uninterrupted speech is the total speaking time for a subject, excluding any and all pauses.) For FT, approximately 4.3 minutes of uninterrupted speech are available, comprising a total of 870 words; for FK 9.5 minutes comprising 2045 words; for MP 10.8 minutes comprising 2554 words; and for FS 10.3 minutes comprising 2401 words.

6 Some remarks on the transliteration provided with the recordings

The data are provided with a word-level transliteration (word labelling). The transliteration was performed by the author, a non-native (albeit competent) speaker of Swedish. Researchers that wish to make use of the data may make use of this transliteration, possibly using it as the basis for searches or subject it to automatic text processing. Therefore, the rationale behind the transliteration conventions will be outlined here.

The aim of the transliteration was to facilitate lexical look-ups rather than to indicate or reflect the segmental content. For instance, the function word *det* is always indicated simply as "det" in the transliteration, without regard for any variability in its production (e.g. [de:t], [de:], [dɛ], [ɾe:] or [d̪e]). This approach was also applied in the labelling of minimal responses and lexical fillers. For example, lexical fillers of the "eh" or "er" type are indicated with a semicolon ; in the transliteration, irrespective of their segmental content (schwa-like, [e]-like, [œ]-like, creaky, nasalized, etc.)

A prominent feature of the transliteration is that contiguous pieces of speech (i.e. stretches of speech which contain no silence pauses) are demarcated at the onset and offset with a period (full stop) symbol. Thus the transliteration does not attempt to reflect the syntactic structure of an utterance, but instead only the presence of silence pauses. Note, also, that the transliteration provides no evaluation of or amendment to the grammaticality of an utterance.

7 Format and availability

The sound files are 16-bit stereo (with one speaker on each channel) and have a sampling rate of 16 kHz. The files are provided in the uncompressed Wave PCM format (i.e. *.wav). The word label files are provided as text files in WaveSurfer format. The data are made available as is, with no guarantee of groundbreaking research results. To obtain the data, please e-mail a request to the author to obtain a web address from which to download the data.

Acknowledgements

The author would like to thank Sven Björsten, Peter Branderud and Hassan Djamshidpey for their assistance with the recording of the data.

References

Anderson, A.H., M. Bader, E.G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. Thompson & R. Weinert, 1991. The HCRH Map Task Corpus. *Language and Speech 34*(4), 351–366.

Edlund, J. & M. Heldner, 2005. Exploring Prosody in Interaction Control. *Phonetica 62*(2–4), 215–226.

Helgason, P., 2002. *Preaspiration in the Nordic Languages: Synchronic and Diachronic Aspects*. Ph.D. Thesis, Stockholm University.

Megyesi, B., 2002. *Data-Driven Syntactic Analysis – Methods and Applications for Swedish*. Ph.D. Thesis, Department of Speech, Music and Hearing, KTH, Stockholm.

Megyesi, B. & S. Gustafson-Čapková, 2002. Production and Perception of Pauses and their Linguistic Context in Read and Spontaneous Speech in Swedish. *Proceedings of the 7th ICSLP*, Denver.

Lund University, Centre for Languages & Literature, Dept. of Linguistics & Phonetics
Working Papers 52 (2006), 61–64

The Relative Contributions of Intonation and Duration to Degree of Foreign Accent in Norwegian as a Second Language

Snefrid Holm
Department of Language and Communication Studies,
Norwegian University of Science and Technology (NTNU)
snefrid.holm@hf.ntnu.no

Abstract

This study investigates the relative contributions of global intonation and global segment durations to degree of foreign accent in Norwegian. Speakers of Norwegian as a second language (N2) from different native languages (L1s) plus one native Norwegian (N1) speaker are recorded reading the same sentence. The N2 utterances' global intonation and global segment durations are manipulated to match the N1 pronunciation. In this way every N2 speaker provides four utterance versions: the original, a duration corrected version, an intonation corrected version and a version with both features corrected. N1 listeners judge the degree of foreign accent between each speaker's four utterance versions. The results show that a) the combined correction of both features reduces the degree of foreign accent for all speakers, b) each correction by itself reduces the degree of foreign accent for all but two of the investigated L1 groups and c) some L1 groups benefit more from intonation correction whereas others benefit more from duration correction.

1 Introduction

When learning a second language after early childhood the resulting speech will normally be foreign accented (e.g. Flege, Munro & Mackay, 1995). The phenomenon of foreign accent is complex and comprises issues regarding the nature of the foreign accent itself as well as the foreign accent's various effects on listeners, for instance regarding social acceptance or the ability to make oneself understood.

A foreign accent may not in itself hinder communication. Although degree of foreign accent is often confounded with degree of intelligibility, a growing body of evidence supports the view that even heavily accented speech may sometimes be perfectly intelligible (Derwing & Munro, 1997; Munro & Derwing, 1995). The relationship between a deviating pronunciation on the one hand and its effect on listener dimensions like intelligibility or perceived degree of foreign accent on the other hand is not clear. There is however a general belief that prosodic deviations are more important than segmental ones, at least for intelligibility, although there are rather few studies to support this view (Munro & Derwing, 2005).

This study aims to establish which of the two pronunciation features global intonation and global segment durations contributes most to perceived degree of foreign accent in Norwegian as spoken by second language learners.

The present paper reports on a study which is part of a larger work where the next step will be to investigate the effect of the same two pronunciation features upon intelligibility. In this

way I hope to shed some light upon the relationship between non-native pronunciation, degree of accent, and intelligibility.

2 Experimental procedure

2.1 Recordings

The speakers were 14 adult learners of Norwegian as a second language with British English, French, Russian, Chinese, Tamil, Persian and German as their L1s. There were two speakers from each of the L1s. The speakers were of both sexes. In addition one male native Norwegian speaker from the South East region was recorded in order to provide an N1 template.

The speech material was recorded in a sound-treated studio using a Milab LSR 1000 microphone and a Fostex D-10 digital recorder. Files were digitized with a sampling rate of 44.1 kHz and later high-pass filtered over 75 Hz. Speech analyses and manipulations were carried out with the Praat program (Boersma & Weenink, 2006).

The N2 speakers and the N1 speaker all read the same Norwegian sentence: *Bilen kjørte forbi huset vårt (=The car drove past our house).*

2.2 Stimuli

Global intonation and global segment durations in the N2 utterances were computer manipulated to match the N1 production of the same sentence. The intonation was manipulated by replacing the intonation contour of each N2 utterance with the stylized intonation contour of the N1 utterance. Because of durational differences between the N1 utterance and the various N2 utterances, the intonation contour had to be manually corrected in the time domain. Because of pitch level differences between the speakers, especially between the sexes, the intonation contour also had to be manually shifted in frequency so as to suit the individual N2 speaker's voice. Manipulation of segment durations required a phonemic segmentation of both the N1 and the N2 utterances. All segment durations were measured and the N2 phonemes were lengthened or shortened to match the segment durations of the N1 utterance.

Three manipulated versions of each speaker's original utterance were generated: one intonation corrected utterance version, one duration corrected utterance version and one utterance version with both features corrected.

The stimuli thus consisted of four utterance versions for each speaker: the original utterance and three manipulated versions. These four versions were put together as pairs. Each pair was put in a separate sound file with a two-second pause in between. These stimulus pairs enabled the direct comparison of each speaker's four utterance versions. Note that each stimulus pair consists of two utterance versions from the same speaker so that utterance versions are always compared within speaker.

2.3 Experiment

13 native Norwegian listeners evaluated the stimulus pairs. None reported experience with N2 speech out of the ordinary and none reported poor hearing. There were 8 listeners from low-tone dialects and 5 listeners from high-tone dialects. The listeners were paid for their participation.

The listener was seated in a sound-treated studio and the sound was presented through loudspeakers. The listener's task was to judge which of the two utterance versions in each stimulus pair sounded less foreign accented than the other. They also had the option to rate the two utterances as equally foreign accented. All stimulus pairs were presented in random order, and they were presented 10 times each.

The listeners were not told that some of the utterances they would hear were altered through computer manipulation. The participants seemed to find the test design comprehensible.

3 Results

The listeners' responses were subjected to statistical testing. However, no statistics will be presented in this paper. The main findings will be presented and discussed in the following.

3.1 Intonation vs. duration

The results show that when both global intonation and global segment durations are manipulated, this correction reduces the amount of perceived foreign accent in the N2 speech. This effect is statistically significant for all N2 speakers across the different L1s.

When the listeners are exposed to speech where only one pronunciation feature is corrected, it is shown that each correction separately contributes to the reduction of foreign accent. This effect is statistically significant for all L1 groups with two exceptions. For the N2 speakers with British English as their L1, the native listeners judge the degree of foreign accent as unaltered despite global intonation correction. For the N2 speakers with German as their L1, the correction of global segment durations does not affect the perceived degree of foreign accent.

It is thus clear that, in general, both global intonation and segment durations are significant contributors to percieved degree of foreign accent. The interesting question is which of these two corrections reduces the degree of foreign accent most effectively. This is shown to vary between the L1s as presented in Table 1 below. The table also shows the relative size of the accent reduction brought about by the corrections.

Table 1. The middle column shows the correction that contributes the most to degree of foreign accent for each of the L1s. The rightmost column shows the relative size of the accent reductions. T= Tamil, C= Chinese, E= English, F= French, G= German. >= larger effect.

L1	Main contribution	Effect size
Tamil Chinese English	Segment durations	T>C>E
French German	Global intonation	F>G
Russian Persian	Equal effect	-

The table shows that the N2 speech produced by speakers with the native languages Tamil, Chinese and English is affected more by the correction of global segment durations than by the correction of global intonation correction for the purpose of foreign accent reduction. Conversely, the N2 produced by speakers with the native languages French and German is perceived as having less foreign accent when the global intonation is corrected than when the global duration is corrected. For the Russian and Persian participants there was no difference between the two pronunciation features. This means that correcting global intonation reduces the amount of perceived foreign accent to the same degree as correcting global segment durations.

The L1s can thus be categorized according to which of the two investigated pronunciation features reduces the foreign accent more than the other. There are however differences within each of these two categories as the degree to which the foreign accent is reduced by a correction differs between the L1s. Native speakers of Tamil, Chinese and English all benefit

most from intonation correction, but the effect of the correction has greater impact on the foreign accent for some L1 groups than for others. The Tamil speakers' N2 is more foreign accent reduced by the correction than the Chinese speakers' N2 and the Chinese N2 speech is more foreign accent reduced than the English speakers' N2. Likewise, speakers with French and German as their native languages benefit most from duration correction, but the foreign accent reduction effect is larger for the French L1 group than for the German L1 group.

The native Norwegian listeners that participated in the experiment represented both low-tone and high-tone dialects. No correlation was found between listener dialect and responses in the perception experiment.

References

Boersma, P. & D. Weenink, 2006. *Praat: doing phonetics by computer* (Version 4.4.17) [Computer program]. Retrieved April 19, 2006, from http://www.praat.org/.

Derwing, T. & M.J. Munro, 1997. Accent, intelligibility and comprehensibility: Evidence from four L1s. *Studies in second language acquisition 19*, 1-16.

Flege, J.E., M.J. Munro & I.R.A. MacKay, 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America 97*, 3125-3134.

Munro, M.J. & T. Derwing, 1995. Processing time, accent and comprehensibility in the perception of native and foreign accented-speech. *Language and Speech 38*, 289-306.

Munro, M.J. & T. Derwing, 2005. Second language accent and pronunciation teaching: A research based approach. *TESOL Quarterly 39*(3), 379-397.

Lund University, Centre for Languages & Literature, Dept. of Linguistics & Phonetics
Working Papers 52 (2006), 65–68

The Filler *EH* in Swedish

Merle Horne
Dept. of Linguistics and Phonetics, Centre for Languages and Literature, Lund University
merle.horne@ling.lu.se

Abstract

Findings from a pilot study on the distribution, function and phonetic realization of the filler EH in interviews from SweDia2000 interviews are presented. The results show that EH occurs almost exclusively after function words at the beginning of constituents. The phonetic realization of EH was seen to be of three basic forms: a middle-high vowel (e.g. [ɐ], [e], [ə]), a vowel+nasal (e.g. [ɐm], [əm], [ən]), and a vowel with creaky phonation (e.g.[ə̰], [ɐ̰]). The vowel+nasal realization occurs as has been shown for English before other delays and is associated with planning of complex utterances. Since creaky phonation is associated with terminality, the creaky vowel realization of EH could be interpreted as signalling the juncture between the filler and an upcoming disfluency.

1 Introduction

The filler, or 'filled pause' *EH* has often been termed a 'disfluency' (e.g. Shriberg, 2001), since it constitutes a delay in the flow of speech associated with referential meaning. However, since it can often be assigned pragmatic functions, such as signalling an upcoming focussed word (Bruce, 1998), or need on the part of the speaker to plan or code his/her speech and thus a desire to hold the floor, *EH* can also be considered to be an integral part of the linguistic system (see e.g. Allwood (1994), and Clark & Fox Tree (2002) who refer to it as a 'word'). In a study on English, Clark & Fox Tree (2002) found that its realization as *Uh* signals a minor delay in speaking, whereas *Um* announces a major delay in speaking.

A number of studies on Swedish have reported some characteristics of *EH* in different speaking styles, but none have focussed on the variation in the phonetic realization of *EH* as far as I know. Hansson (1998), in a study on the relationship between pausing and syntactic structure in a spontaneous narrative, found that the filled pauses (n=22) in her material occurred at clause boundaries after conjunctions and discourse markers and before focussed words. Lundholm (2000) in a study on pause duration in human-human dialogues found that the filler *EH* (n=55) in authentic travel-bureau dialogues occurred in turn non-initial position and had a duration similar to silent planning pauses (mean = 340 ms). Eklund (2004) in a number of studies on simulated human-human and human-machine dialogues found that the most common position of *EH* (n=2601) was utterance-initial before another disfluency and that it was most often followed by *jag* 'I', and *det/den* 'it'. The filled pauses were found to have a mean duration of about 500 ms, thus considerably longer than those found by Lundholm (2000) in authentic task dialogues.

2 Current study

The present study has been carried out to pursue the investigation of *EH* in spontaneous data to get some better idea as to its distribution, function, and phonetic realization in authentic interviews where the speech is basically of a monologue style. Spontaneous speech from the