It seems proper to end this paper by a (translated) quote from Ryszard Kapuściński's book *Ebony* (p. 63 in the Swedish translation *Ebenholts*):

> In many African societies the children get names after an event which has happened at the birthday. The name *Edu* was short for education, as the first school in the child's village was opened the same day.
>
> Where Christianity and Islam were not rooted the richness of names was infinite. The poetic talents of the parents were displayed when they e.g. gave the name 'Bright morning' (if born in the early morning) and 'Acacia shadow' (if born in the shadow of an acacia tree). If a child had been born when Tanganyika had gained its independence the child could have been christined to *Uhuru* (Swahili for 'independent'). If the parents were adherents of president Nyerere the child could be given the name *Nyerere*.
>
> The introduction of Christianity and Islam had reduced these opulent worlds of poetry and history to a few dozens of names from the *Bible* and the *Koran*. Since then there are only long rows of *James* and *Patrick* or *Ahmed* and *Ibrahim*.

References

Allén, S. & S. Wåhlin. 1979. *Förnamnsboken*. Stockholm: Almqvist & Wiksell.

Barnes R. B. 1982. 'Personal names and social classification'. In D. Parkin (ed.), *Semantic anthropology*, 211-26. London: Academic press.

Brylla, E. 2004. *Förnamn i Sverige*. Stockholm: Liber.

Lundström, H. & J.-O. Svantesson (eds.). 2005. *Kammu: om ett folk i Laos*. Lund: Lunds universitetshistoriska sällskaps årsbok 2006.

Kapuściński, R. 2003. *Elfenben* (Engl. *Ebony*). Stockholm: Bonnierpocket.

Modéer, I. 1964. *Svenska personnamn*. Lund: Studentlitteratur.

Rydén, H. 2006. 'Ett av fyra barn har unikt namn'. *Sydsvenskan*, 26 Febr, 2006.

Otterbjörk, R. 1963. *Svenska förnamn*. Stockholm: Norstedts.

Sigurd, B. & M. Eeg-Olofsson. 2004. 'Modeväxlingar hos förnamn i Norden'. *Arkiv för nordisk filologi* 119, 209-25.

Villarsen Meldgaard, E. 2002. *Den store navnebok: Hvilket navn skal vi vælge? Hvad betyder navnet?* København: Aschehoug.

Lund University, Dept. of Linguistics and Phonetics
Working Papers 53 (2008), 145–158

A bidialectal experiment on voice identification

Maria Sjöström, Erik J. Eriksson, Elisabeth Zetterholm and Kirk P. H. Sullivan

1 Introduction

Our voice and speech are part of our identity and we can distinguish different humans through their voices. The voice gives the listener a cue about the speaker's gender, age and regional background, and socio-economic class. We have all successfully recognized friends and family members by voice alone. For example when a friend has called us on the telephone and not said who was calling. False, or incorrect, identifications occur. We have all incorrectly recognized the voice on the telephone as the voice of a friend. False identifications also occur among unfamiliar voices due to voice disguise. This can create problems for the judicial system when relying on earwitness testimonies and voice identification tests as direct evidence in a criminal case. This paper, based on a study by Sjöström 2005, provides an insight in the field of earwitness identification and empirically investigates the power of dialect as an attribute that listeners use when identifying voices and how a switch of dialect affects the success of voice identification.

The voice is unique for each person and is dependent upon the shape of the vocal tract and individual phonetic habits. Further, we do not always produce exactly the same sounds: the situation as well as our health and emotional state affect our voices. This is referred to as intraspeaker variability. Differences between speakers are referred to as interspeaker variability. It is generally assumed that interspeaker variability is greater than intraspeaker variability. This makes voice recognition and identification possible. Research on voice imitation, has demonstrated the flexibility of the human voice and that it is possible to alter one's voice and speech to near that of another person (Zetterholm 2006). This might cause problems in an identification task.

In the process of recognizing a voice, humans attend to certain features of the individual's speech, such as speaking fundamental frequency (f0), articulation, general voice quality, prosody, vocal intensity, dialect/sociolect, speech impediments, idiosynctratic pronunciations and language pattern and unusual use of linguistic stress, affect and emphasis (e.g. Gibbons 2003, Hollien 2002). The listener may use all or only a few of these features when trying to identify a person, depending on what information is available. Which of these features serve as the most salient ones when recognizing a voice is unclear.

In the literature there is a distinction between the terms recognition and identification. Voice recognition often refers to an overall feeling of familiarity with a voice, while voice identification is the more specific ability of being able to discriminate a particular person's voice from one or more voice samples (Yarmey 1995). This is of importance in a forensic situation when a witness is going to identify the perpetrator amongst a number of suspects. Earwitness identification parallels the procedure used for eyewitness identification: voice line-ups, where the witness can listen to a number of voices, are comparable to the face line-ups that are used in the area of eyewitness identification. However, there are differences due to way in which humans process auditory and visual memories. Further, the analysis of patterns and features might affect the identification task (Hollien 2002).

2 Factors that could affect the voice identification

There are a number of factors involving the listener that can affect a listener's degree of success when trying to identify a voice. These include voice familiarity, memory, emotional state and language or dialect familiarity.

Based on the results of a number of experiments (Hollien, Majewski & Doherty 1982) it was shown that listeners, under normal speaking conditions, who were familiar with the speaker, correctly identified speakers 98% of the time whereas listeners, who were unfamiliar with the speaker correctly identified the speakers in only 32.9% of the cases. As Yarmey et al. 2001:285 wrote: "Van Lancker el al., (1985a, b) suggested that the cognitive processes underlying voice recognition of familiar and unfamiliar speakers are different. Whereas recognition of familiar voice depends more on pattern recognition than on feature analysis, the discrimination of unfamiliar voices is more likely to be a feature analysis task than pattern recognition. Because feature analysis, such as the recognition of specific pitch, pitch variability, inflections, and so on, are more difficult to execute than pattern recognition

high familiar speakers should yield more accurate identification responses than less familiar speakers." However, it is not clear how reliable we are in recognition texts; Rose & Duncan 1995 reported a success rate that varied from 31% to 83% depending on the test material used.

Our memory is another factor that may affect the identification task. Different experiments show that the ability to identify an unfamiliar voice decreases after time. McGehee 1937 performed a number of experiments where listeners were to listen to an unfamiliar voice and after time-intervals ranging from 1 day to 5 months try to recognize the voice. The results show that the recognition rate dropped to 68.5% as early as after two weeks. Clifford & Denot (reported by Yarmey 1995) reported even more dramatic deterioration with a decrease of 50% in recognition rate after one week.

Stress, fear and other emotions of the event itself are other factors that affect the witness of a real crime situation that have to be taken into consideration in an earwitness identification task. This is, as Yarmey 1995 pointed out, impossible to investigate in a controlled experiment situation.

The listener's language/dialect familiarity may also affect the identification of a voice. Yarmey 1995:799 reported that: "Just as people of another race may all look alike, people are more likely to hear the voice of speakers of a foreign or accented language as homogeneous or similar to each other, but perceive diversity or distinctive forms of speech among speakers of their own regional area" (p.799). In an experiment conducted by Doty 1998 native English and U.S. citizens were tested on their ability to recognize voices belonging to their own nationality and voices belonging to other nationalities. The results show that participants were highly accurate in recognizing fellow citizens (88%) but poor at recognizing foreign voices (13%), regardless of racial background. Studies with both monolingual and bilingual speakers also clarify the critical role of language familiarity in voice identification (Thompson 1987, Goggin et al. 1991). One possible explanation for these difficulties is the listener's use of schemata and prototype in speaker identification. Schemata are based on the exposure to voices in the neighbourhood and prototype-matching is an approach in which voice features are compared to mentally stored prototypes. This can lead to difficulty in identifying unfamiliar dialects/languages, according to Doty 1998.

3 Voice disguise

If you want to hide your identity by changing your voice and speech there are a number of ways to disguise your voice. Whispering is one of the easiest ways since it conceals the most salient characteristics of the voice, such as pitch and intonation (Yarmey et al. 2001). Another effective disguise method is using a hypernasal voice (Reich & Duke 1979). In a forensic context, criminals may purposely try to hide their identity by using a voice disguise. Rodman 2000:1 defines voice disguise as "any alteration, distortion or deviation from the normal voice, irrespective of the cause". Statistics from the German Federal Police Office shows that 15-25% of the annual cases involving speaker identification involve at least one type of voice disguise (Künzel 2000). Some of the perpetrators' 'favourites' included: falsetto, pertinent creaky voice, whispering, faking a foreign accent and pinching one's noise.

Imitation is another useful type of voice disguise that can effectively employed when trying to hide one's identity. When the goal of imitation is entertainment, the imitation of the well-known person or target speaker is often more like a caricature with pertinent individual characteristics out of the voice highlighted for effect (Zetterholm 2003). The use of imitation in criminal acts focuses more on group identity, such as a regional or a social dialect. Markham 1999 investigated if it was possible for speakers to produce convincing readings in various Swedish dialects. Both the speaker's ability to consistently keep a natural impression and to mask his other native dialect were investigated. His findings suggest that individual speakers differ greatly in these abilities. Some speakers succeeded in both masking their own dialect and fooling the listeners that they were native speakers of another dialect. Other speakers were only able to mask their own dialect; they failed in creating a natural sounding accent that convinced the listeners that they were native speakers of another dialect. Markham pointed out that although it is important to hide one's identity when using voice disguise, it is also important to create an impression of naturalness in order to avoid suspicion.

4 Bidialectalism

The term bidialectalism is adopted from the term bilingualism. A bidialectal speaker controls two dialects and uses them in special circumstances, but often the speaker mixes the features from the two dialects (Chambers & Trudgill 1980). Humans have the ability to perceive and understand different dialects of their native language, but most often only produce one dialect.

Hazen 2005 argued that we switch some language variation patterns to those of another social or regional group in some situations, e.g. code switching, but that we do not usually have command of two dialects with quantitative and qualitative accuracy (www.as.wvu.edu/dialect/varres/bidiupen.pdf). Hazen argued that there is no evolutionary motivation for humans to be productively bidialectal as long as humans are receptively bidialectal. We can learn a second dialect, but age of learning impacts upon the learner's ability to sound like a native speaker of the 'new' dialect. There is a relationship to principles of dialect acquisition and first- and second-language acquisition, e.g. the critical period and acquisitional order (Chambers 1992).

In the present study we have recorded a Swedish bidialectal speaker. He uses two different Swedish dialects on a daily basis and he shifts dialect depending on the audience. He does not only change some segments or patterns when speaking, he changes his dialect without mixing features. A brief auditory and acoustic analysis is presented in Section 6.1.

5 A bidialectal experiment

The study presented here provides an insight in the power of dialect as an attribute that listeners use when identifying voices and how a switch of dialect affects the success of voice identification. Dialect is one of the first things forensic practitioners look at when trying to establish the speaker's identity (Hollien 2002). Despite this, little research has focused on dialect's role as an attribute that people listen to in the voice identification process.

To assess whether dialect is a feature that listeners attend to and that distracts them from the physiologically grounded features of the voice, we constructed four perception tests to investigate whether a bidialectal speaker is equally recognizable regardless of the dialect in which he is speaking. We also wanted to test if there is a difference in the ability to identify the bidialectal speaker dependent on the listener's native dialect.

5.1 Differences between the two Swedish dialects

The two dialects spoken by the bidialectal speaker used in this experiment are Scanian, spoken in the Southern part of Sweden, and a variety of the Stockholm dialect. These two dialects differ in many aspects and are not considered similar. Elert 1997 has listed some dialectal differences, both those relating to voice quality and pronunciation. Scanian speakers have a lower fundamental frequency (F0) than speakers of the Stockholm dialect, and this lower F0 is often accompanied by a creaky voice quality. Also, speakers of the Stockholm dialect tend be more nasal than Scanian speakers. One of the most salient characteristics of south Swedish dialects is the long vowel diphthongization, where the vowel sound moves from a central position and out towards the target vowel such as: [ᵉi(j)] for [iː]. Diphthongization also occurs in the Stockholm dialects, although the opposite takes place, the vowels are centralized such as: [iː(j)ᵊ] for [iː]. Another difference between the dialects is the pronunciation of the r-sound. Scanian speakers use a uvular trill [ʀ] or a uvular fricative [ʁ] while Stockholm dialects use an alveolar trill [r] or a retroflex fricative [ʐ]. Further, there is an absence of retroflex consonants in Scanian and the sound combinations *rt, rd, rs, rn* and *rl* are pronounced with a uvular [ʀ] and a dental consonant. In Stockholm dialects the above sound combinations would be pronounced with the retroflex consonants [ʈ], [ɖ], [ʂ], [ɳ] and [ɭ]. Difference between the dialects in prosody is described by Bruce 1998 and is apparent in timing and use of tonal accent I and II.

6 Material and method

6.1 The target speaker

The target bidialectal speaker is a male Swede who speaks Scanian and a variety of Stockholm dialect on a daily basis. He was born near Stockholm but moved to Scania as a five-year old. He reports that he continued to use only Stockholm dialect, despite the change of region, until middle school when he started using Scanian with his friends. At the time when recorded, the speaker said he used Scanian with his wife and most of his colleagues, and Stockholm dialect with most relatives, including his children although they speak Scanian.

An acoustic analysis of the speaker's both dialect voices was performed by a phonetician (Dr Zetterholm), which confirmed that the speaker is fluent and that his two voices carry the typical characteristics of the two dialects, see 5.1. The speaker is also consistent in his use of the two dialects.

6.2 The voices

Two recordings of the story *The princess and the pea* were made by the target speaker. In one of them he read the story using the Stockholm dialect, and in the other he read it using his Scanian dialect. These recordings will hereafter be referred to as TargetSC (Scania) and TargetST (Stockholm).

Four more recordings of the same text were used. Two foils with a Stockholm dialect and two foils with Scanian dialect were used for voice

identification tests. The foils were all chosen with regard to their similarities with the target voice in dialect, age, and other voice features such as creakiness. These four voices will be referred to as foil 1-4. Foil 1 and 3 speaks Scanian, foil 2 and 4 speaks with a Stockholm dialect.

6.3 The identification tests

To test if regional belonging affects the ability to identify the bidialectal speaker used in this experiment, listeners from two different regions in Sweden were chosen to participate. One listener group originates from the same region where one of the native dialects of the bidialectal speaker is spoken, Scania. The other region is Northern Sweden where neither of the bidialectal speaker's dialects are used, and where therefore the familiarity of the two dialects should be low. Only participants who were native Swedish speakers and who did not report any hearing impairment were selected. Ten participants in each tests, from each of the two regions, gives 80 participants in total. All participants were randomly assigned one of the four tests.

Four different earwitness identification tests were constructed for participants to listen to. The stimuli used in the test were three phrases selected from each recording, the target speaker and the four foils. Each test began with the entire recording of *The princess and the pea* as the familiarization voice, and was followed by a voice line-up. Each voice line-up contained the four foil voices and one of the target's two dialect voices. There were three repetitions of the three phrases spoken by all voices in the line-up. Hence the voice line-up contained $3 \cdot 3 \cdot 5 = 45$ voice stimuli. The line-up order of the stimuli was randomized. For a description of the composition of the tests, see Table 1. 'SC' is used for Scanian, and 'ST' for Stockholm dialect. The test 'SC-ST' uses the target's Scanian voice as the familiarization voice and the target's Stockholm dialect voice in the line-up. A pre-made test CD was used and the participants listened to the test over loudspeakers. A tone of 280 Hz was inserted between every block of ten stimuli to help participants keep track of where they were in the test. Participants were to decide, for each stimuli, if the voice was the same as the voice they heard in the familiarization reading by circling either 'yes' or 'no'. Test SC-SC and ST-ST were created as control tests. They investigate if the target speaker is recognized to the same degree when speaking with his two different dialects. In test ST-SC and SC-ST the question is if the target speaker is recognized even when changing dialect.

Table 1. The composition of the voice identification tests including which of the target's voices was used as familiarization voice and which were included in the voice line-ups.

Test	*Familiarization voice*	*Line-up voices*
SC-SC	TargetSC	Foil 1-4 + TargetSC
ST-ST	TargetST	Foil 1-4 + TargetST
ST-SC	TargetST	Foil 1-4 + TargetSC
SC-ST	TargetSC	Foil 1-4 + TargetST

Table 2. Possible responses in a two-alternative forced choice identification test.

	Response	
Stimulus	'Yes'	'No'
Target	Hit	Miss
Non-Target	False alarm	Correct rejection

6.5 Data analysis

The method of analysis chosen for this experiment is the same used and described by Schiller & Köster 1998. In a two-alternative forced choice test such as this, there are four different types of responses for each voice stimulus in the line-up: hit (when the listener correctly assigns a 'yes' to the target stimulus), miss (when the listener assigns a 'no' to a target stimulus), false alarm (when the listener assigns 'yes' to a non-target stimulus) and correct rejection (when the listener assigns 'no' to a non-target stimulus) (see Table 2).

By calculating the hit and false alarms rates as proportions of the total number of possible hits and false alarms, the listeners' discrimination sensitivity can be determined. This sensitivity is a measure of one's ability to discriminate between targets and non-targets. The measure used is d' (see Green & Swets 1966), which is the difference between the hit rate (H) and the false alarm rate (F), after first being transformed into z-values. The d'-equation is: $d' = z(H)-z(F)$. A d'-value of 0 occurs when listeners are not able to discriminate between targets and non-targets.

7 Results

Participants of both control tests, SC-SC and ST-ST, show positive mean d'-values (1.87 and 1.93). It was shown through a two-tailed Student's t-test that there was no significant difference in identification of the two dialects and they can therefore be considered equally recognizable ($t(38) = 0.28$, $p >$ 0.05). By conducting a one-sample t-test it was shown that the d'-values for both tests are highly distinct from 0 ($t(39) = 18.45$, $p < 0.001$) and therefore high degree of identification of both dialects can be concluded. There was no significant difference found between the two test groups in response bias, ($t(38) = -1.91$, $p > 0.05$), and thus the participants of both groups were equal in tendency to answer 'yes' or 'no'.

For the test ST-SC the mean d'-value was 0.44, and for test SC-ST it was –0.07. Further, a two tailed t-test showed these d'-values did not differ significantly ($t(38) = 1.93$, $p > 0.05$), and therefore the two tests can be considered equal in identification difficulty. A one sample t-test was conducted and showed that the mean d'-value of the two tests were not significantly separated from 0 ($t(39) = 1.36$, $p > 0.05$). This suggests that there is a tendency to answer randomly. No significant difference was found between the two groups of listeners in response bias ($t(38) = 0.24$, $p > 0.05$).

As there was no significant difference between the test SC-SC and ST-ST nor the test ST-SC and SC-ST, it is possible to group these tests into 'control tests' and 'dialect shifting tests' and look at the overall performance results. Mean d' for the control tests was 1.90 and for the dialect shifting tests 0.18. It was found that the difference in performance between the two types of tests was significant. ($t(78) = -10.11$, $p < 0.001$). The same grouping can be done with the response bias measure showing a bias rate of 0.26 for the control tests and 0.84 for the dialect shifting tests. A significant difference was found in response bias between the tests ($t(78) = 5.97$, $p < 0.001$). By looking at the bias values a higher tendency to answer 'no' in the dialect shifting tests than the control tests can be observed.

7.1 Regional differences between listeners from Northern Sweden and Southern Sweden

In the test SC-SC, listeners from Northern Sweden had a mean d'-value of 1.49 and listeners from Scania had a mean d'-value of 2.26. A one-sample t-test found that both d'-values were statistically different from 0, and thus it could be concluded that both regions were able to identify the target voice using Scanian. The Scanian participants' d'-value was significantly higher than the d'-value of the participants from Northern Sweden ($t(19) = -2.81$, $p < 0.02$). The response bias was different for the both listener groups, for Northern Sweden listeners the bias value was 0.34 and for Scanian listeners it was –0.04. The difference between the two groups was statistically significant ($t(18) = 2.16$, $p < 0.05$) and indicates that while participants from

Northern Sweden had a slight preference to answering 'no', Scanian participants did not show any 'yes' or 'no' preference.

In the test ST-ST, the mean d'-value of the listeners from Northern Sweden was 1.87 and from Scania 2.00. There was no significant difference between these mean d'-values ($t(18) = -0.47$, $p > 0.05$). The two listener groups' mean d'-value were statistically different from 0 according to one-sample t-tests, so both regions were able to identify the target speaker using his Stockholm dialect. The response bias for Northern Sweden listeners was 0.39 and for Scania listeners 0.36. Both were statistically separated from 0, according to one-sample t-tests. The difference in bias between the two regions was not significant ($t(18) = 0.18$, $p > 0.05$).

For the dialect shifting tests, the results for the test ST-SC show that there was no significant difference between the mean d'-values of listeners from Northern Sweden, 0.64, and from Scania, 0.24, $t(18) = 0.99$, $p > 0.05$). In test SC-ST, the mean d'-value of Northern Sweden was 0.90 and of Scania 0.60. This difference was not significant $t(18) = 0.33$, $p > 0.05$). Taken together, there was no significant difference between the tests ST-SC and SC-ST in mean d'-value for listeners from Northern Sweden ($t(18) = 1.76$, $p > 0.05$), or the listeners from Scania ($t(18) = 0.96$, $p > 0.05$). A one-sample t-test showed that the mean d'-value for both listener groups was not statistically different from 0, and thus identification in both these tests were random. In test ST-SC response bias was significantly distinct from 0 for both listener groups, 0.56 for the Northern Sweden listener group and 1.16 for the Scanian listeners. This difference in bias between the two regions is significant. In test SC-ST the participants from Northern Sweden showed a response bias of 0.62, which is statistically distinct from 0. The listeners from Scania showed a response bias of 1.02, which is also statistically distinct from 0. Once again, the participants from Scania had a significantly higher bias rate ($t(18) = -2.12$, $p < 0.05$) and thus had a higher preference for answering 'no'.

8 Discussion

The results indicate that a switch of dialect can be used as a kind of voice disguise since the listeners find it much more difficult to identify the speaker when a shift of dialect takes place. High identification rates were found for all listeners in both control tests (SC-SC and ST-ST), see Figure 1. However, there were significant differences between listeners from Northern Sweden and Scania in the test SC-SC (see Figure 2). The Scanian listeners were better at identifying the Scanian target voice; this is in accordance with the

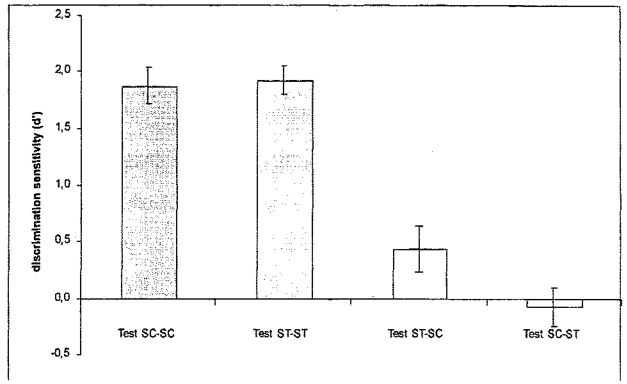

Figure 1. Mean discrimination sensitivity rates for all tests: SC-SC, ST-ST, ST-SC and SC-ST. Standard error of the mean is also displayed.

hypothesis that familiarity with the dialect increases identification ability. The participants from Scania have greater exposure to the Scanian dialect and thus have developed recognition strategies which are superior to speakers of other dialects. The listeners from Northern Sweden have no such extensive exposure and therefore their strategies are not as suited for the Scanian dialect. There was no difference in identification between the two regions in the dialect shifting tests and both listener groups had low identification rates.

There was a significant preference for answering 'no' for all participant groups except for the Scanian listeners in the SC-SC test. The tendency to answer 'no' was stronger in the dialect shifting tests compared to the control tests. This may reflect a stronger uncertainty amongst the listeners when there was a switch of dialect. Also, significant differences in response bias between the two regions were found. In the control test SC-SC, participants from Northern Sweden were more likely to answer 'no' than participants from Scania, who were equally likely to use either 'yes' or 'no'.

It is apparent in this study that dialect is a strong attribute and has a high priority in voice identification. There are distinct differences between the two dialects, Scania and the Stockholm dialect. It is possible that the listeners pay more attention to regional markers than individual voice features. It might not occur to the listeners that a dialect shift was possible as, it is not common that one speaker switches between two different dialects and sounds like a native speaker of both dialects (Chambers 1992, Hazen 2005; see Eriksson et al. (forthc.) for a more detailed discussion of this point).

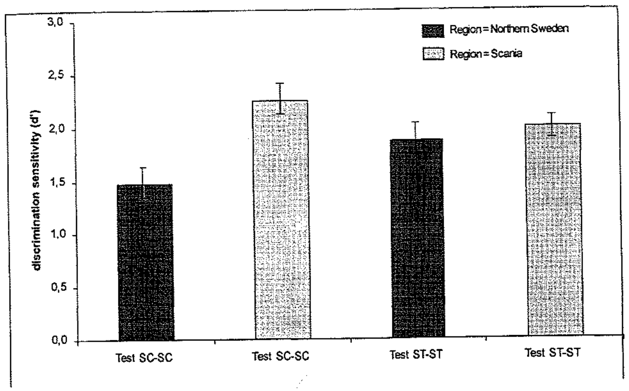

Figure 2. Mean discrimination sensitivity for test SC-SC and test ST-ST within region Northern Sweden and Scania. Standard Error of the mean is also displayed.

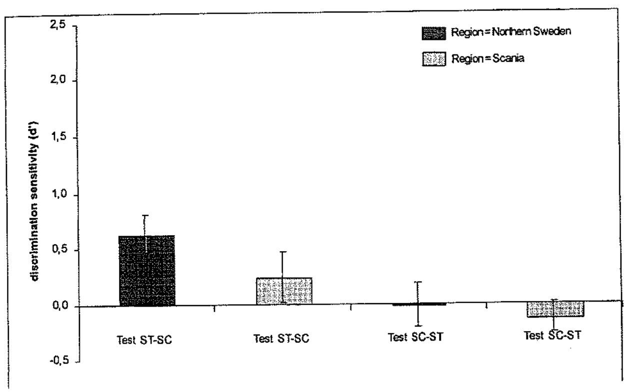

Figure 3. Mean discrimination sensitivity for test ST-SC and test SC-ST within region Northern Sweden and Scania. Standard Error of the Mean is also displayed.

One has to be aware that the conditions of this experiment differ from a real forensic earwitness situation. The participants were not in a stressful or hostile situation, they listened to the target voice for two and a half minutes, it was the same text sample in both the familiarization phase and the identification test, it was studio recordings with a minimum of background noise, they were instructed to remember the voice heard, which could enhance their

memory, and there was a far-shorter delay between the time of the encoding and the identification test that would be the case in a forensic situation.

If a perpetrator uses different dialects when committing a crime and during a criminal investigation and court case, it could result in a miscarriage of justice if the witness is unable to recognize the perpetrator's voice.

Acknowledgements

This paper was partly supported by *Riksbankens Jubileumsfond* through their funding of the project *Imitated voices: a research project with applications for security and the law*. Dnr K2002-1121:1-4, awarded to Umeå University.

References

Bruce, G. 1998. *Allmän och svensk prosodi*. Lund: Dept. of Linguistics and Phonetics.

Chambers, J.K.P. 1992. 'Dialect Acquisition' *Language* 68, 673-705.

Chambers, J.K. & P. Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.

Doty, N.D. 1998. 'The influence of nationality on the accuracy of face and voice recognition' *American Journal of Psychology* 111, 191-214.

Elert, C.-C. 1997. *Allmän och svensk fonetik*. Stockholm: Norstedts.

Eriksson, Erik J. 2007. *That voice sounds familiar: factors in speaker recognition*. PhD Thesis, (*Umeå Studies in Cognitive Science* 1). Umeå: Department of Philosophy and Linguistics, Umeå University, Sweden.

Gibbons, J. 2003. *Forensic linguistics*. Oxford: Blackwell.

Goggin, J.P., C.P. Thompson, G. Strube & L.R. Simental. 1991. 'The role of language familiarity in voice identification' *Memory & Cognition*, 19:5, 448-458.

Green, D.M. & J.A. Swets. 1966. 'Signal detection theory and psychophysics' New York: John Wiley.

Hazen, K. 2005. www.as.wvu.edu/dialect/varres/bidiuopen.pdf, accessed 2006-04-04.

Hollien, H. 2002. *Forensic voice identification*. San Diego: Academic Press.

Hollien, H., W. Majewski & E.T. Doherty. 1982. 'Perceptual identification of voices under normal, stress and disguise speaking conditions' *Journal of Phonetics* 10, 139-148.

Künzel, H.J. 2000. 'Effects of voice disguise on speaking fundamental frequency'. *Forensic Linguistics* 7, 1350-1771.

Markham, D. 1999. 'Listeners and disguised voices: the imitation and perception of dialectal accent'. *Forensic Linguistics* 6, 289-299.

McGehee, F. 1937. 'The reliability of the identification of the human voice' *Journal of General Psychology* 17, 249-271.

Reich, A. & J. Duke. 1979. 'Effects of selected voice disguises upon speaker identification by listening'. *Journal of the Acoustical Society of America* 66, 1023-1028.

Rodman, R. & M. Powell. 2000. 'Computer recognition of speakers who disguise their voice'. *The international conference on signal processing applications and technology ICSPAT2000.* http://www.csc.ncsu.edu/faculty/rodman/Computer%20Recognition%20of%20Speakers%20Who%20Disguise%20Their%20Voice.pdf (accessed 19 April 2008).

Rose, P. & S. Duncan. 1995. 'Naïve auditory identification and discrimination of similar voices by familiar listeners'. *Forensic Linguistics* 2, 1-17.

Schiller, N.O. & O. Köster. 1998. 'The ability of expert witnesses to identify voices: a comparison between trained and untrained listeners'. *Forensic Linguistics* 5, 1-9.

Sjöström, M. 2005. 'Earwitness identification: can a switch of dialect fool us?'. Masters paper, Dept of Philosophy and Linguistics, Umeå university.

Thompson, C.P. 1987. 'A language effect in voice identification'. *Applied Cognitive Psychology* 1:2, 121-131.

Van Lancker, D., J. Kreiman & K.T.D. Wickens. 1985. 'Familiar voice recognition: patterns and parameters, Part 2: Recognition of rate-altered voices' *Journal of Phonetics* 13, 39-52.

Yarmey, A.D. 1995. 'Earwitness speaker identification' *Psychology, Public Policy, and Law* 1, 792-816.

Yarmey, A.D., A.L. Yarmey, M.J. Yarmey & L. Parliament. 2001. 'Commonsense beliefs and the identification of familiar voices'. *Applied Cognitive Psychology* 15, 283-299.

Zetterholm, E. 2003. *Voice imitation: a phonetic study of perceptual illusions and acoustic success*. Lund: Dept. of Linguistics and Phonetics, Lund University.

Zetterholm, E. 2006. 'Same speaker – different voices: a study of one impersonator and some of his different imitations'. *Proceedings SST2006*, 70-75. Auckland, New Zealand.

Lund University, Dept. of Linguistics and Phonetics
Working Papers 53 (2008), 159–177

Deductive chart parsing in Haskell

Marcus Uneson

1 Introduction

Given a formal grammar and a string of tokens, the problem of *parsing* amounts to deciding whether the string is recognized by the grammar; and, if so, returning some suitable representation of its structure. Parsing has ubiquitous applications as a preprocessing step in computational linguistics, for instance in speech recognition, machine translation, and information extraction.

Grammars describing natural language (as opposed to formal grammars, explicitly designed to minimize ambiguities) are notoriously ambiguous, and naive parsing algorithms often have time complexity $O(a^n)$ in the length of the input sequence. *Chart parsing*, originally proposed by Earley (Earley 1970) and Cocke, Kasami, and Younger (Kasami 1965, Younger 1967), is a family of widely used dynamic programming algorithms which achieve $O(n^3)$ running times, by saving partial parses, *items*, in a *chart*, or lookup table. Chart parsing has been generalized (Shieber, Schabes & Pereira 1995) to *deductive parsing*, where a simple, dedicated natural deduction prover allows the parsing process to be described declaratively. In this framework, a particular parsing algorithm corresponds to a particular logic with a particular set of inference rules and axioms; thus, imperatively rather diverse top-down and bottom-up algorithms can be expressed relatively uniformly.

The present paper describes an attempt to transfer (the chart parsing part of) the deduction engine of Shieber et al. from the logical into the functional programming paradigm, with the hope of reaping well-known functional benefits such as referential transparency and higher-order functions without sacrificing either too much declarativity or speed in the process. We draw the outlines of a reasonably efficient deduction engine in the purely functional language Haskell.

The paper is organized as follows. In Section 2, we present the notation and mechanics used for grammar, chart, and parsing logic. The bulk of the paper is made up by the implementational notes in Section 3, where we