The Irish Journal of Education 1973 vii, 1 pp 40 52

AN ANALYTIC COMPARISON OF LISTENING IN TWO LANGUAGES

JUDITH A SCHWARTZ, JOEL SINGER and JOHN MACNAMARA*
McGill University

The skills of bilinguals in listening to speech were analysed with a view to revealing difficulties related to listening to a language which has not been fully mastered Only materials which were known to all subjects who participated were employed The main conclusions are (a) that the meaning of words was determined more slowly in the weaker language, (b) that syntax was interpreted more slowly in the weaker language, (c) that subjects were more inclined to fall behind in interpreting speech in the weaker language

The studies here reported, like those reported in Macnamara, Feltin, Hew and Klein (2), are inspired by the desire to understand the problems of students who follow courses in a language which they have not yet mastered The earlier studies examined the reading skills of such students, the present ones examine their listening skills When students listen to a teacher who is speaking a language with which they are not familiar they may, of course, be tripped up by words, idioms and syntactic structures which they simply do not know There is scarcely any need to demonstrate that such lacunae constitute a problem We have limited our attention to words and structures which our subjects knew, and explored the speed with which they interpreted them Most people who have had the experience of listening to a foreign language and found that their interpretation lagged behind the speaker's words will understand what we are about We want to explore some of the factors that, as it were, slow down our ears so that we miss parts of what is said There is no need to stress the importance of listening in education, but in addition we were curious to compare listening and reading, because in reading the student sets his own pace, whereas in listening the pace is set for him by the speaker

We analysed listening to speech into several components and we contrasted performance in the second language with performance in the mother tongue We selected students whom we knew to be weak in the

*The work reported in this paper is based upon Judith A Schwartz s M Sc thesis It was funded by a Canada Council grant to John Macnamara Requests for reprints should be directed to John Macnamara, Department of Psychology, McGill University Montreal 110 P Q , Canada

second language so as to magnify the effects we were looking for Psychology, unfortunately, knows of no other microscope We are then more concerned with the existence of effects than with their magnitude We do not suggest that all bilinguals are weaker in the second language by the quantities we set out below, we suggest, rather, that where a bilingual has difficulty in following a speech in his weaker language, the factors to whose existence our data point are likely to be at work

METHOD

Subjects

Twenty-four English-speaking university students 10 men and 14 women, and 24 French-speaking university students, 11 men and 13 women, acted as *S*s and were paid for doing so It was important to include *S*s from both linguistic groups to check whether effects found with one group might be due to some bias associated with the particular linguistic materials If the results for one linguistic group form the mirror image of those for the other one, such bias can reasonably be ruled out

Each *S* was a native-speaker of one of the two languages, had spoken only that language in his home, and had been educated exclusively in it Each *S* had also been taught the other language at least throughout high school, and in addition he was tested for knowledge of all the words in that language which were employed in the experiment *S*s who did not know three or less words were rehearsed in them (a minimum of four times) until they were familiar with them Each *S* participated in all parts of the experiment

Materials

Materials vary in a rather complicated manner To begin with, twenty-eight French words and their twenty-eight English counterparts were selected all naming common objects or actions of which pictures could easily be drawn The idea was to present a picture and a word together and measure the time *S*s took to decide whether or not the word described the picture Twenty-two of the French words were feminine nouns and six were verbs The English words were flower, hand, clock, cup, lamp, wheel, leg leaf cow house mouth, bicycle, light bulb, door, apple, car, ball, hen, girl tail, mouse, wing, sing, read, paint, drink, write, throw The French ones were *fleur, main, horloge, tasse lampe, roue, jambe, feuille vache, maison, bouche bicyclette,*

ampoule, porte, pomme, auto, balle, poule, fille, queue, souris, aile, chanter lire, peindre, boire, ecrire, lancer Verbs were represented pictorially by stick figures performing some easily identifiable action, nouns were represented by the objects they named Two filmstrips were made each containing all the pictures in different randomly determined orders Four tape recordings of the words were made, two in English and two corresponding ones in French One English and one French tape recording was used with one of the filmstrips only, the other two tape recordings were for use with the other filmstrip In all pairings, half the words matched the pictures and half did not The order of matches and mismatches was randomized

In order to examine syntactic effects present-tense sentences were composed from the same words Four new words were added in each language—'a', 'can', 'has', 'not'—so as to increase the possibilities for composing sentences From the logical point of view each sentence was a universal, e g , 'A cow has a wing', 'A hen doesn't have a tail' Half the sentences were positive and half were negative, and within each group, half were true and half were false *S*'s task was to say whether the sentence was true or false

These sentences, as presented, formed two distinct groups In one, all sentences (half true and half false) were of the same syntax (henceforth, fixed-syntax sentences) In the other, syntax varied from one sentence to the next (varied-syntax sentences) The reason for this was that the words-and-pictures task did not form a good basis for comparison with tasks which involved the interpretation of sentences It seems possible that the words-and pictures task, at least in a *S*'s weaker language, constituted a recognition task rather than a recall one The picture might have given him an extra cue to the meaning of the word which accompanied it For this reason the words-and-pictures task was treated largely as separate from the rest of the experiment in which *S* certainly had to recall the meaning of the words he heard In order to set up a condition in which *S*s would be required to recall the meanings of words without being required to analyze the syntax, the fixed-syntax sentences were prepared *S*s were well prepared for them by a set of examples, and so we expected them to ignore the syntax which was quite superfluous In other words, the task they were set was to interpret the sentence's main lexical items and compare them for a predetermined relationship In responding to the varied syntax sentences *S*s had to analyze the syntax Comparisons between response times to

these sentences and those with fixed syntax was expected to reveal the effect of syntactic processing

There were six fixed-syntax sentences, three true and three false All six were positive and all involved a part whole relationship, e g , 'A cow has a tail ' These were preceded by eight practice sentences of a similar type, but none of the twenty-eight key words were used in them There were eight varied-syntax sentences in each language, four true and four false Among the sentences of each truth value, two were positive and two were negative The varied syntax sentences of each language were preceded by four practice sentences in which none of the twenty eight key words were employed The sentences were recorded in four separate blocks fixed-syntax English, fixed-syntax French, varied-syntax English, varied syntax French There was a ten-second pause between each pair of sentences in a block

We come now to the last set of sentences which were constructed to examine the effects of longer contexts on the perception of single sentences It seemed likely that the effects of context might be twofold On the one hand a supporting context might facilitate perception of the 'test' sentence because it might prime *S* for the test sentence On the other hand, in as much as the attempt to perceive each sentence in the series which formed the context might overload the perceptual system, the context might cause *S* to drop behind and so come poorly prepared to the test sentence Let us call these two the 'priming effect' and the 'overloading effect'

Our plan for testing these effects was to construct two different sorts of paragraphs, one which matched the test sentence and one which did not In constructing paragraphs several common words were used together with some of the key words All the sentences in a particular paragraph were related in that all referred to the same topic For example, a paragraph about a cow would have the word 'cow' as subject or object in each of its sentences A test sentence can now be defined as being in context if its subject is the same as the key word in its paragraph (in the previous example, 'cow'), otherwise it is out of context The number of sentences in a paragraph varied from five to seven, so that *S* could not predict which was the test sentence Test sentences were signalled by a 0 25 sec tone which sounded immediately at the end of the sentence Let us call test sentences which were in context, 'matching' sentences, and ones which were out of context, 'non matching' sentences Of each type there were four sentences in

each language, two true and two false, one of each pair being negative and one being positive All were in the present tense and were similar to the varied-syntax sentences

This is perhaps the point at which to note that the paragraphs, including the test sentences, were not the same in meaning across languages This is in contrast to all other sections of the experiment in which English and French stimuli were *equivalent in meaning* The reason for changing the practice in connection with the paragraphs was to reduce learning effects We suspected that whole paragraphs were much more memorable than isolated sentences or isolated words

To study the effects of priming and overloading we had to be able to compare performance on test sentences (after paragraphs) with that on isolated sentences of varied syntax This raised a problem Our varied-syntax sentences were presented without signal tones, presumably the task of responding to them differed considerably from that of responding to sentences marked with such signals At least, *Ss* knew as they heard the varied-syntax sentences that they had to respond to that particular sentence, whereas in listening to one of the sentences of a paragraph they knew only that they might have to respond to that one To overcome this difficulty we composed eight extra varied-syntax sentences in each language and placed signal tones at the end of each These sentences we mixed at random with the paragraph tasks so that *Ss* could not be sure which condition they were about to encounter, a varied-syntax sentence a test sentence in context, or a test sentence out of context

The sentences on which *Ss* were tested varied in length from five to seven syllables English and French counterparts never differed in length by more than one syllable Perhaps, after so long a description, it will clarify the various conditions if we list them

i words and pictures,
ii fixed-syntax sentences (no signal tones),
iii varied-syntax sentences (no signal tones),
iv test sentences in context (with signal tones),
v test sentences out of context (with signal tones),
vi varied syntax sentences (with signal tones)

Procedure

The experiment was divided into eight sections, four in each language Tasks 1, 2 and 3 in the above list constituted three of the sections the

fourth consisted of tasks 4, 5 and 6 taken in combination Instructions corresponding to each section were read aloud while *S* read them silently, just before a section was presented The instructions were presented in *S* s native language

In task 1 the pictures were presented by means of a Dukane Projector (Model 576 47B) which completed the operation of presenting a stimulus in one-twentieth of a second The words were presented binaurally through earphones The tape-recorded word and the filmstrip were automatically synchronized by means of a diapilot An impulse had been placed on the second channel of the tape to coincide with the beginning of a word The diapilot picked up the impulse and simultaneously moved the projector forward and started a clock *S* compared word and picture and responded by pressing one of two keys, one marked (+) if they matched, the other marked (—) if they did not Either key stopped the clock The position of keys was varied across *Ss* for half, the (+) key was under the right index finger, and for the other half it was under the left one Reaction times for each stimulus were recorded for each *S* The order of presentation of filmstrips and languages was completely counterbalanced across *Ss*, though the words-and-pictures task came first for all *Ss*

In responding to all other stimuli, *Ss* were instructed to press the (+) key if the sentence was true, the (—) key if it was false An impulse at the end of each sentence was picked up by a diapilot which started a clock which was stopped when *S* pressed a response key Reaction times were recorded for each sentence

There were three tasks in each of two languages which means that there were 720 possible orders of tasks A representative sample of twenty-four orders were chosen one for each *S* Each of the six tasks was presented four times in each of the six possible serial positions This ought to have reduced to negligible proportions any undesirable order effects The position of the two response keys was varied as described in connection with task 1 Throughout the experiment a ten second delay was allowed between presentation of consecutive stimuli to allow *E* to record reaction time

Six measures were obtained for each *S* in each language Each measure was the *S*'s average for performance on a particular task

Before the experiment began a polygraph printout of each tape was made The polygraph showed the relative positions of the stimulus materials and the impulse which started the clock The object was to

determine whether there was a systematic bias in favour of one language A set of criteria for determining the end of speech was established which for two independent judges yielded consistent judgements of the space between the end of speech and the beginning of the impulse The measures thus obtained revealed no appreciable differences across languages associated with the individual words of task 1 There were, however, differences across languages between sentences Each *S*'s response times for individual sentences were adjusted for such differences

RESULTS

A preliminary inspection of the data revealed numerous outlying scores, due perhaps to inattention or temporary confusion in *S* To lessen the effect of such scores, the natural logarithm transformation was applied to all scores based on correct responses Times associated with erroneous responses were eliminated, as were the times for the corresponding sentence in the other language or in the other condition with which comparisons were to be made However, times which were omitted in one set of comparisons were included in others where there was no danger of resulting bias This enabled us to make the maximum use of the data It does, however, make the presentation of the data confusing, so we will report only means associated with the principal interaction times, which after all are the only quantities of interest in the interpretation of results

In analysing the data the method of paired differences was used throughout *In* all twelve cross-language comparisons (six for each linguistic group of *Ss*) the means differed significantly in the expected direction However, with one exception we are not directly concerned with these comparisons The exception is task 1, words and pictures, in which *Ss* consistently responded more rapidly to words in their mother tongue (L1) than to words in their weaker one (L2) This suggests that even when *Ss* knew a word and its meaning, they arrived

TABLE 1

MEANS OF TRANSFORMED MEASURES FOR WORDS-AND-PICTURES TASK TOGETHER WITH RESULTS OF *T*-TESTS

	English	French	*t*	*df*	*p*
English *Ss*	—0 08	0 06	4 29	23	0 01
French *Ss*	—0 06	—0 14	2 11	23	0 025

at that meaning more quickly if the word was in L1 than if it was in L2 The fact that the difference in response times was in the opposite direction for the French and English groups indicates that the finding was not an artifact of the stimulus material

For the rest the main interest centres on the interactions between conditions and languages The relevant figures are set out in Tables 2 and 3

TABLE 2

MEANS OF TRANSFORMED MEASURES ASSOCIATED WITH THE PRINCIPAL INTERACTIONS, TOGETHER WITH RESULTS OF *T*-TESTS (ENGLISH *Ss*)

Tasks		English	French	*t*	*df*	*p*
A	Fixed-syntax sentences	—0 38	0 05	1 37	23	NS
	Varied-syntax sentences (without tones)	—0 25	0 04			
B	Sentences in context	—0 15	0 28	1 10	23	NS
	Sentences out of context	0 11	0 48			
C	Sentences out of context	0 14	0 47	2 52	23	0 01
	Varied syntax sentences (with tones)	—0 03	0 11			

TABLE 3

MEANS OF TRANSFORMED MEASURES ASSOCIATED WITH THE PRINCIPAL INTERACTIONS, TOGETHER WITH RESULTS OF *T*-TESTS (FRENCH *Ss*)

Tasks		English	French	*t*	*df*	*p*
A	Fixed syntax sentences	—0 26	—0 45	0 66	23	NS
	Varied-syntax sentences (without tones)	—0 05	—0 19			
B	Sentences in context	—0 04	—0 13	2 95	23	0 01
	Sentences out of context	0 27	—0 12			
C	Sentences out of context	0 28	—0 03	2 68	23	0 01
	Varied-syntax sentences (with tones)	—0 03	—0 12			

A significant interaction shows that when allowance was made for a difference in response times across languages in one task, there still

remained an additional difference across languages in the next task This enables us, for example to focus attention on speed of processing syntax when allowance was made for speed in interpreting lexical items Incidentally, the *t*-test which compares two differences is equivalent to the interaction term in an analysis of variance

In estimating significance a one-tailed test was used throughout The justification for doing so comes from the unidirectional hypotheses on which the study was based These hypotheses had the support of the parallel work on reading by Macnamara *et al* (2)

To help refine and understand the results a simple linear model was devised While time may show that a linear model is unsatisfactory in the analysis of such data, we felt that it was fully justified in the present circumstances Our study is largely exploratory and the measures contain a large error component Furthermore, the model only spells out in explicit form the linear function which underlies all *t*-tests and analyses of variance in other people's work

The model attempts to isolate five components which are likely to be important in the comprehension of speech The components might be described as (a) perception and understanding of single words, (b) perception and understanding of words in sequence, (c) interpretation of syntax (d) ability to use previously heard material to facilitate the interpretation of a sentence (priming effect), and (e) processing lag due to the encumbrance of having to interpret preceding material (overloading effect) The various response times contain these components in the manner displayed in Table 4

TABLE 4

COMPONENTS OF BILINGUAL LISTENING COMPREHENSION

Conditions	Components
1 Words and pictures	(a)
2 Fixed-syntax sentences	(b)
3 Varied syntax sentences	(b) + (c)
4 Sentences in context	(b) + (c) — (d) + (e)
5 Sentences out of context	(b) + (c) + (e)

(a) = perception and understanding of single words
(b) = perception and understanding of strings of words
(c) = interpretation of syntax in single sentences
(d) = ability to use previously heard material to facilitate the interpretation of sentences (priming effect)
(e) = processing lag due to the encumbrance of having to interpret preceding sentences (overloading effect)

Notice that component (a) is not contained in conditions 2-5. This is the component which we have already seen to be significant. For reasons explained in the methods section it does not feature in the later tasks. Note also that component (d) is entered as a negative quantity because it contributes to *S*'s speed of listening and reduces his response times. Finally, there are only five components even though there are six tasks. The reason is that two of the tasks, varied-syntax sentences with and without tones, are linguistically equivalent, and so do not yield different components for linguistic processing time. The value of each component in each language can be arrived at by simple arithmetic and interactions consist in differences between components across languages.

The results of statistical analyses are given in Tables 2 and 3. In the analysis relating to component (c), the interpretation of syntax, only positive varied-syntax sentences were compared with fixed-syntax sentences. All fixed-syntax sentences were positive, and it is known that response times to negative sentences tend to be slower than those positive ones. It was necessary, then, to exclude negative fixed-syntax sentences from the analysis.

The interaction of (e) components across languages was not significant either for English or for French *S*s. This was contrary to predictions, but further discussion will be left to the appropriate section. The interaction of (e) components across languages (overloading effect) was significant for both linguistic groups. In both cases the difference was in the expected direction; overloading caused a greater lag in L2 than in L1. In other words *S*s had greater difficulty in keeping up with the speaker in L2 than in L1.

The interaction of (d) components across languages (priming effect) was not significant for English *S*s. This was contrary to expectation in that *S*s did not make more use of context in L1 than in L2 to anticipate what was coming in the test sentences. On the other hand, for French *S*s the interaction was significant in the wrong direction. At first sight this suggests that French *S*s made more use of context in L2 than in their mother tongue, but we will return to this finding in the next section.

DISCUSSION

The basis for interpreting the results, it must be remembered, is that all *S*s, both French and English, understood the meanings of all the

words and the uses of all syntactic constructions found throughout the experiment On this assumption, *Ss*' difficulties must be accounted for on grounds other than ignorance The first major finding is that *Ss* seemed to interpret the meanings of single words more quickly in L1 than in L2 The first question to be answered about this finding is whether the observed difference was due, in part at least, to a difference in perceptual speed between the two languages Macnamara *et al* (2), in a study of reading skill, observed no significant difference between L1 and L2 in the perceptual threshold for words They suggest, as a possible reason, that *Ss* were highly familiar with the materials because of the repeated use of a small set of stimuli In the present study, also, *Ss* were repeatedly tested with the same small set of stimuli and, therefore, they were probably very familiar with the materials This is not to say that perception of words in general is not significantly faster in L1 than in L2 Crothers and Suppes (1) have shown that a perceptual difference can exist, both in the visual and in the auditory modes, between English and Russian In their experiment, *Ss* were asked to press a key corresponding to what they thought was the graphemic representation of a syllable heard from a tape-recorder Results indicated that, even after many trials, *Ss*' latencies were greater for Russian items than for English ones Since their experiment involved a recognition task, it is very similar to the task in this experiment of matching words and pictures Therefore, it does seem likely that a corresponding difference would exist in the auditory perception of English and French words However, it is clear from other studies (2), that in visual tasks, *Ss* do process the meanings of words faster in L1 than in L2 It is thus highly likely that in this auditory task, *Ss* processed the meanings of isolated words faster in L1 than in L2 It is hoped that a further study aimed specifically at distinguishing the auditory perception of words from the determination of their meanings will settle whether or not the results of this task are partly due to a perceptual factor

Turn now to the fixed-syntax sentences Although we have not reported it in the results section, the interaction between the words-and-pictures task (a) and fixed-syntax sentences (b) was significant for both groups The differences were in the expected direction On the assumption that the significant result for words-and-pictures arises from the fact that *Ss* take longer to determine the meaning of isolated words in L2 than in L1, it follows that the interaction across languages between the two

tasks is probably due to some other factor. The most likely factor is the time taken to process syntax. This interpretation would help to explain the absence of a significant cross-language interaction between fixed-syntax sentences and varied-syntax ones. Component (b) already contains the syntactic factor. This conclusion suggests that we failed to get *S*s to ignore syntax in the fixed-syntax sentences.

Component (e), the overloading effect, was arrived at by subtracting times for varied-syntax sentences (with tones) from those for sentences out of context. It seems highly plausible that all components associated with varied-syntax sentences without tones should also be found in varied-syntax sentences with tones. All varied-syntax sentences were composed together and randomly assigned to the different conditions, subject to the restraints relating to truth value and affirmative/negative aspect which are explained in the materials section. The interaction across languages relating to component (e) was significant for both linguistic groups. This is strong support for the over-loading effect. That is, when due allowance is made for the time taken to determine the meaning of isolated words and also for the time taken to interpret the syntax of isolated sentences, there remains a further factor which differentiates the time for processing sequences of sentences in the two languages. In other words there is a cumulative effect related to sequences of sentences; *S*s tend to become overloaded more readily in L2 than in their mother tongue.

The final analysis relates to component (d), the priming effect, which as we have seen did not meet with our expectations. In fact the results for English *S*s tended to go in the unexpected direction, while those for French *S*s were significant in the unexpected direction. Our interpretation is not that *S*s are generally able to make better use of context in L2 than in their mother tongue, but rather that our experimental conditions were not sensitive enough to reveal normal functioning. It is necessary to remember that every sentence in each paragraph could be regarded as a complete unit in itself. This limits the role of context in the experiment. Results might have been better if knowledge of the truth or falsity of the test sentence was dependent upon the preceding material in the paragraph. In L1 *S*s must have had no difficulty in keeping up with the speaker, and, therefore, context was not so important. In L2, it was possible for *S*s to sacrifice meaning in order to keep up with the speaker and suffer no ill effects in interpreting the test sentence. It seems, however, that French *S*s adopted a strategy in

L2 which they did not employ in L1 One such strategy would be to pick out the key word in a paragraph and use it as a clue to the topic of the test sentence This would have reduced their task with the test sentence in L2 to a search for the syntax and for the final word In L1 perhaps such a strategy was not felt to be necessary Moreover, English *Ss* may have adopted the same strategy, but not applied it quite so effectively Indeed there are several suggestions in the data that English *Ss* had a relatively poorer grasp of French than French ones had of English This would explain why the results for English *Ss* were similar in trend to those for French ones, but did not reach significance

The results then are best interpreted as follows Our *Ss* when working in L2 were less efficient than when working in L1 at interpreting single words and at interpreting the syntax of isolated sentences, while they were more susceptible to the listening lag due to overloading What is more, our analysis suggests that these effects were cumulative, their total effect being greater than that of the individual difficulties Thus we may expect that students who experience difficulties when working in a second language may have defects more hidden and more all pervasive than ignorance of particular lexical items and of syntactic structures At the very least, this should constitute for their teachers a warning to be slow to attribute such students' failures and incomprehension to stupidity Unfortunately we are not at the present time in a position to suggest measures to overcome these difficulties It does not seem in principle to be beyond our ingenuity to invent and test out such measures and in the interest of millions of students who might thereby benefit we hope that the research will not be long delayed

REFERENCES

1 CROTHERS E, and SUPPES P *Experiments in second language learning* New York Academic Press, 1967

2 MACNAMARA J, FELTIN M, HEW, M, and KLEIN, M An analytic comparison of reading in two languages *Irish Journal of Education* 1968 2 41 53