Chapter one

Introduction

Now the whole earth had one language and few words. And as men migrated from the east, they found a plain in the land of Shinar and settled there. And they said to one another, ‘Come, let us make bricks, and burn them thoroughly.’ And they had brick for stone, and bitumen for mortar. Then they said, ‘Come, let us build ourselves a city, and a tower with its top in the heavens, and let us make a name for ourselves, lest we be scattered abroad upon the face of the whole earth.’ And the LORD came down to see the city and the tower, which the sons of men had built. And the LORD said, ‘Behold, they are one people, and they have all one language; and this is only the beginning of what they will do; and nothing that they propose to do will now be impossible for them. Come, let us go down, and there confuse their language, that they may not understand one another’s speech.’ So the LORD scattered them abroad from there over the face of all the earth, and they left off building the city. Therefore its name was called Babel, because there the LORD confused the language of all the earth; and from there the LORD scattered them abroad over the face of all the earth. (Genesis 11:1-9)

1.1 English as a lingua franca

It is suggested in the Bible that the ideal state of the world would be, and at some stage was, one in which all mankind spoke the same language. However, God punished mankind for its arrogance with the multiplicity of languages, or the ‘confusion of tongues’. Although a blessing for professional linguists, language teachers, translators and interpreters alike, the fact that there exist some 6,000 languages on the face of this earth which are mutually unintelligible, has been a matter of enormous financial consequences. It has been estimated, for instance, that the cost of having all documents translated in all the languages spoken in the European Community for the transactions of the European Parliament are in excess of 1 billion Euros a year.¹

Over a century ago the Polish ophthalmologist Zamenhof devised the artificial language Esperanto, in an attempt to provide the world with a common language

¹ James Owen in London in the National Geographic News (February 22, 2005): ‘The European Union has been operating in 20 official languages since ten new member states joined the legislative body last year. With annual translation costs set to rise to 1.3 billion dollars (U.S.), some people question whether EU institutions are becoming overburdened by multilingualism’.
that would be easy to learn and use, so that the confusion of tongues could be overcome. Although Esperanto has had numerous speakers, it never rose to the status of a lingua franca of the world. If any language may aspire to that status today, it would have to be English.

Indeed, English has become the language of international politics, trade, finance, and science. This comes with mixed blessings. On the one hand it brings the convenience of global communication, but the downside is that we now face a bewildering variety of forms of English (‘Engishes’) with foreign accents characteristic of the various nations on this earth, which are difficult to understand – for native listeners of English and even more so for non-native listeners. These varieties of English are sometimes mockingly referred to by portmanteau designations such as Spanglish (Spanish English), Dunglish (Dutch English), and Chinglish (Chinese English). Often the problem of non-native communication is no more than a mild nuisance, but human lives may be at stake when, for instance, an air-traffic controller is a native speaker of Spanish and has to understand English messages spoken by a Dutch airline pilot (and vice versa) in a noisy cockpit.

1.2 Topic of the dissertation

The topic of the present dissertation is the mutual intelligibility of speakers of English from diverse native-language backgrounds. As will be explained in greater detail in Chapter two, when a person speaks a language that is not his mother tongue, the language produced deviates in many respects from that of its native speakers. The most noticeable deviation of this so-called interlanguage is in the way the foreigner pronounces the target language. In fact, the foreign speaker’s approximation to the target language will have a large number of sound properties, not only in the pronunciation of the vowels and consonants but also in the realization of the speech melody and rhythm, that seem to be copied from the speaker’s mother tongue. Generally, this native-language interference is so strong that the foreign speaker’s mother tongue can be established just by listening to his pronunciation of the foreign language. Native listeners are sensitive to the deviation from the native norm in the speech of foreign learners but, normally, communication does not break down on account of this. However, it has been shown that foreign-accented speech is highly vulnerable to background noise; it is clearly a less optimal code than speech between two native speakers (see Chapter two). Be this as it may, the native listener is normally able to cope with deviant speech and reconstructs the foreign speaker’s intentions in spite of the suboptimal signals. The communicative problems will be severely aggravated when both interactants, i.e. speaker and listener, are non-native speakers, especially when they do not share the same mother tongue. In such situations the speaker produces distorted sound patterns (reminiscent of his mother tongue) which the listener cannot interpret because they do not conform to the patterns needed for the target language nor to the patterns in the mother tongue of the listener.
1.3 Approach

The basic problem that this thesis addresses, then, is to establish how difficult is it for speakers and listeners to understand each other when using English as a lingua franca, when the interactants do not share the same mother tongue. We will compare the results with several ‘control’ conditions. In one, both speakers and listeners are native users of English – which, of course, is the situation where optimal communication is expected. In a second control condition, either speakers or listeners, but not both, use English as a foreign language, and in a third control condition the subjects are neither native speakers nor native listeners of English but share the same mother tongue.

These conditions were obtained by having Chinese, Dutch and American speakers of English produce English words and sentences and offering the recordings to listeners with the same three native-language backgrounds. This yields nine combinations of speaker and listener nationalities:

<table>
<thead>
<tr>
<th>Native language of speaker</th>
<th>Native language of listener</th>
</tr>
</thead>
<tbody>
<tr>
<td>American</td>
<td>American</td>
</tr>
<tr>
<td>American</td>
<td>1</td>
</tr>
<tr>
<td>Chinese</td>
<td>4</td>
</tr>
<tr>
<td>Dutch</td>
<td>7</td>
</tr>
</tbody>
</table>

Communicative problems are expected to be greatest in combinations 6 and 8, which involve non-native speakers and listeners with different mother tongues. Optimal communication is predicted for combination 1, which contains native speakers and listeners of English. Comparisons will be made of combinations 4 and 7, with native listeners and foreign speakers as opposed to 2 and 3, with native speakers and foreign listeners. This comparison will tell us whether non-native communication is better when the speakers are native or when the listeners are native. A possible ‘interlanguage benefit’ (see below) in non-native communication can be tested in the combinations 5 and 9, where non-native speakers and listeners of English have the same native-language background, i.e. Chinese in 5 and Dutch in 9.
1.4 Goal of the study

This type of research has not been done before. To be true, there has been a wealth of research on the intelligibility of foreign-accented speech for native listeners of the target language (for a survey with emphasis on English as the target language, see Chapter two) and on the intelligibility of English for foreign listeners relative to native listeners. The point is, of course, that in all these studies there is always one party that uses English as the native language. The problem we address in the present study is more complicated, viz. the mutual intelligibility in English of non-native speakers with different source-language backgrounds. In fact, I am aware of just one (recent) study that addresses part of the issues raised here. Bent and Bradlow (2003) determined sentence intelligibility scores for a large number of foreign learners of American English of diverse linguistic backgrounds (Chinese, Korean, Japanese, Rumanian and many other nationalities). Not only did their results bear out that intelligibility was best between American speakers and listeners, but they also showed the existence of what they called an interlanguage speech intelligibility benefit, that is, that intelligibility between foreign learners of English sharing the same mother tongue was demonstrably better than between learners with different native languages.

In this thesis I want to study these matters in greater detail, using a much smaller variation of language backgrounds of speakers and listeners, but targeting the intelligibility not only at the sentence level but also at the lower levels of individual vowels and consonants, and of consonant clusters. Such a detailed study might allow us to pinpoint the problematic sounds, separately for speakers and for listeners, and from that to understand why intelligibility at the sentence level is successful to the degree that it is.

Concretely, we have asked the following questions for each of the nine combinations of speaker and listener nationality (or rather: native language backgrounds):

1. How well are English vowels identified in /hVd/-sequences (and what is the structure in their perceptual confusions)?
2. How well are English consonants and C-clusters identified in intervocalic position (and what is their confusion structure)?
3. What is the intelligibility of words in various types of sentences?
4. Which linguistic aspect (vowel identification, consonant identification, cluster identification, word recognition) provides the most sensitive measuring tool to determine differences in intelligibility?

The non-native speakers and listeners used by Bent and Bradlow (2003) differed considerably in their English proficiency. The Korean learners of English, for instance, were much better than the Chinese learners. It is unclear in their study, however, if the difference between the two Korean and two Chinese speakers was due to longer length of residence in the USA, to younger age of learning, or whether Korean learners have an edge over Chinese learners because the Korean sound system is more like that of English than the Chinese sound system is. In our study we have made an effort to select learners of English in the Netherlands and in China...
that were representative of their populations. Specifically, we targeted young adult learners of English as a foreign language, i.e. in a situation of supervised learning in an environment where English is not the dominant language, nor the language of instruction. The learners were university students who do not specialize in English language and/or literature, and they did not have any regular contact with native speakers of English. The speakers we selected were in the middle of their peer groups, and represent the English proficiency of the typical young academically trained user of English as a foreign language in China and in the Netherlands.

1.5. Effect of linguistic distance

We have studied the mutual intelligibility of Chinese, Dutch and American (native) speakers of English. Dutch and English are West Germanic languages which are genealogically quite close and typologically similar. The two languages share a large number of cognates in their vocabularies, have many similarities in word and sentence structures, use comparable prosodic systems (both languages are of the stress accent type) and have highly similar segmental sound systems (phonetics and phonology). Chinese is a completely different language, typologically a polysynthetic language with simple syllable structures, a complex lexical tone system, and a smaller vowel inventory than English. Detailed comparisons of the segmental sound systems of the three languages will be given in Chapter three. In our thesis we test the – obvious – hypothesis that mutual intelligibility in a lingua franca situation increases as the native languages of the interactants are more similar. We predict, accordingly, that Dutch-accented English is more intelligible than Chinese-accented English. This will not only be the case when the listeners are American but also when the listeners are Chinese or when they themselves are Dutch (the latter advantage would be due to the interlanguage benefit).

Potentially, comparing mutual intelligibility of non-native speakers and listeners of English may be used as a method to establish linguistic distance between any two languages in the world. There has been an upsurge of research activity in dialectometry on establishing the degree of similarity (and by implication linguistic distance) among dialects of a language or among languages within a language family. The research methodology does not rely on linguists’ (or even naïve language users’) intuitions of linguistic distance between varieties, but quantifies linguistic distance in terms of the number of symbolic operations, i.e. deletion, addition or substitution of phonemes in a transcription of word pairs (Levenshtein distance metric, see for instance Heeringa and Nerbonne, 2001; Heeringa, 2004; Gooskens and Heeringa, 2004). The method works quite well, even when there is a fair number of non-cognate word pairs between the two languages under comparison. The method has been verified against both judged and functionally determined communicative distance measures, such as intelligibility scores (percentage of correctly translated words) and opinion scores on intelligibility. The results indicate that the distance metric makes an accurate prediction of subjective and objective intelligibility scores.

However, the method breaks down when two unrelated languages are compared. First of all, when there are no cognate word pairs shared between the languages, then the number of symbolic operations that have to be carried out to map a word
onto its counterpart in the other language is determined by chance (and will be very large). Secondly, when two languages are not related to each other mutual intelligibility will be zero, so that no correlation can be established between the distance metric and the practical intelligibility measures. And yet linguists will readily agree that some sound systems are more like each other than others, even if all the languages belong to different families. Here, I would argue, we could fruitfully turn to the mutual intelligibility of these languages if their speakers and listeners use English. The more distant two languages, the smaller the mutual intelligibility when these speakers and listeners use English.

1.6 Contrastive analysis

As matters stand today, it is not possible to express the distance (difference) between two languages such as Dutch and English, or between Chinese and English, numerically. The differences are multidimensional and it is unclear how the various dimensions should be weighed against each other. All we can say, or rather assume, is that Dutch and English are closer than Chinese and English, but not how much closer. Nor would the (differences in) distance measures allow us to make a prediction of specific learning problems. Assuming that non-native communication is more problematic than communication between native speakers, can we predict specific difficulties from a comparison of the two sound systems? If the communication is between two non-native speakers of English who do not share the same source language, can specific problems be predicted by comparing the sound systems of all three languages involved? In this thesis we will attempt to make such predictions, based on various models of positive and negative transfer from the mother tongue to the foreign language, and test these against the observations in our experiments.

A literature survey in Chapter two will show that generally, contrastive analyses of the sound systems of source and target language have not been very successful in predicting learning problems. Sounds and contrasts that should be problematic proved easy in practice, and unexpected learning problems have been observed where the contrastive analysis predicted none. We will use the contrastive analysis only as a frame of reference in order to facilitate the presentation of the results. At best, it will allow us to show that certain views on native language interference in the foreign language provide better explanations than some other views. A second benefit of contrastive analyses is that may fulfill a useful role in interpreting findings post hoc, and in classifying types of errors (confusion patterns) post hoc.

1.7 Structure of the dissertation

After this short introductory chapter, the thesis is structured as follows.

Chapter two provides extensive background on the production and perception of non-native speech, models of second language acquisition, and techniques for measuring intelligibility at the full range of linguistic levels.
Chapter three contains a rather traditional overview of the sound systems of the three languages involved in the study, viz. Chinese, Dutch and English, as well as a comparative analysis of the languages in order to predict specific pronunciation and/or perception problems for the various combinations of speaker and listener nationalities.

In Chapter four I will outline the overall setup of the experimental work undertaken in the thesis, and provide a motivation for the choices we made. The chapter then describes the basic materials we collected from groups of 20 speakers for each language background, and how two optimal speakers (one male, one female) were selected from each set of 20 for the definitive tests.

In Chapter five I will present an acoustic analysis of English vowels spoken by Chinese, Dutch and American speakers, and consider how distinct the vowels in the English inventory are from each other, in terms of spectral and temporal properties. We will do this by applying a statistical technique called Linear Discriminant Analysis. The results of the analysis may be used as a prediction of perceptual confusions in the English vowel system as produced and perceived by the three groups of speakers.

In Chapters six through nine I present detailed results for the production and perception of vowels (Chapter six), simple consonants (Chapter seven), consonant clusters (Chapter eight) and for words in meaningless as well as meaningful sentences (Chapter nine).

In Chapter ten I will consider the relationships between the lower (word) and higher (sentence) levels, and try to establish which of the six tests we used affords the clearest separation of the various groups of speakers and listeners. I will then summarize the results, and draw overall conclusions.