sweden’s role in Scandinavia has been regarded as that of a stereotypical older brother: arrogant, annoying and somewhat boring, but also successful, influential, and economically stable. This may be a reason for Danes to be more positive about Swedish than Swedes about Danish and for Danes to be less positive about their own language than Swedes.

The inherent-value hypothesis claims that attitudes to language are (at least partly) triggered by qualities that are intrinsic in language (Garrett 2010: 228). It argues that some languages (or language varieties) are intrinsically more aesthetically pleasing due to their sound characteristics than other languages. The definition of ‘inherent values’ refers to values that are not socially or culturally imposed. This includes universal values of language characteristics that are found cross-culturally, such as individual sounds or classes of sounds but some researchers also include values that depend on the language of the listener (e.g. Van Bezooijen 2002).

Positive evidence in favor of the inherent-value hypothesis has been found by Van Bezooijen (1996). She asked Dutch lay subjects to aesthetically evaluate a number of European languages. Phoneticians made global phonetic descriptions by rating the same languages on phonetic scales. The attributed degree of beauty proved largely predictable from a combination of judged ‘melodiousness’ and ‘softness’. Also fast
tempo, precise articulation and fronted articulation were positively correlated with the aesthetic judgments. Although it is unknown to what extent ratings are based on social connotations by the listener the correlational study suggests that aesthetic evaluations of languages may indeed have a phonetic basis, that includes both segmental and suprasegmental features. However, so far the inherent-value hypothesis has not been tested experimentally.

In this paper we look for experimental evidence for the fact that the inherent-value hypothesis plays a role in the aesthetic evaluations of Swedish and Danish. We will present the results of a follow-up study of the investigation by Schüppert et al. (2015). As discussed above, it is tempting to assume that the asymmetric attitudes found in this investigation should be explained by imposed norms and social connotations. However, on the basis of the results it is not possible to exclude that inherent values may also play a role.

3. Method

This study investigates evaluative reactions to recordings (a matched guise test) of a Danish-Swedish bilingual speaker to a German as well as a Chinese group of listeners from Oldenburg and Chonqing Jiaotong University, respectively. We assume that both groups of listeners will act as outsiders and therefore not be affected by imposed norms and social connotations in the way that Scandinavian listeners would be.

3.1. Stimulus material

Speakers

For our matched-guise material we recorded a bilingual Danish-Swedish speaker. The speaker is a young female Dane who has grown up in Southern Sweden but consistently speaks Danish with her Danish parents and siblings at home. Much care was taken to choose a balanced bilingual speaker, i.e. a speaker that spoke both Danish and Swedish in such a way that it sounded native to speakers of both languages. To test this we created two so-called ‘voice parades’. We presented native listeners with five recordings of native speakers, one by the bilingual speaker and four by distracters, and instructed them to pick out one speaker that sounded non-native (forced choice). We assumed that if the bilingual speaker was not chosen as the foreigner more often than on chance level, he or she sounds native.

Two voice-parades were conducted, a Danish and a Swedish one. Five recordings were presented to 30 Danish and 15 Swedish listeners. For the Danish version, the four distracter recordings were produced by native female Danish speakers from the greater Copenhagen area, the same geographical area that the bilingual hailed from. The distracter recordings in the Swedish version were all recordings of female speakers
from Southern Sweden. In both voice parades, the bilingual speaker was presented as the third speaker of five. The results of the voice-parades are shown in Table 1, which demonstrates that the bilingual speaker was not judged as sounding less native than the distracters. In the Danish voice parade, the bilingual speaker was selected by none of the listeners as having a foreign accent; in the Swedish voice parade she was chosen by 10% of the listeners, which is still clearly below chance level. Table 1 demonstrates that the recordings of the bilingual speaker were not rated significantly less native sounding than the other recordings by neither Danish nor Swedish listeners. We therefore conclude that our speaker sounds native both when speaking Danish and when speaking Swedish. Crucially, our speaker speaks standard Swedish rather than Southern Swedish regiolect: She used a an alveolar flap [ɾ] instead of a voiced uvular fricative [ʁ], the tonal distinction between accent 1 and accent 2 followed the standard Swedish pattern and no rising diphthongs were produced, such as [eu] for /u:/ and [aa] for /ɑː/.

Table 10 Results of the voice parades for the Danish-Swedish bilingual speaker. Grey shaded cells indicate speakers that were picked at or above chance level. Note that each distracter represents two different speakers, i.e. a Danish and a Swedish one.

<table>
<thead>
<tr>
<th>Bilingual</th>
<th>Distracter 1</th>
<th>Distracter 2</th>
<th>Distracter 3</th>
<th>Distracter 4</th>
<th>Chance</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Swedish listeners</td>
<td>3 10 5 17</td>
<td>0 0</td>
<td>13 43</td>
<td>9 30</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>Danish listeners</td>
<td>0 0 0 0</td>
<td>3 20</td>
<td>12 80</td>
<td>0 0</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

Text

The text used in the matched-guise experiment consisted of six sentences from the children’s book *Can’t you Sleep, Little Bear?* by Waddell & Firth (2005). In addition to recordings of Danish and Swedish, which were made by one and the same speaker, recordings of the same sentences, translated specifically for this study, were also made by bilingual speakers of the following languages: Finnish, German, Norwegian, Dutch (three different recordings), Frisian and Danish.¹ In total, the stimulus material thus comprised ten different audio fragments representing six different languages. The recordings were 22.97 and 23.97 seconds long for Danish and Swedish, respectively. The distracter fragments in the test ranged between 19.56 and 31.55 seconds in length, see Table 2.

¹ In fact our study contained 5 different matched guises. Attitudes towards Frisian and Dutch are presented in Hilton & Gooskens 2013.
Table 2 The Languages and the mean durations of the recordings of the Danish-Swedish bilingual (speakers A) and the speakers who read the distracter fragments (B to E).

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Language</th>
<th>Duration (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Danish</td>
<td>22.97</td>
</tr>
<tr>
<td></td>
<td>Swedish</td>
<td>23.97</td>
</tr>
<tr>
<td>B</td>
<td>Danish</td>
<td>20.63</td>
</tr>
<tr>
<td></td>
<td>Norwegian</td>
<td>19.56</td>
</tr>
<tr>
<td>C</td>
<td>Dutch</td>
<td>26.50</td>
</tr>
<tr>
<td></td>
<td>German</td>
<td>31.55</td>
</tr>
<tr>
<td>D</td>
<td>Dutch</td>
<td>28.85</td>
</tr>
<tr>
<td></td>
<td>Finnish</td>
<td>30.27</td>
</tr>
<tr>
<td>E</td>
<td>Dutch</td>
<td>26.39</td>
</tr>
<tr>
<td></td>
<td>Frisian</td>
<td>30.45</td>
</tr>
</tbody>
</table>

3.2. Procedure

All fragments were played to the listeners twice with an inter-stimulus interval of 6 seconds. The ten recordings were presented in two different orders to the listeners. 85 Chinese and 35 German informants heard the Danish recording as the first fragment and the Swedish recording as the 6th, while 56 Chinese and 38 German informants heard the test with the Swedish recording played 5th and the Danish played 10th. T-tests conducted within the German-speaking listener group (N = 73) shows that there are no significant differences in attitude ratings between listeners of the different playing orders (Danish ratings: $t(58) = -0.16, p = .88$; Swedish ratings: $t(58) = 1.60, p = .12$).

The listeners were provided with rating questionnaires consisting of semantic differential scales (Osgood, Suci & Tannenbaum 1957). They were asked to evaluate their opinion on a five-point scale where two bipolar adjectives were extreme values. The adjective pairs were ‘old-fashioned – modern’, ‘stupid – smart’, ‘unattractive – attractive’, ‘strange – normal’, ‘unfriendly – friendly’ and ‘poor – rich’. These adjectives can be classified into the three categories dynamism (‘old-fashioned – modern’ and ‘strange – normal’), attractiveness (‘unattractive – attractive’ and ‘unfriendly – friendly’) and superiority (‘stupid – smart’ and ‘poor – rich’) following the framework for language attitude testing in Zahn & Hopper (1985). Figure 1 is an image of an English translation of the questionnaire used in the investigation.

After completing the language evaluations, informants were asked to provide biographical information regarding their age, gender, region of origin, academic background, language(s) spoken at home and brief language learning histories.
3.3. Participants

In total 73 German-speaking and 141 Mandarin-speaking listeners participated in the test. All informants were either students at a higher education institution or pupils in the two highest classes of secondary schools. The German listeners (33 males and 40 females) all lived in Oldenburg, approximately 320 kilometers from the Danish border and were native speakers of German. Five German informants reported using an additional language at home, but no Scandinavian language. They were on average 17.6 years old (ranging between 16 and 20 years). The Chinese listeners (41 males and 100 females) were from The Chongqing Jiaotong University in the Nan’an district in central China and were native speakers of Mandarin. Their mean age was 19.6 years (ranging between 17 and 24 years).

4. Results

4.1. German listeners

To investigate whether the attitudes towards Danish are less positive than towards Swedish also outside Scandinavia we first carried out an experiment similar to the one reported by Schüppert et al. (2015), but this time with a German group of listeners.

Overall results

Similarly to the study reported in Schüppert et al. (2015), the data was coded by assigning the lowest score (‘1’) to the judgments ‘strange’, ‘ugly’, ‘stupid’, ‘old-fashioned’, ‘unkind’, and ‘poor’, and the highest score (‘5’) to the judgments ‘normal’, ‘beautiful’, ‘smart’, ‘modern’, ‘kind’ and ‘rich’. The remaining scores were given for any of the points between the extremes, which means that we interpret the semantic differential scale as a linear scale.

Figure 2 shows the mean ratings on seven scales by the German listeners of the bilingual speaker when she spoke Swedish (black bars) and Danish (white bars). The bar graphs show that the speaker received higher scores for all judgments when she spoke
Swedish than when she spoke Danish, except for the judgment ‘modern’. For example, the same speaker is on average judged as being smarter and more beautiful when speaking Swedish than when speaking Danish.

![Figure 2](image-url) Mean ratings on seven scales by the German listeners when the bilingual speaker spoke Swedish (black bars) and when she spoke Danish (white bars).

**Do the evaluations measure seven different attitudes? – data reduction with PCA**

To test whether the bilingual speaker is rated significantly more positively when she speaks Swedish than when she speaks Danish, we reduced the data by conducting a principal component analysis (PCA) on the overall ratings on the six personality traits ‘normality’, ‘beauty’, ‘smartness’, ‘modernity’, ‘kindness’ and ‘richness’, and on the judgment of the beauty of the language. These seven variables served as input for the analysis.

The PCA revealed that most of the seven variables were significantly interrelated. The two exceptions were ‘modernity’-‘kindness’ and ‘modernity’-‘richness’. The correlation coefficient for ‘modernity’-‘normality’ is only marginally significant ($p = .05$). This suggests that the anticipated modernity of a speaker might be an extracted component in its own in a PCA. The correlation coefficients and their significance levels are displayed in Table 3. It can be seen from the table that not all input variables are significantly interrelated. In addition, the correlation coefficients never exceeded $r = .67$. The determinant for the data is 0.17, which is greater than the necessary value of 0.00001 (Field 2005), and suggests that multicollinearity is not a problem.
Table 3  Correlation matrix of the seven input variables into the PCA, indicating the correlation coefficient and the significance level (*p < .001, **p < .005, ***p < .01, ****p < .05).

<table>
<thead>
<tr>
<th></th>
<th>Modernity</th>
<th>Smartness</th>
<th>Beauty (person)</th>
<th>Normality</th>
<th>Kindness</th>
<th>Richness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smartness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beauty (person)</td>
<td>.25***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normality</td>
<td>.14*</td>
<td>.25****</td>
<td></td>
<td>.56****</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kindness</td>
<td>.07</td>
<td>.28****</td>
<td>.40****</td>
<td>.30****</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Richness</td>
<td>.10</td>
<td>.39****</td>
<td>.23***</td>
<td>.17*</td>
<td>.15*</td>
<td></td>
</tr>
<tr>
<td>Beauty (language)</td>
<td>.17*</td>
<td>.42*****</td>
<td>.67*****</td>
<td>.47****</td>
<td>.40****</td>
<td>.40****</td>
</tr>
</tbody>
</table>

The anti-image correlations between all seven variables were sufficiently high (all $r \geq .72$). For all seven variables taken together, Bartlett’s test of sphericity resulted in $\chi^2(21) = 244.53$ ($p < .001$), indicating that our dataset is not an identity matrix. The Kaiser-Meyer-Olkin measure of sampling adequacy was .79, which is a sufficient value according to Kaiser (1974). These descriptive results indicate that a principal component analysis is an appropriate way of reducing the data.

As indicated by the low correlation coefficients for ‘modernity’ with some of the remaining personality traits, the PCA extracted two principal components. The first component had an eigenvalue of 2.98, which indicates a sound data reduction. The second component had an eigenvalue of 1.02 and was therefore included as well. All subsequent components were excluded since their contribution to the model was not sufficient. To illustrate this, Figure 3 displays the seven eigenvalues for the maximum number of extractable components as a function of these seven components (components 1 to 7). It can be seen that the ‘elbow’ of the graph is located at the second extracted component. This suggests that this component might be excluded from further analysis (Field 2005), but since its eigenvalue is above 1, we decided to include it into the model. Components 3 to 7 were excluded.
Figure 3  Scree plot showing eigenvalues per component for the German listeners.

After extraction, we applied a Promax rotation, because we found that the ratings for most input variables were significantly correlated (19 of 21 correlation coefficients were significant).

The correlation coefficients between the two extracted components and the seven input variables after Promax rotation are given in Table 4. The correlation between the relevant extracted component is moderate for ‘smartness’ ($r = .68$) and high for the variables (all $r \geq .70$). The extracted components can therefore be assumed to represent the seven input variables well. Component 1 seems to measure ‘attractiveness’ while component 2 seems to measure ‘status’. The two components explain 57% percent of the variance in the data.

Table 4  Component matrix with correlation coefficients between the two extracted principal components and the scores on the seven input variables after Promax rotation.

| Beauty (person) | .81 |
| Normality      | .81 |
| Kindness       | .76 |
| Beauty (language) | .71 |
| Richness       | .73 |
| Modernity      | .72 |
| Smartness      | .68 |

Now the data from the seven input variables have been reduced to two components that seem to represent two rather independent fields of personality traits, namely ‘attractiveness’ and ‘success’. These components consist of standardized values (z-scores), which means that the mean value for all 60 listeners is 0 and the standard deviation for all listeners is 1. The two components form the basis of the remaining
analyses and represent the ratings of all seven judgments of the language of the bilingual speaker.

*The difference in Germans’ ratings of Swedish and Danish language*

The question was whether the German participants rate the bilingual speaker significantly more positively when she speaks Swedish rather than Danish. To investigate this, a pairwise t-test was conducted on the two extracted components, i.e. on the z-scores.

The test confirmed our hypothesis that the speaker is judged as being more attractive ($t(139) = 3.35, p = .001$) when she spoke Swedish ($M = 0.27$) than when she spoke Danish ($M = -0.27$). In addition, the speaker was perceived as being more successful ($t(139) = 2.46, p = .02$) when she spoke Swedish ($M = 0.20$) than when she spoke Danish ($M = -0.20$).

**4.2. Chinese listeners**

The results of the ratings by the German listeners show that there is an asymmetry in attitudes toward Danish and Swedish even outside Scandinavia. However, by choosing a group of listeners that is likely to recognize the two languages that they are asked to judge we are not able to exclude the influence of stereotypical attitudes towards the languages. We wanted to investigate whether the asymmetry is due to imposed norms and stereotypical attitudes toward the languages only or whether any of the two languages have inherent qualities that sound aesthetically more pleasing to people in general. To examine this we needed to test a group of listeners who did not know the languages and thus could not have developed stereotypical opinions about the languages. Therefore, we repeated our experiment with a group of Chinese listeners.

**Overall results**

The Chinese data was coded in the same way as for the German listeners reported in section 3.1. Figure 4 shows the mean ratings of the bilingual speaker given by the Chinese listeners on the seven scales when she spoke the Swedish (black bars) and Danish (white bars). It can be seen that the speaker received higher scores for all judgments when she spoke Swedish rather than Danish.

---

1 Note that by having converted the positive scores from a 5-point-scale into z-scores, the mean of all scores is 0 and the standard deviation is 1.
Do the evaluations measure seven different attitudes? – data reduction with PCA

Again, we reduced the data by conducting a PCA on the overall ratings on the six personality traits and on the judgment of the beauty of the language. The PCA revealed that all seven variables were significantly interrelated with correlation coefficients never exceeding $r = .67$. Anti-image correlations between all seven variables were even slightly higher than for the German listeners (all $r \geq .85$). Bartlett’s test of sphericity resulted in $\chi^2(21) = 711.64$ ($p < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy was .88, which is a more than sufficient value (Kaiser 1974).

The analysis extracted one principal component with an eigenvalue of 3.82. The second component had an eigenvalue of 0.86. Figure 5 displays the seven eigenvalues as a function of the seven potential components. Just as for the data from the German listeners, the ‘elbow’ of the graph is located at the second extracted component, which led us to exclude all subsequent components and conduct further analyses on the first extracted component.

Figure 4 Mean ratings of 7 judgments by the Chinese listeners when the bilingual speaker spoke Swedish (black bars) and when she spoke Danish (white bars).
The extracted component correlates very highly \((r > .80)\) with the two scores for ‘beauty’ (language and person), highly with the scores for ‘richness’, ‘modernity’, and ‘normality’ (all \(r > .70\)), and moderately with the scores for ‘smartness’ and ‘kindness’ \((r > .50)\). The correlation coefficients are given in Table 5. This extracted component can therefore be assumed to represent the seven input variables well. The extracted component explains 54.6% percent of the variance in the data and served as independent variable in the subsequent analysis.

<table>
<thead>
<tr>
<th>Extracted Component</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Beautiful (language)</td>
<td>.85</td>
</tr>
<tr>
<td>Beautiful (person)</td>
<td>.80</td>
</tr>
<tr>
<td>Rich</td>
<td>.77</td>
</tr>
<tr>
<td>Modern</td>
<td>.76</td>
</tr>
<tr>
<td>Normal</td>
<td>.75</td>
</tr>
<tr>
<td>Smart</td>
<td>.68</td>
</tr>
<tr>
<td>Kind</td>
<td>.52</td>
</tr>
</tbody>
</table>

By conducting a PCA, the data from the seven input variables has again been reduced to one component representing ‘attractiveness’ and consisting of z-scores. This component forms the basis of the remaining analyses and represents the ratings of the bilingual speaker when she speaks Danish and Swedish with regard to six personality traits.

The difference in Chinese listeners’ ratings of Danish and Swedish
To test the hypothesis that the bilingual is rated significantly more positively when she speaks Swedish than when she speaks Danish, a pairwise $t$-test was conducted on the extracted component, i.e. on the z-scores. It revealed that the bilingual speaker was rated significantly ($t(127) = 5.34, p < .001$) more negatively when she spoke Danish ($M = -0.30$) than when she spoke Swedish ($M = 0.30$). This means that even in the case where the listeners do not know the languages and therefore do not have any preconceived opinions about them, the attitudes towards Swedish are more positive than towards Danish. This must be explained by characteristics of the languages themselves since imposed norms and speaker characteristics cannot have influenced the judgments.

5. Discussion and further research

The results presented in the previous section show that listeners from outside of Scandinavia have similar attitudes toward Danish and Swedish as the speakers who are very familiar with these languages. Both German and Chinese listeners judged Swedish more positively than they judged Danish when it came to the seven evaluation scales. It has generally been assumed that Danes are more positive towards Swedish than Swedes towards Danish because of extra-linguistic factors such as imposed norms and social connotations. We cannot be sure that this is not the case for the German listeners. However, only 2 of the 73 Informants could correctly identify the Swedish sample as such, and only 3 informants correctly labelled the Danish sample as Danish, but most informants did not fill in an answer to this question. The Chinese listeners did not identify the languages they heard and must have based their judgments on the recordings themselves. This investigation therefore provides clear evidence that inherent language characteristics can play a role in aesthetic evaluations.

A question that still remains to be answered is what these characteristics are and whether they are in fact universally pleasing, or whether these evaluations also come about through socialization. Previous indications that inherently pleasing language characteristics exist have been found for different linguistic levels (Van Bezooijen 1996). Examples of potential relevant features are syllable structure (e.g. presence or absence of consonants clusters), rhythm (regular alternation of accented and unaccented syllables), pitch level, contour and variation, tempo, tonality or place of articulation (in the front or the back of the mouth). Danish and Swedish are known to be prosodically rather different. Danish is often referred to as a monotonous language while Swedish is a pitch accent language with a more lively intonation as word tones can have lexical meaning (Bruce 1977, Elert 1972, Gårding 1977). It can be hypothesized that monotonous speech is generally considered less attractive than intonationally varied speech. Experimental research with listeners from varying language backgrounds (Dutch, British, Kenyan, Mexican, and Japanese) has shown that a ‘lively’ manner of speaking, with varied pitch patterns, a great number of pitch movements, a wide pitch range, and
many stresses, is cross-culturally strongly associated with positive personality characteristics, such as willpower, self-confidence and openness (Van Bezooijen 1988). These vocal stereotypes of personality might generalize to the aesthetic evaluation of languages.

To find evidence that the prosodic differences between Danish and Swedish are indeed at least part of the explanation for the findings in the present investigation we had a closer look at the Danish and Swedish recordings used for our experiment. To investigate whether Danish in general is a more monotonous language than Swedish we extracted the pitch contour of the two recordings. Measurements showed that the mean pitch is almost the same for the two languages (240.1 Hz for Swedish and 238.2 Hz for Danish), but the standard deviation is larger for Swedish (46.3 Hz) than for Danish (37.6 Hz). The larger standard deviation in Swedish could give a universal impression of a more lively manner of speaking and this may result in more positive judgments.

A way to test the role of intonation for the perception experimentally would be to present the same groups of German and Chinese listeners with monotonized versions of the Swedish and Danish recordings together with monotonized distracters. If intonation plays an important role in the explanation of the different attitudes towards Danish and Swedish the differences are expected to disappear or become smaller when the judgments are based on monotonous versions where there are no differences in pitch range or number and nature of pitch movements. It would also be a possibility to present versions to the listeners where the Danish intonation has been put on the Swedish recording and vice versa. Attitude judgments towards the Danish recording with Swedish intonation should be more positive than towards the Swedish recording with Danish intonation if intonation is the decisive factor in the different judgments.

We offered the listeners the possibility to give comments on their evaluations. Some of the comments show that not only the intonation may cause the negative evaluations of the Danish speaker but that also other segmental factors may play a role. Here are a few examples of comments by the Germans about the Danish recording: ‘ungewohnte Sprache’ (unusual language), ‘merkwürdige Silben’ (strange syllables), ‘einfach absurd’ (simply absurd), ‘klingt wie ein betrunkener Schwede’ (sounds like a drunk Swede). Even though these comments do not show precisely what characteristics of the Danish language course these impressions, they do seem to be rooted in the sounds of the language. It is possible, for example, that the Danish glottal stop (sted) and the fact that Danish exhibits particularly many reduction phenomena and consonant vocalizations contribute to the negative impression of Danish.

Our results have provided evidence for the inherent-value hypothesis, i.e. that there are qualities intrinsic to language that can trigger attitudes towards them. A next step would be to find out whether these qualities are in fact universal, or also created by social connotations in listeners. As explained in the introduction, the inherent characteristics of languages can be found cross-culturally but may also depend on the
language of the listener. Studies on aesthetic evaluations in other scientific areas have found evidence for positive effects of familiarity. In our case, it is possible that linguistic characteristics of Swedish cause more positive evaluations both among the German listeners and the Chinese listeners. However, it is also possible that Swedish sounds aesthetically pleasing to listeners speaking Chinese (a tonal language) because it is a pitch accent language. In that case the judgments by the German listeners may still be based on stereotypes and social connotations since German is not a tone language. To test whether this is the case we should repeat our experiment with a group of listeners who will not recognize the languages (like the Chinese listeners) and speak a language without tones (like German).

Our results suggest that Swedish is a more beautiful language than Danish, at least to some groups of listeners. However, on the basis of our results it is not possible to say whether Swedish is a more beautiful language compared to other languages or whether Danish is an aesthetically displeasing language. To draw such a conclusion we would need to carry out matched-guise experiments with a larger number of language pairs including Danish and Swedish as one of the guises. In fact recordings of a Norwegian-Danish balanced bilingual were used as distracters in our experiment (see Table 2). An analysis of the judgments of this speaker showed that in this case Danish and Norwegian are judged as equally beautiful. This provides us with evidence that it is Swedish that is a beautiful language and not Danish that is an ugly language. More controlled experiments must be set up to draw stronger conclusions about the beauty of the Danish and the Swedish language.

Bibliography


Bibliografie em. prof. dr. Godelieve Laureys

Boeken


**Boekhoofdstukken**


offered to Harry Perridon (Amsterdam contributions to Scandinavian studies 8). Amsterdam: Universiteit van Amsterdam, 113–138.


Artikelen in tijdschriften


