Mimicked accents — Do speakers have similar cognitive prototypes?

Niklas Torstensson[1,2], Erik J Eriksson[2] & Kirk P H Sullivan[2]

[1]Department of Communication & Information, University of Skövde, Sweden
[2]Department of Philosophy & Linguistics, Umeå University, Sweden

Abstract

There are several possible situations in which perpetrators might want to disguise their voices in order to avoid identification and to deflect the search for them to another person or group of individuals. One possible manner that can be used for voice disguise is the adoption of another accent. This paper examines the mimicking of the British-English Swedish accent, that is mimicking of the Swedish spoken by native British English speakers, by native Swedish speakers. It was found that the speakers selected similar and prominent features of the British-English Swedish accent in their first spontaneous recordings, and that the mimicked accent was impacted upon at word and suprasegmental level after having listened to a native speaker.

1. Introduction

The range of situations in which people might be interested in concealing their voices to avoid being identified and/or deflecting the search for them to another specific individual or group of individuals includes when making threatening phone calls, when orally blackmailing people and when threatening people when masked. Many strategies are available to individuals attempting to disguise their voices and includes:

- Changes in the voice source (i.e. change in pitch, creaky voice or whisper)
- Resonance features (i.e. bite blocks or foreign objects in the vocal tract and hypo / hypernasality)
- Manner of speaking (i.e. reduction of pitch variation, change of speech tempo, change of stress patterns)
- The imitation of a specific known individual
- The imitation of a foreign accent.

Künzel (2000) presents and explains many of these different strategies in detail and he also lists the most frequent methods of voice disguise as the use of falsetto speech, whispered speech, and creaky voice, the creation hypo-nasalized speech by holding the nose, and the adoption of imitation of a foreign accent.

Schiller and Köster (1996), Goggin et al. (1991), Tate (1979) and Thompson (1987) have demonstrated that foreign accent, real or faked, has a serious impact on the individual's ability to recognize and identify a speaker. The recognition accuracy not only deteriorates seriously when a subject speaks in a language not understood by a listener (see: Sullivan and Schlichting, 2000), but also when a subject speaks in a foreign accent. There are several possible reasons for this; a foreign accent, or strong regional dialect, tends to distract the listener. The accent has a tendency to mute speaker-specific idiosyncrasies so that subtle nuances of the language itself can be lost, and it tends to reduce the number of segmental contrasts (Hollien 2002).

Research on speaker identification and speaker recognition has been carried out on a range of languages, language pairs and accents, e.g. German (Schiller and Köster 1996), English/German and Spanish/English (Goggin et al. 1991) and English/Swedish (Sullivan and Schlichting, 2000, and Sullivan and Kügler, 2001). These investigations have contributed to our understanding of the listener side of speaker recognition, but not much about the speaker and speaker strategies relating imitated accents.

This paper, therefore, focuses on the speaker, and speaker strategies when adapting, or faking, a foreign accent normally not used by the speaker. By 'normally not used' it is meant that speakers do not use the particular faked accent for humor or effect in their personal or professional lives; the mimicking of the British English Swedish (BEngSe) accent, that is mimicking of the Swedish spoken by native British English speakers, by native Swedish speakers is the focus of this paper.

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after abstract only review

2. The British English Swedish Accent

At both the phonetic and the phonological level, English is relatively similar to Swedish. For example, Bannert (1990) in his study of the acquisition of a Swedish accent by second language learners described only the mutually understandable Scandinavian languages (Danish and Norwegian), Dutch, German and Icelandic as being more similar to Swedish than English is to Swedish. The result of the distinct set of differences between English and Swedish is that a person speaking Swedish with a British English accent typically exhibits a set of characteristic features. Bannert (1990, 2004) described the differences between Swedish and English, and the characteristics of the British English Swedish accent listed below as based on his work. The presentation of these characteristics is divided into vowel, consonant and prosodic features. Based on these characteristic features it is possible to delimit the features one can expect to define a general British English Swedish accent.

2.1. Vowel features

- The backed rounded vowel, [uː], is more closed in Swedish than in British English. In words such as *sol* [suːl] <sun> and *stor* [stuːr] <big> [uː] is pronounced as [ʊ].
- The vowels in Swedish words that are phonetically close to the English equivalent tend to be pronounced as in English, e.g. the Swedish *republik* [rɛpəbliːk] <republic> is pronounced as in British English i.e. as [rɛpʌblɪk].
- Vowels, especially the more open ones, in word-final and in unstressed positions tend to be reduced to schwa or to be deleted. For example *mattan* [matːan] <the mat> is pronounced as [matːən] and *maten* [mɑːten] <the food> is pronounced as [mɑːtn].
- The rounded front vowels [yː] [øː] and [ɯː] have a tendency to be realized as approximations of [iː], [ə] and [ʌ].
- The Swedish /r/ that in central standard Swedish is normally produced as a trilled [r] is produced as a non-rhotic [ɹ].
- The Swedish long [eː] and [uː] vowels are often diphthongized by British English speakers of Swedish, for example as in *ben* <leg>, pronounced [beɪn] instead of [beːn].
- The phonemic long-short vowel distinction in Swedish tends to normalize towards a vowel length that is somewhere in between the Swedish long and short vowel length; the distinction is generally not present in the British English Swedish accent.

2.2. Consonant features:

- The British English Swedish speaker has a tendency to delete initial consonants that occur in word-initial consonant clusters that do not occur in British English. For example *knä* [knɛː] <knee> is pronounced as [nɛː].
- Consonants can become syllabic: vowel deletion results in consonants becoming syllabic. For example: *segel* [seːgəl] <sail> being pronounced as [seːgl].
- Occurrence of linking 'r' that is not a feature of Swedish, as in *att gå efter* <to go for> pronounced [atgoːɹæfteɹ] instead of [atgoː efter].
- S-sounds are often pronounced as voiced instead of unvoiced. For example: *näsa* [nɛːza] <nose> instead of [nɛːsa].
- Voicing errors due to the non-application of the Swedish devoicing rule. For example in the compound word /företagsekonomi/ <business adminstration> where the normally voiced [g] in *företag* <business> gets devoiced in the position before the unvoiced [s] at the word juncture. In this example, the voiceless [s] is moreover voiced incorrectly by British English Swedish speakers

2.3. Prosody features:

- The Swedish long/short distinction is not maintained by British English Swedish speakers.
- British English Swedish speakers have problems with the Swedish word-stress distribution patterns both for words with one tonal peak, for example, fattade <understood> [ˈfatːade] that is pronounced as [fatˈaːde] and for words with two stressed tonal peaks such as *arbetslivet* <the working life> [ˈarbeːtslˈivɛt] that is pronounced as [ɑɹbeːtslˈivɛt].
- The British English Swedish speaker does not maintain the difference between Swedish accent 1 and accent 2, as in *anden* [ˊanden] <the duck>, pronounced with one tonal peak on the first syllable, and *anden* [ˋanden] <the spirit>, pronounced with two tonal peaks. Primary stress on the first syllable, more prominent than the secondary stress on the second syllable. This difference in word accent is extremely important as a word's meaning is sometimes dependent on accent and use of the incorrect accent contributes strongly to the impression of "Swedish with an accent"

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after abstract only review

3. Research questions

Three research questions were posed:

1. Do Swedes when mimicking a British English Swedish (BEngSe) accent capture the elements of this accent that phoneticians such as Bannert (1994, 2004) describe in their descriptions of this accent?
2. Do Swedes, when mimicking a BEngSe accent, mimic similar features? If so, this could indicate a common cognitive prototype, that may also indicate that the selection of these features who increase the changes of voice disguise.
3. Can the mimicked BEngSe accent be impacted upon by exposure to recordings of a BEngSe speaker?

4. Method

In order to remove the possibility for grammar mimickery and delimit the range of BEngSe features that could be mimicked it was decided to use read speech for this study.

4.1. The passage

The passage read by the experimental participants was a text from the newspaper Dagens Nyheter (Snaprud, 2003) about the hibernation of bears. The text read was:

Dags att gå i ide

Skansens björnar gäspar så att tungorna krullar sig. Det är dags att sova. Ny forskning visar att människan har en hel del gemensamt med djur som går i ide, vilket kan leda till säkrare transplantationer av mänskliga organ.

Björnhonan Enter kliar sig i sidan med högerfoten så att bilringarna dallrar. Skansens björnar är feta nu, redo att slumra bort den kalla och mörka tiden som kommer. Enter gäspar stort med hängande ögonlock.

- Nu dröjer det nog bara en vecka eller två innan björnarna går och lägger sig, säger Hans-Ove Larsson, intendent på Skansen i Stockholm.

Svenska brunbjörnar sover normalt ungefär halva året. Under sömnen sjunker kroppstemperaturen från dryga 37 grader till omkring 33 grader. Samtidigt faller pulsen till ungefär tio slag per minut, en fjärdedel av björnens vilopuls i vaket tillstånd.

Björnarna sover sig igenom de hårda tiderna när maten är knapp. Fettet från sensommarens frossardagar räcker till våren tack vare att förbränningen går ner till mindre än hälften.

Andra djur sjunker ännu djupare ner i vintervila, till en nivå där kroppen bara förbränner några procent av den energi som går åt i vaket tillstånd. Rekordet innehas av en arktisk jordekorre arktisk sisel som i sin djupaste dvala kan ha en kroppstemperatur på ner till minus tre grader.

Människan hör till de djur som dör av kyla. En sänkning av kroppstemperaturen med tio grader orsakar hjärtflimmer. Om kroppen blir ännu kallare är risken stor att hjärtat slutar slå.

Men även våra kroppar har en förmåga att åtminstone delvis sjunka ner i en mycket djup dvala. En infarkt i hjärnan stryper blodflödet så att nervceller får brist på syre, vilket i värsta fall kan vara livshotande. Syrebristen startar samtidigt ett försvarssystem som effektivt skyddar mot nya infarkter. Ämnesomsättningen faller och blodets förmåga att levra sig minskar.

- Reaktionen påminner starkt om vad som händer i ett djur som går i ide, säger Roger Simon, neurologiprofessor vid Oregon Health and Sciences University i USA.

Hans slutsats är att hjärnan har samma slags skydd mot syrebrist som ett djur i vinterdvala, vilket han nyligen rapporterade i medicintidskriften Lancet. Roger Simon hoppas att de nya rönen ska leda till mediciner som försätter hjärnan i dvala utan att skada den. En sådan medicin skulle kunna skydda hjärnan mot syrebrist som ibland uppstår i samband med en operation.

4.2. The participants

There were two groups of speakers: A group of native Swedish speakers and a native British English Speaker.

4.2.1. The Swedish participants

The Swedish participants were three male aged between 30 and 40 years. All were native speakers of Swedish. All spoke with the same regional dialect that was not Central Standard Swedish. All the speakers were advanced learners of English.

4.2.2. The native British English Speaker

The native British English Speaker was a male speaker, aged 39 years who has lived in Sweden for the past 10 years. He is an advanced learner of Swedish. He returns frequently to the United Kingdom and listens to the British Media daily. He speaks with a Southern English modified RP accent. His Swedish accent has many of the expected characteristics described in Section 2.

4.3. Procedure

The Swedish participants were not informed of the exact nature of the task prior to the first recordings. They were asked if they minded being recorded for a phonetic research purpose. On arrival each participant was presented with the text to be read. After they had read the text and felt able to read it aloud without problem, they were asked to read the text as if they wanted a listener to believe that they were English.

After the first recording, a second recording was immediately made with the Swedish participants reading the text in their natural Swedish voices. Thereafter, they were given a recording of the text made by the British English Speaker reading the same text. The participants were told to listen to the text as often as possible during the following week to improve their accent imitation. All of the Swedish participants confirmed that they had done so when they returned one week later to make a second recording of the text.

All of the Swedish participants recordings were made in Skövde using a Røde NT 3 microphone and a Waveterminal U24 A/D-converter, and recorded in 16 bit mono at 22 000 Hz. The speech files were normalized using WaveLab 4.0 software.

5. Acoustic and auditory analysis

An acoustic and an auditive analysis were preformed. These compared and contrasted the recordings to find both intra- and inter-speaker variations and similarities. The recordings were analysed for the characteristic features of the BEngSe accent as described by Bannert (1990) and outlined in Section 2. The first analysis concerns the unprepared, intuitive variant of accented speech compared to the reference recording and earlier research in the area and the second analysis the changes that occurred in the imitated BEngSe accents after a repeated listening to a recording of the same text by a native BEngSe speaker.

5.1. Intuitive strategies

Auditive analyses of the first recordings showed similar imitation strategies were used by all three Swedish subjects. All the Swedish subjects had, for example, made distinctive changes to their pronunciations of /a/ and /r/ in all possible contexts in the given text. There are, however, dissimilarities in the realization of these changes. The auditive and acoustic analysis of the subjects in presented subject by subject.

5.1.1. Subject 1

Subject 1's general strategy is presented prior specific details of his strategy are presented.

5.1.1.1 General strategy

Subject 1 altered his intonation pattern and the quality of a few specific sounds that are detailed below. Compared to his normal pattern of intonation, his imitation of a BEngSe accent has much larger fluctuations in F_0. This is combined with changing word stress pattern so that the first syllable of every word has the word stress.

5.1.1.2 Specific details

All /r/-sounds are pronounced as retroflex, fricative or non-rhotic [ɹ] instead of [r] by subject 1 regardless of the position of the /r/. The pronunciation of /a/-sounds were changed to a more fronted [a] instead of the subject's usual variation between [a] and [ɑ]. Pronunciation of non-rhotic laterals are also affected, so that the normal Swedish [l] was often realized as the retroflex [ɭ].

Some realizations of word initial grapheme combinations like /dj/ and /hj/ that are pronounced [j] in Swedish were realized by the affricate [dj] and an aspirated [ḫj] respectively. An example is the word /djur/ (animal) normally pronounced [juːr], but in subject 1's first recording it was realised as [djɵːɹ]. Another example is subject 1's imitation of a BEngSe speaker's pronunciation of /hjälp/ (help), that is pronounced as [jɛlp] by native Swedish speakers, yet in his BEngSe imitation it is rendered as [ḫjælp].

Occasional use of English word pronunciation is also found, like in *puls* (pulse), pronounced like the English [pʌls] instead of the Swedish [pɵls].

5.1.2. Subject 2

Subject 2's general strategy is presented prior specific details of his strategy are presented.

5.1.2.1 General strategy

Compared to this subjects' manner of reading in Swedish, the first general impressions upon listening to his BEngSe imitation are that speech tempo is substantially lower and that a majority of the vowels are diphthongized. The general articulation pattern is reminiscent of a southern Swedish, Scanian, variety of Swedish with /r/-sounds varying between a back tremulant [ʀ] and a back fricative [ʁ].

All pronunciations of /a/ are more fronted, with a clear tendency to be realised as [a] / [aː] rather than the Swedish variation between [a] and [ɑ].

5.1.2.2 Specific details

Subject 2's most prominent feature is the diphthongization of the vowels. Examples include the imitation of *ide* <winter lair> as [ɛiːdə] (the canonical Swedish pronunciation is [iːdeˀ], and *sover* <sleeps> that in Swedish is pronounced as [soːvɛr] but was pronounced as [səuvɔʁ] in Subject 2's imitation of a BEngSe accent.

Other intuitive features found for this speaker are English pronunciation of some words, for example, the Swedish *procent* <percent> is pronounced in English instead of the standard Swedish [prʊsɛnt].

5.1.3. Subject 3

Subject 3's general strategy is presented prior specific details of his strategy are presented.

5.1.3.1 General strategy

Subject 3 changed his intonation pattern dramatically from the intonation pattern he used in his recording of the text in his unaccented speech. Larger fluctuations of F_0 were combined with a change in stress.

The vowel /a/ is more fronted in this speakers BEngSe imitation than in his Swedish accent and the /r/-sounds are pronounced as retroflex, fricative or non-rhotic [ɹ] rather than as in his Swedish accent where he uses [r].

Accepted after abstract only review

5.1.3.2 Specific details

This speaker uses the English pronunciation of words for several words in his BEngSe imitation. For example, *system* and *pulse* are pronounced as English words. This subject also used the English translation for phonetically close words. For example *from* instead of *från*, [froːn], and *percent* instead of *procent*, [prʊsent].

The word-initial grapheme combination /dj/ in words like /djur/ (animal) and /djupare/ (deeper) was consistently pronounced as an affricate in subject 3's BEngSe accent, for example, [dʒɵːɹ] instead of [jɯːr] and [dʒɵːpaɹɛ] instead of [jɯːpareˌ].

5.2. The revised strategies

Table 1 presents a summary of the strategies used by the subjects in their imitations of a BEngSe accent and the changes to their strategies after having listened a number of times over a week to a recording of the text by a native speaker of English who speaks Swedish with a BEngSe accent. None of the features focused upon were removed between recording 1 and recording 2. However, a number of features were added: these are indicated by square brackets in Table 1.

Table 1: Feature appearing in the intuitive recordings are indicated by an X; features that were added after training are indicated by [X]. As all of the features used in the first recordings continued to be used in recording 2, there is no indication of features present in recording 1 but not recording 2.

Feature	Reference	Speaker 1	Speaker 2	Speaker 3
backed rounded vowel, [uː]				X
English equivalent word		X	X	X
reduction to schwa	X	[X]	[X]	[X]
vowel deletion	X	X		X
approximations of [yː] [øː] and [ɯː]	X			X
[r] produced as [ɹ]	X	X	X	X
diphtongized [eː] and [uː]	X	X	X	X
aspiration between stop and vowel	X			
voiced /s/	X			
vowel quantity errors (long / short)				
word-stress errors	X	X		X
accent 1 / accent 2 errors				
word-initial consonant deletion				
syllabic consonants	X	X	[X]	X
intrusive /r/	X	[X]		X
Word-initial consonant clusters (dj)		X		X
deviant pronunciation of /a/	X	X	X	X

6. Discussion

The inter-speaker comparisons show significant similarities in terms of the features that were changed; this can be interpreted as suggesting that these three speakers hold a common view of the characteristic of a BengSe. Although, it is clear that certain features are seen both as distinctive, typical and necessary for this accent and that together these features could be both necessary and sufficient to create the illusion of a genuine BEngSe, it is unwise to delimit a BEngSe prototype upon data from three male speakers.

The fact that none of the speakers used the voiced-/z/ feature probably depends on source-language factors. As this is a speech sound that is part of the British English phoneme inventory, but not the Swedish, it is likely to be not perceived, ignored, or forgotten. However, since this feature is described as being prominent in the accent description and present in the reference recording, this suggests that it is one of the necessary features for an accent to be universally perceived as BEngSe. The lack of such a feature in combination with the presence of other prominent ones could thus indicate that an accent is mimicked and not authentic.

This manner of reasoning lends support to the idea that rather than defining a prototypical accent, the

definition should be of distinctive features, each necessary and together sufficient, to form the cognitive impression of a genuine accent.

7. Conclusions

After evaluating the results with respect to the research questions posed, the following conclusions can be drawn. Speakers, when imitating a well-known accent, capture many of the prominent features of such an accent. Compared to earlier phonetic research, there is a high degree of correspondence between expected and actual findings. Many of the elements described as being prominent, e.g. by Bannert (1994, 2004) appear in the mimicked accent recordings.

Inter-speaker comparisons show, as mentioned earlier, some differences but at a broader level, a large number of similarities; many of the prominent features are shared between all the subjects that took part in this pilot study on foreign accent imitation. It, thus, appears possible that speakers of Swedish form a cognitive prototype for accents that are to a high degree similar between individuals.

As for the third research question, if the mimicked accent can be impacted upon by exposure to a genuine BEngSe accent, the answer must be affirmative. Our experiment has shown that exposure over a relatively short time period has effects on all levels, prosodic as well as vowel- and consonant features.

8. Acknowledgements

The authors thank the participants for the enthusiasm with which they undertook the task.

9. References

Bannert, R. (1994, 2004) *På väg mot svenskt uttal.* Second edition. Studentlitteratur. Lund, Sweden.

Goggin, J., Thompson, C., Strube, G. & Simental, L. (1991). The role of language familiarity in voice identification. *Memory and Cognition* 19: 448–458.

Goldstein, A. G., Knight, P., Bailis, K. & Conover, J. (1981). Recognition memory for accented and unaccented voices. Bulletin of the Psychonomic Society 17: 217–220.

Hollien, H. (2002). Forensic Voice Identification. San Diego, CA, Academic Press.

Hollien, H., Majewski, W. & Hollien, P. A. (1974) Perceptual identification of voices under normal stress and disguised speaking conditions. Journal of the Acoustical Society of America 1974, 56.

Künzel, H. (2000) Effects of voice disguise on speaking fundamental frequency. Forensic Linguistics 7, 149-179

Schiller, N. O. & Köster, O. (1996). Evaluation of a foreign language speaker in forensic phonetics: A report. Forensic Linguistics 3: 176–185.

Snaprud, Per (2003) Dags att gå i ide. *Dagens Nyheter*, Web edition Nov. 2:nd 2003

Sullivan, K. P. H.; and Kügler, F. (2001). Was the knowledge of the second language or the age difference the determining factor? *Forensic Linguistics.* 8.2, 1-8.

Sullivan, K. P. H. and Schlichting, F. (2000) Speaker discrimination in a foreign language: First language environment, Second language learners *Forensic Linguistics: The International Journal of Speech, Language and the Law* **7**: 1350-1771.

Tate, D. A. (1979) Preliminary data on dialect in speech disguise. Current Issues in the Phonetic Sciences. H. Hollien and P. A. Hollien. Amsterdam, John Benjamins. 9: 846–850.

Thompson, C. (1987). A language effect in voice identification. Applied Cognitive Psychology 1: 121–131.