Paper accepted for publication in *English for Specific Purposes*

Speaking rate and information content in English lingua franca oral presentations

Rebecca Hincks
Unit for Language and Communication
Department of Speech, Music and Hearing
School of Computer Science and Communication
Royal Institute of Technology, KTH, Stockholm
hincks@kth.se
+46 8 790 6918
+46 8 790 6030 fax

Rebecca Hincks is an Associate Professor of English at The Royal Institute of Technology (KTH) Sweden. Her research interests include the use of speech technology in computer-assisted language learning and the study of oral presentations as a spoken genre.

Abstract

This paper quantifies differences in speaking rates in a first and second language, and examines the effects of slower rates on the speaker's ability to convey information. The participants were fourteen fluent (CEF B2/C1) English L2 speakers who held the same oral presentation twice, once in English and once in their native Swedish. The temporal variables of mean length of runs and speaking rate in syllables per second were calculated for each language. Speaking rate was found to be 23% slower when using English. The slower rate of speech was found to significantly reduce the information content of the presentations when speaking time was held constant. Implications for teaching as European universities adopt English as a medium of instruction are discussed.

Keywords: L2 speech, oral presentations, lectures, rate of speech, lingua franca English

1 Introduction

As English continues its growth as a lingua franca[1], more and more speakers across the world find themselves in front of an audience that needs to hear the speaker's message in a language that neither speaker nor listener is entirely comfortable with. One reason for the discomfort can be traced to the extra time it takes to formulate one's message in a second language (L2). Slower English speakers in business meetings have inhibitions about taking the floor from native speakers (Rogerson-Revell, 2007), and international students may be frustrated by their ability to formulate responses quickly enough to contribute to classroom discussion (J. Jones, 1999). Though researchers have begun to explore the effect of L2 language use in interactive situations such as the meeting or the seminar, the ramifications of slower L2 speaking rates when holding an instructional monologue, such as a presentation or a lecture, have not been explored.

Understanding differences in speaking rate is important for many reasons, one of which is the changing linguistic situation in universities across Europe. By facilitating the movement of students between countries, the Bologna Process has instituted a dramatic increase of the use of L2 English in the university classroom (Wilkinson, 2004). For example, at Sweden's Royal Institute of Technology (KTH), the balance between native Swedish students and foreign students has changed greatly in recent years. As many as 70% of its Master's programs are now being given in English to serve the needs of the growing numbers of students who don't speak Swedish. More than 20% of the

[1] The data for this study were gathered in an English-language-learning context, and in that respect do not meet all the criteria established by researchers in the field of English as a Lingua Franca (Jenkins, 2007). However, the student engineers who participated were soon to leave the classroom and enter the lingua franca environment of Northern European industry, and were practicing one of the most critical and high-stakes tasks they would need to perform in their future careers. For that reason, I believe it is appropriate to use the term 'English as a lingua franca' regarding the study.

university's students come from outside Sweden. For teachers, this of course means a switch from teaching in one's native language to teaching in a lingua franca medium. Neither teachers nor students are entirely satisfied with this new linguistic situation. Teachers complain that they lose spontaneity in their teaching; students complain about the quality of their teachers' English. English courses for teachers have been instituted at many northern European universities, but teachers often do not have time to attend them. The pedagogical implications for students of the shift to English-language instruction have also been studied. Klaassen (2001) concluded that, at least after the first year of instruction, a teacher's pedagogical skill was more important than the language used. Airey & Linder (2006), on the other hand, found that when lectured in English, "students asked and answered fewer questions and reported being less able to follow the lecture and take notes at the same time," (2006, p. 558) even though the students themselves had not anticipated differences in the learning situation.

The cognitive demands of using a second language result in a slower rate of speech for most speakers. When time is limited, as it usually is when one is to deliver a lecture or an oral presentation, a slower rate of speech must affect the content of the lecture in one way or another. The best-case scenario would be a more concisely delivered L2 lecture; the worst-case scenario would be that important information was omitted for lack of time. The purpose of the research reported on in this paper was to first quantify differences in speaking rate when speakers hold presentations in their native language and in fluent English, and then to examine the effect of different speaking rates on the information content of the two presentations per speaker.

1.1 Previous work on speaking rate

Temporal aspects of speech can be explored from the perspective of both first and second language speech. As a pioneer in the field, Goldman-Eisler (1968) established that speaking rate in L1 was largely determined by the number and length of pauses rather than by changes in the articulation of individual words. Rate of speech in an L2 was examined in a number of papers in the proceedings of a 1978 workshop on the temporal variables of speech (Dechert & Raupach, 1980). Some of the reported data can be used to calculate the kind of relative (L1-L2) speaking rates that are of interest for the present study: for 20 non-fluent French learners of English, speaking rate (SR) dropped by 33% in the L2, and mean length of runs (MLR, or the number of syllables between pauses) dropped by 43% (calculations based on data published by Deschamps, 1980). Speakers tend also to transfer their pause patterns from their L1 to their L2 (Raupach, 1980, p. 270). The L2 work of Deschamps and Raupach was continued by Towell, Hawkins & Bazergui (1996) who carried out a longitudinal study of 12 English-speaking natives who studied French over a four-year period of time. They found that the variable that best characterized the students' developing fluency was the MLR, the average value of which increased by 22% after a 6-month stay in a French-speaking country. Mean SR increased by 15%. Kormos and Dénes (2004), who studied Hungarian learners of English, also found that the MLR was the variable that best corresponded to judges' ratings of fluency.

The above-mentioned studies, like many studies of learner speech, used elicitation techniques in a controlled laboratory environment to obtain relatively short speech samples that could be examined in terms of a number of phonetic variables. The speech collected in this way—generally by having the subject describe a series of pictures—is spontaneous but at the same time somewhat artificial in that it is not an authentic

communicative situation. Indeed, some of the laboratory studies deliberately created a situation where there was no interlocutor with whom the subjects could interact. Adding a listener component to the study of temporal variables allows for a richer analysis. Lane, Grosjean, LeBerre and Lewin (1973) studied the effect of manipulating temporal variables on listener comprehension, finding that slowing down speech greatly increases comprehension. Griffiths (1990) studied speech comprehension in relation to temporal variables for Japanese learners of English. An interesting finding was that a group of six experienced lecturers were unable to modify their speaking rate to accommodate the linguistic proficiency of their audiences (Griffiths & Beretta, 1991). The single lecturer studied by Camiciottoli (2005), on the other hand, was able to adjust his speaking rate when addressing Italians instead of native English speakers, decreasing it by 32%. Anderson-Hsieh and Koehler (1989) studied the reverse situation from the temporal perspective, looking at an L2-English-speaking teaching assistant addressing L1-English-speaking students, and finding that comprehension rates of heavily accented speech improved as speaking rates dropped from 4.4 to 2.5 syllables per second.

1.2 The present study

Temporal variables have thus been explored from the L1 perspective, the L2 perspective, and various interfaces between them. The work that has been done has been carried out in an attempt to study the cognitive processes underlying linguistic production (Goldman-Eisler, 1968), to understand language typology (Grosjean & Deschamps, 1973), to support a theoretical model for the process of second language acquisition (Towell et al., 1996) or for tools in language assessment (Chambers, 1997; Rekart & Dunkel, 1992). The present study is motivated by other needs that could be described as pragmatic rather

than theoretical. We are now in a situation, at least in Europe, where more speakers than ever before are carrying out their daily business in a second language, English. The fact that speakers speak more slowly in a second language may be obvious but it is not trivial in the globalizing world. The questions asked here is therefore *how much* are speakers slowed down? Does the slower rate of speech mean that when time is limited, parts of an intended message may be left out?

This study has gathered data about the temporal characteristics of not only L2 but also L1 speech. This is not only necessary for the comparative nature of the study, but also because one goal of research into the temporal aspects of instructional speech could be to establish target speaking rates for lecturers and presenters. Studies have shown that comprehension both for L1 and L2 users improves as rates slow (Anderson-Hsieh & Koehler, 1989; Conrad, 1989; C. Jones, Berry, & Stevens, 2007; Zhao, 1997). Native speakers in particular often need to learn to slow down their rate of speech, even when addressing other native speakers (Lynch, 1994). This can be difficult to achieve. Griffiths & Beretta found "no evidence of an intuitively shared feel for a rate at which to pitch ... deliveries" to student groups of varying ability in English (1991, p. 9). Native-speaking teachers need to learn what speaking rates are appropriate.

As noted above, much temporal research has focused on laboratory speech samples, and the little work that has focused on naturally occurring speech has looked at the lectures of university professors, although not in a situation where the same lecture has been delivered in two languages. The genre that is examined in the present study, the oral presentation, differs from laboratory speech in a number of ways. First of all, it can be said to be neither 'read,' nor 'spontaneous' but rather 'guided', 'planned' or 'semi-

spontaneous.' Secondly, it reflects the kind of task that many people regularly meet, with a real communicative need and active listeners. Finally, the seven to ten minutes per speaker used in this study are longer speech samples than have been previously examined in the L2 temporal studies. Instead of mining a small amount of speech for a wide variety of features, the study focuses on the two variables, MLR and SR, that have shown to be most salient in previous research.

Another difference from previous L2 studies is that the speakers in this study, though students, were relatively proficient speakers of English. They represent the upper ranges of English proficiency that can be encountered in northern Europe today, and northern European English speakers are generally seen to be at the forefront of proficiency in the lingua franca context[2] (Erickson, 2004). Using the Council of Europe (Europe, 2001) descriptors, the speakers would be placed in either the B2 or C1 categories regarding oral production. The B2 descriptor for spoken production is as follows: "I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options." The C1 descriptor reads: "I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and rounding off with an appropriate conclusion." The speakers should be seen as representatives of the types of people who need to use English on a daily basis to carry out their work, and who can be frustrated by the extra cognitive load that it entails, despite their skill in the language. The slower speaking rates in L2 that were found in the study should therefore be seen as potentially minimal differences; speaker groups

[2] Readers who would like to make judgments for themselves regarding speaker ability are directed to (journal's website) where sound files containing 30-second samples of each speaker have been placed.

generally less fluent than Swedes are likely to show even larger differences in speaking rate.

A possible objection to simply describing speakers in terms of their speaking rates is that such an approach cannot evaluate qualitative differences in the presentations. A competent speaker could theoretically still include the same basic content in an L2 presentation, even at a slower pace, by eliminating superfluous detail or repetition and being more direct. Therefore, the study of temporal measures (Study 1) was followed by a second study, which examined the information content of the pairs of presentation and quantified the points of difference between them (Study 2).

1.3 Variables used in the study

Speaking rate can be counted in both words and syllables and by both minutes and seconds. Working at the syllable level is a necessity for any kind of comparative study, for two reasons. First, speakers can vary in their preferences for short or long words, and indeed when using a second language are likelier to use shorter, high-frequency words. Over a long stretch of speech, a difference in average word length will affect a speaking rate that is measured in words. Secondly, this is a cross-linguistic study, and although Swedish and English are closely related languages genetically, they use different orthographic conventions that would skew a comparison based at the word level. For example, 'environmental impact report' in English is the compounded '*miljökonsekvensbeskrivning'* in Swedish; the English version contains nine syllables and three words, while the 'one'-word Swedish version contains eight syllables. It can be assumed that the information content of syllables is equivalent when comparing genetically related languages such as English and Swedish. For the purposes of precision,

this study uses the second rather than minute as the length of time, and speaking rate (SR) is thereby expressed in syllables per second (sps). This approach was used in another study of instructional monologue in the L2 environment (Griffiths & Beretta, 1991); other work has measured syllables by minute (Kormos & Dénes, 2004; Towell, Hawkins, & Bazergui, 1996).

Another variable that has been found (Kormos & Dénes, 2004; Towell, Hawkins, & Bazergui, 1996) to be highly relevant in the study of second language fluency is what is known as the mean length of runs (MLR)—what could also be called utterance length, or the amount of speech, in syllables, between pauses. The MLR will generally be shorter in L2 speech than in L1 speech, and in that way give an indication of the frequency of pauses in the speech. Different pause lengths have been used to define the boundaries of the runs, but most studies have used a length between 200 and 300 milliseconds. This study uses a length of 250ms, or one quarter of a second.

The information content of the presentations has been compared using the concept of the ‘idea unit’ (Mayer, 1985). In Mayer’s work, the idea unit was used to compare the information content and structure in expository texts with subjects’ ability to recall what they had read. An idea unit often, but not always, corresponds to a single verb clause, and expresses an action, event or state related to the topic of the text. It is further explained in section 2.3.2.

2 Method

2.1 Speech material

2.1.2 Participants

The fourteen participants in the study, six women and eight men, were Master's students of Engineering at Sweden's Royal Institute of Technology, KTH, taking an elective course in Technical English. All were native speakers of Swedish and had gone through the regular Swedish school system, which begins teaching English at an early age using communicative pedagogy. Young Swedish speakers of English have been shown to be among the best in Europe (Erickson, 2004). Their fluency is generally attributed not only to the success of school instruction, but also to other factors such as the use of subtitles rather than dubbing in foreign language television and films, the linguistic similarities between English and Swedish, and the motivation to learn a second language generated by being a native speaker of a relatively minor European language (Berg, Hult, & King, 2001).

The students were all roughly 24 years old and completing their third or fourth years of engineering studies in a variety of disciplines. They had taken a written diagnostic test upon application to the language department, and had been placed in either the upper intermediate (B2+) (10 subjects) or advanced classes (C1) (4 subjects). The oral presentations were recorded in the second half of the 56-hour courses, so that students had had plenty of time to warm up their spoken English. In addition, the four advanced students had previously held a shorter oral presentation for their classes. In summary, the subjects were fairly fluent speakers of English who can be seen as representative of many Europeans who need to use English regularly as part of their work. In the researcher's opinion, their English was also perceptually on a par with many of their teachers at KTH.

2.1.2 Data collection

Instruction in presentation skills is an important component in the Technical English classes at KTH. The students practice what they have learnt by holding their own presentation and by analyzing other presentations in the form of written peer review comments. The presentation is to be about ten minutes in length though teachers generally do not enforce the time restrictions as strictly as is commonly done in conference settings. The topic for the presentation is up to the student but is to be of a basically technical nature. Table 1 lists the topics chosen by the students in this study.

Table 1. Topics of the oral presentations by speaker

Speaker	Topic	Speaker	Topic
S1(M)	Cell Phones: Health Risks?	S8(M)	Geothermal Heating
S2(M)	How Cats Always Land on their Feet	S9(M)	Steel Production
S3(M)	Cell Phone Security	S10(M)	Treatment of Pollen Allergies
S4(F)	High-speed Trains	S11(M)	Augmented Reality
S5(M)	Quantum Teleportation	S12(M)	Muscles
S6(F)	Hybrid Cars	S13(F)	Flat TVs: LCD or Plasma?
S7(M)	The DARPA Grand Challenge Car Race	S14(F)	Biological Weapons

All students but one used presentation software. S8 used overhead transparencies. None of the speakers used a manuscript. Teachers and classmates gave the students feedback on the class presentation.

During the two different terms when these subjects were studying English, they and many of their classmates allowed their classroom presentations to be audio recorded as part of a larger effort to collect a presentation corpus, presently consisting of about 100 recordings. All Swedish natives were also asked whether they would like to present their presentation again, this time in Swedish, for a small sum of money. Though many students expressed an initial willingness, scheduling difficulties and time pressures in the

end narrowed down to a group of five students in Fall 2004[3], and nine students in Spring 2006. The students were told that they could use the same visual material as they had used for the English material and were assembled in small groups, so that an audience would be present to hear the presentation.

2.1.3 Equipment and software

Audio recordings were made directly into a computer using a small clip-on microphone. Analysis was done using WaveSurfer (Sjölander & Beskow, 2000) to present the speech waveform and enable the measurement of pause length.

2.1.4 Transcription

The 28 presentations were carefully transcribed in a two-step process. First the entire presentation was orthographically transcribed, including filled pauses. Speech recognition was a helpful tool in the English transcriptions. The speaker-dependent dictation software Dragon NatSpeak 9 was trained to the researcher's voice, who then repeated exactly what the speakers said into a microphone—a task that required concentration, but made the transcription process very efficient. A complete, though imperfect, transcription could be produced in real time—10 minutes for a 10-minute presentation. Listening to the presentation two or three more times allowed for correction of the inaccuracies and addition of the filled pauses that the speech recognition is trained to ignore. The vocabulary of the dictation software was impressive, including Swedish place names and rare words such as types of pharmaceuticals and phenomena (e.g. *quantum teleportation*). The following conventions were observed in transcribing, in order to facilitate calculating speech rate in syllables: numbers were written out as words and words that are often

[3] Data regarding the speaking rate and pitch variation of these five speakers was published in (Hincks, 2005a).

reduced, such as *nå-gon* (‘someone’) uttered as *nån*, were transcribed as they were pronounced.

The second phase of transcription, which allowed further correction to any eventual inaccuracies, was to break the transcriptions into ‘runs’, using pauses as boundaries. The speech waveform was used to locate all silent or filled pauses longer than 250 milliseconds. A filled pause is not readily visible in the waveform, and so it was necessary to listen carefully and make run breaks for most of the filled pauses as well, unless they were extremely short ones. The run breaks appear as line breaks in the transcription, but the length of the pauses themselves was not collected as data.

2.2 Calculating temporal measures for Study 1

To calculate speaking rate and mean length of runs, all orthographic representations of filled pauses were removed so that they would not be counted as actual syllables. Each individual run, or line of the transcript, was broken into syllables by inserting spaces to represent syllable boundaries. By marking one line at a time and having the word processor do a word count, syllables per run could be quickly counted and transferred to a spreadsheet to calculate the mean length of runs and the total number of syllables per presentation. The total number of syllables was divided by the length of the presentation in seconds (as shown by the digital recording) to find the speaking rate.

2.3 Analyzing differences in content for Study 2

In order to determine in what ways the two versions of the presentations differed in content, the pairs of transcripts, divided into runs, for each speaker were combined into a single version, represented as a two-column table where equivalent runs were matched to the greatest possible extent with each other in the same row. An excerpt for one speaker

is shown in Figure 1. Recall that the run is a phonetic unit, representing an utterance between two pauses. Where no equivalent run existed, or where a speaker required several runs in one language to express what was expressed in one run in the other, blank cells were left in the table. Because all speakers but one were speaking with prepared slides as support, the order of content was largely the same, and where it was not, it was still possible to match the transcripts at many junctures. Since the transcripts were aligned with each other according to content, rather than time, a 9-minute presentation in English could be aligned with an 8-minute presentation in Swedish.

Swedish (L1) transcript, with English translation	English (L2) transcript
'Screen burn-in', eller *(screen burn-in, or)*	eh Screen burn is a phenomenon when you have a static picture on the plasma screen
skärminbränning *(screen burn-in)*	
	and you have
är något som existerar enbart på plasmaskärmar, och det är när man har en statisk bild på skärmen *(is something that only exists on plasma screens, and it is when you have a static picture on the screen)*	it's no movement at all then the picture can burn into the screen and stay there permanently.
under en längre period, *(for a longer period of time)*	
så kan det bli en liten skugga efter den, man hade problem med det för framförallt de här loggarna som finns på TVkanalen, så har du på den väldigt länge och sen byter du kanal då står det fortfarande en fyra där uppe fast du tittar på trean till exempel, kan du se en skugga bakom den. *(you can get a little shadow from it, there used to be problems primarily from the logos you have from the tv channels, if you have it on for a long time and then you change the channel there's still a 4 up there even though you're watching channel 3, for example, you can see the shadow behind it)*	this was a problem before when you like you viewed a channel for really long time
	and they have like a logo up in the corner and then
	the next day you switch the channel and it still stands TV 4 up here?
eh Plasmatillverkarna idag då de har då skydd mot de här att de skiftar pixlarna lite emellanåt med algoritmer, **men man kan inte skifta de för mycket för då skulle det ju synas,** *(eh the plasma manufacturers today have protection against this they shift the pixels a little between algorithms, **but you can't shift them too much or you'd see it**)*	eh Today eh
	all the plasma screens have screen burn in protection
	it's eh
	just an algorithm that switch
så det är fortfarande ett problem, och **de räknar med att ungefär tio timmar på en statisk bild,** *(so it's still a problem, and **they figure about ten hours for a static picture**)*	eh changes and shifts the pictures a little bit but still it's a problem
	eh for the plasma screens but
på en bra plasma idag, innan det skulle bli någon typ av inbränning. *(for a good plasma today, before there would be any kind of burn-in)*	

Figure 1*. A small portion of the aligned transcripts for S13. Each cell is filled by a run, or a stretch of speech between pauses >250 ms. Aligning transcripts in this way allows the location of information that was omitted or added in one or the other presentation, shown in bold in this example.*

2.3.1 Meta-discourse items

All runs that were devoted primarily to comments about the presentation or the surrounding environment, rather than to the topic of the presentation, were color coded as meta-discourse. These included introductory, summarizing and transitional utterances, such as 'Today I will be talking to you about cell phones'. The meta-discourse items are not considered in the comparative study of informational content. While they clearly contribute to the effectiveness of the oral presentation (Chaudron & Richards, 1986; Jung, 2006), reliable comparisons cannot be safely made between the Swedish data and the English data because the students had been provided in their English classes with an extensive set of meta-discourse items to use in their presentations.

2.3.2 Differences in content: Points of difference (PODs)

The remainders of the presentations, i.e. excluding the runs devoted to meta-discourse, were coded for differences in informational content. To achieve consistency in assessment, this process was repeated twice for each speaker, and the results of the second analysis used. This process moved the analysis from the run, a phonetic division, to the idea unit, a semantic extraction from the transcript. Any piece of information related to the topic of the presentation that appeared in one version but not the other was highlighted. The bolded phrase from Figure 1, *"de räknar med att ungefär tio timmar på en statisk bild,"* ("they figure it takes about ten hours for a static picture") is an example of a fact about screen burn-in presented in Swedish but not anywhere in the presentation in English. This type of discrepancy was termed a 'point of difference' (POD), in this case in favor of the L1. It was expressed as an idea unit in a separate list of this speaker's L1 PODs, appearing as point number 11: "screen burn-in can happen after about 10

hours." Lists were also made for L2 PODs, points that were made in English but not in Swedish. A sample list of points of differences for one speaker is shown in Appendix 1. The lists were carefully reviewed to ensure they followed the concept of the 'idea unit' (Mayer, 1985) by consisting of single verb clauses rather than complex sentences. This meant that conditional sentences, sentences with relative clauses, and sentences of cause and effect were broken into two or more idea units. Semantic criteria overrode syntactic, however. It was important that the idea unit reflect a piece of information that a listener might be expected to recall from the presentation, again following Mayer's work on the recall of expository writing. For example, the relative clause in "antibodies are things *that catch a dangerous substance*" was not counted as a separate idea unit because of the emptiness of the word 'thing', while "then they attach to other cells *that destroy the organism*" was counted as two idea units. Coordinated items were counted as more than one idea unit only if the speaker repeated the subject and verb each time, so that "it's used in skyscrapers, stadiums and megastructures" was one idea unit, while "cortisone is a steroid, it affects your fluid balance, it affects your growth, it affects your bone mass" was counted as four idea units. This choice was made to reflect the amount of speech that was used to express the idea.

2.3.3 Time-normalized points of difference

The total number of points of difference itself, however, is not sufficient to answer the question of how speaking rate affects the informational content of the presentations. Most of the L1 presentations were of course shorter than their L2 equivalents, though a few of them were longer as the students took the opportunity to expand on the topic. It was therefore necessary to normalize the two presentations for time. The length of the shorter

presentation in a pair was used as the baseline, and its time position marked in the aligned transcripts. Instead of counting only points of difference after this point, idea units themselves were counted. If time had been kept constant, all idea units of the longer presentation would never have been uttered, so they were subtracted from the total. Likewise, after this point the idea units of shorter presentation had already been uttered (in time, not in the content-based alignment) so they needed to be added to the totals.

3 Results

3.1 Study 1

Table 2 presents the temporal data collected in Study 1. MLR and SR will be discussed below. The table clearly shows the extent to which speakers produced more speech in a shorter amount of time in their first language. The mean total syllables in L1 was slightly longer than that for L2: 2137 vs. 1879 (a significant difference, t (13) = 1.92, $p < .05$, one-tailed.) The mean speaking time, in contrast, was one minute shorter for L1 than L2; 538 s. vs. 596 (roughly 9 vs. 10 minutes). This difference was also significant: t (13) = -1.92, $p < .05$, one-tailed.

Table 2. Temporal data gathered for first study.

Speaker	Mean length of runs (MLR) (syllables)		Total Syllables		Total Time (seconds)		Speaking Rate (SR) (syll/s)	
	Swedish L1	English L2	Swedish L1	English L2	Swedish L1	English L2	Swedish L1	English L2
S1(M)	11.57	6.40	2244	2272	588	836	3.82	2.72
S2(M)	10.98	6.94	1383	1534	439	648	3.15	2.37
S3(M)	8.94	7.09	1225	1318	370	475	3.31	2.77
S4(F)	10.55	8.00	1889	1568	535	568	3.53	2.76
S5(M)	11.49	8.23	2367	1844	545	576	4.34	3.20
S6(F)	8.84	8.45	1538	1369	477	472	3.22	2.90
S7(M)	9.27	9.14	1660	1773	506	597	3.28	2.97
S8(M)	11.14	9.23	1437	1938	355	580	4.05	3.34
S9(M)	11.37	9.87	2411	1797	602	521	4.00	3.45
S10(M)	14.79	10.36	3934	2487	840	700	4.68	3.55
S11(F)	14.26	10.99	2737	2571	676	752	4.05	3.42
S12(F)	12.96	11.16	1361	1296	411	425	3.31	3.05
S13(F)	20.36	12	3502	2519	722	696	4.85	3.62
S14(F)	19.73	15.34	2229	2025	464	491	4.80	4.12
SD	**3.62**	**2.38**	**820**	**446**	**138**	**119**	**.61**	**.46**
Mean	**12.59**	**9.51**	**2137**	**1879**	**538**	**596**	**3.89**	**3.12**

3.1.1 Mean Length of Runs

Figure 2 illustrates the mean length of speaker runs (MLR) in syllables. In the figure the speakers are ordered according to the MLR in L2, which has been shown to correlate best with ratings of fluency (Kormos & Dénes, 2004; Towell, Hawkins, & Bazergui, 1996). The speakers on the left are thus generally less fluent in English than those on the right. The speakers have been assigned identification numbers according to this variable.

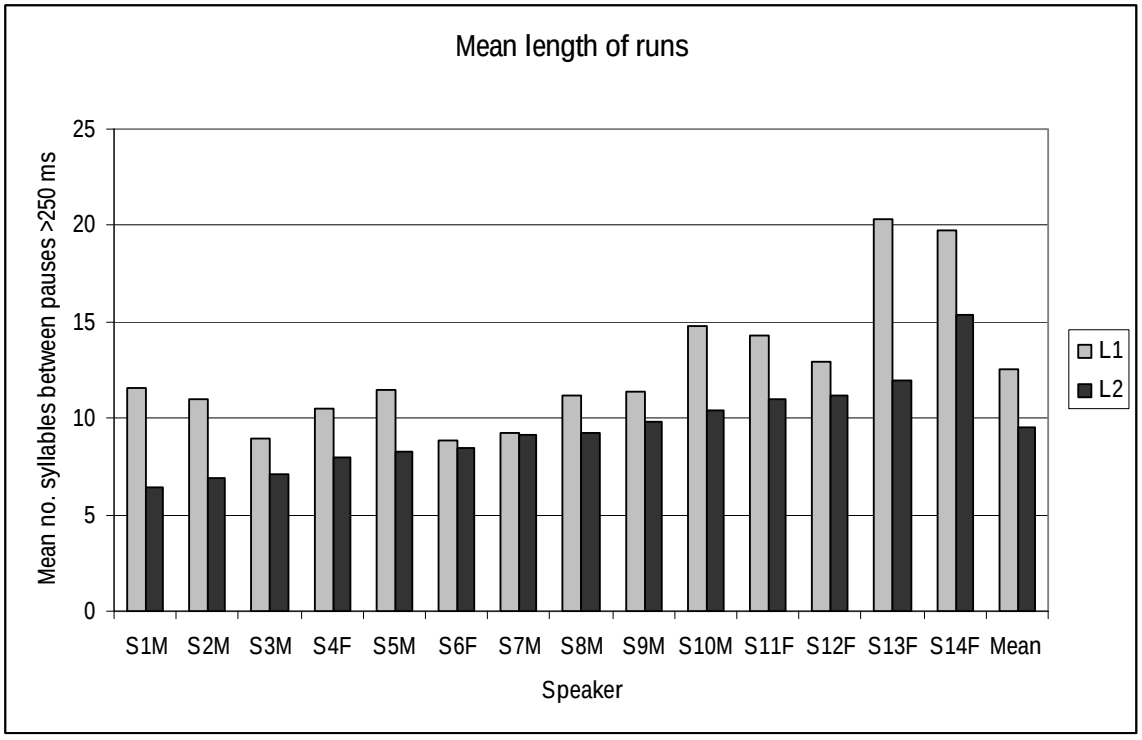

Figure 2*. Mean length of runs (MLR) per speaker for L1 (Swedish) and L2 (English). All speakers have a shorter MLR in their second language.*

All speakers produced shorter run lengths in L2 than in L1. The means were 12.59 syllables per run in L1 and 9.51 syllables per run in L2, a mean difference of 3.08, SD 2.15. This shorter run length in L2 is statistically significant: $t\,(13) = 5.35$, $p < .01$, two-tailed. Two female students, S13 and S14, had exceptionally long MLRs of about 20, indicating that they barely paused at all as they held their presentations in Swedish. S13 was a successful salesperson of the subject of her presentation, flat-screen TVs; note that her MLR drops by nearly 50% when she is operating in English. The run lengths by speaker correlate strongly between languages: R=0.82, indicating that speakers with short or long runs in their L1 maintained that relative characteristic in the L2.

3.1.2 Speaking rate

The mean SR for L1 was 3.89 sps, and in L2 3.12 sps. The slower speaking rate in L2 is statistically significant: t (13) = 8.39, $p < .01$, two tailed. This can also be expressed as a mean difference of 20.8%, where L2 is 23% slower than L1, and L1 is 18.7% faster than L2. Figure 3 illustrates the individual speaking rates, once again ordered according to the English SR. As expected, all speakers spoke more quickly in L1: at least 3 sps, with three speakers approaching a speaking rate of 5 sps. In L2 the rates range from a low of 2.37 sps to a high of 4.12 sps. As with MLR, the SRs between languages correlate strongly, R=0.85, so that speakers who speak relatively slowly in their L1 also tend to do so in the L2. They also naturally correlate by speaker with mean length of runs (Figure 1): 0.82 for L1, and 0.89 for L2.

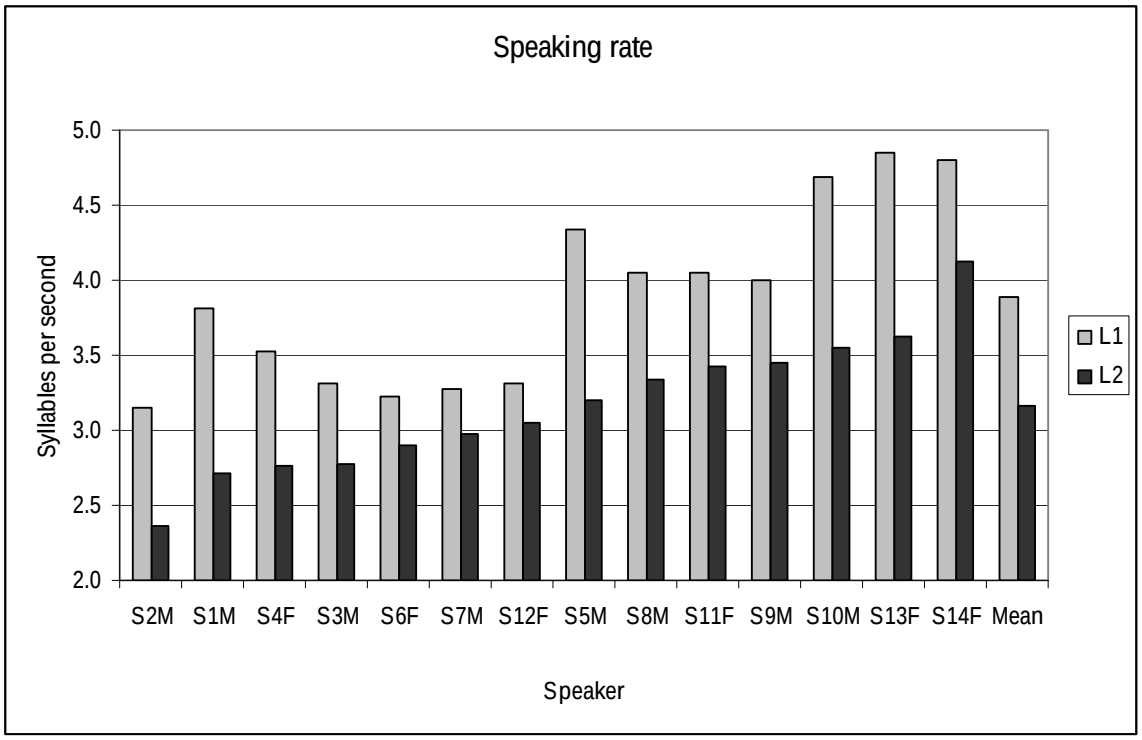

Figure 3. *Speaking rate in syllables per second by speaker for L1 (Swedish) and L2 (English). The mean speaking rate is 23% slower in L2. Subject identification numbers indicate their rank order in MLR, but subjects are ordered by L2 SR.*

3.2 Study 2

The second study investigated differences in informational content in the presentations in English and Swedish. Results, in the form of total numbers of points of difference (PODs), are presented first for the complete presentations, which in most cases differed in the total speaking time. Because it is misleading to compare the informational content of, for example, a 10-minute presentation with a 7-minute presentation, results are then presented normalized for time.

3.2.1 Total points of difference

When speakers complete their presentations as planned, they are able to convey about as much information in one language as in the other. Figure 4 illustrates the variation between speakers, where the bars indicate the number of items of information about the topic of the presentation that were mentioned in one presentation but were not mentioned in the other presentation, as described in section 2.3.2.

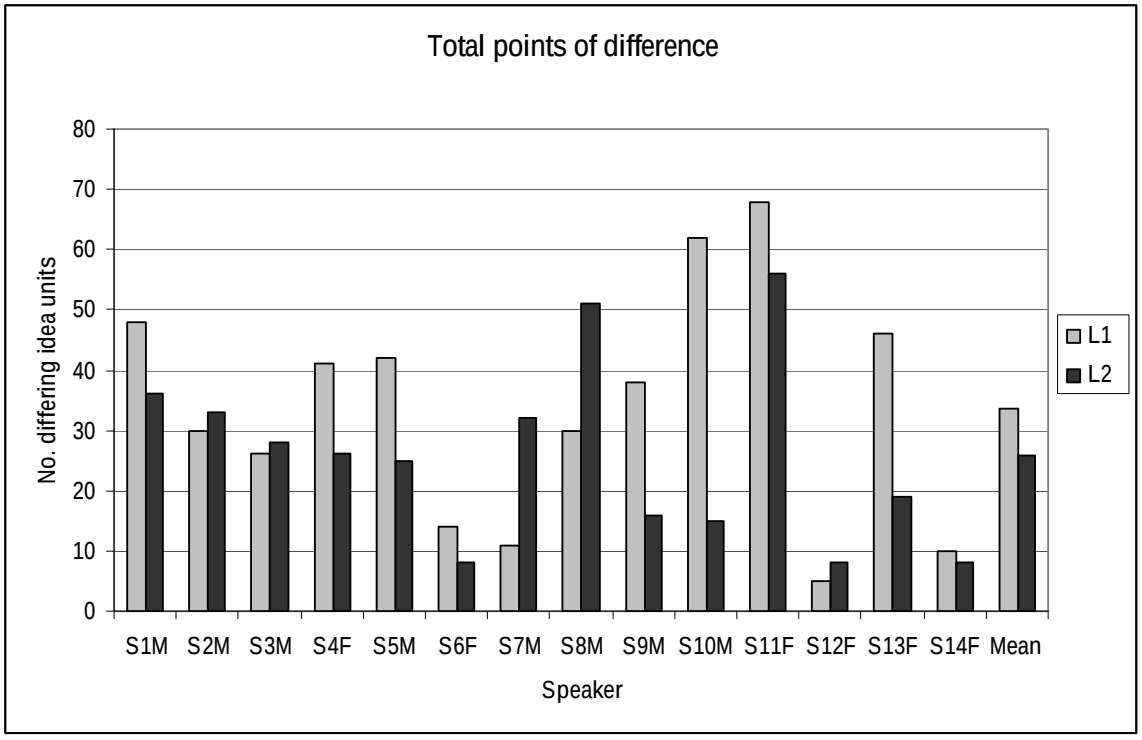

Figure 4*. Points of difference in favor of L1 and L2, over the entire presentation as delivered. The differences are insignificant.*

Speakers with low differences, such as S6, S12, and S14, have closely followed an internal plan as to what they wanted to say in the two languages, and have diverged little from the plan in their presentations. They have also spoken for about the same amount of time in the two languages. Other speakers, such as S10 and S13, have added a lot of new information to their L1 presentations, perhaps because they were very knowledgeable about their topics and took the opportunity to elaborate on their L2 preparations. Speaker 11, who has high POD values for both L1 and L2, spoke freely and relatively spontaneously in her presentation, which accounts for the differences between them. While most speakers favor the L1 in terms of introducing new information, this is not the case for S7 and S8. S7 admitted to not having thought about how to explain his topic, the American DARPA Grand Challenge, in Swedish, and was unable to convert information such as *miles per hour* into *kilometers per hour* on the spot as he spoke. S8, who spoke

without the use of presentation software, quite simply held a much shorter presentation in Swedish than in English. Though speakers on average contributed eight more pieces of information in their L1, the differences in informational content over the length of the whole presentation are insignificant t (13) = 1.61, $p < .05$.

3.2.2 Time-normalized points of difference

When the presentations are normalized for time, however, the picture changes. Normalization meant, in most cases, subtracting a number of PODs from the L2 presentations, points that were made at the end of the longer presentation when the speaker was in effect speaking in overtime (see section 2.3.3). It also meant adding a number of points to the L1 presentation, which had not previously been counted as PODs because the information had been presented in the full versions of both presentations. Because the parallel information was given at a point after which the longer presentation had been cut off, it became a POD, because in the longer presentation the speaker would never been able to explain these points if time had been equal. Averaged across all the speakers, normalizing meant that five points were subtracted from the L2 presentations and as many added to the L1 presentations. However, some speakers needed the converse, points added to their L2 presentations, but here the average across all speakers was only 2.

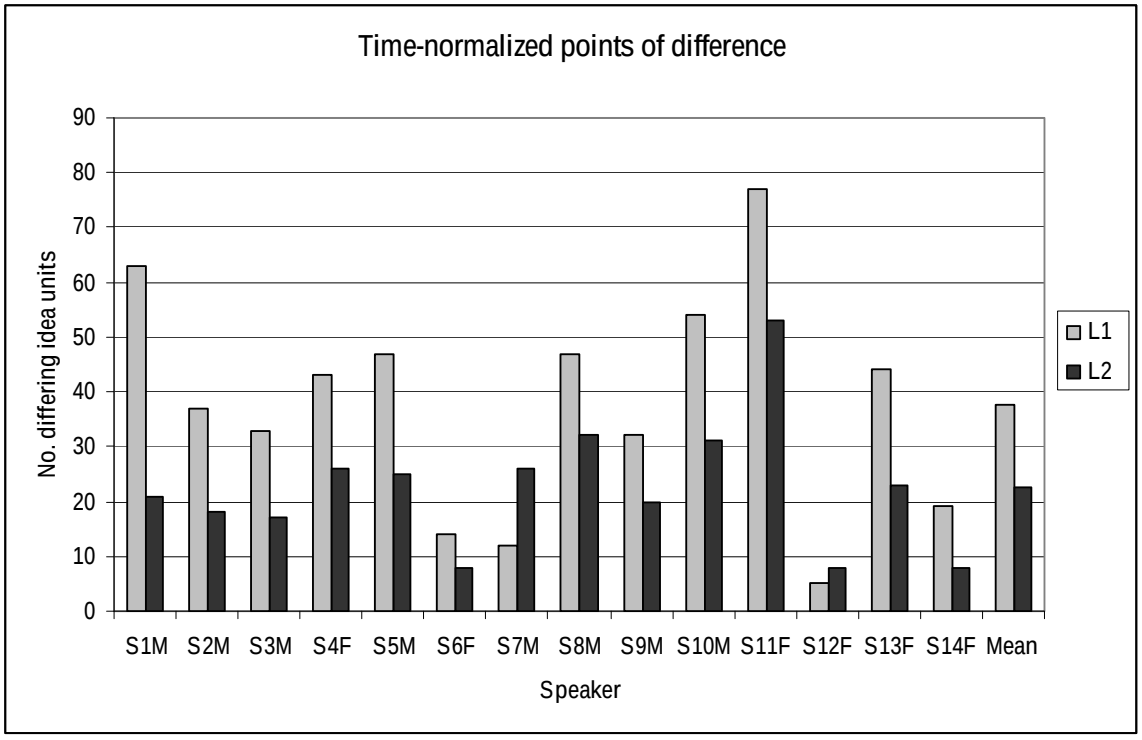

Figure 5*. Points of difference between presentations, when time is kept constant. Differences are significant. Only speakers 7 and 12 say more about their topics in the L2.*

Figure 5 gives a much more consistent picture than Figure 4. While we see that one speaker, the relatively unprepared S7, has lower values in L2 than in L1, nearly all of the other speakers produce many more PODs in relation to time in their first languages. Here the effects of the different speaking rates are translated into a more efficient communication of information. On average, the speakers contribute 15 more pieces of information about their topic when they have the opportunity to speak their own language and time is kept constant. The differences are significant: t (13) = 4.28, $p < .01$, two-tailed.

4. Discussion and conclusions

4.1 Temporal measures

Both similarities and differences between the L1 and L2 presentations have been revealed by this examination of two temporal variables: the amount of speech uttered between

pauses (MLR), and the speaking rate (SR), including pauses, over 6-14 minutes. To begin with the similarities, there is a strong effect of individual speaking style between the performances in the two languages. The correlations between L1 and L2 of 0.82 (SR) and 0.85 (MLR) show that those speakers who used shorter run lengths and slower rates of speech in one language used them in the other language as well, confirming previous work done on laboratory speech (Deschamps, 1980; Raupach, 1980; Towell, Hawkins, & Bazergui, 1996). Though other researchers have suggested using run length to automatically measure fluency in second languages, it is important to recognize that run length differs in one's first language as well—that fluency is a concept with relevance to our command of our native languages.

The main research issue addressed in Study 1 was an attempt to quantify the effect on speaking rate of using an L2 in the oral presentation situation. Using English instead of their native language meant that all speakers had shorter run lengths and slower rates of speech. On average, using English slowed the speakers down by 23%. The difference can be attributed to the frequent short pauses—as evidenced by the shorter run lengths—that are necessary for the speakers to find the formulations they need in L2. A long run length shows that that linguistic knowledge has been proceduralized (Levelt, 1989; Towell, Hawkins, & Bazergui, 1996), as exemplified in this study by S13, whose high MLR in L1 can be attributed to the fact that she was a practiced salesperson of her subject. When using a second language, however, the participants in this study, though they were speaking about material they themselves had prepared and were fluent speakers of English, show the degree to which operating in a second language affects the cognitive processes underlying speech production.

In terms of this study's relationship to previous research on temporal variables of speech in the L2 context, it is evident first of all that the mean MLR and SR values found in the present study are much higher than those found in Towell et al.'s (1996) longitudinal study of learners of French, both for the L1 and the L2. This indicates the different natures of the speaking tasks: presenting a planned talk is very different than describing material the researcher has prepared for you. However, the discrepancies of the speaking rates in the Towell et al. study are in line with the results found here. Those students spoke 27% more slowly in L2 than in L1 at the beginning of their study, and then 16% more slowly after their six-month stay in the L2 country.

This study has compared larger amounts of speech per person than are found in many laboratory speech studies (cp Griffiths 90 sec., Kormos & Dénes 2-3 minutes, Towell et al. 1-2 minutes), between 6 and 14 minutes of monologue per person, resulting in temporal variables that are very reliable indicators of an individual's speaking style in a real communicative situation. However, there are a few factors that call for caution when comparing the L1 and L2 presentations with each other. First of all, the L1 presentations were all made following the L2 presentations, for practical reasons. A randomized order would be preferred to prevent any practice effects. Any possible effect could be mitigated by the facts that a period of several weeks elapsed between the two presentations, and the likelihood that the speakers practiced more before the L2 than the L1 presentation. The L2 presentation was a high-stakes event, where the speakers were in effect being examined for a significant part of the final grade for their English course. In coming to be recorded in their L1, they were as likely to be driven by monetary motivations as by a desire to contribute to the understanding of the dual-language

situation in the university world. The transcripts reveal that a number of the speakers did not know how to express some concepts in Swedish, indicating that they had not thought through their presentations before coming to be recording. This discrepancy between the two presentation situations is mirrored in real-life: speakers who must work in a second language are likely to practice ahead of time to make sure that they have command of the vocabulary and expressions they must use to communicate. If these speakers had been better prepared for their L1 presentations, even larger differences between L1 and L2 would have been found, so these methodological imperfections should not negatively impact the validity of the results.

4.2 Differences in content

This study might dispel any illusions that L2 speakers can manage to deliver the same amount of information despite their slower rate of speech in an L2. When time was not controlled, there were some differences in information content, but they were not large, indicating that the speakers were proficient in English and well-prepared for their task. When time was kept constant, however, the slower speaking rate meant that information was left out. The least-fluent English speaker, S1, was so hampered by L2 that his L1 presentation included three times the PODs when time was normalized. However, the gap between the two languages was also apparent for more fluent speakers, such as S13 and S10. These two participants were exceptionally knowledgeable about their topics (one was a salesperson and one went on to do a PhD related to the subject of his presentations), raising the worrisome concern that the more one knows about a topic, the bigger the differences that appear when one is required to talk about it in an L2. If this could be

shown to be true, it would have serious implications for lecturing in a second language. Content could be eliminated due to linguistic constraints.

The study has used the idea unit as a unit of measurement adequate to establish quantifiable differences in content. However, there are other differences between the presentations that would be trickier to measure. One of these might be the use of metaphor or the frequency of adjectives, aspects that add important detail for the listeners. For example, as can be seen in Figure 1, S13 in her L1 talks about screen burn-in as being like a 'shadow', mentioning the word *'skugga'* twice in her lively description. She doesn't use the image at all in English, and the presentation is the poorer for it. Yet she has managed to define burn-in, and so the differences here were not considered as being worthy of classification as a POD. Thus, a finer-grained in-depth analysis could possibly uncover even stronger differences in content than the already significant differences revealed by using idea units.

Another phenomenon revealed by the study is that of domain loss in the first language. Though several speakers at times had to search for terminology in Swedish, one speaker, S7, was at such a loss to explain an American road race in Swedish that the information content of his L1 presentation suffered.

4.3 Implications

The central purpose of this study has been to raise awareness of some of the measurable effects of using an L2 to do a task one normally does in an L1. In the rush to attract foreign students, European universities are shifting their mediums of instruction at an unprecedented rate. Course plans and schedules need to be adapted to accommodate this. Let us consider what would happen if the results of this study were extrapolated from a

ten-minute presentation to a 45-minute lecture. If the rate of a delivery of a 45-minute lecture is slowed down by 25%, then the lecture will take closer to an hour to finish. If information is omitted from the L2 lecture at the same rates as were found in this study, then a 45-minute lecture could lack as much as 60 pieces of information that would have been mentioned in the lecturer's first language. Time constraints become even further tested when students who are cautious about using English save all their questions for after class, as found by Airey and Linder (2006). If it is not possible for university administrations to schedule more time for a course that is being given in a lingua franca, then teachers need to adapt the course content. The challenges faced by L2 speakers extend beyond the classroom—other measures that could be considered to accommodate them could include variable speaker time at conferences and other gatherings.

The slow-down effect of 20-25% that was found in the study needs to be seen as a conservative estimate, given the facts that the students were relatively fluent speakers of English and had prepared and practiced for their English presentations. Airey (2009, p. 91) compared speaking rates of students being interviewed in both English and Swedish about their learning experiences, and found a whopping 45% difference in speaking rates.

Faster lecturing is generally not better, far from it. While teachers using an L2 may be constrained by combinations of their own speaking style and their L2 proficiency, L1 teachers have at least the theoretical possibility of choosing a speaking rate that is appropriate for the audience and context. Yet this can be extremely difficult to do. Griffiths & Beretta (1991) found that native-speaking English teachers lectured at about 3.5 sps (a rate that they felt was too fast) regardless of the English proficiency of their audiences. When using their L1, the Swedish speakers in the present study spoke even

more quickly: 3.9 sps. The slower L2 rate of 3.1 sps is likely to be more appropriate for an audience that is also composed of L2 speakers—testing this hypothesis would be an interesting area of future research. Griffiths & Beretta concluded that people can adapt their speaking rate when they are in conversational interaction with a less proficient user, but that they lose their sense of what is appropriate when the feedback is taken away, for example in a monologue. Therefore, "training in rate perception and modification should be more rigorously incorporated into teacher training programs" (1991, p.16) so that teachers can learn to slow down their speech when necessary. Speech engineers could contribute to the pedagogy of public speaking by developing applications that give online feedback on rate of speech, so that speakers could be warned when they begin to speak too quickly (Hincks, 2005b). Indeed, present-day dictation software could give this kind of information after the fact, by calculating the words transcribed in relation to the time spent speaking. Future developments in speech processing may enable the creation of live feedback, by means of, for example, a light that flashed when speech rate exceeded 3.5 sps.

English native speakers have an advantage in the global economy, an advantage that brings with it an obligation to make allowances for their colleagues who are not operating in a medium of their own choice. Acknowledging the simple need for more time is one way of meeting that obligation.

References

Airey, J. (2009). *Science, Language and Literacy.* Uppsala University, Uppsala.

Airey, J., & Linder, C. (2006). Language and the experience of learning university physics in Sweden *European Journal of Physics 27*, 553-560.

Anderson-Hsieh, J., & Koehler. (1989). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning, 38,* 561-613.

Berg, E. C., Hult, F. M., & King, K. A. (2001). Shaping the climate for language shift? English in Sweden's elite domains. *World Englishes, 20*(3), 305-319.

Camiciottoli, B. (2005). Adjusting a business lecture for an international audience: a case study. *English for Specific Purposes, 24*, 183-199.

Chambers, F. (1997). What do we mean by fluency? . *System, 25*(4), 535-544.

Chaudron, C., & Richards, J. (1986). The Effect of Discourse Markers on the Comprehension of Lectures. *Applied Linguistics, 7*(2), 113-127

Conrad, L. (1989). The effects of time-compressed speech on native and EFL listening comprehension. *Studies in Second Language Acquisition, 11*(1), 1-16.

Dechert, H., & Raupach, M. (Eds.). (1980). *Temporal Variables in Speech* The Hague Mouton

Deschamps, A. (1980). The syntactical distribution of pauses in English spoken as a second language by French students In *Temporal Variables in Speech* (pp. 255-262): Mouton.

Erickson, G. (2004). *English: here and there and everywhere. En undersökning av ungdomars kunskaper i och uppfattningar om engelska i åtta europeiska länder*. Stockholm: Skolverket.

Europe, C. o. (Ed.). (2001). *Common European Framework of Reference for Languages*: Cambridge University Press.

Goldman-Eisler, F. (1968). *Psycholinguistics. Experiments in Spontaneous Speech* London: Academic Press.

Griffiths, R. (1990). Speech rate and NNS comprehension: a preliminary study in time-benefit analysis. *Language Learning 40*(3), 311-336.

Griffiths, R., & Beretta, A. (1991). A controlled study of temporal variables in NS-NNS lectures. *RELC Journal, 22*(1), 1-19.

Grosjean, F., & Deschamps, A. (1973). Analyse des variables temporelles du francais spontané II. Comparison du francais oral dans la description avec l'anglais (description) et avec le francais (interview radiophonique). *Phonetica 28*, 191-226.

Hincks, R. (2005a). *Computer Support for Learners of Spoken English.* Royal Institute of Technology, Stockholm.

Hincks, R. (2005b). Measures and perceptions of liveliness in student oral presentation speech: a proposal for an automatic feedback mechanism. *System, 33*(4), 575-591.

Jenkins, J. (2007). *English as a lingua franca: Attitude and Identity*. Oxford: Oxford University Press

Jones, C., Berry, L., & Stevens, C. (2007). Synthesized speech intelligibility and persuasion: Speech rate and non-native listeners. *Computer Speech and Language, 21*, 641-651.

Jones, J. (1999). From Silence to Talk: Cross-Cultural Ideas on Students' Participation in Academic Group Discussion *English for Specific Purposes, 18*(3), 243-259.

Jung, E. H. (2006). Misunderstanding of academic monologues by nonnative speakers of English. *Journal of Pragmatics, 38*, 1928-1942.

Klaassen, R. (2001). *The International University Curriculum: Challenges in English-Medium Engineering Education* Delft University of Technology Delft, The Netherlands

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*, 145-164.

Lane, H., Grosjean, F., LeBerre, J., & Lewin, E. (1973). Exploring some properties of foreign-language utterances that control their comprehension. *Linguistics 112*, 15-22.

Levelt, W. (1989). *Speaking: from Intention to Articulation*. Cambridge: MIT Press.

Lynch, T. (1994). Training lecturers for international audiences. In J. Flowerdew (Ed.), *Academic Listening: Research perspectives* (pp. 269-289). Cambridge: Cambridge University Press.

Mayer, R. E. (1985). Structural Analysis of Science Prose: Can We Increase Problem-Solving Performance? . In B. K. Britton & J. B. Black (Eds.), *Understanding Expository Text: A Theoretical and Practical Handbook for Analyzing Explanatory Text* (pp. 65-89). Hillsdale, New Jersey: Lawrence Erlbaum Associates

Raupach, M. (1980). Temporal variables in first and second language speech production In *Temporal Variables in Speech* (pp. 263-270): Mouton

Rekart, D., & Dunkel, P. (1992). The Utility of Objective (Computer) Measures of the Fluency of English as a Second Language. *Applied Language Learning, 3*, 65-85.

Rogerson-Revell, P. (2007). Using English for International Business: A European case study. *English for Specific Purposes, 26*, 103-120.

Sjölander, K., & Beskow, J. (2000). *WaveSurfer: An open source speech tool.* Paper presented at the International Conference on Spoken Language Processing 2000, Beijing.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The Development of Fluency in Advanced Learners of French. *Applied Linguistics, 17*(1), 84-119.

Wilkinson, R. (Ed.). (2004). *Integrating Content and Language: Meeting the challenge of multilingual education* Maastricht University of Maastricht

Zhao, Y. (1997). The Effects of Listeners' Control of Speech Rate on Second Language Comprehension. *Applied Linguistics, 18*(1), 49-68.

Appendix 1: Points of difference for S3

In English but not in Swedish

1. the cell phone sends out signals
2. from the base station you're connected to a cell phone company
3. you're connected to another base station
4. if you're calling another cell phone
5. that's the basic principle of communicating
6. 3g is to gsm as gsm is to nmt
7. you have converted your message into a series of zeros and ones
8. you have a carrier signal
9. that is at some frequency
10. it's a simple and beautiful method
11. the method is also a liability
12. because it's easy to detect with proper equipment
13. you're not allowed to use the base station either
14. the fact that the signal goes everywhere
15. means that it gets into public places
16. this is legal
17. gsm has 140 bit encryption

Cutoff point for English. The following PODs subtracted for time normalization.

18. every gsm phone has a receiver
19. that can break codes
20. it does this every time you speak
21. if you can record the signal
22. you can translate it back
23. 3g hasn't been built out all the way
24. gsm is safer than 3g
25. you can still record and analyze the signals
26. you don't have that luxury in gsm
27. don't put out your credit card number
28. normal conversation is not a problem)

In Swedish but not in English

1. surveillance means that someone is listening to
2. and understanding what you say
3. the base stations sense that the phone has started
4. you always have a primary base station
5. the difference between nmt and gsm is that gsm is digital
6. you use phase shifting
7. 3g uses the same principle
8. but with more complicated patterns
9. you're not allowed to break in anywhere in the hardware
10. you're not allowed to bug people
11. the path from the base station to the phone is not regulated
12. there you can do whatever you want
13. to break the code you need to know what kind of phone one has
14. you can break the code immediately with the right software
15. 3g is more complicated than gsm
16. it has several levels
17. where you can determine the coding level
18. but lots of people still use gsm
19. the same is true for sending data via the cell phone
20. it usually isn't well encrypted
21. if you use a regular computer with fiber optics
22. the legislation protects you
23. the legislation forbids listening
24. it is also possible to detect this in other ways
25. if you send out waves in space
26. anyone can listen

Idea units after cutoff point. These points did appear in both presentations, but since they were 'squeezed in' to the shorter presentation in the portion of the parallel transcripts that was coded as overtime, they need to be counted for normalization.

27. with Bluetooth you can structure the security
28. depending on whether the environment is secure or insecure
29. with gsm the security is not as strong
30. be aware of what you are saying
31. there is a big risk of being listened to
32. be careful especially with doing financial business
33. be careful with anything you don't want listened