

ARTICLE

Is Swedish more beautiful than Danish? Matched guise investigations with unknown languages

Nanna H. Hilton^{1*} , Charlotte Gooskens¹ , Anja Schüppert¹  and Chaoju Tang²

¹Center for Language and Cognition, University of Groningen, The Netherlands;

Emails: n.h.hilton@rug.nl, c.s.gooskens@rug.nl, a.schueppert@rug.nl

²School of Foreign Languages, University of Electronic Science and Technology of China, Chengdu, Sichuan, P. R. China; Email: 1006946669@qq.com

*Email for correspondence: n.h.hilton@rug.nl

(Received 22 May 2020; revised 11 March 2021; accepted 11 March 2021; first published online 01 July 2021)

Abstract

Are some languages universally seen as beautiful? And if so, what are the linguistic traits that make some languages sound more pleasing than others? This paper addresses these two questions. We do so with results from two listener experiments that use speech samples from a bilingual Danish–Swedish speaker in a matched guise test, where the listeners are ‘previously unexposed’ students from central China. Our results indicate that listeners from Central China with no previous exposure find Swedish more pleasing sounding than Danish. This finding provides evidence that there could be features of language that sound more beautiful to listeners cross-culturally. In a follow-up experiment we remove the intonation contours of the speech to see whether this prosodic trait plays a role for evaluations. The results show that the difference in evaluations between Swedish and Danish disappears when both speech samples are monotonised. We discuss the importance of our findings for language attitudes research.

Keywords: Danish; language attitudes; prosody; Swedish

1. Introduction

Language attitude research shows that people have strong and consistent aesthetic associations with different languages. This appears from empirical experiments, qualitative investigations, as well as from surveys: We hold attitudes towards languages and accents other than our first. Speakers of English, for example, often perceive Romance languages like Italian, French and Spanish to be examples of beautiful languages, while German or Arabic are seen as less attractive (Giles & Niedzielski 1998). Since the advent of experimental research into language attitudes in the 1960s, a number of hypotheses have been put forward about how our views of out-group speech varieties are created and developed.

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

In theoretical approaches to language and social psychology it is assumed that the vocal cues in speech activate stereotypical information for the listener, and that the social traits associated with the stereotype are then attributed to the speaker (e.g. Dovidio & Gluszek 2012:93). It is generally agreed, then, that attitudes toward languages other than our own reflect previous experience with the variety, or knowledge about the variety imposed upon the listener (see Edwards 1999 for an overview). Most previous language attitude studies find support for this argument, formulated in the imposed-norm hypothesis (Giles, Bourhis & Davies 1975) for the creation of language attitudes. It seems agreed upon that the development of evaluative responses towards language cannot occur in a social vacuum, and that evaluations are created and become dynamic in a social context (e.g. Giles & Ryan 1982).

The question whether language attitudes could be expressed without any contact with speakers, i.e. purely on the basis of traits that are inherent to the language was formulated early on as the inherent-value hypothesis (Giles et al. 1975). The argument behind this hypothesis is that some languages (or language varieties) could be intrinsically more aesthetically pleasing due to their sound characteristics only. Previous studies have found little evidence that certain linguistic varieties are universally, or inherently, more pleasing than others. Yet, indications exist in literature that there are attitudes that can be said to fall between the scope of the two hypotheses above, i.e. that there are certain linguistic traits that can elicit evaluative responses without listeners having been previously exposed to the variety in question. Such factors can include, for example, features that make the signal processing more difficult on the side of the hearer, and that may hence be experienced, universally, as displeasing (see Dragojevic & Giles 2016).

The basic process of stereotype creation starts when we group traits and features that other humans display, and make generalisations on the basis of these. Upon meeting with a new person, the recognition of certain traits can activate a set of other expected traits that a social category is thought to have. We know that language plays a role both as a trait that we form opinions about, as well as the instrument that allows us to label the different groups and sets of characteristics that we create (see Burgers & Beukeboom 2020). An interesting question is what happens when we are met with a foreign speech signal for the first time, whether listeners activate stereotypes, and if so, which ones. It is possible that listeners can still have preferences for specific characteristics of a language, for example due to a similarity with features in another, familiar, variety they have been previously exposed to.

Studies of language attitudes and speech intelligibility have found consistently that there is an interaction between how much we understand of a language and how positive we are towards the speakers of the variety as well as the variety itself (e.g. Boets & De Schutter 1977, Van Bezooijen 1994, Schüppert, Hilton & Gooskens 2015). In this paper we address the question of how a group of previously unexposed listeners evaluate two unknown speech varieties, and whether certain linguistic features can still be subject to preference in this situation.

2. Language attitudes towards known and unknown varieties

In the 1970s, Giles et al. (1975) introduced two hypotheses to explain why some language varieties are regarded as more pleasing than others. They demonstrated that listeners do not make socially meaningful aesthetic judgements about varieties they have never heard before, providing evidence for the imposed-norm hypothesis (formulated in Giles et al. 1975), that emphasises the importance of social connotations and cultural norms for the creation of attitudes. In an experimental setting, British listeners who were unfamiliar with the Greek language did not differ in their evaluations of Athenian and Cretan accents (Giles et al. 1975). Other studies find evidence that a language variety is considered attractive when its speakers are socially privileged, explaining for example why English listeners locate Received Pronunciation (RP or BBC English) at the top of an aesthetic hierarchy, regional English accents in the middle, and urban English accents at the bottom (e.g. Giles 1970, Milroy & McClenaghan 1977, Trudgill & Giles 1978). RP is evaluated the highest because of its 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, then, applies in a social context where languages are overtly associated with particular social categories. The widely accepted theory of language attitudes, then, assumes that evaluations of speech are related to our ability to stereotype other humans on the basis of their social belonging, in a process of social categorisation. Listeners attribute the social attributes of the stereotype to the linguistic features and speakers they are hearing, thus forming 'language attitudes' (see Dovidio & Gluszek 2012).

The inherent-value hypothesis (in Giles et al. 1975), on the other hand, postulates that attitudes to languages could be triggered by qualities that are intrinsic to them. Inferring this hypothesis would mean that some languages (or language varieties) are intrinsically more aesthetically pleasing due to their sound characteristics than other languages (Garrett 2010: 228). The definition of 'inherent' refers to values that are not socially or culturally imposed. This includes characteristics that are found cross-linguistically, such as individual sounds or classes of sounds. While the hypothesis that language users create language attitudes on the basis of social connotations is undoubtedly true, only a few studies following Giles et al. (1974) have specifically looked into the evaluation of unknown speech varieties to see whether any linguistic features can be said to have such an 'inherent' value.

To date there are no findings of linguistic features that are universally seen as more pleasing than others. However, quite a number of studies have indicated that there are evaluative responses that can be made to linguistic varieties that could be seen as independent of the creation of stereotypes of the group in question. Mays (1982) finds that American listeners can correctly identify the social class of Arabic speakers well above chance level, even with no previous exposure to Arabic. Similarly, Moreau et al. (2014) find in their study of the perception of education level in 54 Wolof speakers that listeners who have no previous exposure to the language (students in Barcelona, Spain) are just as good at perceiving the education level of these speakers as listeners who may have had previous exposure (students in Senegal) with mean scores of 63.7% and 62.5% correct answers respectively,

which is significantly above chance level. Moreau et al. (2014) ask whether there may be some universally valuable traits in speech, such as fluency and vocal intensity, that indicate prestige throughout cultures, and call for more research to the topic.

Another study that finds some positive evidence in favour of certain linguistic factors being universally pleasing or displeasing, is Dragojevic & Giles (2016). They find that a noisier surrounding makes listeners evaluate speech more negatively, probably due to the increased difficulty in processing speech. Similarly, Van Bezooijen (1996) asked Dutch lay subjects to aesthetically evaluate several European languages. At the same time, she asked phoneticians to make global ratings of the languages on phonetic scales. The attributed degree of beauty proved largely predictable from a combination of judged 'melodiousness' and 'softness'. A fast tempo, precise articulation, and fronted articulation were positively correlated with the aesthetic judgments. Although it is unknown to which extent ratings are based on social connotations by the listeners, the correlational study suggests that aesthetic evaluations of languages may indeed have a phonetic basis, that includes both segmental and supra-segmental features.

Van Oostendorp (2004) suggests that place of articulation may play a role for evaluations across cultures. Articulation in the front of the mouth may be aesthetically more pleasing than articulation in the back of the mouth. Van Oostendorp assumes that this may be related to language acquisition in childhood. Irrespective of their mother tongue, children generally learn front consonants earlier than back consonants (Jakobson 1968:297; Menn & Vihman 2011: 263). Another hypothesis by Van Oostendorp (1996) is that the perceived beauty of sounds is related to the ease with which they can be sung (sonority). Sonorant consonants would thus give rise to positive impressions because the vocal cords vibrate spontaneously when they are pronounced. In the same context, the structure of the syllable could be important. Syllables ending in a vowel are easy to sing and would therefore be perceived as beautiful.

As far as prosody is concerned, it could be hypothesised 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 stressed syllables, 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 generalise to the aesthetic evaluation of languages. A recent study of aesthetic evaluations of European languages show that sonority and timing are among factors that can explain variation in speech evaluations of foreign languages (Reiterer et al. 2020), but the study was done on European listeners to an array of other varieties spoken on the continent, hence not languages the listeners can be said to be entirely unfamiliar with.

We need only minor linguistic details to create associations in our cognitive system, and these can be reliant on the linguistic context that they find themselves in. Studies in experimental sociolinguistics find ample evidence that relatively small differences in vowel or consonant production trigger evaluative responses

(e.g. Niedzielski 1999, Clopper & Pisoni 2004). Intriguingly, a particular phonetic detail may carry entirely different social meanings to listeners when used in two different registers, even of the same language. This is the case, for instance, of fronted /s/ that hold different social meanings to listeners in Modern Copenhagen Speech versus the Copenhagen Street Language (Pharao et al. 2014), where the context is likely to activate different stereotypes and expected behaviours in the listeners. It seems obvious that listeners must be highly familiar with the speech communities at hand to make these consistent fine-grained evaluative distinctions. We do not know, however, to which extent listeners ignore their capacity to align evaluations with linguistic information when listening to previously un-heard speech varieties.

In this paper we investigate this question with evaluations from listeners without former exposure to a language. We examine the language attitudes towards a Swedish and Danish bilingual speaker in a matched-guise test held with Chinese listeners who have had no previous exposure to Scandinavian languages. We further investigate to which extent supra-segmental features play a role for these evaluations in a follow-up experiment where we remove intonation from the speech signal to see how this affects evaluations by previously unexposed listeners.

3. Scandinavian languages, intelligibility, and attitudes

In language attitude testing, a robust finding has been an asymmetry in evaluative responses to speakers of Danish and Swedish in Scandinavian listeners. In particular Danish is evaluated more negatively by Scandinavian neighbours than vice versa (Haugen 1966, Delsing & Lundin Åkesson 2005, Schüppert & Gooskens 2011). The asymmetry in attitudes towards Swedish and Danish has often been explained anecdotally with the idea that Swedish culture is held in higher general regard compared to the other Scandinavian cultures, and that Sweden is likened to the big brother of the three Scandinavian nation states (Delsing & Lundin Åkesson 2005:136; Sletten 2005:71). The label big brother refers to Sweden's relative political and economic influence in the past, compared to that of 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.

In Ladegaard & Sachdev (2006:95) the idea that a positive relationship exists between regard for culture and regard for language is formulated in the language–culture consonance hypothesis. While they find some evidence that this relationship exists, there is more evidence in Ladegaard & Sachdev (2006:102) of the language–culture discrepancy hypothesis. They find that more youngsters state a preference for another culture than for the culture that speaks the accent they are aiming to acquire. The hypothesis that there must be a relationship between regard for culture and regard for language can therefore not univocally be inferred.

Another explanation has been put forward to explain the asymmetry in attitudes in Scandinavia: namely the asymmetry in intelligibility. Intelligibility can be influenced by the extent to which the speaker can decode the linguistic information in the signal they hear (Van Bezooijen 1996, Dragojevic & Giles 2016). Swedish and

Danish are linguistically very similar and the speakers of the two languages can communicate using their own languages. In the past, several studies have been 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 almost invariably show that the intelligibility levels between Swedes and Danes are asymmetric when it comes to spoken language: Danes understand Swedish better than Swedes understand Danish. It could be the case, then, that Swedish listeners regard Danish as a less attractive speech variety simply because they do not understand it well. This argument is part of a chicken and egg discussion, however, as several studies have indicated that the existence of negative language attitudes is a potential obstruction for successful inter-group 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). That intelligibility and attitudes stand in relationship to one another thus seems likely, but the direction of influence has never been established.

Furthermore, the attitudinal investigations of Danish and Swedish that have been conducted so far may have used methodologies that can be viewed as somewhat problematic. Either direct measures, i.e. explicit questioning about language, or speaker evaluation techniques with fragments from different people have been used to measure evaluations. Direct questioning may elicit one set of opinions, or even language ideologies, that can be different to subconsciously held associations and language attitudes (see Kristiansen 2009). In speaker evaluation techniques, evaluations of languages may be affected by individual speaker characteristics such as voice quality, mean pitch level and intonation (e.g. Zuckerman & Driver 1989). To avoid this methodological problem a larger number of speakers can be used for speaker evaluation, averaging out effects of variability between speakers ('verbal guise').

Yet, a more straightforward method to collect less consciously held attitudes and neutralise 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 et al. 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 (distractors) 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. By eliciting evaluations about the speakers rather than the languages themselves, the listeners are less likely to base their evaluation on overtly held stereotypes, and possibly instead on privately held opinions. In addition, since the two varieties spoken by the bilingual are in fact produced by the same speaker, language usage is the only feature that is being evaluated (and not voice characteristics, for instance).

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 seven and 16 years of age judged the Swedish and Danish recordings and four other languages on five-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 neighbouring language. Furthermore, the speaker was rated more positively by the Danes when she spoke Swedish than by the Swedes when she spoke Danish. These results thus confirm the results of the previous Danish–Swedish attitude investigations discussed above. As the matched-guise technique was used, it is unlikely that the asymmetric attitudes are caused by differences in voice quality. A question that remains is where the aesthetic perceptions of linguistic features stem from.

In this paper we re-use the speech samples employed by Schüppert et al. (2015) and test a population that is not influenced by imposed norms nor social connotations as there has been no previous exposure: listeners in their late teens and early twenties in central China who have never visited Europe. Furthermore, we look for experimental evidence of linguistic features playing a role in the asymmetry of aesthetic evaluations of Swedish and Danish and present a second experiment in which we test the effects of a manipulation of a prosodic trait, intonation, on the evaluations of listeners.

4. Study 1

The first part of the study investigates evaluative reactions in a matched guise test of a Danish–Swedish bilingual speaker in a Chinese group of listeners from Chongqing Jiaotong University. The answers that the listeners gave to questions in a questionnaire showed that the listeners had had no previous exposure to either language. Furthermore, they were not given information about which languages they were listening to, and can therefore be assumed to be unaffected by imposed norms, social connotations or (varying levels of) intelligibility.

4.1. Stimulus material

4.1.1. Speech samples

The Danish and the Swedish speech samples were produced by the same speaker: a young female Dane who had grown up in Southern Sweden but consistently spoke Danish with her Danish parents and siblings at home. A crucial factor in using the matched-guise technique is that reactions are attributable to the language itself. Therefore, much care was taken to ensure that the bilingual speaker sounded natively Danish and Swedish. This was done by organising two so-called voice parades that explored whether the bilingual speaker sounded as native to listeners with both language backgrounds as other native speakers of the two languages did. The voice parades involved presenting native listeners (none of who participated in the subsequent matched-guise experiment that was conducted for Schüppert et al. 2015) with recordings of native speakers, including one recording by the bilingual speaker, and instructing them to identify one speaker that sounded non-native. The varieties that the bilingual speaker and the speakers in the voice parades spoke could all be characterised as regional standard as spoken in Copenhagen (Danish) or

Table 1. Results of the voice parade for the Danish–Swedish bilingual speaker. Shaded cells indicate speakers that were picked at chance level or above.

	Bilingual speaker		Distracter 1		Distracter 2		Distracter 3		Distracter 4		Chance level
	N	%	N	%	N	%	N	%	N	%	%
Swedish listeners	0	0	0	0	3	20	12	80	0	0	20
Danish listeners	3	10	5	17	0	0	13	43	9	30	17

Southern Sweden (Swedish). We assumed that if the bilingual speaker is not chosen as the foreigner more often than at chance level, she sounds sufficiently native for our purpose.

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, four other recordings were produced by native female Danish speakers from the greater Copenhagen area, the same geographical area that the bilingual speaker 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 tests are shown in Table 1, which demonstrates that the bilingual speaker was not judged as sounding less native than the distracter recordings. In the Swedish voice parade, the bilingual speaker was selected by none of the listeners as having a foreign accent; in the Danish voice parade she was chosen by 10% of the listeners, which is still clearly below chance level.

As the table shows, the bilingual speaker was not rated significantly less native sounding than the other recordings by Danish and Swedish listeners, respectively. This suggests that all stimuli recorded for the experiment are perceived as native Danish and Swedish.

4.1.2. Text and test

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

Sound Fragment 1: Danish guise – original version (Supplementary Material 1; <http://soundcloud.com/user-191305898/danishguise>)

Sound Fragment 2: Swedish guise – original version (Supplementary Material 2; <http://soundcloud.com/user-191305898/swedishguise>)

Recording 1

What impression does this speaker make? This speaker sounds:

		1	2	3	4	5	
a	old-fashioned	O	O	O	O	O	modern
b	stupid	O	O	O	O	O	smart
c	unattractive	O	O	O	O	O	attractive
d	strange	O	O	O	O	O	normal
e	unfriendly	O	O	O	O	O	friendly
f	poor	O	O	O	O	O	rich

What do you think of this language?

ugly	O	O	O	O	O	beautiful
------	---	---	---	---	---	-----------

Could you elaborate on your answer? (optional)

Do you know which language you have just heard. If so, which one?

Figure 1. Questionnaire used in the matched guise experiment.**4.2. Procedure**

All fragments were played to the listeners twice with an inter-stimulus interval of six seconds. The ten recordings were presented in two different orders to the listeners. The listeners were provided with rating questionnaires consisting of semantic-differential scales (Osgood, Suci & Tannenbaum 1957). They were asked to evaluate their opinion about the speaker 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). In addition, they were also asked to evaluate the beauty of the language on a scale from 1 = ‘ugly’ to 5 = ‘beautiful’, and they were asked whether they recognised the language that was being used. Figure 1 is an image of an English translation of the questionnaire used in the investigation. The participants completed the questionnaires on personal computers in Microsoft Word.

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. In particular we were interested in knowing whether the listeners had any previous exposure to Danish or Swedish.

4.3. Participants

In total, 141 Mandarin-speaking listeners (41 males and 100 females) participated in the test. All participants completed the study and therefore data from all participants is included in the analysis. They were from The Chongqing Jaotong

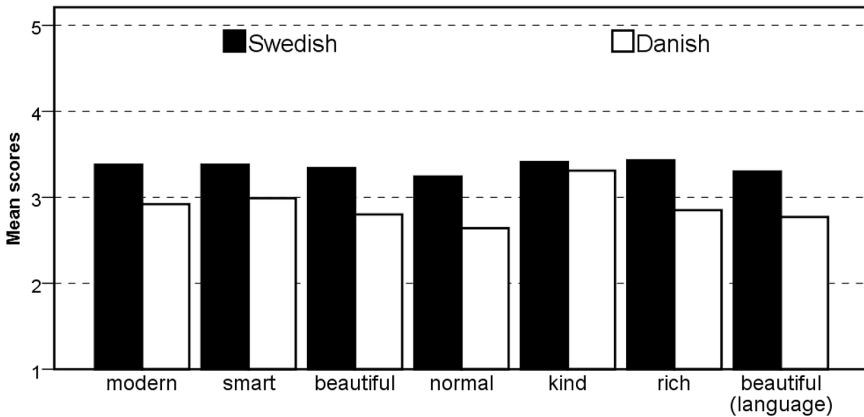


Figure 2. Mean ratings of 7 judgments when the bilingual speaker spoke Swedish (black bars) and when she spoke Danish (white bars).

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). None of the listeners had previous exposure to Danish or Swedish, nor did any of them recognise the languages. To the question whether the listeners knew which language they were listening to no one answered with a particular language group, responses were simply 'I do not know' or blank.

Eighty-five Chinese informants heard the Danish recording as the first fragment and the Swedish recording as the 6th, while 56 informants heard the test with the Swedish recording played 5th and the Danish played 10th. *T*-tests conducted within the listener group ($N = 141$) on the ratings of the perceived friendliness of the person shows that there are no significant differences in attitude ratings between listeners of the different playing orders (Danish ratings: $t(134) = -1.681$, $p = .096$; Swedish ratings: $t(135) = -0.026$, $p = .98$).

4.4. Results

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.

4.4.1. Raw results

Figure 2 shows the mean ratings of the bilingual speaker given by the Chinese listeners on the seven scales when she spoke Swedish and Danish. It can be seen that the speaker received higher scores for all judgments when she spoke Swedish rather than Danish.¹

4.4.2. 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 all seven variables were significantly inter-related with correlation coefficients never exceeding $r = .67$.

The analysis extracted one principal component with an eigenvalue of 3.82 (a measure of the covariance in the data, where the highest number indicates the most significant principal component). The second component had an eigenvalue of 0.86, which led us to exclude all subsequent components and conduct further analyses on the first extracted component. This 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 s $> .70$), and moderately with the scores for 'smartness' and 'kindness' ($r > .50$) and can therefore be assumed to represent the seven input variables reasonably well. The extracted component explains 54.6% percent of the variance in the data. By conducting a PCA, the data from the seven input variables has 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.

4.4.3. 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 the attitudes towards Danish. The explanation for this must be explained by characteristics of the languages themselves, as imposed norms and speaker characteristics cannot influence the judgments.

4.5. Discussion

Schüppert et al. (2015) reported that the bilingual Danish–Swedish speaker was rated more positively by Danish listeners when she spoke Swedish than by Swedish listeners when she spoke Danish. The results presented in the previous section of the current paper confirm that listeners with no previous knowledge about Danish and Swedish hold similar attitudes towards these languages: 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, or through

an asymmetry in intelligibility. The Chinese listeners could not identify the languages they heard and could therefore not be subject to imposed norms about Swedish and Danish. They must have based their judgments on the recordings themselves. This investigation therefore provides evidence that linguistic characteristics play a role in aesthetic evaluations.

Previous studies indicate that universally pleasing, and thereby also displeasing, language characteristics could exist at 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 both phonologically and prosodically different. One often commented upon difference, however, is intonation. Danish is often referred to as a monotonous language. Grønnum (1990:207–208, 2003:129) notes that some languages have larger prosodic inventories than others and that the rather poor inventory of Danish prosody may result in the impression of Danish being prosodically little expressive. On the other hand, Swedish is often perceived to have a more lively intonation. This perceived difference may be because Swedish uses a larger pitch range and is a pitch-accent language where word tones can have lexical meaning (Elert 1972, Bruce 1977, Gårding 1977). It could be 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 generalise to the aesthetic evaluation of languages. To investigate whether intonation differences between Danish and Swedish may have played a role for our results we set up a follow-up experiment.

5. Study 2: Follow-up experiment with intonation manipulation

To find evidence of whether the intonation differences between Danish and Swedish could be part of the explanation for the findings in the present investigation we had a closer look at the 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 livelier manner of speaking and this may result in more positive judgments. To test the role of intonation for the perception experimentally the follow-up experiment presents Chinese listeners with monotonised versions of the Swedish and Danish recordings together with monotonised distracters. If intonation plays an important role in the explanation of the different attitudes towards Danish and

Swedish, the differences are expected to become smaller or disappear when intonation is removed.

5.1. Method

In the follow-up experiment we conducted a study with a new group of Chinese listeners from the Chongqing Jiaotong University, using manipulated recordings of the test used in our first experiment. All recordings were monotonised.

5.2. Material

The material for the follow-up experiment was identical to that used in our first study (described in Section 4.1 above) save for the fact that the total of ten speech samples were monotonised using the Praat function that removed f0 variations and set its value to 224 Hz throughout.

Sound Fragment 3: Danish guise – monotonised version (Supplementary material 3; <http://soundcloud.com/user-191305898/28danishswedish-mon>)

Sound Fragment 4: Swedish guise – monotonised version (Supplementary material 4; <http://soundcloud.com/user-191305898/13swedish-mon>)

5.3. Procedure

The procedure used for the study was identical to that in the initial study, described in Section 4.2 above.

5.4. Participants

In total 316 Mandarin-speaking listeners participated in the test. Their background is comparable to that of the listeners in Study 1. The listeners (108 males and 198 females as well as eight informants who did not specify their gender) were from The Chongqing Jiaotong University in the Nan'an district in central China and were native speakers of Mandarin. Their mean age was 21.4 years (ranging between 17 years and 43 years).

One hundred and seventy-seven informants heard the Danish recording as the first fragment and the Swedish recording as the sixth, while 139 informants heard the test with the Swedish recording played fifth and the Danish played 10th. *T*-tests conducted within the listener group ($N = 316$) on the ratings of the perceived beauty of the person shows that there are no significant differences in attitude ratings between listeners of the different playing orders (Danish ratings: $t(313) = -3.409$, $p = .92$; Swedish ratings: $t(313) = -4.311$, $p = .16$).

5.5. Results

5.5.1. Overall results

The data was coded in the same manner as described in Section 4.4. Figure 3 shows the mean ratings on seven scales by the Chinese listeners for the bilingual speaker

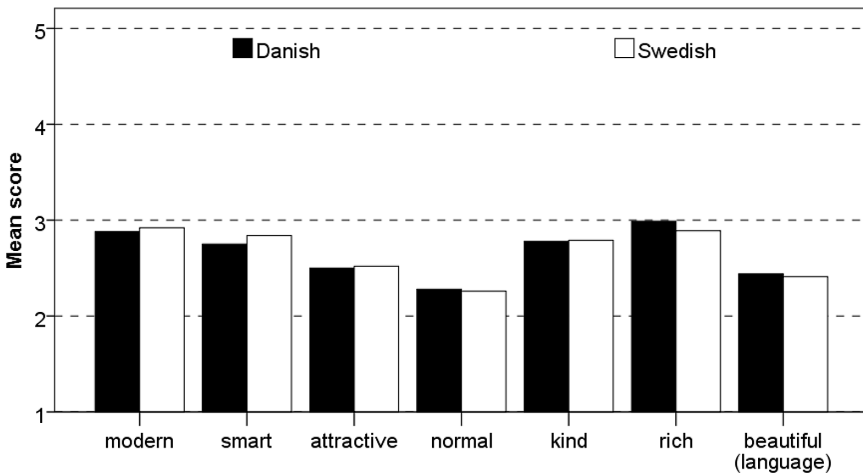


Figure 3. Mean ratings on seven scales when the bilingual speaker spoke Swedish (black bars) and when she spoke Danish (white bars).

when they heard her speak monotonised Swedish (black bars) and monotonised Danish (white bars). The bar graphs show that the speaker received almost the same scores for all judgments when she spoke Swedish and when she spoke Danish.

5.5.2. 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 'modernity', 'smartness', 'beauty', 'normality', '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 all the seven variables were significantly interrelated and thus extracted one principal component with an eigenvalue of 4.39, which indicates a sound data reduction. The second component had an eigenvalue of 0.75 and was therefore excluded, as were all subsequent components, since their contribution to the model was not sufficient. After extraction, a Promax rotation was applied, as we found that the ratings for the input variables were significantly correlated. The correlation between the relevant extracted component is high for all seven variables (all $r \geq .75$). The extracted component can therefore be assumed to represent the seven input variables well. It seems to measure the overall 'attractiveness' of the speech samples and it explains 63% percent of the variance in the data. The component consists of standardised values (z -scores), which means that the mean value for all 299 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.

5.5.3. Ratings of Monotonised Danish and Swedish

To investigate whether the Chinese participants still rate the bilingual speaker significantly more positively when she speaks Swedish rather than Danish, even when they are confronted with monotonised material, a pairwise *t*-test was conducted on the extracted component, i.e. on the *z*-scores. The test confirms what can be seen quite clearly in Figure 3: the asymmetry in judgments between Danish and Swedish disappears if pitch is removed from the speech. The judgments of the monotonised Danish sample ($M = 0.01$) do not differ significantly ($t(282) = 0.2$, $p = .84$) from the judgments of the monotonised Swedish sample ($M = 0.02$).

6. Discussion

Imposed norms and social connotations have repeatedly been shown to form our attitudes towards languages (Dragovic 2018) and there is no doubt that these factors explain a large part of how we develop emotive responses towards the languages that we are exposed to. However, in a setting where listeners are unfamiliar with the languages they are listening to, they may still be able to consistently express attitudes towards the sound of the language. If the speech material consists of unknown (or non-existent) languages, the linguistic correlates of aesthetic judgments in the investigation will necessarily be phonetic and phonological in nature, as syntactic, morphological and lexical factors cannot play a role in the evaluation of languages unknown to the listener.

Our results show that Chinese listeners prefer Swedish over Danish speech, while not having had any previous exposure to these varieties. This means that even without social associations specifically to the labels ‘Swedish’ and ‘Danish’, the listeners consistently rate the speaker as sounding friendlier, wealthier, and more intelligent when she speaks Swedish than when she speaks Danish. Social categorisation and stereotyping are developed over time and we know that the listeners in question are unlikely to have developed any specific associations towards the varieties heard. However, we cannot know which other possible stereotypes the speech samples used in this experiment may have activated in the listeners. While this is an interesting question for future research, it remains a fact that the listeners consistently prefer one variety over another, namely Swedish over Danish.

We do not necessarily want to claim here that linguistic features in Swedish speech are universally pleasing – it could be that Chinese listeners find features in Swedish attractive due to a larger degree of familiar sounds in the language, for instance. The higher degree of variation in pitch contours in Swedish could be an attractive trait to Chinese listeners due to the relative importance of tone in their native language. Studies on aesthetic evaluations in other scientific areas have found evidence for positive effects of familiarity (e.g. Peskin & Newell 2004).

A next step would be to find out whether these qualities are in fact held cross-linguistically, or even universally. A logical follow-up experiment would be to test whether we find the same evaluations in an experiment with a group of listeners who will not recognise the languages (just like the Chinese listeners) and whose native language is more similar to Danish in its prosody and segmental properties.

Only through further studies will we gain comprehensive insight to the role that articulation and tone bears for evaluations of speech across cultures.

In a follow-up study it would also be recommendable to add speech samples from more speakers from other regions to the experiment to be able to draw conclusions about the generalisability of the results to more speakers and to other regions of Sweden and Denmark. Also, the gender of both the speakers and the listeners could be a variable to be investigated since previous research has shown a difference in the evaluations of languages depending on gender (and other social identities) (Cargile et al. 1994).

It should be noted that the monotonisation of the fragments in our experiment resulted in somewhat unnatural sounding speech. The relatively low scores that the guises achieve on the 'strange-normal' dimension in the test could support an argument that oddness of the samples in Study 2 makes it more difficult for the listeners to evaluate the fragments than in the original Study (1). To explore this further one could consider doing another intonation experiment to crystallise the role of intonation for the ratings. This could be done, for instance, by manipulating a more Danish-like intonation pattern on to the Swedish speech data, and a Swedish-like intonation pattern on to the Danish pattern and replicate the experiment. Yet also in such an experiment would the resulting speech samples sound unnatural. An alternative could be to get bilingual speakers to perform different prosodic patterns on different samples of speech, yet speakers who have this ability could be difficult to come by.

As discussed in Section 2, additional linguistic characteristics of languages may play a role in our evaluations of languages. These could be tested experimentally by systematically varying specific acoustic cues by means of computer-manipulated speech to determine their effects on aesthetic judgments of listeners. Speech resynthesis techniques allow us to create more 'neutral' natural voices and systematically change different cues via digital manipulation of the sound waves. If, for example, we wish to test the hypothesis that Arabic is perceived as an ugly language due to its pharyngeal sounds, we can remove these sounds from Arabic samples (or non-existing languages constructed by means of speech synthesis to exclude the role of imposed norms or social connotations) or replace them by other consonants. If the manipulated version is perceived as more beautiful than the original version by listeners with no knowledge about the language, the pharyngeal sounds must be causally related to the negative attitude towards Arabic. Other possibly relevant parameters are for example syllable structure (e.g. presence or absence of consonant 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).

7. Conclusion

The results from our studies show that when speech samples are presented to listeners who have previously not been exposed to a specific language, they can make consistent evaluative ratings. We see in our results that Chinese listeners evaluate Danish as less aesthetically pleasing than Swedish, and that the previously

attested asymmetry in reciprocal language attitudes between Swedish and Danish speakers is found also among listeners with no previous exposure to the language. We therefore conclude that the asymmetric evaluations between Swedes and Danes can only partially be explained by asymmetry in cultural esteem or intelligibility. They could also be triggered by properties of the linguistic features. In that sense we find support for more explanatory factors for the formation of language attitudes than just the imposed-norm hypothesis. Part of the explanation for the differences in evaluations of Swedish and Danish that have been consistently found in the literature can lie in certain linguistic features being more universally pleasing features than others.

In a follow-up experiment we have shown that when intonation contours are removed from the speech signal, the difference in evaluative responses by unfamiliar listeners to the two speech samples disappears. This is an indication that especially the difference in Danish and Swedish intonation causes the different evaluations and it supports the impression by Danish linguists that Danish is a monotonous and rather inexpressive language (Grønnum 1990:208, 2003:129) and ‘the popular opinion among Danes, according to which Danish is ‘flat’ and dull, whereas Swedish is admirable and enviably melodious’ (Grønnum 1992:70).

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/S0332586521000068>

Acknowledgements. We would like to give a heartfelt thanks to Jason Kobelski and Dymphi van der Hoeven for coding the data for us.

Note

1. For the other five matched guises (see Section 4.1.2) no significant differences were found.

References

- Bø, Inge.** 1978. *Ungdom og naboland. En undersøkelse av skolens og fjernsynets betydning for nabospråkførståelsen* [Youth and neighbouring country: An investigation of the influence of school and TV on inter-Scandinavian comprehension]. Stavanger: Rogalandsforskning.
- Boets, H. & Georges de Schutter.** 1977. Verstaanbaarheid en appreciatie. Nederlandse dialecten uit België zoals inwoners van Duffel die ervaren [Intelligibility and appreciation. Dutch dialects from Belgium as experienced by inhabitants of Duffel]. *Taal en Tongval* 29, 156–177.
- Börestam Uhlmann, Ulla.** 1991. *Språkmöten och mötesspråk i Norden* [Language meetings and meeting languages in the Nordic countries] (Nordisk språksekretariats rapporter 16). Oslo: Nordisk språksekretariat.
- Bruce, Gösta.** 1977. *Swedish Word Accents in Sentence Perspective* (Travaux de l'Institut de linguistique de Lund 12). Lund: CWK Gleerup.
- Burgers, Christian & Camiel J. Beukeboom.** 2020. How language contributes to stereotype formation: Combined effects of label types and negation use in behavior descriptions. *Journal of Language and Social Psychology* 39(4), 438–456.
- Cargile, Aaron C., Howard Giles, Ellen B. Ryan & James J. Bradac.** 1994. Language attitudes as a social process: A conceptual model and new directions. *Language and Communication* 14, 211–236.
- Clopper, Cynthia G. & David B. Pisoni.** 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32(1), 111–140.
- Delsing, Lars-Olof & Katarina Lundin Åkesson.** 2005. *Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska* [Does the language keep together the Nordic

- countries? A research report of mutual comprehension between young speakers of Danish, Swedish and Norwegian]. Copenhagen: Nordiska ministerrådet.
- Dovidio, John F. & Agata Gluszek.** 2012. Accents, nonverbal behavior, and intergroup bias. In Howard Giles, Cynthia Gallois, Jake Harwood, Miles Hewstone, Michael. Hogg, Scott A. Reid & John C. Turner (eds.), *The Handbook of Intergroup Communication*, 87–99. London: Routledge.
- Dragojevic, Marko.** 2018. Language attitudes. In Howard Giles & Jake Harwood (eds.), *Oxford Research Encyclopedia of Intergroup Communication*, vol. 2, 179–192. New York: Oxford University Press.
- Dragojevic, Marko & Howard Giles.** 2016. I don't like you because you're hard to understand: The role of processing fluency in the language attitudes process. *Human Communication Research* 42(3), 396–420.
- Edwards, John.** 1999. Refining our understanding of language attitudes. *Journal of Language and Social Psychology* 18(1), 101–110.
- Elert, Claes-Christian.** 1972. Tonality in Swedish: Rules and a list of minimal pairs. In Evelyn Scherabon Firchow, Kaaren Grimstad, Nils Hasselmo & Wayne A. O'Neil (eds.), *Studies for Einar Haugen*, 151–173. The Hague: Mouton de Gruyter.
- Field, Andy.** 2005. *Discovering Statistics Using SPSS*. London: Sage.
- Gårding, Eva.** 1977. *The Scandinavian Word Accents* (Travaux de l'Institut de linguistique de Lund 11). Lund: CWK Gleerup.
- Garrett, Peter.** 2010. *Attitudes to Language*. Cambridge: Cambridge University Press.
- Giles, Howard.** 1970. Evaluative reactions to accents. *Educational Review* 22, 211–227.
- Giles, Howard, Richard Y. Bourhis, Peter Trudgill & Alan Lewis.** 1974. The imposed norm hypothesis: A validation. *Quarterly Journal of Speech* 60, 405–410.
- Giles, Howard, Richard Y. Bourhis & Ann Davies.** 1975. Prestige speech styles: The imposed norm and inherent value hypotheses. In William C. McCormack & Stephen A. Wurm (eds.), *Language in Anthropology IV: Language in Many Ways*, 589–596. The Hague: Mouton.
- Giles, Howard & Nancy Niedzielski.** 1998. Italian is beautiful, German is ugly. In Laurie Bauer & Peter Trudgill (eds.), *Language Myths*, 85–93. Hamondsworth: Penguin.
- Giles, Howard & Ellen Bouchard Ryan.** 1982. Prolegomena for developing a social psychological theory of language attitudes. In Ellen Bouchard Ryan & Howard Giles (eds.), *Attitudes Toward Language Variation*, 208–223. London: Arnold.
- Gooskens, Charlotte.** 2006. Linguistic and extra-linguistic predictors of inter-Scandinavian intelligibility. In Jeroen van de Weijer & Bettelou Los (eds.), *Linguistics in the Netherlands* 23, 101–113. Amsterdam: John Benjamins.
- Grønnum, Nina.** 1990. Prosodic parameters in a variety of regional Danish standard languages, with a view towards Swedish and German. *Phonetica* 47, 182–214.
- Grønnum, Nina.** 1992. *The Groundworks of Danish Intonation: An Introduction*. Copenhagen: Museum Tusulanum Press.
- Grønnum, Nina.** 2003. Why are the Danes so hard to understand? In Henrik Galberg Jacobsen, Dorthe Bleses, Thomas O. Madsen & Pia Thomsen (eds.), *Take Danish, for Instance: Linguistic Studies in Honour of Hans Basbøll Presented on the Occasion of his 60th Birthday, 12th July 2003*, 119–130. Odense: Odense University Press.
- Haugen, Einar.** 1966. Semicommunication: The language gap in Scandinavia. *Sociological Inquiry* 36(2), 280–297.
- Jakobson, Roman.** 1968. *Child Language, Aphasia and Phonological Universals*. The Hague: Mouton.
- Kaiser, Henry F.** 1974. An index of factorial simplicity. *Psychometrika* 39, 31–36.
- Kristiansen, Tore.** 2009. The macro-level social meanings of late-modern Danish accents. *Acta linguistica hafniensia* 41(1), 167–192.
- Ladegaard, Hans J. & Itesh Sachdev.** 2006. 'I like the Americans... But I certainly don't aim for an American accent': Language attitudes, vitality and foreign language learning in Denmark. *Journal of Multilingual and Multicultural Development* 27(2), 91–108.
- Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner & Samuel Fillenbaum.** 1960. Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology* 60, 44–51.
- Maurud, Øivind.** 1976. *Nabospråksforståelse i Skandinavia. En undersøkelse om gjensidig forståelse av tale- og skriftspråk i Danmark, Norge og Sverige* [Neighbouring language comprehension of spoken and written language in Denmark, Norway and Sweden]. Stockholm: Nordiska rådet.

- Mays, David V.** 1982. Cross cultural social status perception in speech. *Studies in Second Language Acquisition* 5(1), 52–64.
- Menn, Lise & Marilyn Vihman.** 2011. Features in child phonology: Inherent, emergent, or artefacts of analysis? In Rachid Ridouane & G. Nick Clements (eds.), *Where do Phonological Features Come from? Cognitive, Physical and Developmental Bases of Distinctive Speech Categories*, 261–301. Amsterdam: John Benjamins.
- Milroy, Lesley & Paul McClenaghan.** 1977. Stereotypical reactions to four educated accents in Ulster. *Belfast Working Papers in Language and Linguistics* 2, 1–11.
- Moreau, Marie-Louise, Ndiassé Thiam, Bernard Harmegnies & Kathy Huet.** 2014. Can listeners assess the sociocultural status of speakers who use a language they are unfamiliar with? A case study of Senegalese and European students listening to Wolof speakers. *Language in Society* 43(3), 333–348.
- Niedzielski, Nancy.** 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1), 62–85.
- Osgood, Charles E., George J. Suci & Percy H. Tannenbaum.** 1957. *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.
- Peskin, Melissa & Fiona Newell.** 2004. Familiarity breeds attraction: Effects of exposure on the attractiveness of typical and distinctive faces. *Perception* 33(2), 147–157.
- Pharao, Nicolai, Marie Maegaard, Janus S. Møller & Tore Kristiansen.** 2014. Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43(1), 1–31.
- Reiterer, Susanne M., Vita Kogan, Annemarie Seither-Preisler & Gašper Pesek.** 2020. Foreign language learning motivation: Phonetic chill or Latin lover effect? Does sound structure or social stereotyping drive FLL? *Psychology of Learning and Motivation* 72, 165–205.
- Schüppert, Anja & Charlotte Gooskens.** 2011. Investigating the role of language attitudes for perception abilities using reaction time. *Dialectologia – Special Issue II*, 119–140.
- Schüppert, Anja, Nanna H. Hilton & Charlotte Gooskens.** 2015. Swedish is beautiful, Danish is ugly? Investigating the link between language attitudes and intelligibility. *Linguistics* 53(2), 375–403.
- Sletten, Iben Stampe.** 2005. *Nordens sprog med rødder og fødder* [The languages of the Nordic countries with roots and feet]. Copenhagen: Nordisk Ministerråd.
- Trudgill, Peter & Howard Giles** (1978). Sociolinguistics and linguistic value judgments: Correctness, adequacy, and aesthetics. In Frank Coppieters & Didier L. Goyvaerts (eds.), *Functional Studies in Language and Literature*, 167–190. Gent: Story-Scientia.
- Van Bezooijen, Renée.** 1988. The relative importance of pronunciation, prosody, and voice quality for the attribution of social status and personality characteristics. In Roeland van Hout & Uus Knops (eds.), *Language Attitudes in the Dutch Language Area*, 85–103. Dordrecht: Foris.
- Van Bezooijen, Renée.** 1994. Aesthetic evaluation of Dutch language varieties. *Language and Communication* 14, 253–263.
- Van Bezooijen, Renée.** 1996. Aesthetic evaluation of Dutch: Comparison across dialects, accents, and languages. In Daniel Long & Dennis R. Preston (eds.), *Handbook of Perceptual Dialectology*, vol. 2, 13–31. Amsterdam: John Benjamins.
- Van Oostendorp, Marc.** 1996. *Tongval: hoe klinken Nederlanders?* [Accent: How do the Dutch sound?]. Amsterdam: Prometheus.
- Van Oostendorp, Marc.** 2004. Welluidend desalniettemin. De mooiste woorden van onze taal [Melodious desalniettemin: The most beautiful words in our language]. *Onze taal* 73(4), 76–79.
- Waddell, Martin & Barbara Firth.** 2005. *Can't You Sleep Little Bear?* Cambridge: Candlewick.
- Zahn, Christopher J. & Robert Hopper.** 1985. Measuring language attitudes: The speech evaluation instrument. *Journal of Language and Social Psychology* 4, 113–123.
- Zuckerman, Miron & Robert E. Driver.** 1989. What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior* 13(2), 67–82.

Cite this article: Hilton NH, Gooskens C, Schüppert A, and Tang C (2022). Is Swedish more beautiful than Danish? Matched guise investigations with unknown languages. *Nordic Journal of Linguistics* 45, 30–48. <https://doi.org/10.1017/S0332586521000068>