



Universiteit
Leiden
The Netherlands

English as a lingua franca: mutual intelligibility of Chinese, Dutch and American speakers of English

Wang, H.

Citation

Wang, H. (2007, January 10). *English as a lingua franca: mutual intelligibility of Chinese, Dutch and American speakers of English*. LOT dissertation series. LOT, Utrecht. Retrieved from <https://hdl.handle.net/1887/8597>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/8597>

Note: To cite this publication please use the final published version (if applicable).

Summary

English as a lingua franca: Mutual intelligibility of Chinese, Dutch and American speakers of English

In the last century, English has developed into the lingua franca of the world. It is now the language of international business, trade, commerce, politics and science. This development has led to a large variety of non-native Englishes, i.e. varieties of English spoken by learners whose native language differs from English. Such varieties are sometimes disparagingly referred to as, for instance, Chinglish (Chinese-accented English), Dungleish (Dutch-accented English), Spanglish (Spanish-accented English), and so on. In these non-native varieties, English is spoken with a distinct foreign accent. Such accents not only allow listeners to identify the non-native speaker's mother tongue, they may also reduce the non-native speaker's intelligibility. Also, a non-native listener's perception of English may be less effective, due to imperfect knowledge of the English sound system, lexicon and morpho-syntax. There is a large body of research on the production and perception of English by non-native learners. Very little, however, is known at this time about the specific problems that arise when non-native speakers communicate in English, if these speakers do not share the same native language. Such situations are found, for instance, when a Dutch airline pilot has to communicate in English with the control tower at an airport in Spain. In our research, we address the problems that come up when Chinese and Dutch speakers communicate with each other in English.

Specifically, I aim to determine the mutual intelligibility of Chinese, Dutch and American speakers in English. Dutch and English are related West-Germanic languages, which share a large part of their vocabularies and whose sound systems do not differ greatly. Standard Chinese (Mandarin), being a Sino-Tibetan language, has a structure that is very different from either Dutch or English, and shares none of the vocabulary. As a first approximation, we test the hypothesis that Chinese speakers of English are more difficult to understand by Dutch (and American) listeners than Dutch (and American) speakers are for Chinese listeners. Secondly, we ask whether non-native English is easier to understand when the speaker and the listener have the same native language. Do Chinese listeners understand Chinese-accented English better than either Dutch-accented English or even American native English? Similarly, do Dutch listeners have less difficulty in understanding a fellow Dutch speaker of English than when listening to a Chinese (or American) speaker of English? This so-called inter-language benefit has only recently begun to receive attention. My study is probably the first to attempt a full-scale investigation of this phenomenon. An itemized list of specific research

questions is included at the end of this summary, together with the conclusions that can be drawn from the experiments, which serve as the answers to the questions.

Intelligibility is tested by determining how well listeners recognize the words a speaker utters, in the order intended by the speaker. Intelligibility is a prerequisite for comprehension (or speech understanding) but differs from the latter in that it does not explicitly involve meaning. In my research I establish the intelligibility of words in meaningless and in meaningful sentences. In order to understand why word recognition is problematic in non-native communication, I also test the ability of Chinese, Dutch and American listeners to identify individual vowels, consonants and consonant clusters in English spoken by Chinese, Dutch and American speakers of English, in all nine possible combinations of speaker and hearer nationalities (or rather: native language backgrounds). American, rather than British, speakers of English were used as controls as the norm of English teaching for my Chinese speakers is American, and Dutch-accented English does not seem to differ more from the American than from the British pronunciation of English.

After my introductory chapter, in which I formulate these research questions, Chapter two presents relevant literature on the topics of intelligibility testing, foreign-language acquisition and the effect of non-nativeness on the production and perception of a language. Chapter three contains a detailed contrastive analysis of the sound systems of Chinese (Mandarin) versus English and of Dutch versus English. Potential problems in the production and perception of English sounds by Chinese and by Dutch learners of English are identified in the analysis, and supported by claims made in the pedagogical literature.

In Chapter four I describe the procedures followed to obtain the materials needed for the experimental part of the research. I attempted to find optimally comparable speakers of Chinese-accented and of Dutch-accented English, one male and one female speaker for each group. These optimal speakers were selected from larger groups of ten male and ten female speakers in each country, such that the optimal speakers were right in the middle of their peer group. In both countries, the speakers targeted were young academic users of English, who had not specialized in English and had never lived in English-speaking environments.

Chapter five, as in intermezzo, presents a detailed acoustical analysis of the vowels produced by the three groups of 20 speakers (ten males, ten females per language background). The results show that Chinese and Dutch speakers keep the English vowels less distinct than the American native speakers do. Nevertheless, the Chinese and Dutch-accented vowels can be identified quite successfully by Linear Discriminant Analysis (LDA). This automatic classification procedure revealed that there is substantial acoustic detail in the foreign-accented vowel tokens that may serve to identify the vowel tokens but is not used by human listeners.

Chapters six, seven and eight present the results of the vowel, consonant and consonant cluster identification tests, respectively, by 36 Chinese, 36 Dutch and 36 American listeners. Vowels were presented in /hVd/ contexts and had to be identified with forced choice from the 20 vowels of English. All onset consonants (24) were presented intervocally in /a:Ca:/ contexts, as was a selection of 21 CC and CCC clusters. Results are first presented in terms of percent correctly identified

targets. In the second part of each chapter an error analysis is presented in terms of confusion structure, using confusion matrices (in appendices) and confusion graphs (in body of text) highlighting the most important vowel and consonant confusions for each of nine possible combinations of speaker and hearer backgrounds. No confusion analysis is given of the consonant clusters as these structures proved relatively easy to identify for all speaker-listener combinations, so that there were not enough errors to make a confusion analysis worthwhile. The overall results show that success in communicating vowels, consonants and clusters depends primarily on the language background of the listener rather than that of the speaker. American speakers/ listeners are generally more successful as speakers and as listeners than are the Dutch subjects, who in turn are more successful than the Chinese speakers and listeners. In spite of these overall effects, however, I find a systematic interaction between speaker and listener language background, revealing a clear effect of the inter-language benefit.

Chapter nine tests word recognition, first in so-called Semantically Unpredictable Sentences (SUS), and second in a selection of sentences taken from the Speech-in-Noise (SPIN) test. In SUS sentences, words appear in six grammatical frames but do not make up a meaningful sentence, e.g., *The state sang by the long week* or *Why does the range watch the fine rest?* In such sentences, later words do not benefit from correct recognition of earlier words. Listeners wrote down all the content words in these sentences, while function words were pre-given on the answer sheets. In the SPIN materials, the listeners wrote down the final word in each sentence, which was either unpredictable from the earlier words in the sentence (as in *We should consider the **map***) or highly predictable (as in *Keep your broken arm in the **sling***). The results show that effects of speaker and listener language background are generally stronger in these word-recognition tasks than in the earlier sound identification tests. But again, the native-language background of the listener exerted a stronger effect than that of the speaker, and again substantial interlanguage benefit could be shown.

Chapter ten presents a summary of findings and then systematically tries to answer the research questions that were identified in the introductory chapter. These questions and answers are summarized below.

1. Is it true that speaker/hearers with an L1 that is close to the target language have an edge over learners with a more distantly related L1? My results show that, indeed, Dutch learners are more successful as both listeners and speakers of English than Chinese learners, even with both groups are selected from young academic users of English as a foreign language.
2. To what extent do separate tests at the lower levels (vowels, consonants, clusters) and at the higher levels (word recognition in nonsense sentences, and in low/high predictability meaningful sentences) contribute independent information to the measurement of mutual intelligibility? It turns out that vowel, consonant, and cluster identification scores are only moderately intercorrelated so that each subskill may contribute independent information to the higher-order word-recognition skill. For Chinese listeners the intercorrelations

are smaller (r-values between .25 and .60) than for either Dutch or American listeners (r-values between .51 and .72).

3. Can word recognition be predicted from success in identification of vowels, consonants and clusters at the lower level? What, more generally, is the correlation between the various types of test results? Generally, the results on the higher-order word recognition tests cannot be predicted with great accuracy from the lower-order phoneme and cluster identification tests. Multiple R is never better than .70, so that maximally 49 percent of the variance in the word recognition scores is accounted for by the lower-order skills. Interestingly, word-recognition can be predicted better from phoneme identification scores when the listeners are either American or Dutch (R-values between .25 and .70) than when they are Chinese (R-values between $-.27$ and $+.25$).
4. Which tests are most successful in discriminating the better from the poorer listeners? Generally, higher-order skills (word recognition) discriminate better between the three listener groups than lower-order (phoneme identification) skills. The best separation of the three groups (Chinese, Dutch, American listeners) is obtained for the high-predictability SPIN sentences, in which the sentence-final word can be recognized more successfully if the listener has also recognized the earlier words in the sentence. Interestingly, this type of test is also closest to real-life intelligibility tasks.
5. Can vowel and consonant errors/confusions be predicted from a contrastive analysis of the sound systems of source and target language? Trivially, sounds that are (almost) the same in the learner's native language (source language) and in English (target language), were transmitted more successfully between speaker and listener than sounds that differ between source and target language. This is as predicted by Lado's classical transfer model. However, within the class of sounds that differ between source and target language further predictions fail. So-called new sounds (target sounds that differ substantially from any sounds in the source language) are not transmitted any better than so-called similar sounds, which differ more subtly between source and target language. Here, the predictions made by the more recent Speech Learning Model fail.
6. Can vowel perception and confusion structure be predicted from an acoustical analysis? Does an LDA on F1, F2 and duration measurements yield the same types of errors as in human perception? Our results indicate that cross-linguistic human perception of vowels can be predicted, with varying success but invariably (much) better than chance, from the acoustic properties of the vowel tokens as produced by native speakers and foreign learners, using Linear Discriminant Analysis. The technique may also be used to predict (part of) the confusion structure of (English) vowels in non-native communication with either or both speaker and hearer having a different language than English and even different native languages.
7. Which factors contribute most to mutual intelligibility? Is the quality of the speaker more or less important to the effectivity of the communication process than the quality of the listener? My results show unequivocally that the effect of listener nationality (or native-language background) is stronger than the effect

of speaker nationality. The overriding importance of the listener effect is found in each of the six tests administered in the test battery.

8. Is the native listener always the best performer? It turns out that, in terms of absolute scores, the American native listeners generally, but not always, obtain the best results. In three tests, Dutch listeners were more successful than the American control listeners but only if the speakers were also Dutch. This, then, is an example of what we may call absolute interlanguage benefit.
9. Do our results support the hypothesis that native/interlanguage benefit exists? Although interlanguage benefit was found in the test results even when absolute scores were used as the criterion (see 8 above), I argue that the phenomenon of interlanguage benefit is more insightfully studied in relative terms. We should first compute an expected intelligibility score based on the mean performance of the listener group and of the speaker group. Relative to this expected score, combinations of same speaker and listener nationality yield higher scores in 16 out of 18 test situations. The overall conclusion, then, is that the interlanguage benefit is pervasive.

