社会イノベーション研究
第３巻第１号（91-112）
2008年１月

2007年10月16日掲載承認

Phonological familiarity and short-term verbal memory: Implications for teaching English as a foreign language (TEFL)

Dennis Riches

Introduction

This paper describes a small-scale experiment in speech perception and short-term memory, the results of which suggest some shortcomings in the conventional pedagogy of teaching English as a foreign language (TEFL).

The experiment used a common practice of speech perception research, that of measuring perception of 'nonwords' that are fabricated words that obey the phonology of a language, but are contrived to be words without meaning. Only one of the respondents was multilingual, and all of them had lived in English Canada since childhood and considered English as their native language. They were asked to listen to ten two-second utterances and repeat them as accurately as possible immediately after hearing them. Five of the utterances were in the African language, Hausa, which none of the volunteers had any familiarity with. The other five utterances were strings of nonword English.

The experiment tested the hypothesis that the volunteers would be able to repeat the familiar English-like utterances more accurately than utterances from a language with an unfamiliar phonology. All the volunteers reported that the English-like utterances were easier to hold in memory and repeat, and the experimenter's

observations agree with the volunteers' subjective responses.

Although this result is unsurprising, it is surprising that this difficulty does not get more attention in TEFL, particularly when native speakers teach their language to non-native learners. Native speaker teachers of English tend to focus on top-down listening strategies because their bottom-up knowledge of phonological rules has been operating below conscious awareness since their early childhood. They never had to learn it the way their students have to.

Language has been described as a unique system of communication in the animal world because it has a duality of patterning. One system of this patterning is the systematic arrangement of sounds that have no meaning, and the other, which language pedagogy tends to concentrate on, is the arrangement of these sounds into meaningful elements. Teacher training programs and the literature of TEFL reflect this bias. However, if we assume that long-term memory and comprehension will be impaired when short-term memory does not function, the results of the experiment described herein add to previous work that suggests that more attention should be paid to developing methods that improve speech perception and comprehension at the fundamental level of phonological encoding of the speech signal.

Previous research in verbal short-term memory

The studies covered in this section underline the importance of short-term memory and phonological processing in comprehension, learning and the formation of long-term memory. The experiment was inspired by my work teaching English to Japanese university students. Listening skills are usually taught to these students by methods that emphasize top-down mental processes such as employing extra-linguistic cues and the context set up in pre-listening activities. Many courses and textbooks designed to teach listening skills intentionally withhold transcripts from students, based on the belief that a resort to analysis will divert attention from the need to comprehend holistically.

Listening comprehension is a particularly acute problem for Japanese learners because the phonology of Japanese is so unlike the phonology of English. The Japanese English education system solves this problem by avoiding it and defining Eng-

lish ability largely through visual media – that is, book-based teaching and measures of proficiency based on reading ability.

The communicative teaching theory of recent decades, popular outside of the Japanese education system, is a one-size-fits-all approach, largely concerned with the naturalistic learning of syntax and lexicon, with phonology mostly an afterthought. It cannot take account of typological differences between English and other languages on an individual basis. It has been formed and promoted mostly by native speakers of English working in Britain and the United States (ESL settings). Their students are from linguistically diverse backgrounds, immersed in English-speaking cultures, so attention to L1 differences is not feasible.

For teachers in EFL settings, it is more practical and advisable to take account of the typological differences between English and the common L1 of the learners. In the case of Japanese, it would be hard to find a language that, in terms of phonology, would give greater disadvantage to learners of English. It is a notoriously sound-poor language, with a small inventory of phonemes that includes only five vowels. English is stress-timed, but Japanese is a syllable-timed 'mora' language, in which syllables are, with a few exceptions, made up of a single vowel or consonant vowel pair (V or CV). There are no consonant clusters that can give complex English single-syllable word with this structure: CCVCCCC (*twelfths* – /twɛlfθs/), so the inventory of possible syllables in the Japanese is only a few hundred, as opposed to a few thousand in English. This contrast makes English fast speech and word segmentation particularly difficult for Japanese learners (including the Japanese people who teach English) to master. The first six years of formal English education in Japan consist of learning vocabulary and grammar via the written language. A Japanese high school graduate may have an English vocabulary of thousands of words (an admirable feat of visual memory of English orthography), but he has no idea what this vocabulary sounds like in speech because his training included learning kana transcriptions above English orthography, and these distort pronunciations and insert vowels between each consonant in consonant clusters.

In communicative language teaching, there are few methods and materials to assist learners with the difficulty of training their ears to the sound system of a new language. One common approach in teaching listening is to simplify the speech by

lowering the rate of speech and making the sound quality ideal. However, research by Bradlow & Bent on the clear speech effect for L1 and L2 listeners had a startling result that should give pause to teachers who think simplified speech improves the chances of comprehension by L2 learners. They found that slowing the rate of speech, improving the signal to noise ratio, and using a female voice *did* improve the comprehension of L1 listeners, *but not* L2 listeners. They concluded, "...clear speech is essentially native-listener oriented, and therefore is only beneficial to listeners with extensive experience with the sound structure of the target language."

Another element of the conventional approach for teaching listening is to have students answer comprehension questions about simplified news reports or academic lectures. Learners can often seem to master these tasks because they answer the questions correctly, but in my own experiments with these advanced learners in Japan, they have been unable to accurately transcribe what they apparently comprehended based on responses to comprehension questions. When the exercises are de-contextualized, stripped of the illustrations, pre-listening exercises, and multiple-choice questions that often lead one to the answers, the illusion of comprehension disappears. In pointing this out, I do not deny the importance of contextualization in listening comprehension. Nor do I deny that the exercise at least provides useful exposure to the target language. The problem is just that it cannot be confused with comprehension. Of course, both L1 and L2 listeners use context, and teachers should make good use of it, but it should not be over-emphasized and relied on so heavily that it gives a false impression of language proficiency, to the teacher or the learner himself. Methods that take learners to this 'advanced' level still relying heavily on guesswork and context clues do not really help them. In their free time Japanese learners do not casually listen to the sorts of English media they study in class because it is still too mentally taxing to be an enjoyable way to get information or entertainment. The elementary level of understanding – phonological encoding – still does not function in these students, even if they have been labeled 'advanced' according to TOEIC scores and other measures.

Modern theorizing about short−term memory began with Miller's famous 1956 paper *The magical number seven, plus or minus two: Some limits on our capacity for processing information.* Its influence on education and on culture in general can

be seen in numerous ways. It influenced the standard length of telephone numbers, and the preferred structure of self-help pop psychology books with their seven steps, seven chapters or seven subheadings, or lists of seven essentials that form easily remembered seven-letter acronyms.

Another foundational work was Baddeley et. al.'s 1975 paper *Word length and the structure of short-term memory.* They stated that verbal short−term memory is only as long as the number of words perceived in approximately 1.6 seconds (the precise duration is controversial, but this is figure is generally accepted, plus or minus a few fractions of a second). It is constrained by time rather than volume of input. They proposed *a tape-loop* metaphor for short-term verbal memory, while others have compared it in computer terms as a *slave system or a memory cache* that is constantly dumped and refreshed with new input. Because these are metaphorical models, it is not easy to prove their validity or explain how short-term memory interacts with pre-existing knowledge and other cognitive functions to form lasting memories and understandings of language input. However, it is not controversial that there is a clear difference between long-term and short-term memory. Studies of subjects with brain pathologies show that they are distinct systems. Numerous studies have also indicated that the listener's acquired phonological system is key to the working of short-term verbal memory, so much so that it is called the *phonological loop* or *phonological memory.*

In research on dyslexia, deficits in phonological processing are now viewed as the underlying cause of difficulties in learning to read. Lundberg notes the difficulty of diagnosing dyslexia in L2 learners because their difficulties in learning language and reading in the target language are, on the surface, the same as that of the L1 dyslexic learner. They both have trouble with accurate phonological encoding.

Lundberg gives a good overview of the importance of phonological encoding and memory in language acquisition. Common experimental techniques rely on nonwords that conform to the phonology of the target language but have no semantic content. Both L1 and L2 language proficiency and aptitude have been linked to proficiency in recalling nonword samples. Finally, Lundberg makes an interesting point that is seldom recognized in the debates within TEFL about the learning of grammar – about whether it can be be taught or can only be left to natural acquisition

processes:

> There is also a possibility that the acquisition of syntax is related to the phonological loop. Syntactic rules are abstracted on the basis of language patterns consisting of strings of words. These word strings must first be held in phonological working memory. A low capacity will impede the construction of more permanent, long-term memory representations.

Thus when we read in TEFL literature the view that the L2 learner will acquire grammar in a natural way that cannot be altered, we must keep in mind that these theories were put forward by native speakers of English who seem to have taken the phonology of their own language as self-evident to all learners. If grammar rules are not evident in the output of learners, this is usually taken to be a sign of the undeveloped output system, a weakness in the monitoring of patterns already internalized. They assume the input has been accurately perceived. However, the acquisition problem may be rooted in the fact that grammatically salient phonemes (such as English plural markers /s/ /z/ and /əz/, past tense markers /d/ /t/ and /əd/, or all the phonemes in unstressed function words) are not perceived accurately, if at all. The problem persists with advanced grammar structures because there too the important elements are unstressed. In this example: *I'd 'a' helped you if I'd known you were in trouble* the unstressed underlined elements contain important auxiliary verbs that are barely perceptible to an ear that is not attuned to them. The counterargument is that every human ear hears the same sound wave, no matter what the native language of the hearer is, but we must remember that the human perceptual system easily deceives itself with its own expectations, a fact exploited by every magician. Weber concluded from research with English speakers learning German that it is likely that they listened to the new language with the same sort of unconscious perceptual prejudice, which in this case was the tendency to listen to German while applying the phonotactic rules of English.

Chunking is one recent trend in TEFL in which verbal short-term memory research seems to be implicitly acknowledged. It is concerned mostly with the learning of lexicon, and it is the idea that words are not remembered in isolation but in

groups that, coincidentally, are spoken in 1 – 2 second bursts. Chunks are common phrases that collocate verb to object, like *ride the train,* or they might be a few words that cover a complex concept like deciding to do something that will not have a negative outcome (*I might as well*). Although this may be stored in memory as a single item, the learner still has to segment it and recognize the four words it comprises. Many chunks consist of a mix of grammatical markers, lexical items and function words that come in packages of stressed and unstressed elements. In order to be learned, they need to be perceived by an ear that is attuned to English phonology. Japanese learners, for example, are poor at perceiving the unstressed elements of English sentences. In the chunk *that's as good as it gets,* they are likely to hear only the underlined stressed words. They are often exhorted by teachers to 'listen for the key words' but it is usually the case that the key words are easy because they are stressed, but they are useless without the unstressed elements that surround them. *That's good gets* would be baffling to anyone.

The teaching of phonology may be ignored precisely because it is so complex that it defies attempts to teach it. Rost writes, "Although efficient auditory perception underlies effective listening, it would be oversimplifying .... to suppose that learning to listen involves massive practice with phonological decoding alone... it is doubtful that 'fast speech rules' can be learned deductively and consciously applied in real time... Rather we should expect that learners will acquire gradually a phonological sensitivity to the new language in contexts of actual use" (Rost, 57). However, on the same page he goes on to add that in pedagogy we should focus on the importance of "developing selective attention" to features of English that speakers of other languages would not focus on.

More recent studies of infants support Rost's view that phonology is too complex to learn deductively. Studies of infants (see Johnson & Jusczyk) suggest that they quickly acquire all the arcane rules that can be uncovered by linguists: segmentation rules that indicate word boundaries, prosodic features, phonotactic rules that limit which phonemes co-occur within a syllable, and where they can occur, allophonic variability, systematic alternation, and syllable length and structures. All of this extremely complex pattern recognition and computation of statistical probabilities starts to happen in the first years of life, so it is obvious that humans begin life

with a highly specialized ability to acquire language at first by recognizing sound patterns and their relative frequency of occurrence in their mother tongue. This developmental period is a window of biological development that older foreign language learners do not have an opportunity to exploit. Even those scholars who doubt the existence of a critical period for language acquisition do allow that there is strong evidence of the critical period for acquiring a native-like accent, which we can infer to also mean the ability to acquire the entire phonological system in terms of both perception and production. The point made here is that the acquisition of phonology is the prerequisite for the acquisition of all aspects of language. It is wrong to think of a separation between a critical period for phonology and a critical period for syntax and lexicon. These can be learned effectively later in life with the compensating strategy that detractors of the critical period fail to acknowledge: resort to learning via the written language. The written word is a crutch that children do not need during the years when they acquire language, before learning to read in their formal education.

As much as Rost advocated for a pedagogy that would focus selective attention on the salient features of English phonology, he states, “understanding spoken language is essentially an inferential process based on perception of cues rather than straightforward matching of sound to meaning” (Rost, 33), and it seems that no one in TEFL followed up on his contrary recommendation to develop a pedagogy of selective attention to sound patterns. What is needed is perhaps something more sophisticated than the familiar minimal pair drills found in pronunciation textbooks, or indeed a better strategy than consigning problems phonological to marginal pronunciation exercises tossed into to chapters arranged around grammar points, themes and communicative functions (see any of the contemporary EFL course books, or Rost 2005, p. 37 for a specific example). Advances in technology have led to some valuable methods produced by the makers of language learning software and computer language labs, but they are far from being a mainstay in TEFL. These technological approaches have learners do a variety of complex perception and production exercises, giving them spectrographs of their own output to compare with that of the model. However, these new technologies are unlikely to receive much attention in the field of TEFL because of the profession’s bias toward communicative language

teaching and humanistic approaches. The technological approach to the problem implies a redefinition of the role of the teacher, one which those trained in CLT are unlikely to embrace. It is highly likely that time in the computer lab is more valuable for low proficiency learners of English than time in the communicative language classroom, but language teachers have little motivation to conduct studies on this question.

TEFL is made up mostly of humanities graduates who are interested in creatively applying their education in the classroom, so the wonky fields of phonological pedagogy (which does not exist), or research in sound perception, would rate even lower interest levels than grammar pedagogy. Graduates in the humanities of the last few decades have been much more interested topics such as whole language, socio−cultural theory, holistic education, or teaching content and social awareness in the language classroom. Thus they have been pre-disposed to viewing listening mostly as a top-down cognitive process, something that is learned in the process of having attention focused elsewhere. Furthermore, we cannot ignore the preferences of learners, or the market for EFL instruction, that expect methods to be entertaining and enjoyable, but perhaps not the most effective.

In communication theory there is the well−known bromide that 90% (or is it 80% or 70% or 60%?) of a message is communicated non-verbally. In extreme views, utterances have no intrinsic meaning except for that which is in the mutually agreed hallucination shared by producer and perceiver. In milder views, meaning is dependent mostly on semantic schemata and extra-linguistic factors that are in play as people communicate. Besides, it is believed that there are no standards in the numerous contexts in which English is used globally, so usage and meanings should be free of standardizing prescriptions.

These views have held allure because, having been revealed by experts, you are an expert too if you can learn them and pass on the counter-intuitive revelations that overturn the misguided common sense understanding of conventional thinkers. However, rational inquiry often leads one right back to restating the common sense that everyone believed in the beginning. For example, no matter how little of the message is conveyed by words, it is the most crucial part. People do not take satisfaction from watching films with the audio track missing, no matter how well the actors

express themselves non-verbally. It is easy to imagine other situations in which we are frustrated when spoken language fails. We lose patience in trying to communicate with someone who has language impairment from a stroke, or feel annoyed when background noise makes speech hard to listen to. Furthermore, yes, of course, our words have meanings only because we mutually agree on them, but our multi-word utterances have meaning because of the formal properties of language and the evolved design of the human language faculty which uses two formal systems (a duality of patterning) – one for assembling meaningless sounds, and another for arranging those sounds in meaningful patterns. Finally, as for the problem of agreeing on a standard English to learn, it is in some aspects of phonology that commonalities have been found, in spite of the common belief that unfamiliar accents cause incomprehensibility across varieties of both native and non-native English. Jenkins' (158-159) research on the history of this issue led to her tabling *The Lingua Franca Core* (see Table 5) of features of native or non-native varieties of English that provide comprehensibility to both native and non-native listeners.

Researchers who have done recent studies in speech perception, with the advantages of brain imaging technology, place more emphasis on the smooth functioning of the phonological loop than to higher level cognitive processes. For example, Natsume et. al. gave intense instruction in English prosodic patterns to Japanese learners, then hooked them up to an electroencephalogram. Compared with those who had not received the instruction, they had greater occurrence of theta waves while listening to short lectures in English, and theta waves have been linked in other research to states of attentiveness and memory formation.

A similar line of inquiry is found in the therapy for language learning based on the Tomatis method (see Gerritsen). This method is based on the assumption that you cannot produce a sound which you cannot hear, and it states that many of the overtones of a foreign language are inaudible to learners. Base tones produced in the larynx are universally the same, but they are altered differently in each language by the overtones laid on them as they pass through the oral cavity. English, on average, has more high-frequency overtones, and this presents a challenge to the ear of the learner accustomed to a language with lower frequency overtones. The Tomatis therapy focuses on strengthening the muscles of the hammer and stirrup in the middle

ear, and it does this by making the subject listen to sounds which are in the frequencies which need to be heard better. The therapy also pays attention to the problem of differing syllable length between languages (for example, English: 75 milliseconds, French: 50 milliseconds). Empirical studies, for example the Coomen experiment in 1976, have demonstrated the beneficial effect of the therapy on overall foreign language skills. What is notable about the Tomatis method is its aim to strengthen fundamental weaknesses rather than compensate for them, and it is also notable that Dr. Tomatis, being French, has left no lasting influence in the field of TEFL where his work is almost unknown.

Poeppel & Hackl, in *The Architecture of Speech Perception,* conclude their chapter thus:

> Speech perception is the process of extracting information from an acoustic signal and constructing the appropriate representation that can interface with the stored items in your mental lexicon and the linguistic computational system... Speech perception is hard – for example because there is no one-to-one mapping from stretches of sound to phonemes and because there are no (obvious) invariant properties in the signal. That these difficulties are not trivial is attested by the fact that automatic speech recognition technology is not particularly far along. Nevertheless, the human brain deals with the problems effectively. We suggest that the efficacy of the system derives from at least three properties of the speech processor. First, a speaker's *knowledge* of phonology significantly helps the process. Second, the problem is broken down in space: multiple areas [of the brain] contribute to different aspects of the problem (much like vision). Third, the problem is broken down in time by analyzing signals on different time scales. The prerequisite for the development of a model of the cognitive neuroscience of speech is theoretical agreement on what the appropriate linguistic units of study are. Here it is built on the assumption that the basic unit of speech that makes sense of neuronal data is the distinctive feature [phoneme]. It is the concept that best connects linguistic theory to biological data."

McQueen (1998) sums up by stating, “The legality of sound sequences appears to be computed on-line during recognition. It is argued that segmentation is achieved via competition between candidate words, that competition is modulated by knowledge about where in the input candidates are unlikely to begin or end, and that phonotactic constraints are one of several information sources used in this segmentation process.” Davis & Marslen-Wilson found that “...acoustic differences in embedded syllables assist the perceptual system in discriminating short words from the start of longer words. The ambiguity created by embedded words is therefore not as severe as predicted by models of spoken word recognition based on phonemic representations.”

These findings summarized above are notable for their basis in physiology and their use of reliable technology as opposed to teacher intuitions and ideological agendas. It is significant that these researchers, being disinterested in language teaching, did not start out trying to prove, for example, a hypothesis about the crucial role of teacher-learner interaction in learning outcomes. Their goal is “a model of the cognitive neuroscience of speech.” Research agendas aimed at explaining inter-personal interaction, and learner motivation and affect have exerted a strong influence on education in recent decades, but it is seldom acknowledged that they are in the realm of moral philosophy and may be impossible to verify with empirical data. As much as these issues are worthwhile to consider in teacher training, teachers can get a refreshing perspective from research that is concerned simply with the basic neurology of information processing. Once this is understood, one can more effectively hypothesize about how learning is affected by the complex variables of cultural differences, teacher-learner interaction, socio-economic background, and individual variations in temperament and intelligence.

For native speakers to appreciate how much they depend on their native phonological system, or bottom up processing, to understand spoken language, the research findings described above can be supplemented with some simple experiments that control either for context or for clarity of the speech signal. Items 1 – 3 in the list below have high context but poor clarity, and they effectively give the native speaker of English the handicap of a low-proficiency non−native speaker.

1. Worsen the signal to noise ratio. Listen to a broadcast of breaking news while children play in the room and someone else washes dishes loudly in the adjacent kitchen. Take notes, listen attentively for five minutes, then answer multiple choice questions on the content. Keep the volume at a reasonable level that does not disturb the children's play.
2. Listen to a speech, written coherently in standard English, but delivered by someone making numerous errors outside the Lingua Franca Core (see Table 5). That is, in lay terms, someone with a very 'thick' accent. Pay attention for fifteen minutes, take notes, and accurately summarize the views of the speaker afterwards.
3. Listen to an American airline pilot speaking to passengers on the public address system while you are exposed the ambient noise of the cabin.

In the opposite way, numbers 4 – 7 illustrate how much you can understand when context is absent, but clarity is good. Most native speakers can successfully complete these simple experiments, and this disproves the notion that the speech signal is inherently ambiguous without context.

4. Ask a friend to randomly read aloud a few words from the dictionary. Repeat each word.
5. Ask a friend to randomly read aloud a few sentences from randomly chosen books, pausing every 2 – 4 seconds. Repeat after your friend at each pause.
6. On your DVD player, play a stretch of randomly chosen dialog or narration from a film you have never seen before. Pause the playback every 2 – 4 seconds and repeat the words you hear.
7. Flip the radio dial every few seconds between a few different talk radio programs. Is each snippet of speech incomprehensible because it is de-contextualized?

The experiment described below highlights the importance of familiar phonology in speech perception and short−term verbal memory. Like the experiments in the list above it underlines the importance of the smooth functioning of the phonological

system in listening comprehension, and the results support the argument that present teaching methods could be supplemented with methods that strengthen this complex array of perceptual skills that operates below the conscious awareness of native speakers.

The experiment

The experiment used a common practice of speech perception research, that of measuring perception of nonwords. These nonwords are fabricated words that obey the phonological rules of the subject's native language, but they are nonsense words that convey no meaning. Ten volunteers were recruited in Vancouver, Canada. Only one of the volunteers was multilingual, while the others had lived in English Canada since childhood and considered English as their native language. They were asked to listen to ten two-second utterances and repeat them as accurately as possible immediately after hearing them. Five of the utterances were in the African language, Hausa, which none of the volunteers had any familiarity with. The other five utterances were strings of nonword English.

The experiment tested the hypothesis that the volunteers would be able to repeat the familiar English-like utterances more accurately than utterances from a language with an unfamiliar phonology. All the volunteers reported subjectively that the English-like utterances were easier to hold in memory and repeat, and the experimenter's objective observations agree with the volunteers' responses.

The nonword samples of English sentences are listed in Table 1, with a corresponding real English sentence beside each nonword sentence. Readers can notice that the nonword sentence is a mask of the English sentence beside it. The word count is the same, the total number of syllables is the same, and number of syllables and the syllable stress is the same in each word. For example, in the first pair of nonword and real sentences, the first six words are monosyllables, and the seventh word is a two−syllable word with stress on the second syllable. The nonword sentences also obey English phonotactic rules that give listeners cues for word segmentation. For example, in English phonotactic rules (or sonority hierarchy rules) the phoneme pair /sl/ can come at the beginning of a word or syllable and the pair /ls/

Table 1

Respondents listened to each nonword sentence then immediately tried to repeat it.

	Sentences of English nonwords (two-second utterances)	**Sentences of English nonwords correspond to these sentences in syllable structure, stress and rhythm**
1	sə ɛnt ʌb və doi ɪg məʒɛd	The end of the road is ahead.
2	fai tolb rəm ʤɔd bə lʌm	I told him not to come.
3	wɛdz neit lʌn ni dʌp di nʌb	Lets wait and see what he does.
4	læz ni nai suk gliz	Pass me my book please.
5	sə nɪgəm nɔsts nə məgə nɪkti	The ticket costs a dollar fifty.

can come at the end of a word or syllable, but not vice versa. When an English speaker hears *Neil sees her often* the phonotactic rule can be applied, in addition to knowledge of male names, syntax and vocabulary, to assist with word segmentation. Though the correct segmentation seems obvious to a speaker of English, a foreign learner might wonder if it concerns a person named Nee “lseeing” her often.

Thus, the nonword sentences are English−like in every respect except that phonemes have been substituted to make them meaningless. All the phonemes used are English phonemes, and the sentences were recorded by a native speaker of mid-Atlantic American English who practiced them until he could read them as connected speech with the same prosody as the corresponding real sentence.

There are various conventions in use for the phonemic transcription of English because of the variety of dialects and the variety of opinions about how the standard language should be transcribed, so the symbols used here are a best approximation of the forty or so phonemes that linguists generally agree are used in mid-Atlantic American English. Readers who need a guide to the symbols used in the samples can refer to the iconic sample words beside each phoneme in Table 2.

The two-second samples of Hausa were obtained from T*he Handbook of the International Phonetic Association.* This handbook contains samples of various languages and their transcriptions into phonetic script. Each sample in the handbook is a translation of the fable *The North Wind and the Sun.* The Hausa transcription of the story is in Table 3, and an audio file of the sample can be obtained from the handbook’s website (see University of Victoria Department of Linguistics).

Table 2

Pronunciation key of the less familiar phonemic symbols used in the nonword sentences.

ɘ	elementary		æ	cat
ɛ	egg		o	low
ʌ	cup		ai	eye
ɔ	hot		oi	toy
ʊ	look		ei	play
i	see		ʒ	vision
ɪ	fit		ʤ	jam

For each subject, the experiment took only a few moments. She listened to a recording of Set A, five two-second segments of the Hausa version of *The North Wind and the Sun*, and Set B, five nonword English sentences that were each approximately two seconds in duration. After listening to each sentence, she was asked to repeat the sentence immediately, and each response was recorded. No attempt was made to balance male and female respondents, or look for male−female differences in responses, but through the random selection process it turned out that seven of them were female and three were male. All of them except one described themselves as proficient only in English. The one multilingual respondent learned Hindi and Farsi in childhood, and she did the best of anyone in repeating accurately both the nonword and the Hausa sentences. Her participation in the survey provided a useful contrast with the monolingual English speakers.

Before summarizing the results, it must be admitted that they are impossible to quantify perfectly. The subjects rated their own success at repeating the sentences they heard, and the experimenter's judgments, made with human ears, are also unavoidably subjective. The experiment could be redone by comparing acoustic spectrographs of the originals and the responses, but the subjective evaluation is considered sufficient here because the experiment was designed mostly to illustrate a common sense truth that is easily forgotten in language teaching. Furthermore, when the test subjects listened to the Hausa samples, in most cases they shrugged, smiled and gave up the attempt, but they did not do this with the nonword samples. Thus, there was little ambiguity or room to wonder about the possible errors of subjective judgments. There was a clear difference between the reactions to Set A and to Set B. The familiarity of a language's sound system is a crucial element in verbal memory

Table 3

Transcript of *The North Wind and the Sun,* Hausa version

Source: *The Handbook of the International Phonetic Association*

IPA Transcription:

wətə raːnaː, də ʔɪskər hʊ̀ntuːrùː tə ʔərèːwəʔ | də raːnaː sʊkə̀ ji gə̀rdəmàː ʔə kə̂n koːwàːtʃeːtʃɛ̀ː də̀gə̀ tʃɨkɪnsù tə ɸi kʼərɸiː. tò sʊnàː tʃɨkɪn wənnə̀ŋ gə̀rdəmàː | səj gàː wəni mətə̀fɪjiː, jaː zoː sə̀ŋɛʔ də̀ rìːgər səŋiː. tò ʃiːkɛ̀ːnən, səj ʔɪskə̀r hʊ̀ntuːrùː də̀ raːnaː sʊkə̀ ji jə̀rdʒeːdʒeːmìjaː, ʔə̀ kə̀n jə̂w, zaːʔə̀ ji kʷʼʊreʔ, dɔn ʔə̀ gə koː wàː zə̂j ʔijə̀ sâː mətə̀fɪjĩn jə̀ tuːɓɛ̀ː rìːgərsə̀ʔ ʔə̀ kə̂n tiːlə̀s. tò ʃiːkɛ̀ːnən, səj ʔɪskə̀r hʊ̀ntuːrùː tə buːsoː səŋɪntə̀ʔ, də̀ kʼərɸiː | ʔijə̀ jɪntə̀ | ə̀mmaː ìnâː. səj tə kaːsə̀ sâː ʃiː wənnə̀m mətə̀fɪjiː | jə̀ kʷəɓɛ̀ː rìːgərsə̀ʔ, dɔn kʷʊ̀wa, jaː dəndə̀nneː rìːgə̂r gə́mgə́m ʔə̀ jɨ̀kɪnsə̀ʔ. tò də̀gə̀ kʼə̀rʃeː dəj, səj ʔɪskàː tə səllə̀maː, də̀gə̀ nən nɛ̀ː kʷʊma, ʔɪtə raːnaː | tə ʃɨ̀gə naːtə̀ ʔajkɪ̀ŋ. ʔaj kòː raːnaː | səj tə buːɗoː həskʲə̀ntə̀ də̀ zaːfɪntə̀ wə̀rwə́r. hə́ɓə̀ː | kə̀n kə̀ tʃê: mɛ̀ː, səj gàː ʃi wənnə̀m mʊ̀tʊ̂m | mətə̀fɪjiː | jaː kʷə̀ɓɛ̀ː rìːgərsàː, bâː gʲɪrmaː də̀ ʔərzɨ̀kiː. gənɪ̀ŋ həkə̀ kʷʊ̀wa | səj ʔɪskàː dóːlɛ̀ tə ʔəmìntʃeː ʔə̀ kə̂n tʃêːwaː, lə̀lleː | raːnaː taː fiː tə̀ kʼərɸiː.

Orthographic version:

Wata rana, da iskar hunturu ta arewa da rana suka yi gardama a kan kowace ce daga cikinsu ta fi ƙarfi. To, suna cikin wannan gardama, sai ga wani matafiyi ya zo sanye da rigar sanyi. To, shi ke nan, sai iskar hunturu da rana suka yi yarjejeniya a kan yau, za'a yi ƙure don a ga ko wa zai iya sa matafiyin ya tuɓe rigarsa a kan tilas. To, shi ke nan, sai iskar hunturu ta buso sanyinta da ƙarfi iya yinta, amma ina?! Sai ta kasa sa shi wannan matafiyi ya kwaɓe rigarsa, don kuwa ya dandanne rigar gamgam a jikinsa. To daga ƙarshe dai, sai iska ta sallama, daga nan ne kuma, ita rana ta shiga nata aikin. Ai ko, rana sai ta buɗo haskenta da zafinta warwar. Haba! Kan ka ce me, sai ga shi wannan mutum, matafiyi, ya kwaɓe rigarsa, ba girma da arziki. Ganin haka kuwa, sai iska dole ta amince a kan cewa, lalle rana ta fi ta ƙarfi.

English Orthographic version:

The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shined out warmly, and the immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

and learning. The results are exactly what a layman or linguist would expect. The subjects did better at accurately repeating the nonword English sentences than they did at repeating the Hausa sentences.

In the questionnaire, all ten respondents said the Hausa sentences were more difficult to repeat. They rated their own accuracy in Set A with an average score of 19, and in Set B with an average score of 32, 1.68 times higher. They rated difficulty on a scale of 1 to 5, with 1 being “impossible to repeat” and 5 being “able to

repeat accurately." Thus a numerical value was placed on their impressions of their performance.

The respondents could usually make a reasonable approximation of the non-word English samples, and some of them were very accurate parrotings, but when they attempted to repeat the Hausa samples they often faltered and took the option to pass completely on an attempt. When they tried, their responses seldom resembled the original, in a way that was obvious even to someone unfamiliar with the Hausa language. The fact that one respondent was multilingual turned out to be a fortunate addition to the experiment because she was the only person who was able to quite accurately repeat both Set A and Set B. One of the monolingual respondents did noticeably better than eight of the others on both sets, which seemed to confirm research in foreign language learning aptitude that found a high degree of variability in individual phonological sensitivity (see Caroll & Sapon).

Conclusion

Almost anyone would find it impossible to repeat utterances of a an unknown language with an unfamiliar phonological system. This is intuitively obvious, so this experiment serves as an illustration of a point to be taught to language teachers, rather than as a new contribution to linguistics, and it is not something that people lacking the 'benefit' of training in communicative language teaching need to be told. Linguists and cognitive scientists already know well about the role of phonological memory in language comprehension. The argument made in this paper is that TEFL professionals have not paid enough attention to this matter, so it is hoped that this small-scale experiment conducted with English native speaker subjects will make an impression on English native speaker TEFL teachers. The point is that their native phonological system works below conscious awareness but is crucial to language comprehension. Because it is below conscious awareness, they are likely to be unaware of how much the lack of this phonological system impedes language ability in L2 learners of English. More attention to strengthening this fundamental skill is essential to the development of the higher level skills that TEFL professionals have customarily been more concerned with.

Table 4

Respondents' Questionnaire – With Results Compiled and Summarized in Italics

1. Which set was more difficult to recall and repeat accurately?
   total responses: 10, the number of responses for each option is in parentheses
   A = 5 two-second utterances of Hausa B = 5 two-second nonword English sentences

   i) A more difficult than B. (*10*)
   ii) B more difficult than A. (*0*)
   iii) A and B were equal in difficulty. (*0*)

2. Rate your accuracy in doing Set A
   5 = I was able to repeat the sentences accurately. → 1 = I found it impossible to repeat the sentences.

   total responses: 10, the number of responses for each value is in parentheses

   Circle a number: 5(*0*) 4(*0*) 3(*3*) 2(*3*) 1(*4*)

   Difficulty rating for Set A: 19 = [(5 x 0) + (4 x 0) + (3 x 3) + (2 x 3) + (1 x 4)]
   (the lower score indicates greater difficulty in repeating the sentences)

3. Rate your accuracy in doing Set B
   5 = I was able to repeat the sentences accurately. → 1 = I found it impossible to repeat the sentences.

   total responses: 10, the number of responses for each value is in parentheses

   Circle a number: 5(*1*) 4(*2*) 3(*4*) 2(*3*) 1(*0*)

   Difficulty rating for Set B: 32 = [(5 x 1) + (4 x 2) + (3 x 4) + (2 x 3) + (1 x 0)]
   (the higher score indicates less difficulty in repeating the sentences)

   Ratio of respondents' subjective rating of difficulty between Set B and Set A: 32/19 = 1.68

Table 5

Jenkins' Lingua Franca Core (Jenkins, 158-159) – Proposed description of requirements for mutual intelligibility among native and non-native speakers of English.

Phonological error involves an error in producing any of the following (not in any order of priority):

1. The consonantal inventory with the following provisos:
   - rhotic [ɹ] rather than other varieties of /r/ permissible
   - intervocalic /t/ rather than [ɾ] permissible
   - most substitutions of /θ/ /ð/ and /ɬ/ permissible
   - close approximations to core consonant sounds generally permissible
   - certain approximations not permissible (i.e. where there is a risk that they will be heard as different consonant sound from that intended)
2. Phonetic requirements:
   - aspiration following the fortis plosives /p/ /t/ and /k/
   - fortis/lenis differential effect on preceding vowel length
3. Consonant clusters:
   - initial clusters not simplified
   - medial and final clusters simplified only according to L1 rules of elision
4. Vowel Sounds:
   - maintenance of vowel length contrasts
   - L2 regional qualities permissible if consistent, but /ː/ to be preserved
5. Nuclear stress production and placement and division of speech stream into word groups

References

Caroll, J. B., Sapon, S. M. (1959). *Modern Language Aptitude Test.* New York: Psychological Corporation.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the Structure of Short-Term Memory. *Journal of Verbal Learning and Verbal Behavior,* 14, 575-589.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America,* 112, 272-284. Retrieved August 12, 2007, from http://asa.aip.org/

Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden-path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance,* 28, 218-244.

Gerritsen, J. (2001). How the Tomatis Method Accelerates Learning Foreign Languages. *The Tomatis Method.* Retrieved August 29, 2007, from http://www.tomatis.com/English/Articles/languages.htm

Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI Study of Bilingual Sentence Comprehension and Workload. *NeuroImage, 15,* 647-660. Retrieved August 24, 2007, from www.idealibrary.com

Hurford, J. R. (1991). The Evolution of the Critical Period for Language Acquisition. *Cognition, 40,* 159-201. Retrieved August 5, 2007, from http://www.sciencedirect.com/science

International Phonetic Association. (1999). *Handbook of the International Phonetic Association.* New York, NY: Cambridge University Press.

Jenkins, J. (2000). *The Phonology of English as an International Language.* Oxford: Oxford University Press.

Johnson, E. K., & Jusczyk, P. W. (2001). Word Segmentation by 8-Month-Olds: When Speech Cues Count More Than Statistics. *Journal of Memory and Language,* 44, 548-567.

Kahneman, P. (1973). *Attention and Effort.* New York: Prentice-Hall.

Kolinsky, R. (1998). Spoken Word Recognition: A Stage-processing Approach to Language Differences. *European Journal of Cognitive Psychology, 10,* 1-40.

Lundberg, I. (2002). Second language learning and reading with the additional load of dyslexia. *Annals of Dyslexia.*

McQueen, J. M. (1996). Transitional Probability and Phoneme Monitoring. *Proceedings of the 4th. Annual Conference of Spoken Language, 2,* 20-32.

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language, 39,* 21-46.

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review, 63, 81-97.

Natsume, K., Yoshida, N., Nakano, H. (2007, August). *EEG during L2 English rhythm acquisition and an English instruction robot using the rhythm.* Poster session presented at the annual meeting of the Japan Association for Language Education and Technology, Nagoya, Japan.

Poeppel, D., & Hackl, M. (2007). The architecture of speech perception. In J. Pomerantz (Ed.), *Topics in Integrative Neuroscience: From Cells to Cognition.* Cambridge University Press.

Rost, M. (1990). *Listening in Language Learning.* New York: Longman.

Rost, M. (2005). *WorldView 2 Student Book.* New York: Longman.

Staff, P. T. (2006). The talking cure. *Psychology Today.* Retrieved August 7, 2007, from http://www.psychologytoday.com/

University of Victoria Department of Linguistics. Retrieved September 1, 2007, from http://web.uvic.ca/ling/resources/ipa/handbook_downloads.htm

Weber, A. (2000). The Role of Phonotactics in Segmentation of Native and Non-native Continuous Speech. *Proceedings of SWAP: Workshop on Spoken Word Access Processes* (pp. 143-146). Nijmegen, The Netherlands: Max Plank Institute for Psycholinguistics.