Proceedings of The Institute of Acoustics

THE PERCEIVED RHYTHM OF ENGLISH AND FRENCH AS ASSESSED BY THE TAPPING TASK

D.R. Scott(1), S.D. ISARD(1) & B. de Boysson-Bardies(2)

(1) Lab. of Exp. Psychology, University of Sussex.
(2) Lab. de Psychologie, C.N.R.S., Paris.

It is a widely held view that language has a rhythmic structure, organized around regularly occurring phonetic events. Pike[1], Abercrombie[2] and many others divide all languages into two classes: syllable-timed languages, in which the regularly occuring phonetic event is the syllable, and stress-timed languages, in which the regularly occuring phonetic event is the stressed syllable. French is said to belong to the class of syllable-timed languages, and English to the class of stress-timed languages.

Measurements of recorded samples of English speech reveal that the placement of stresses is not perfectly regular; far from it, in fact. One attempt at reconciling these measurements with the subjective impression of regularity has been a hypothesis that listeners to stress-timed languages are subject to a sort of perceptual illusion. They hear the stresses as more regularly spaced than they really are. The question remains open whether the illusion resides in the speech itself, or in the competence of the native speaker, or in some interaction between the two.

The most convincing piece of evidence for this perceptual illusion is provided by Donovan and Darwin[3,4], who found that when subjects are asked to tap out the rhythm of the stressed syllables of English utterances, the rhythm of their taps is more regular than that of the stress beats of the target utterances, although they are remarkably accurate at tapping to the same rhythms presented as noise bursts in silence.

If the explanation of Donovan and Darwin's results lies either in the fact that the subjects were English speakers, or that they were listening to English speech, then we might expect different results using French stimuli, which are supposed to be syllable-timed, or French subjects, who should have a different rhythmic competence. The following experiment tests these predictions.

We constructed sets of English and French sentences that were matched for number of syllables, syntax and general meaning. These sentences are shown in Table 1.

We followed Donovan and Darwin in beginning all the stressed syllables of a given English sentence with the same stop consonant, to make it easy to explain to the subjects what they were supposed to tap to. The target phonemes are underlined in the table. There are 4 in each sentence.

PERCEIVED RHYTHM OF ENGLISH AND FRENCH

Table 1. Stimulus sentences

ENGLISH

1. the Dame Doubts that her Debts have reDoubled
2. the Beer Is always Bad in that Bar where he Boozes
3. the Turk takes his Tiger from the Tower
4. the Play will Please both Peter and Paul
5. the Prince surPrized his Page in the Park

FRENCH

1. la Dame se Doute que ses Dettes reDoublent
2. la Bière est toujours Bonne dans ce Bar où il Boit
3. le Turc Tient son Tigre près de la Tour
4. la Pièce va Plaire à Pierre et à Paul
5. le Prince surPrend son Page dans le Parc

We recorded each set of sentences, spoken by a female native speaker. The resulting utterances each contained one tone group.

Sixty-four subjects listened to ten repetitions of all sentences and were instructed to tap out the rhythm of the target phonemes after each presentation. Thirty-two of the sixty-four subjects were native speakers of British English who could neither read nor speak French. The remaining thirty-two subjects were native speakers of French and were tested in Paris. We could not get French subjects who had as little exposure to English as our English subjects had to French, but none of the French subjects could converse in English.

Half of the English subjects and half of the French heard the English sentences before the French sentences and the other half heard them the other way around. When they had finished listening to all the sentences, the subjects performed the same task in a control condition where they heard 5 sets of 4 noise bursts in silence. The rhythm of each set of noise bursts was taken from one of the sentences.

To measure the irregularity of a set of taps, we calculated

$$|\log(foot1/foot2)| + |\log(foot2/foot3)| + |\log(foot3/foot1)|$$

The value of this measure is 0 if the taps are perfectly regular, and increases as the ratios of the intervals between the taps diverge from 1. That is, a 3:1 ratio counts as more irregular than a 2:1 ratio, regardless of the actual durations of the intervals. We compared the deviation from regularity of each set of taps with that of the stimulus sentence, to decide whether the sentences had been regularized or not.

PERCEIVED RHYTHM OF ENGLISH AND FRENCH

Figure 1 shows a typical example of responses by English listeners to matched English, French and noise-burst stimuli. The top graph shows their responses to the English sentence, and the bottom graph to the matched French sentence. The solid black line shows the actual durations of the three feet of the sentence (ie. the intervals between the P's in Prince and Prize, Prize and Page and Page and Park). The dotted-line shows their responses to speech and the dashed-line represents their responses to the matched noise bursts.

Our results for English listeners tapping to English and noise bursts replicates Donovan and Darwin's finding; the rhythm of their taps is more regular than that of the stressed syllables of English target utterances. However, English listeners also tap more regularly to French sentences, and there is no significant difference between their responses to English and French. Like the listeners in Donovan and Darwin's experiments, our subjects did not regularize the rhythm of the matched noise-bursts. As the figure indicates, the slope for speech is flatter than that of the actual durations, which means that regularization has occurred. On the other hand, the slope for the noise stimuli follows the actual durations.

Figure 2 shows the responses of French listeners to the same set of stimuli. Again, this example is typical of their responses. As you can see, they respond in the same fashion as the English listeners; they too regularize both English and French utterances, and not the non-speech control stimuli. The one significant difference between English and French listeners was that the French tapped more slowly than the English when responding to speech stimuli.

So it appears that the phenomenon of perceptual isochrony which Donovan and Darwin describe is not specific to English, nor to English listeners. What is not clear is why listeners should regularize their taps to speech stimuli, but not to the noise stimuli. One possible explanation lies in the differing complexity of the two tasks: tapping to speech is a more difficult task than tapping to noise-bursts. The speech stimuli are acoustically more complex than the corresponding sequences of noise-bursts, and the memory load in remembering a sentence is greater than that in remembering four noise-bursts.

We have tested this hypothesis in a second experiment where nine English listeners tapped to the English utterances and matched noise-bursts used in the previous experiment and to a third set of matched stimuli which were not speech but which were similiar to speech in terms of acoustic complexity. This new set of stimuli was constructed by distorting the English utterances in such a way that the target phonemes remained undisturbed, but the spectral envelope of the intervening unstressed syllables was averaged over a 200 millisecond window. The end result of this distortion gives

the impression of slurred speech heard under water, and is unintelligible.

In this experiment, listeners regularized both the intact speech stimuli and the distorted speech but not the noise-bursts. Figure 3 shows listeners' responses to three versions of the same English sentence shown before. Remember, that it is the shape of the functions that is important.

So, the phenomenon of regularization is not even specific to speech, but extends to other unintelligible noises with some speech-like properties. A hypothesis that must be strongly considered is that it is simply a response bias that occurs when the task gets too difficult.

We think it would be unwise to simply jettison the strong intuition that English and French have different sorts of rhythmic organization on the basis of results like these. However, the tapping task has shown itself too insensitive to "tap" the source of this difference, whatever it may be.

Acknowledgement
This work was supported by a grant from the European Science Foundation.

REFERENCES

[1] K.L. Pike, The Intonation of American English. Ann Arbor: University of Michigan Press, 1945.

[2] D. Abercrombie, 'Syllable quantity and enclitics in English', In: In Honour of Daniel Jones, Abercrombie, Scott & Trim (eds.), London:Longmans, 216-222, 1964.

[3] A. Donovan and C.J. Darwin, 'The perceived rhythm of speech', Proceedings of the 9tn International Congress of Phonetic Sciences, Copenhagen, 1979.

[4] C.J. Darwin and A. Donovan, 'Perceptual studies of speech rhythm: isochrony and intonation.' Proceedings of Nato ASI on Spoken Language Generation and Understanding, J.C. Simon (ed.), Dortrect:Reidel, 1980.

PERCEIVED RHYTHM OF ENGLISH AND FRENCH

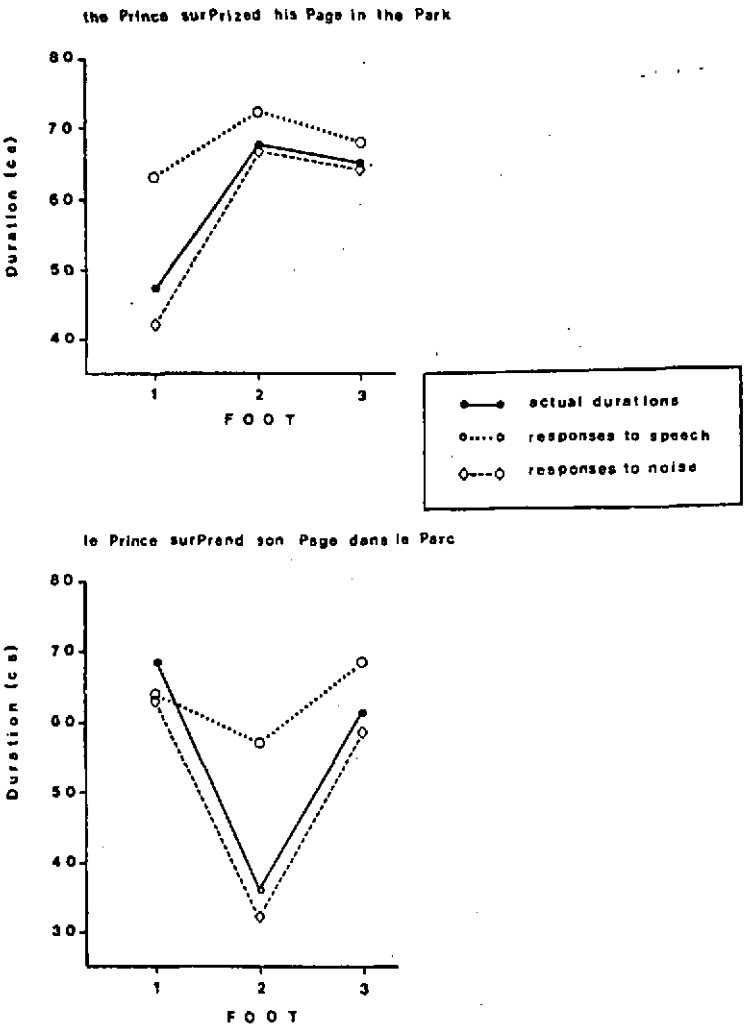

Figure 1

PERCEIVED RHYTHMS OF ENGLISH AND FRENCH

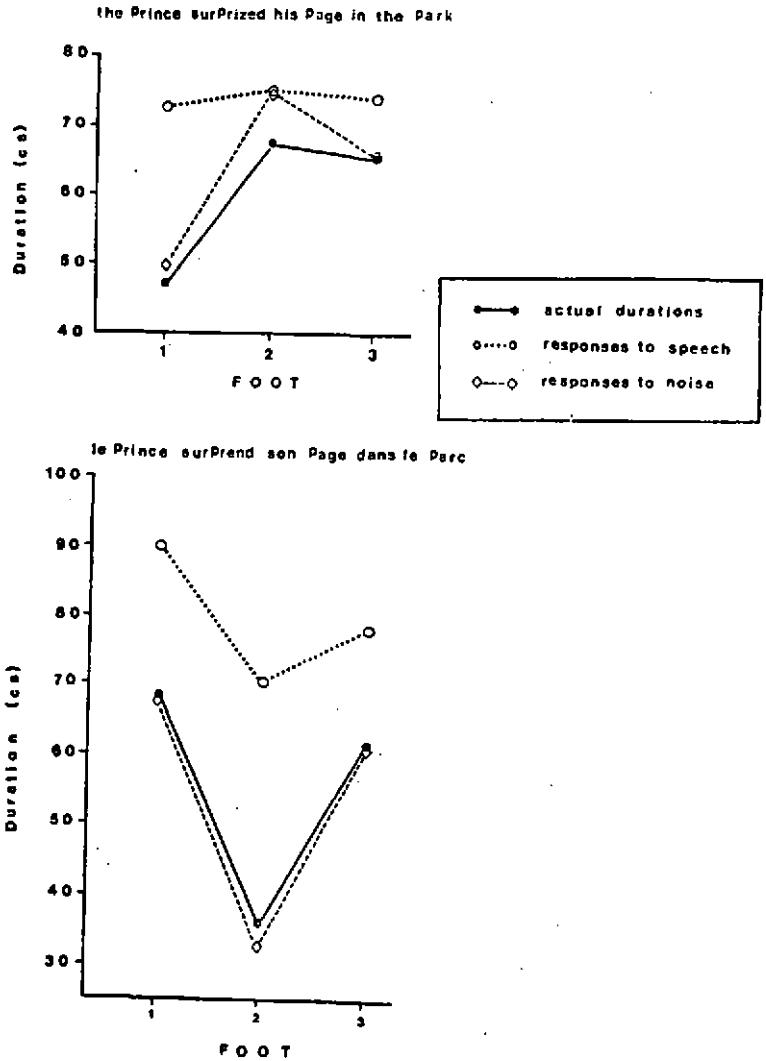

Figure 2

PERCEIVED RHYTYMS OF ENGLISH AND FRENCH

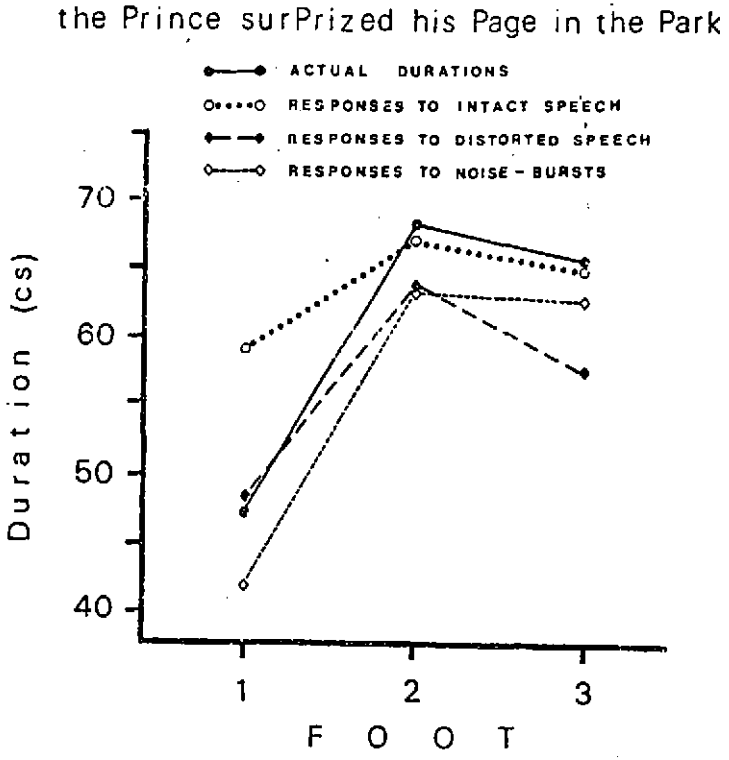

Figure 3