Foreign accent, cognitive load and intelligibility of EMI lectures

Christian Jensen and Jacob Thøgersen, University of Copenhagen

Abstract
This study investigated the effect of foreign accent on the understanding of spoken texts in two different contexts: (1) when listeners extract surface level meaning of simple utterances, labelled "intelligibility in simple tasks" (IS) below, and (2) when they answer content questions to a complex text, labelled "intelligibility in complex tasks" (IC) below. We hypothesised that foreign accented speech would require more cognitive processing in all situations, but that it would have little detrimental effect on intelligibility in the simpler of the two tasks. We expected decreased intelligibility as a result of combining the increased cognitive workload of the foreign accent with the higher cognitive demands of the second task. In other words, the study investigated an interaction effect between task complexity and processing difficulties caused by accented speech. In Experiment 1, IS and processing times were measured in a sentence verification task with ten native and non-native speakers of English. Two speakers, with similar intelligibility but yielding different reaction times, were selected for Experiment 2, which measured IC using simulated university lectures. The results indicate the hypothesised interaction between context and the understanding of accented speech. We discuss the theoretical and methodological implications of this, as well as the relevance of our results for English-medium instruction at Nordic universities.

Keywords: intelligibility, comprehensibility, accent, cognitive load, English-medium instruction

1. Introduction

The effect of (often foreign) accent on lecture comprehension is an interesting aspect of English-medium instruction in higher education and has been commented on by university students in several studies, e.g. Hellekjær (2010). Previous studies on the intelligibility of accented speech have usually focussed on simple tasks, where the significance of a foreign accent has often been found to be quite small. However, we believe that the effect of processing difficulties associated with accent can be considerably larger in the context of university lectures in which English is used as a lingua franca, which is an increasingly common scenario in universities across the Nordic countries.

1.1. EMI in Nordic universities

Recent decades have seen a steady increase in English-medium instruction, EMI, in European universities (Wächter and Maiworm 2008; Wächter and Maiworm 2014; Dimova, Hultgren and Jensen 2015). Wächter and Maiworm (2014: 16) report that "the numbers of identified ETPs [English-Taught Programmes] went up from 725 programmes in 2001, to 2,389 in 2007 and to 8,089 in the present study" (p. 16). The Netherlands, Germany and the Nordic countries, Denmark, Sweden, Norway, Finland and Iceland, have been at the forefront of this development. Within the Nordic countries, this development has spawned a large scholarly interest, both in the ideologies underlying the development and in the new multilingual pracitices it gives affordances to (see e.g. the NJES special issue on "English in Academic and Professional Contexts", NJES vol. 12, No. 1, 2013; and Hultgren, Gregersen and Thøgersen 2014).

Two linguistic consequence of the increasing amount of EMI is that on the one hand more and more teaching is being conducted in the teachers' L2 and on the other that more and more students are being taught in their L2. Hellekjær (2010) investigated students' listening competence in English and found that many students had problems following EMI lectures. Others (e.g. Airey 2009) have studied whether students taught in English acquire the same disciplinary knowledge as those taught in their L1 and have found no discernible difference. Hincks (2010) and Thøgersen and Airey (2011) took a quantitative approach and demonstrated systematically lower rates-of-delivery in the L2, while Thøgersen (2013) used qualitative approaches to highlight the affordances given by the two languages.

One particular concern which has been voiced in both the popular and academic debates on EMI in the Nordic countries is whether students' learning would be negatively affected by the fact that neither the teachers nor the students are typically native speakers of the language of instruction—English. In particular, many commentators have been worried that the teachers' level of proficiency in English would not be sufficiently high for this demanding task. This question has been investigated in quite a few studies, though mostly through questionnaires asking the teachers or their students about any problems they experienced with EMI (Bolton and Kuteeva 2012; Hellekjær 2010; Jensen et al. 2013).

The lecturers' pronunciation of English, or "accent" as it is commonly referred to, has been identified as an important issue in several of these studies (Bolton and Kuteeva 2012; Hellekjær 2010), in the sense that students comment that they often find it difficult to understand the lecturers' accent. The studies have not documented that accent has been the cause of actual comprehension problems—misunderstandings or an overall lower understanding of what has been covered in the lectures (see Airey 2009 mentioned above)—but rather that difficulties subjectively experienced by students are often attributed to accent. Early studies on the intelligibility and comprehensibility of foreign-accented speech suggest that foreign-accented speech does not necessarily cause difficulties. For example, Munro and Derwing (1995a) found that even fairly heavily accented speech can be highly intelligible. However, there is also evidence that foreign-accented utterances require more time to evaluate than utterances by native speakers (Munro and Derwing 1995b; Trude, Tremblay and Brown-Schmidt 2013). To sum up the point, even though foreign accent has been shown to have little effect on intelligibility in some—usually fairly simple—task types, accented speech has also been shown to require increased processing time, or cognitive load. In EMI lectures, the cognitive load is already quite high because of the demanding task itself: to acquire new knowledge based on a presentation of typically complex information. This leads to the main question which we investigate in this paper: Will students learn less from a complex lecture if the lecturer's accent of English can be shown to increase the cognitive load as measured by response time on simpler listening tasks?

To our knowledge little attention has been paid to the question of the students' processing of the linguistically and conceptually challenging speech they are being presented with in EMI. This paper addresses that gap and attempts to contribute to the field of intelligibility and foreign accent by focussing on L2 listeners in a context where both parties in the exchange are (more often than not) non-native speakers.

2. Intelligibility

The influence of accent on intelligibility has been studied extensively within fields such as English as a second or foreign language (ESL/EFL) and World Englishes. Some of the most influential work within ESL/EFL

has been done by Tracey Derwing and Murray Munro (sometimes with other colleagues). Derwing and Munro have established a framework that operates with three concepts which they have argued to be "related but partially independent dimensions" (Munro and Derwing 1995a: 90), namely *accentedness*, *comprehensibility* and *intelligibility*. Within this framework accentedness "refers to how strong the talker's foreign accent is perceived to be"; comprehensibility "refers to listeners' perception of difficulty in understanding particular utterances", and intelligibility "refers to the extent to which an utterance is actually understood" (Munro and Derwing 1995b: 291). The first two are thus perceptual measures obtained by asking listeners to rate a sample on a scale from *no accent* to *very strong accent* for accentedness and on a scale from *extremely easy to understand* to *extremely difficult or impossible to understand* for comprehensibility. Intelligibility is measured through the listeners' performance on some task, for example a sentence verification task, where intelligibility is measured as the number of correctly verified sentences, or a transcription task, where intelligibility is measured as the number of correctly transcribed words. The three measures tend to correlate, though the correlation between accent and intelligibility is often not very strong. In particular, it has been found that even strongly accented speech can be highly intelligible (Derwing and Munro 1997: 11). Munro and Derwing (1995b) showed that while respondents' error rate was only slightly higher for Mandarin-accented speakers than for native English speakers in a sentence verification task, the verification times were longer for the Mandarin-accented utterances. Munro and Derwing suggest that the longer verification times are caused by an increase in processing time for the Mandarin-accented utterances, which, although largely insignificant for the level of intelligibility, leads listener judges to evaluate accented speech as more difficult to understand (Derwing and Munro 1997: 12).

While accentedness, or the degree of (a particular) foreign accent, does not seem to correlate well with intelligibility, the listener's familiarity with that accent has been shown to impact on intelligibility (Gass and Varonis 1984). Similar results have been found for native (regional) accents (Adank et al. 2009; Adank and McQueen 2007). Greater processing costs, measured as longer reaction times in various tasks, have also been shown for both non-native and unfamiliar native

accents (Munro and Derwing 1995b; Adank and McQueen 2007; Adank et al. 2009).

As mentioned above, the methods most commonly used to measure intelligibility within the Derwing & Munro framework are transcription tasks and (less commonly) sentence verification tasks. These methods measure the effect of accent on intelligibility in cognitively relatively simple tasks; answering comprehension questions to a new complex text on the other hand, can be said to measure the effect of accent on intelligibility in a difficult, cognitively more challenging task. The question is, then, whether the added cognitive load associated with processing accented speech will have a greater impact on intelligibility when the task already poses higher cognitive demands on the listener than in simple (or even trivial) tasks. An accent which is measured to be "fully intelligible" in a simple transcription task may lead to less than full understanding of text content in a more demanding communicative situation. Our aim is thus to investigate whether the increased processing time found by Munro & Derwing (1995b) is more strongly associated with intelligibility in more challenging tasks which more closely approximate the context of an EMI lecture.

In the experiments reported on in this paper we investigate and compare two aspects of understanding an accented speaker, namely understanding of a complex new text (Experiment 2) which is compared with understanding of the same speaker in a simpler context, namely assessing the truth values of short, syntactically simple utterances (a sentence verification task, Experiment 1). The sentence verification task requires (and measures) only understanding of the surface meaning of an utterance composed of common words and does not require understanding of speaker intentions or of the relationship between utterances, and it does not require the listener to infer any implicit meaning. The task in Experiment 2, however, requires interpretation on a deeper level in that the content of the message is new and complex with regard to both the concepts used in the text, the connections between them and the sheer volume of new information. This is more cognitively challenging for the listener who must be able to not only decode surface meaning, but also draw his or her own conclusions from the information presented. The considerable length difference between the texts (a few seconds vs. several minutes) makes one task far more demanding on working memory. We believe that the task in Experiment 2 approaches

what students in an EMI lecture will have to do to a greater extent than the more conventional transcription and sentence verification tasks. There are of course still a number of "un-natural" elements introduced by the fact that this is a controlled laboratory experiment. The problems associated with our methods will be addressed in the Discussion.

In order for us to be able to refer more easily to the outcome of the effect on accent on intelligibility, we have assigned different labels to the measures of understanding in the two task types.

1. The abbreviation IS will be used for intelligibility in simple tasks (transcription, sentence verification)
2. The abbreviation IC will be used for intelligibility in complex/demanding tasks (comprehension questions to new complex text)

As stated, this distinction is motivated primarily by a desire to have a convenient way of distinguishing between understanding in simple and complex tasks. It does not imply that IS and IC could or should be understood as two separate constructs. One way to conceptualise the difference would be to treat it as an interaction effect between task complexity and the cognitive load imposed by the accent—the more complex the task is, the more significant may be the effect of processing difficulties on intelligibility.

Following the terminology above where IS and IC denotes different contexts of understanding, our research question can be stated as follows: If accents X and Y are fully IS intelligible but accent Y poses a greater cognitive load on the listener, will accent Y then be less IC intelligible?

3. Research design

The study consists of two experiments, where the first, Experiment 1, feeds Experiment 2. In Experiment 1 we explore the connection between the accentedness, comprehensibility, intelligibility and associated cognitive load of a range of speakers. This is in part a replication of the study by Munro and Derwing (1995b), though we do not generalise to any particular accent(s). Our primary purpose is to find speakers who are more or less equally IS intelligible but require different processing time (taken as evidence of different cognitive load on the part of the listener)

and ideally have very different comprehensibility and accent ratings. Experiment 2 tests intelligibility in a specific type of complex text, namely (simulated) university lectures with the two speakers found in Experiment 1 and thus tests the IC of these speakers.

The test material for Experiment 1 consisted of true/false statements (adapted from Munro and Derwing, 1995b), while the texts for Experiment 2 were two simulated "university lectures", each with six multiple-choice questions.

In an early stage of the study, the true/false sentences and the lectures were recorded by 14 speakers, of which ten were selected for Experiment 1. In the event of misreadings or severe disfluencies, speakers were asked to repeat sentences or, for the lectures, repeat from the last pause. All recording were done in a soundproof studio using a DPA 4066 microphone fed into a Sound Devices 722 portable hard-disk recorder at 24-bit quantisation and 48KHz resolution.

4. Experiment 1

4.1. Method

4.1.1. Speakers
Ten speakers, 30 to 55 years of age participated in Experiment 1. They had seven different L1s. They were chosen to cover accents with which the Danish listeners were assumed to have varying degrees of familiarity. The seven L1s were: English (one speaker of General American, one speaker of Estuary English), Danish (two speakers), Spanish (two speakers), Italian, Japanese, German and Swedish. The listeners were expected to be very familiar with the two native accents and the Danish accents, less familiar with the Swedish and German accents and least familiar with the Italian, Spanish and (especially) Japanese accents. All eight non-native speakers of English were advanced users of English and at the time of the recording employed as PhD students or academic staff at a Danish university. All of them were experienced university teachers, and all of them regular users of academic English as a lingua franca. Most, though not all of them, were experienced in teaching EMI at university level.

114 Christian Jensen and Jacob Thøgersen

4.1.2. Sentences

A list of 40 true/false statements was used. The statements were, for the most part, taken from Munro and Derwing (1995b), in some case slightly adapted to suit the Danish context. The sentence lengths were between 4 and 8 words (average 5.9) and 6 and 10 syllables (average 7.8). Almost all words were among the 2,000 most frequent words (analysed by *VocabProfiler* at http://www.lextutor.ca/), with six words in the 3,000-8,000 range and nine off-list words (Italy, Europe, Japan, swimsuits, Shakespeare, Danish, England, Washington, McDonald's), none of which were judged to be unfamiliar to either speakers or listeners.

4.1.3. Recordings

The 400 individual utterances were extracted from the recordings and saved as individual sound files (16-bit, 48kHz), cutting exactly at the onset and offset of each utterance. Intra-utterance pauses of more than 200 ms (typically after the subject) were reduced to 100 ms to reduce variation in utterance length between the speakers. The audio files were then high pass filtered at 80Hz to reduce differences in low frequency energy and RMS normalised across speakers, so that utterances of each sentence were of equal loudness.

4.1.4. Listeners

Twenty Danish listeners, 17 women and 3 men between the ages of 19 and 26, participated in the experiment. They were all students of Danish at the University of Copenhagen with Danish as their L1 and English as their first foreign language. All of them also had some knowledge of other languages, including (one or more of) French, Spanish, German, Swedish, Norwegian and Italian.

4.1.5. Procedure

The ten speakers' recordings of the 40 true/false statements were distributed over ten sets, or versions, with four utterances per speaker in each version—two true and two false. The experiment was coded in the psycholinguistic software package OpenSesame (Mathôt, Schreij and

Theeuwes 2012). The presentation order of utterances was randomised by the software.

The 20 participants listened through Sennheiser HD201 headphones to the 40 utterances, each played once, and were asked to indicate the truth value of each utterance as quickly and accurately as possible by pressing the <z> key for TRUE and the <m> key for FALSE. No time limit was imposed, so the experiment only proceeded on key press. A practice run with 20 different true/false statements recorded by the authors preceded the actual experiment to familiarise the participants with the procedure.

After the sentence verification task, the participants heard the utterances again, this time presented in 10 blocks of four (one for each speaker), and were asked to rate comprehensibility and accent on 9-point scales based on the questions

- "how easy to understand was this speaker?" (1 = *extremely difficult*, 9 = *very easy*)
- "how native-like was the speaker's accent?" (1 = *strong non-native accent*, 9 = *very native accent*)

In addition, informants were asked to indicate what they thought was the speaker's native language.[1] Responses to this part of the experiment were indicated on a separate response sheet. Basic background information was collected about the informants' sex, age, L2s and experiences living abroad.

The experiment took place in a computer lab with identically configured iMac computers and conducted in sessions of 5–10 participants each.

4.2. Results

Intelligibility was measured as the proportion of correctly evaluated true/false questions (out of a total of 80 for each speaker). Response time (RT) was determined for the correctly evaluated sentences only and

---

[1] This was an attempt at getting a measure of "familiarity", but in retrospect we found this task to reflect *recognisability* of an accent rather than exposure to the accent, so the information has not been used in analyses.

measured from the offset of the audio. Values longer than 3 secs. were treated as outliers and removed from the dataset. A Pearson product moment analysis showed a very weak negative correlation between utterance length and (log transformed) response time, which only just failed to achieve statistical significance (r = -0.06, p = 0.063).

Table 1 shows that utterances were evaluated correctly (true/false) in 95.2% of cases. The success rates varied from 90% to 98.8%, which means that all speakers were very intelligible in this task. The scores are comparable to those in Munro and Derwing (1995b) where the success rate was 98 % for native speakers and 93 % for the Mandarin speakers. Linear mixed-effects analysis was applied to the data using the lme4 package (version 1.1-7) in R 3.1.1 (R Core Team 2017), with one of the two highest-scoring speakers as baseline and sentence and subject as random factors. The only significant inter-speaker differences observed were between the two speakers with the highest intelligibility score (Danish A and Swedish) and the speaker with the lowest intelligibility score (Spanish B) (p < 0.05).

Mean reaction times per speaker varied from 482 ms to 746 ms, and were generally in the range of about 500–650 ms. It should be noted that the differences between some of the speakers are quite small, and the variance is quite large, as can be seen from the standard deviations in column three. A linear mixed-effects analysis with speaker English GA (fastest RT) as baseline, utterance length as a fixed factor and sentence and subject as random factors showed significant differences between the baseline and six other speakers (Italian, Estuary English, Japanese, Spanish B, German and Danish B).

Table 1: Response times (RT) and standard deviations (SD_{RT}) in ms, intelligibility = proportion of correct responses (Intell.), and comprehensibility (Comp.) and accent ratings, ordered by RT. N = 80 (20 listeners x 4 true/false statements) for RT and intelligibility, n = 19 for comprehensibility and accent.[2]

Speaker	RT	SD_{RT}	Intell.	Comp.	Accent
English GA	482	479	97.5	8.9	8.9
Swedish	506	450	98.8	8.1	4.8
Danish A	526	490	98.8	8.7	3.7
Spanish A	528	441	96.3	7.0	4.1
Japanese	593	579	95.0	5.7	2.1
German	602	599	91.3	7.4	4.1
English EE	607	462	96.3	8.6	8.6
Danish B	613	461	95.0	8.5	5.7
Italian	631	540	92.5	7.7	4.8
Spanish B	746	558	90.0	5.0	2.2
Average	583	506	95.2	7.56	4.9

As also found in other studies, comprehensibility ratings are somewhat higher than accent ratings. The two native speakers receive ratings in the very high end of both scales, while the Japanese and Spanish B speakers were rated lowest. Interestingly, one Danish speaker was rated very low for accentedness but second highest for comprehensibility. Scatterplots illustrating the correlations between intelligibility, comprehensibility, accent and RT are shown in Figure 1.

---

[2] One listener had to be excluded from this part of the analysis due to missing responses.
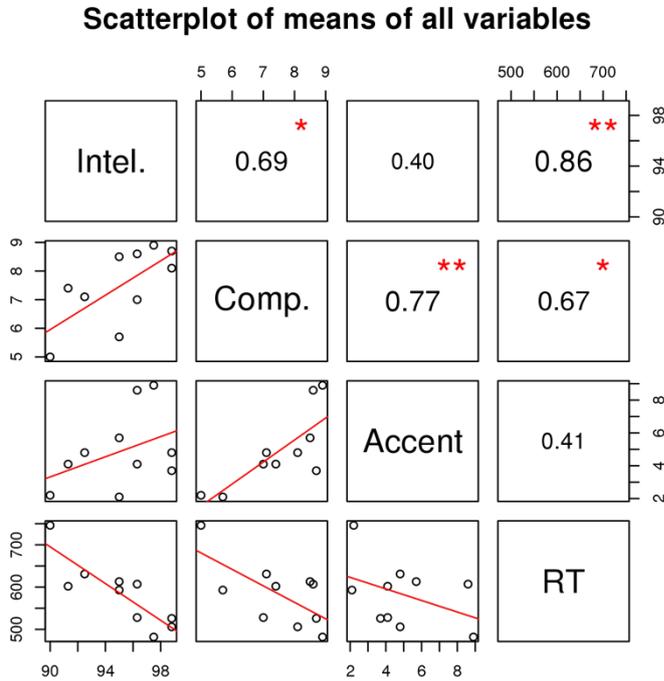
Scatterplot of means of all variables

Figure 1: Scatterplots of mean values for the variables intelligibility, comprehensibility, accent and RT. Correlation values are shown in the upper panel with indication of significance level (**: p< 0.01, *: p < 0.05).

Congruent with findings in previous studies (e.g. Munro & Derwing, 1995a) we found that intelligibility correlated fairly well with comprehensibility ratings (r = 0.69, p < 0.05), while the correlation between intelligibility and accent ratings was lower and non-significant (r = 0.40, p = 0.25). Comprehensibility and accent correlated quite strongly (r = 0.77, p < 0.01). Actual cognitive load, understood as reaction time (RT) showed a moderate negative correlation with comprehensibility (r = 0.67, p < 0.05) but only a weaker and non-significant negative correlation with accent (r = 0.41, p = 0.24). Reaction time was negatively correlated with intelligibility (r = -0.86, p < 0.01), which shows a strong trend for less intelligible speakers to also require more processing time. In other words, speakers whose utterances took

longer to verify were not only perceived to be less comprehensible by the listeners, they were also in general less intelligible in this IS task. However, the perceived strength of accent was not associated with either intelligibility or actual processing costs in this experiment.

4.3. Discussion and conclusion, Experiment 1

All 10 speakers were found to be very intelligible as determined by the sentence verification task, with most of the speakers having an intelligibility score of 95% or above. Although there was an overall effect of speaker on intelligibility, there were only a few significant differences between individual speakers. This may in part be due to a ceiling effect problem. More differences were observed with regard to the cognitive work of understanding the speakers as measured by the reaction time, where the speaker who provoked the shortest reaction times differed significantly from six other speakers. There was an overall correlation of RT with intelligibility. In other words, there is a general trend for intelligibility and cognitive load to correlate negatively. The purpose of Experiment 1 was primarily to find two candidates for Experiment 2. They should have equal intelligibility but differ with regard to RT (one resulting a greater cognitive load for the listener), which is based on the assumption that speakers can be "fully intelligible" in IS tasks but still require added processing time. We cannot claim conclusively to have found two such speakers in Experiment 1, partly because speakers with significantly different RTs did not have exactly the same intelligibility scores and partly because the lack of significant differences in intelligibility could be the result of a ceiling effect or insufficient experimental stimuli (sentences) or listeners. In addition, the correlation between RT and intelligibility suggests that the two measures are indeed not entirely independent. However, we do have speakers which differ only marginally when it comes to intelligibility in this simple task but differ significantly when it comes to RT, accent and comprehensibility. We can therefore proceed to test a weakened version of our hypothesis, namely that differences in intelligibility in complex tasks are larger than differences in intelligibility in simple tasks.

Based on the results from Experiment 1 we selected two speakers for inclusion in Experiment 2, namely the American GA speaker with 97.5% correctly evaluated utterances and the shortest RTs and the Japanese

speaker with 95% correctly evaluated utterances and RTs that were significantly longer than for the American speaker ($t = 2.46$ in an lmer model with the Japanese speaker as baseline and sentence and subject as random factors). Furthermore, the Japanese speaker was rated the second least comprehensible and most accented of the ten speakers (see Table 2).

5. Experiment 2

5.1. Speakers
The speakers in the second experiment were the two speakers selected from Experiment 1, both female, one a native speaker of American English, the other a native speaker of Japanese. Both speakers are experienced in teaching EMI at university level, and as such their recordings can be said to mimic an ecologically valid experience with accented English for the listener judges, although the task deviates from ecological validity in other respects (explained below) in order to control for extraneous influences.

5.2. Texts
The texts for Experiment 2 are two simulated "university lectures", each with six multiple choice questions and four response choices per item. The topics of the lectures are animal behaviour (mimicry)[3] and palaeontology (dinosaurs as warm-blooded or cold-blooded animals).[4] The texts were originally designed as practice material and made freely available by the website www.english-test.net. The texts and questions are based on the first part of the TOEFL test and were found to be suitable for our purposes, as the topics would be unfamiliar to our listeners and the comprehension questions required more than mere recollection of facts stated in the text (see below).

---

[3] See http://www.english-test.net/toefl/listening/A_university_lecture_on_ Animal_Behavior_by_a_professor_of_Biology.html

[4] See http://www.english-test.net/toefl/listening/A_university_lecture_by_a_ professor_of_Paleontology.html

As a training round we used an abbreviated version of a third "university lecture" on black holes[5] read by one of the authors and followed by two questions. It was essential for our choice of texts that they were about themes which are relatively unfamiliar to students in the humanities to eliminate background knowledge as a factor in responding to the multiple-choice questions. A topic like "modern history" or "linguistic theory" would potentially tell more about the students' skills within their own field than their ability to gain information from a lecture on an unfamiliar subject.

Six multiple-choice questions were asked after the reading of each text. The intention of the task was to increase the cognitive demands on the part of the listener under the assumption that the more complex task would increase the effect of the readers foreign accent. In terms of cognitive load, the increased response times documented in Experiment 1 are indications of higher processing costs demanded by the accent. Therefore, fewer cognitive processing resources are available to reflect on the textual input the listeners receive. The comprehension questions, which are all listed in the Appendix, are therefore of a type that they cannot be answered by merely restating information presented in the lecture. Correctly answering the question requires that the listener not only heard the lecture but was able to make inferences based on the information presented, i.e. "interpret" it. For example, one question regarding the palaeontology lecture asks how best to describe the organization of the lecture, e.g. defining scientific terminology or presenting opposing views on the question; and another question asks about the lecturer's own opinion on the presented theories. These facts are not presented in the actual lecture and require the listener to make inferences based on information in the lecture. The multiple-choice questions were not piloted against the target population and/or otherwise tested for equal difficulty. We acknowledge that this is a weakness of the experiment but believe that the crossed design where we control for both text and presentation order minimises any adverse effects of skewness in text or item difficulty.

---

5 See http://www.english-test.net/toefl/listening/A_university_lecture_ by_a_ professor_of_Physics.html

5.3. Recordings

Both texts were read by both speakers using the same equipment as mentioned before; in fact the recordings were made in the same recording session. The recordings were post-processed removing reading errors, false starts and very long hesitations. Our objective was to construct four readings which did not sound manipulated, but were as fluent as possible. The native speaker produced very few mistakes, so very little post-processing was done to these two recordings. The L2-speaker, unsurprisingly, produced more errors and in general two less fluent readings. As mentioned, we tried to minimize the difference by removing false starts and long pauses, but the overall impression of the two speakers—as also noted by the test subjects—is that one is markedly more fluent than the other. We choose to see this as an ecological reality of testing accented speech, since non-native speakers tend to speak less fluently than native speakers.

For both speakers the readings of two texts were of relatively equal length; for the Japanese speaker, 7:36 (mimicry) and 7:06 (dinosaurs) respectively, for the American speaker, 5:39 (mimicry) and 5:26 (dinosaurs) respectively leading to a difference in length smaller than 7% for both speakers. The difference in fluency, however, leads to a noticeable difference of around 25% in length between the two speakers' readings. Coincidentally, this corresponds well with the difference in speaking rates of lectures given in a speaker's L1 and L2, English found in previous studies (Hincks 2010; Thøgersen and Airey 2011).

5.4. Listeners

The listeners in Experiment 2 were 42 second-semester students of Danish at the University of Copenhagen. All 42 listeners had Danish as their L1 and English as their first L2. A further three listeners participated in the experiment but were later removed either because their L1 was not Danish or because they reported extraordinary difficulties in understanding both lectures. Six of the participants had also participated in Experiment 1, but since six months had elapsed between the two experiments, we judged the influence of having heard the speakers (but not the texts) before as negligible.

5.5. Procedure

The experiment was conducted in the faculty's language learning facilities, which allow multiple users to work on their computers in a quiet environment and without being able to see each other (and each other's answers). All computers were identical, and the sound volume was preset at what we judged a reasonably comfortable listening level. Since timing was not a measure in this experiment, we chose to have the lectures and the questions presented on the computer (on a dedicated website) and answers being given on paper.

The experiment has two primary variables, namely Speaker (or accent), our main interest, and Text (and associated questions), since the texts and questions can of course not *a priori* be determined to be of similar difficulty. Additionally, we assumed that the order in which the speakers were heard would be a factor—listeners may lose concentration, or contrast effects may mean that the inherent difficulties of comprehending the two lectures are either enhanced or minimized with task practice. For this reason, four versions of the experiment were produced, and the listeners were randomly assigned to one of four groups, each with a different version of the experiment. All groups listened to both speakers and both texts, but they heard different combinations of speaker, text and presentation order. As chance would have it, more of the listeners that were excluded had answered the C variant than the other versions. This, however, affects only the order in which the lectures were heard. Both speakers, as well as both texts, were heard by an equal number of listeners (21 each). The presentation orders for the four groups can be seen in Table 2.

The listeners' answers to the twelve questions they answered (six per text) were coded as either correct if they corresponded to the answers provided in the key that accompanied the test at www.english-test.net (see the Appendix), or incorrect if they did not, if no answer was given or if listeners had tried to give multiple answers.

Table 2: Presentation order of speech samples for the four groups.

Group	First lecture and speaker	Second lecture and speaker	N
A	Mimicry by Japanese	Dinosaur by American	12
B	Dinosaur by Japanese	Mimicry by American	11
C	Dinosaur by American	Mimicry by Japanese	9
D	Mimicry by American	Dinosaur by Japanese	10

5.6. Results

A total of 504 answers (including the 4 non-answers) were given. There was a marked difference in the number of correct answers to each question, as can be seen in Table 3.

For two questions the success rate was not above chance level (question 5 in both texts). One was found to be misleading, and the other could only be answered based on previous knowledge and not from information actually presented in the lectures. When these two questions are removed, a total of 420 responses were included in the model. Of these, 65% were correct. The dinosaur lecture yielded more correct answers (78% vs. 52%). It may be that the listeners knew more about dinosaur physiology than animal mimicry, or it may be that the questions were easier (for this particular group). In our statistical treatment (a multi-variate analysis) we control for this effect.

Similarly, there were marked differences in the listeners' performance. Only one answered all questions correctly, but a few had 9 out of 10 answers (90%) correct; on the other hand, no-one had no correct answers, but a few had only three (33%), not significantly above chance level. We chose not to exclude more listeners based on these results.

Table 3: The proportion of correct responses to the 12 questions across texts and speakers

Text			% correct responses		
			Japanese	American	Total
Animal mimicry	Question	1	42.9	47.6	45.2
		2	76.2	85.7	81.0
		3	33.3	52.4	42.9
		4	38.1	52.4	45.2
		5	4.8	52.4	28.6
		6	47.6	42.9	45.2
	Total		40.5	55.6	48.0
Dinosaur physiology	Question	1	66.7	66.7	66.7
		2	85.7	100.0	92.9
		3	76.2	90.5	83.3
		4	47.6	61.9	54.8
		5	14.3	33.3	23.8
		6	81.0	100.0	90.5
	Total		61.9	75.4	68.7
Total			51.2	65.5	58.3

Different binomial mixed-effects multi-variate models were fitted to the data using R v3.3.0 and the lme4 package (version 1.1-12). In the models we included listener as a random effect, and thus controlled for individual differences between listeners. As fixed factors, we used Score (Correct-Incorrect) as the dependent variable and as independent variables Text (Mimicry-Dinosaur), Speaker (American-Japanese), PresentationOrder (i.e. the order in which the text was heard) as well as possible interaction effects between Speaker and Text (hypothesizing for example that difficult texts increase comprehension difficulties) and Speaker and PresentationOrder (hypothesizing for example that contrast effects increase difficulties). Neither the interactions nor PresentationOrder proved to be significant. The model that best fit the data is presented in Table 4.

Table 4: Result of mixed-effects multivariate model (lme4, R v3.3.0).

Random effects:					
Groups Name	Variance	Std.Dev.			
Listener (Intercept)	0.2414	0.4913			
Number of obs: 504	groups:	PartID, 42			
Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.1870	0.1960	-0.954	0.3401	
Text: DinoPhys	1.2508	0.2274	5.501	3.77E-08	***
Speaker: American	0.5365	0.2226	2.410	0.0159	*

The results show Text as the most significant factor, but when this factor is controlled for, Speaker is a highly significant factor. The Estimates are log odds. They show that if the Text is the reference text, here the mimicry text, the chance of a correct answer (all else being equal) is the intercept value, i.e. around 45%. If the Text is the dinosaur text, the chance of a correct answer is around 74%. Expressed differently, the statistical odds between the two texts, Exp(1.2508), is 3.5. A similarly but slightly smaller difference is found between the two speakers. If the speaker is the reference level, here the Japanese speaker, the chance of a correct answer is 45% (all else being equal). If the speaker is the American speaker, chances increase to around 59%. Or expressed differently, the statistical odds between the two, Exp(0.5365), is 1.7. Since the data set is fairly well balanced, and since the interaction effects are small and insignificant, the actual number of correct responses given to each speaker's reading is a fair estimate of the effect of accent. On average, the listeners produced 2.55 correct responses (out of 5) when hearing the lecture from the Japanese speaker, against 3.28 when listening to the American speaker.

5.7. Conclusion, Experiment 2
The listeners in Experiment 2 were able to answer significantly more comprehension questions correctly after listening to the American speaker than after listening to the Japanese speaker. This means that the IC of the American speaker was significantly better than that of the foreign-accented Japanese speaker in spite of the (roughly) equal IS which we established in Experiment 1. We also found that the listeners in Experiment 1 not only took longer to evaluate the truth value of the utterances spoken by the Japanese speaker, which is assumed to be evidence of larger cognitive load, but also perceived this speaker to be more difficult to understand, which can also be understood as a sign of larger cognitive load (Munro and Derwing 1995b). These findings support our hypothesis that increased cognitive load associated with accented speech leads to decreased intelligibility in more complex tasks.

6. Discussion and conclusion
In Experiment 1 we tested the intelligibility of selected speakers through a sentence verification task. Results showed that all speakers were fully or almost fully intelligible in this simple task and that even speakers who were judged to be heavily accented could be highly intelligible. There was little difference between the speakers, but this could at least in part be the result of a ceiling effect. We also assessed the processing cost for each speaker by measuring response latencies, or reaction time, and found more variation in this variable, with significant differences between many of the ten speakers. The speakers with the longest reactions times were also generally judged by the listeners to be more difficult to understand.

We selected two speakers with approximately the same intelligibility scores but with significantly different response times: an American speaker with the shortest response times of all ten speakers, and a Japanese speaker with significantly longer response times and low comprehensibility and accent ratings. In Experiment 2 we then tested the IC of the two speakers, understood as the ability of the listeners to answer follow-up question to a short recorded university lecture on an unfamiliar topic. We found that the American speaker was significantly more intelligible than the Japanese speaker, to the extent that listeners

had approximately 0.7 more correct answers (out of 5) when listening to the American speaker.

These findings indicate that speakers who have been measured to be equally intelligible in tasks that require only relatively little cognitive processing, may not be equally easy to understand when the cognitive demands are increased, for example when listening to a much longer and more complex text which require more processing and analytical thinking. The statements in the sentence verification task were simple both in terms of content—the truth value of the sentences was easy to determine if all words had been understood—and in terms of length (only 6–8 words) and with a simple syntactic structure. The lectures, on the other hand, were not only longer but dealt with complex relations in unfamiliar topics. If the difference between IS and IC is indeed caused by the difference in cognitive load in the two tasks, this has some important consequences for the assessment of the effect of foreign accent on overall comprehension (but see below for a discussion of the limitations and potential problems of this study).

First, our results suggest that intelligibility measured as the ability to write down words or sentences or to assess the truth value of short statements, as is done in many studies (e.g. Field 2005; Kennedy and Trofimovich 2008; Munro and Derwing 1995a; Munro and Derwing 1995b; Munro, Derwing and Morton 2006), is not necessarily a fully reliable indicator of intelligibility understood as the extent to which a listener is able to extract and integrate information from a complex, and possibly more ecologically valid, text presented by the speaker. It would appear, in other words, that *difficulty* with the process of word recognition may interfere with the more complex aspects of understanding such as it is grasped in the construct of IC even if the actual success or failure with word recognition is relatively unaffected.

Our results indicate that it can be misleading to draw inferences about a listener's ability to understand a lecturer's accent in complex interaction (i.e. our IC) based on measurements done in more simple interaction. A similar point was made in Munro and Derwing (2015) where it is stated that "a word count approach, for example, focuses strictly on exact word matches, but does not fully address illocutionary force, which would require further probing, perhaps with comprehension questions" (p. 382). Typical reasons for using a transcription task for measuring intelligibility can be that it works with a range of different text

types and yields a high degree of inter-listener reliability but also that it is fairly simple to administer (Munro and Derwing 2015: 382). Interestingly, our results suggest that another simple measure can be useful for extending the results to other contexts, namely (perceived) comprehensibility. The comprehensibility of a speaker (measured by asking judges how difficult the speaker is to understand) may be a better indicator of the processing challenges presented to the listener and thus of the intelligibility in complex tasks, since this construct includes the effort made (and felt) by the listener. This, of course, is consistent with Munro and Derwing's finding that "utterances that were assigned low comprehensibility ratings also tended to take longer to process than moderately or highly comprehensible utterances" (Munro and Derwing 1995b: 289). It is thus possible that the perceptual measure of comprehensibility ratings would more closely reflect actual understanding in situations that require more than merely recognising words or utterances or assigning truth value to semantically and syntactically simple utterances. This could be investigated by measuring the IC of a range of speakers and comparing the results with the IS and comprehensibility scores for the same speakers.

What, then, are the implications of our results for English-medium instruction in higher education in the Nordic countries? The fact that the IC of the speaker who required longer processing time proved to be lower than that of the other speaker certainly indicates that concerns about the students' learning outcome in EMI may be valid: Lectures may be less effective in leading to deeper understanding of a topic if the lecturer's accent itself increases the cognitive load of an already challenging task.

However, the overall significance for EMI cannot be gauged merely on the basis of our results due to certain limitations of our design, some of which relate to Experiment 1 and some to Experiment 2. A central part of our argument relies on the establishment of "equal IS" but different processing times of the two speakers in Experiment 2, and while both speakers were (more or less) fully intelligible in Experiment 1, this may in part have been caused by a ceiling effect, so that more careful screening would have revealed bigger differences between the two, even for simple tasks. We believe that our methods are sufficiently similar to those used in other studies to warrant comparison with these, but it would be useful to test the intelligibility of the two speakers (and

preferably more speakers) using a transcription task or cloze procedure with the texts from Experiment 2 as input. This might provide a more accurate measure of the speakers' IS. Alternatively, adding noise to the stimuli in Experiment 1 might have lowered the overall success rate and revealed larger differences between the speakers.

As for Experiment 2, it is reasonable to say that in spite of our attempts at designing an ecologically valid study, the experimental situation may not reflect accurately on genuine lectures. Firstly, reading a complex text aloud does not correspond to delivering an original lecture in terms of the effect of accent on intelligibility. It is possible that the fluency and prosody of (especially) the Japanese speaker would have been better and more effective in a spontaneous lecture situation and that the difference between the simple text (sentences in Experiment 1) and the complex text ("university lecture" in Experiment 2) has been magnified by the mode of delivery—reading, rather than spontaneous productions. Secondly, non-native speaker lecturers may use various strategies suited to their linguistic competences to increase the effectiveness of a lecture, e.g. repetition, rephrasing, writing of keywords etc. These are factors not easily controlled for in an experimental study. Thirdly, students do not usually listen to lectures on completely unfamiliar subjects given by completely unfamiliar lecturers, and familiarity with subject and/or lecturer is likely to affect the listening process in a positive way. Likewise, lectures are usually organized with visual support—not least the ability to see the lecturer but also visual aids like PowerPoint slides, drawings etc.—and will often, though not always, provide opportunity for the students to ask clarification questions or discuss points made during the lecture. The extent to which these additional factors can alleviate the processing difficulties we have demonstrated here cannot be determined without further research.

Another issue concerns the fact that students may adapt to the lecturer's accent and thus overcome some of the initial processing difficulties. Accent familiarity, and familiarity with a particular speaker, increases with exposure. There is, however, no real consensus on the amount of exposure necessary to overcome processing difficulties. Clarke and Garret (2004: 3647) found that native listeners' "processing speed is initially slower for accented speech than for native speech but that this deficit diminishes within one minute of exposure". However, it has also been found that adaptation can be disrupted by added perceptual

effort, such as processing speech in noise (Ferracane et al. 2015). In addition, to our knowledge very little work has been done on L2 listeners' ability to adapt to unfamiliar accents. While listeners' ability to adapt to unfamiliar accents is very relevant for lecture comprehension in EMI, it is thus beyond the scope of the present paper and will require further investigation.

Finally, listeners' attitudes to speakers' accents may affect understanding (Rubin 1992) and make them trust speakers with a foreign accent less as authorities (Lev-Ari and Keysar 2010; Jensen et al. 2013), and this effect is likely to be stronger for a complex new topic (Experiment 2) than for simple true/false statements (Experiment 1).

It has not been our purpose with this discussion to propose that all university EMI lectures should be given by native speakers of English—or by native speakers and non-native speakers with the same L1 as the students (i.e. speakers with more familiar accents). Our study has had two aims: First, we have wanted to explore certain aspects of intelligibility (in the broadest sense) and accent in order to contribute to understanding of this area in general. To do this, we have examined whether it may be necessary to distinguish between understanding at a more superficial level and deeper understanding of more complex messages, since there would appear to be a kind of interaction effect between processing difficulties caused by the accent and task complexity, due to the different levels of cognitive load. This has practical consequences for the study of understanding of accented speech, not least if we are interested in the understanding of accented speech by real-life listeners in real-life lingua franca situations. We readily admit that the method used in Experiment 2 is "artificial" for the speakers and listeners, but we believe that the discrepancy between levels of understanding that our results exhibit should have methodological consequences. This brings us to the second purpose of the study. We have wanted to cast further light on an area of EMI, viz. accent, which seems to be glossed over rather quickly in the current work on the challenges of introducing a medium of instruction which is often not the L1 of neither lecturers nor students, in spite of the fact that students often comment that pronunciation is one of the important aspects of lecture comprehension. If students report that accent is important to them and can be demonstrated to cause problems, we should

be careful to write off their lived experiences as mere prejudice too quickly.

References

Adank, Patti, Bronwen G. Evans, Jane Stuart-Smith and Sophie K. Scott. 2009. "Comprehension of familiar and unfamiliar native accents under adverse listening conditions." *Journal of Experimental Psychology: Human Perception and Performance* 35(2): 520–529. doi:10.1037/a0013552.

Adank, Patti and James M. McQueen. 2007. "The effect of an unfamiliar regional accent on spoken-word comprehension." *Proceedings of the 16th International Conference of Phonetic Sciences*. 1926–1928. URL: http://pubman.mpdl.mpg.de/pubman/item/escidoc:59662:2/ component/escidoc:59663/Adank_2007_effect.pdf (2 July, 2013).

Airey, John. 2009. "Science, language, and literacy: case studies of learning in Swedish university physics (Doctoral dissertation)." Uppsala: Uppsala University. URL: http://www.diva-portal.org/ smash/record.jsf?pid=diva2%3A173193&dswid=7829 (4 November, 2016).

Bolton, Kingsley and Maria Kuteeva. 2012. "English as an academic language at a Swedish university: parallel language use and the 'threat' of English." *Journal of Multilingual and Multicultural Development* 33(5): 429–447. doi:10.1080/01434632.2012.670241.

Clarke, Constance M. and Merrill F. Garrett. 2004. "Rapid adaptation to foreign-accented English." *The Journal of the Acoustical Society of America* 116: 3647.

Derwing, Tracey M. and Murray J. Munro. 1997. "Accent, intelligibility, and comprehensibility: Evidence from four L1s." *Studies in second language acquisition* 19(1): 1–16.

Dimova, Slobodanka, Anna Kristina Hultgren and Christian Jensen. 2015. *Volume 3 English-Medium Instruction in European Higher Education*. Berlin, Boston: De Gruyter Mouton. (6 September, 2016).

Ferracane, Elisa, Cynthia P. Blanco, Gabriela Cook, Karen Johnson and Rajka Smiljanic. 2015. "Effect of noise on foreign-accent adaptation." *The Journal of the Acoustical Society of America* 137(4): 2435–2436. doi:10.1121/1.4920893.

Field, John. 2005. "Intelligibility and the listener: The role of lexical stress." *TESOL quarterly* 39(3): 399–423.

Gass, Susan and Evangeline Marlos Varonis. 1984. "The effect of familiarity on the comprehensibility of nonnative speech." *Language Learning* 34(1): 65–87. doi:10.1111/j.1467-1770.1984.tb00996.x.

Hellekjær, Glenn Ole. 2010. "Lecture comprehension in English-medium higher education." *Hermes* 45: 11–34.

Hincks, Rebecca. 2010. "Speaking rate and information content in English lingua franca oral presentations." *English for Specific Purposes* 29(1): 4–18. doi:10.1016/j.esp.2009.05.004.

Hultgren, Anna Kristina, Frans Gregersen & Jacob Thøgersen (eds.). 2014. *English in Nordic Universities: Ideologies and practices*. Amsterdam: John Benjamins Publishing Company. URL: http://www.jbe-platform.com/content/books/9789027269591 (28 October, 2016).

Jensen, Christian, Louise Denver, Inger M. Mees and Charlotte Werther. 2013. "Students' attitudes to lecturers' English in English-medium higher education in Denmark." *Nordic Journal of English Studies* 12(1): 87–112.

Kennedy, Sara and Pavel Trofimovich. 2008. "Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context." *Canadian Modern Language Review/La Revue canadienne des langues vivantes* 64(3): 459–489.

Lev-Ari, Shiri and Boaz Keysar. 2010. "Why don't we believe non-native speakers? The influence of accent on credibility." *Journal of Experimental Social Psychology* 46(6): 1093–1096. doi:10.1016/j.jesp.2010.05.025.

Mathôt, Sebastiaan, Daniel Schreij and Jan Theeuwes. 2012. "OpenSesame: an open-source, graphical experiment builder for the social sciences." *Behavior Research Methods* 44(2): 314–324. doi:10.3758/s13428-011-0168-7.

Munro, Murray J. and Tracey M. Derwing. 1995a. "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners." *Language Learning* 45(1): 73–97.

Munro, Murray J. and Tracey M. Derwing. 1995b. "Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech." *Language and Speech* 38(3): 289–306.

Munro, Murray J. and Tracey M. Derwing. 2015. "Intelligibility in research and practice." *The Handbook of English Pronunciation*. Eds. Marnie Reed and John M. Levis. Hoboken: John Wiley & Sons, Inc. 375–396. doi:10.1002/9781118346952.ch21. URL: http://onlinelibrary.wiley.com.ep.fjernadgang.kb.dk/doi/10.1002/978 1118346952.ch21/summary (11 August, 2016).

Munro, Murray J., Tracey M. Derwing and Susan L. Morton. 2006. "The mutual intelligibility of L2 speech." *Studies in Second Language Acquisition* 28(01): 111–131.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: http://www.R-project.org/ (28 November, 2017).

Rubin, Donald L. 1992. "Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants." *Research in Higher Education* 33(4): 511–531.

Thøgersen, Jacob. 2013. "Stylistics and pedagogical consequences of university teaching in English in Europe." *Language Alternation, Language Choice and Language Encounters in International Tertiary Education*. Eds. Hartmut Haberland, Bent Preisler and Dorte Lønsmann. Dordrecht: Springer. 181–199.

Thøgersen, Jacob and John Airey. 2011. "Lecturing undergraduate science in Danish and in English: A comparison of speaking rate and rhetorical style." *English for Specific Purposes* 30(3): 209–221. doi:10.1016/j.esp.2011.01.002.

Trude, Alison M., Annie Tremblay and Sarah Brown-Schmidt. 2013. "Limitations on adaptation to foreign accents." *Journal of Memory and Language*. doi:10.1016/j.jml.2013.05.002. URL: http://www.sciencedirect.com/science/article/pii/S0749596X1300038 7 (4 July, 2013).

Wächter, Bernd and Friedhelm Maiworm. 2008. *English-taught Programmes in European Higher Education*. URL: http://www.lemmens.de/dateien/medien/buecher-ebooks/aca/2014_ english_taught.pdf (6 September, 2016).

Wächter, Bernd and Friedhelm Maiworm. 2014. *English-taught Programmes in European Higher Education: The State of Play in 2014*. Lemmens Medien GmbH.

APPENDIX
Comprehension questions in Experiment 2

Animal   Mimicry   (http://www.english-test.net/toefl/listening/A_
university_lecture_on_Animal_Behavior_by_a_professor_of_Biology.ht
ml)
Questions:

1). Why does the lecturer mention Charles Darwin?
(A) Darwin was the creator of evolutionary theory.
(B) Darwin's theory was used to explain mimicry.
(C) Darwin explained mimicry for the first time.
(D) Darwin separated Batesian and Müllerian mimicry.

2). What is the term used for an organism that is fooled by mimicry?
(A) The recipient
(B) The model
(C) The device
(D) The prey

3). Why does the lecturer mention katydids?
(A) Because they taste bad
(B) Because they resemble bumblebees
(C) Because they can sting
(D) Because they look like leaves

4). Some moths and butterflies have large, owl-like eyespots concealed
on their underwings, which they can suddenly display to a predator.
Which kind of mimicry is this an example of?
(A) Batesian mimicry
(B) Camoflage
(C) Müllerian mimicry
(D) Aggressive mimicry

5). Which organism is presented as an example of Müllerian mimicry?
(A) The Viceroy Butterfly
(B) The Orchid Mantis
(C) The Monarch Butterfly
(D) The Leaf-tailed Gecko

6). Which would be the best title for this lecture?
(A) Animal Behavior
(B) Animal Creativity
(C) Animal Deception
(D) Animal Escapades

CORRECT: 1 B; 2 A, 3 D; 4 A, 5 B; 6 C

Dinosaur Physiology (http://www.english-test.net/toefl/listening/A_university_lecture_by_a_professor_of_Paleontology.html)
Questions:

1). What is this lecture mainly about?
(A) Recent research on dinosaurs
(B) An insoluble problem about dinosaurs
(C) An ongoing argument about dinosaurs
(D) The characteristics of dinosaurs

2). An organism that generates internal heat is called what?
(A) Ectothermic
(B) Endothermic
(C) Homeothermic
(D) Poikilothermic

3). How has the professor organized his lecture?
(A) By listing opposing viewpoints
(B) By suggesting new theories
(C) By describing dinosaur evolution
(D) By defining physiological terms

4). Who found the energy cost of locomotion?
(A) John Ostrum
(B) Charles Darwin
(C) Robert Bakker
(D) Herman Pontzer

5). Which fact suggests that dinosaurs were cold-blooded creatures?
(A) Fossil dinosaurs have been found at high elevations (Alaska and Antarctica, for instance).
(B) Dinosaurs were scaly- they were covered with scales, not hair or fur.
(C) Dinosaurs were the direct ancestors of birds.
(D) Theropod and Ornithopod dinosaurs had larger-than-normal brains.

6). Which best expresses the lecturer's probable opinion on the current state of dinosaur research?
(A) Dinosaurs were most probably cold-blooded animals.
(B) Dinosaurs had some sort of intermediate physiology.
(C) Dinosaurs were actually warm-blooded creatures.
(D) Paleontologists remain divided on this issue.

CORRECT: 1 C; 2 B; 3 A; 4 D; 5 B; 6 D