

A letter to ESL students who want to improve their skills in conversational English.

These remarks were originally prepared for university graduate students who are native speakers of languages other than English, but whose reading and writing skills in English were good-to-excellent. These students' understanding of English spoken at normal conversational speed was quite poor compared to native English speakers. As a result they often say that the native speakers "all speak too fast." These students pronunciation of English is often heavily accented, making them difficult for native English speakers to understand. They are thus typical of many non-native speakers of English who have excellent reading and writing skills in English, and yet have problems in conversing with native speakers. The non-native speakers have trouble understanding the native speakers, and the native speakers also have trouble understanding the non-native speakers. Both native and non-native partners in conversations must constantly ask the other person to repeat, if they are to be successful in understanding each other. In the following paragraphs we explain why and how systematic speech perception training can improve both the perception and the pronunciation of English.

Scientific studies by many research groups, including our own, have shown that one of the main reasons that non-native speakers of English have trouble talking with native speakers is that they have not achieved the ability to rapidly recognize sentences, words, and even individual speech sounds in that new language. Native speakers of any language have developed an ability to recognize the sounds of their language so quickly that no effort at all is required to do that. But when we hear the sounds of a new language we all feel that we need a little more time after each word in order to recognize it. Our short-term memory quickly becomes overloaded when the words seem to come too fast for us to recognize one before the next one has come along. This difficulty in perceiving a new language can only be corrected by learning the same automatic processing skills that are used by the native speakers. What is required to do this is simple training in which errors in recognition are immediately identified and corrected. The training is simple, and if it is repeated in drill sessions over a period of several weeks **it will be successful for the great majority of ESL students.**

Training a single student to recognize the sounds and words of a new language can be done by a human teacher, but the cost of many hours of one-on-one instruction (one teacher working with one student) is very high. The Speech Perception Assessment and Training System for ESL students (SPATS-ESL) is a computer-based program that provides the same sort of individualized speech perception training that you could get from a skilled ESL teacher. The

SPATS-ESL training system finds the English speech sounds with which you have the most trouble and then gives you intensive training on those sounds. After every 15-20 minutes of drill on the individual sounds, SPATS-ESL also gives you training on meaningful sentences, spoken in a natural way by young and old, male and female talkers. Research has shown that in 15-35 hours of training, most ESL students can recognize the sounds and sentences on the SPATS-ESL tasks almost as quickly and accurately as native speakers of English.

If you want to become skilled at having conversations with native speakers of English, here is what we suggest you do.

We have found that the most efficient method is to work through the Speech Perception and Assessment Training System for ESL students (SPATS-ESL) using a fixed sequence of training tasks and tests, called the “Standard Curriculum.” Instead of using the Standard Curriculum, you might want to look through the list of training tasks, and even of specific speech sounds, and then select those on which you believe you need training. We hope you ***will not*** do that, because the ***SPATS-ESL Standard Curriculum*** is designed to identify all of the speech sounds that you do not recognize as quickly and accurately as you should, on the basis of your actual performance scores on the SPATS-ESL training materials. The system will then give you the specific training that you need on those sounds. As you improve, speech sounds of greater difficulty are gradually introduced until you are able to accurately recognize all of the 109 common syllable constituents of English. In addition to this training on the basic sounds of English, SPATS-ESL will also teach you to understand naturally spoken meaningful English sentences.

What is the value of training perception, compared with grammar, vocabulary, or pronunciation? Consider deaf children!

If you are to be very successful at having conversations in English, of course you need to understand the grammatical structures of that language, have a fairly large vocabulary and be able to pronounce English in a way that is easy for native speakers to understand. All three of these, grammar, vocabulary, and pronunciation will improve greatly as a result of perceptual training. Consider children who are born without hearing (congenitally deaf). Such children usually have nothing abnormal about the mechanisms required to produce speech, their lungs, vocal cords, mouth, teeth and tongue. And yet, despite years of training by specialists in the speech of the deaf, such children almost never learn to speak in a way that is easily understood by people with normal hearing. The reason is that they do not have proper perceptual “models” stored in their heads of the sounds they are trying to produce. When such children are given the ability to hear even a small portion of the sounds of speech, through an modern operation called a “cochlear implant,” their speech abilities soon improve to near normal levels. This means their grammar, their vocabulary and their pronunciation. All three improve, *when*

their improved perception gives them the ability hear just some of the details of the speech sounds.

Does the importance of speech perception for deaf children have any meaning for non-native speakers of English who have normal hearing?

Yes it does. Speakers from the Far East who have trouble with the sounds of /r/ and /l/, for example, cannot be expected to pronounce them correctly if they cannot hear the difference between them. And there are many other English speech sounds that are not perceived accurately by many non-native speakers. This is because our ears have become “tuned” to hear the sounds of our native language and, when a non-native speech sound enters our ears, our brain leads us to hear (perceive) it *as the native sound that is the closest to it*. Learning a new language means “re-tuning” our auditory system (ear and brain) to correctly recognize the sounds of that language. Learning the sounds of a new language does not make the sounds of our native language more difficult to understand, because we are able to shift back and forth at will, between the new and the old sets of perceptual models, one set for each language that we know. Again, we have remarkable brains!

Speech scientists have recently learned several things about the perception of a foreign language by non-native speakers (hereafter referred to as L2 speakers or learners). As we said before, the L2 learners almost always complain that native speakers of the language they are learning “speak too fast.” Another is that the speech-to-noise ratio at which words can be identified in a noisy background is much lower for L1 speakers of a language than for L2s. Even if the words spoken are very familiar ones for the L2 speakers, a little noise makes them hard to recognize, while the L1 speaker may have almost no trouble at the same noise level. The threshold for 50% correct recognition of words can be as much as 4-5 decibels lower for the L1 speakers than for the L2 speakers. *Basically this means that more of the total speech sound (waveform) must be audible for the L2 than for the L1 speakers to correctly recognize words.*

Finally, how does this perceptual difficulty experienced by ESL students affect speech communication in everyday situations?

It is important to understand that no one perceives all of the sounds of a spoken sentence accurately, whether it is in our own language or one that we are learning. We have conversations in our own language quite successfully ***despite hearing only portions of the sentences spoken to us***. Language is so predictable that it is often easy to guess the identity of a word in a sentence that was not actually heard. There are research studies in which portions of words in sentences are replaced by meaningless bursts of noise, and the listeners not only

identify the words correctly, but often claim that they really heard the missing portions. The brain plays strange tricks!

Costs of having poor perceptual skills when speaking a new language.

What are the “***hidden penalties***” that must be paid by L2 listeners who are attempting to understand sentences in a new language, spoken at normal conversational rates? Here we need to discuss some situations that are not happy ones, and we regret having to mention them because doing so may be considered to be insulting to some L2 speakers. But these are real consequences that are frequently noticed by L1 speakers of English when conversing with L2 speakers, and we know that the same situations occur when the language used is Chinese and the L2’s native language is English.

As professors in universities in an English speaking country we, and many of our fellow faculty members, have been fortunate to have students from other countries in our classes and as research assistants in our laboratories. Nearly all of these L2 students have had excellent skills in reading and writing English, because they passed written examinations in English in order to be admitted to university programs. However, when we speak in English to these students it is very common that they smile and nod, or keep saying, “yes, yes...”, as though they understand what we are saying. But we learn later, when these very intelligent students make mistakes, *that they did not understand*. They were apparently embarrassed, or for some other reason they were unwilling to admit that they did not understand. This is often a such a serious problem that it would make us hesitate to employ such L2 students in any job in which accurate communication is very important, as for example in laboratory research, or health care.

Some linguists have found it hard to understand why many L2 speakers of English can live in an English speaking country for five to ten years or even more, and still fail to learn to perceive English accurately when spoken at conversational rates. And it is also a puzzle that these students often continue to speak English with very severe foreign accents. Why don’t they learn to understand at normal speech rates and also to correct their pronunciation, when they have had so much experience and opportunity for practice? One reason is that the non-native speaker may spend a great deal of their time in social groups that speaks his or her native language, and thus does not really have as much practice as their 5 or 10 years in the English speaking country might suggest. However another theory that is currently considered likely is that the L2 speaker simply learns the language well enough to achieve most of their routine needs, such as buying items in stores, ordering in restaurants, and sometimes even teaching classes in a university. These things can all be done while still having relatively poor English perception and pronunciation. Scientists who study the development of expert skills suggest that speech perception and production, exactly like a great many other skills (for example, playing tennis, the violin, and chess have been studied) are learned by most people up to a level

at which they can do most of what they need or want to do, and at that point they basically stop making serious efforts to improve any further. On the other hand, a few speakers, tennis players, violinists, and chess players, continue attempting to improve...**and they do**. The difficulty in distinguishing /ra/ from /la/, for example, was once thought to be uncorrectable in L2s listeners/speakers from China or Japan, but research has now shown this to be wrong. What is required for continued learning of any of these complex skills (e.g speech, violin play, chess, tennis), is the correction of errors on the basis of immediate feedback of the correct responses. This approach works for many perceptual, motor and cognitive skills, and we are now sure, on the basis of research, that it works for speech perception as well. Once the perceptual problems have been corrected, it is very likely that improved pronunciation will follow after a period of time. We believe the reason for the delay in learning correct pronunciation is this. If the subtle details of correctly produced English speech sounds are basically inaudible (e.g. /la/ vs /ra/) to an L2 speaker, that person cannot be expected to learn to pronounce them correctly. But after those details become audible, through systematic perceptual training, the speaker can then begin to monitor his own pronunciation and make needed corrections.

The last topic is even more likely to be uncomfortable for the non-native speaker. It deals with what a very famous scientist (von Helmholtz) termed “unconscious inference” long ago, in the 1850’s. The idea is that we arrive at many conclusions about things, often without knowing the basis for our conclusions. As an example, most of us with normal hearing can tell when the source of a sound moves from a few degrees to our left, to a few degrees to our right. But few people know that the human auditory system is capable of making that judgment on the basis of *10-20 microseconds* difference in the time of arrival of the sound wave at the two ears. Our judgment (or perception) of the direction from which the sound came is *immediate and compelling*. In other words, it is made automatically without any thought about how we do it. More complex judgments, or Inferences, are made constantly as we live and interact with people. The specific judgment, or inference, that is important for the L2 student is that which is made when a person you are talking to often does not understand exactly what you are saying. Consider the inference you make about a person who seems to understand only if you *use simple words and speak more slowly than usual*. Without really thinking much about the matter, or perhaps without even thinking about it at all, you are likely to conclude (infer) that a person to whom you have to speak slowly and with simple words **is not very intelligent**. In the case of a person who is using a second language, this is usually an incorrect inference. But it is so much a part of our almost reflex-like, or automatic, inference-drawing cognitive mechanisms, that we have to work quite hard to avoid that erroneous conclusion. This has been mentioned here because it is a very real penalty of having perceptual skills in English that are less accurate than they could be with training. Poor English perception not only reduces the efficiency of a person’s conversations with native speakers, it also is very likely to have a

negative impact on the L1 person's estimate of the overall intelligence of the L2 speaker. Note that this surely does not happen only when the conversation is in English. It must be common as well when the conversation is in Chinese, and a native speaker of English is attempting to converse in that language.

The evidence we have collected on over a large number of L2 listeners in university language classes strongly suggests that most of them could come close to the perceptual skills of native speakers for English speech sounds, after 15-35 hours of training. Some can do it in less time, and a few may require more. But we believe that 90-95% will be hearing English speech sounds as represented in the SPATS-ESL curriculum as accurately as native speakers after about 35 hours of the type of training with that curriculum. These results are in good agreement with a large number of training studies conducted by other investigators, whose experiments used only two (such as /l/ vs. /r/), or a few speech contrasts. The main advantage of SPATS-ESL is that it trains all of the sounds of English that are commonly not perceived correctly by L2 learners of that language, and automatically focuses training on the specific sounds that are of difficulty for the individual student.

As the developers of SPATS-ESL, we hope these thoughts are of some help to you, and we look forward to working with any of you who decide to purchase this training program,

Sincerely,

James D. Miller, Ph.D.

Charles S. Watson, Ph.D.

Communication Disorders Technology, Inc.,
Indiana University Research Park
501 N. Morton St.
Bloomington, Indiana
USA