Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia, July 6-9, 2004

PERCEPTION OF UNATTENDED SPEECH

Marie Rivenez[1,2], Chris Darwin[1], Anne Guillaume[2]

[1]Department of Psychology
University of Sussex
Falmer, Brighton, BN1 9QG, UK
[2]Département Sciences Cognitives
Institut de Médecine Aérospatiale du Service de Santé des Armées
BP 73 91223 Brétigny sur Orge Cedex France

m.rivenez@sussex.ac.uk
cjd@sussex.ac.uk
aguillaume@imassa.fr

ABSTRACT

This study addresses the question of speech processing under unattended conditions. Dupoux et al. (2003) have recently claimed that unattended words are not lexically processed. We test their conclusion with a different paradigm : participants had to detect a target word belonging to a specific category presented in a rapid list of words, in the attended ear. In the unattended ear, concatenated sentences were presented, some containing a repetition prime presented just before the target words. We found a significant priming effect of 22 ms (Experiment 1), for category detection in the presence of a prime compared with no prime. This priming effect was not affected by whether the right or the left ear received the prime (Experiment 2a and 2b). We also found that the priming effect disappeared when there was no pitch range difference between attended and unattended messages (Experiment 3 and 4). Finally, we replicated the priming effect by compelling participants to focus on the attended message asking them to perform a second task (Experiment 5).

1. INTRODUCTION

In auditory display applications, it is common for there to be more than one sound source present at a time. What information do listeners extract from a sound source that they are not attending to? The aim of this study is to investigate whether we process more than one speaker at a time. We ask the question of the human capacity to process spoken words while focusing on another sound source. Most studies agree that simple attributes of unattended speech are processed [1][2][3][4] but there is less unanimity over whether unattended speech can be processed at the lexical level. There is a large amount of data consistent with the idea that unattended speech cannot be lexically processed without attention and a similar number of studies showing that they can (see [5] for a review; and [6] for a more recent review). The usual paradigm to assess this question consists in presenting to the participant two messages at the same time, usually one in each ear. Participants are asked to focus on one ear by doing a specific task (e.g. shadowing task) and ignore the other ear. Experimenters look at the influence of the unattended message on performance on the attended message. All the studies which have shown lexical processing of the unattended message have used isolated words in the unattended message. However in such condition, participants attention could have been attracted by the unattended isolated words. Consequently, the nominally unattended message may be actually attended. Holender (1986) claimed that in order to control these attentional switches, it was very important to use continuous speech in the unattended message. He points out that all the studies that did use continuous speech showed no lexical processing of unattended messages. On the other hand, these studies have always used an indirect measurement of lexical activation.

To our knowledge there is only one recent study (Dupoux, Kouider and Mehler, 2003) which deals with these two points (continuous speech and direct measurement of lexical activation). The authors asked participants to perform a lexical decision task on a target word (they had to say whether the word was a word or a non word) which was presented to the right ear (*attended message*). At the same time, they presented a repetition of the word (a prime) in the left ear (*unattended message*) (Fig. 1a). The hypothesis was that if there is a lexical processing of unattended messages, then participants should react faster when the target word is primed (the same word as the target is presented in the unattended ear) than when it is unprimed (an unrelated word is presented in the unattended ear). The authors found a significant priming effect (participants reacted faster when the target word was primed than when it was unprimed) but participants were conscious of hearing the prime. In order to avoid these attentional switches toward the prime, they embedded the prime in a carrier sentence (continuous speech) (Fig. 1b). In this condition, they found no significant priming effect. Moreover, the participants were unaware of the identity of the prime.

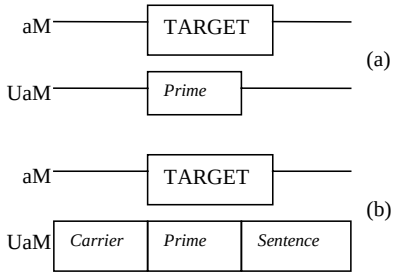

Figure 1: *Experimental design used in Dupoux, Kouider & Mehler (2003). aM refers to the attended message and UaM to the unattended message. Panel (a) shows the design with an isolated word in the unattended message and panel (b) shows the design with continuous speech (carrier sentence) in the unattended message.*

The Dupoux et al. (2003) paradigm is open to three criticisms: First the authors temporally compressed the prime in order to decrease its saliency and reduce the number of attentional switches toward the unattended ear. Such compression may lead to difficulties in the perceptual organization of the prime when it is embedded in a carrier sentence because words boundaries may be harder to detect. Second, Dupoux et al. (2003) always presented the prime to the left ear which is mainly connected to the right hemisphere, whereas the left hemisphere is more specialized in linguistic processing. And third, the authors used the same voice for the attended and unattended messages, which may make the perceptual organization of the prime harder.

The purpose of this experiment is to clarify the lexical processing of unattended speech using the same direct measurement of lexical activation as the one used in Dupoux et al. (2003) study. In Experiment 1, we attempt to test unattended speech processing using continuous and non degraded speech in the unattended message. In Experiment 2, we assess the effect of the ear in which was presented the unattended message on unattended speech processing. In Experiment 3 and 4, we investigate the effect of the pitch range difference between attended and unattended message. Finally in Experiment 5, we do a control experiment to prove our experimental design and check that participants are not paying attention to the unattended message when they do the task.

2. GENERAL METHOD

We presented a rapid list of words in the attended message in one ear. Participants had to detect a word belonging to a specific category. In the other ear, we presented the unattended message which was a series of four nonsense sentences. One of the sentences contained a prime which was a repetition of the word to detect in the attended message. The end of the prime was always synchronized to the start of the target. The hypothesis was that is there is a lexical processing of unattended words, then participants should be faster to detect the target word when it is preceded by a related prime than when it is preceded by an unrelated prime in the unattended message.

2.1. Apparatus and stimuli

The attended messages were a list of monosyllabic words presented at the rate of 2 words per second. All the words were low frequency words (average Kucera-Francis = 20 [7]) in order to make the task impossible if participants made attentional switches toward the unattended message. The target words were nouns belonging to specific categories, chosen from the English data base build by Hampton & Gardiner (1983) [8]. We used 11 different categories: Bird, Clothing, Fish, Food flavoring, Fruit, Furniture, Insect, Sport, Vegetable, Vehicle and Weapon chosen from the same data base. The attended message was recorded by a male speaker (C.J.D.), who spoke the lists of words with as flat a pitch contour as possible, producing an average fundamental frequency of 140Hz. The attended message was always presented at a level of 72dB.

The unattended messages were nonsense sentences in order the make the meaning unpredictable in case participants made occasional attentional switches toward the unattended message. They were spoken with a normal prosody by the same talker as produced the attended lists. They started about 800 ms after the attended message and were reduced in level by 12dB relative to the attended message in order to help listeners direct their attention appropriately. The prime was a repetition of the target word occurring as part of a nonsense sentence. The end of the prime was synchronized to the start of the target. Because the target was pronounced as a monotonous list item whereas the prime was part of a fluent, normally intonated sentence, the two were physically very different.

2.2. Design and procedure

Participants were told to listen to one ear (left or right, depending on the experiment) and detect nouns belonging to specific categories by pressing a button. They were told the categories they were going to listen to and examples were given (‘for instance, if you have the category Bird, then you will have to press the button if you hear Blackbird’). They were also told to ignore the right ear. When the experiment started, they were presented a name of category in the middle of the screen. After 500 ms, the trial started and the category name remained on the screen until the end of the trial. The next trial started automatically after the end of the preceding one. Participants had first a practice session composed of 12 targets (9 trials) and lasting about 2 minutes. Test session followed the practice session and was composed of 60 targets (39 trials) and lasted about 7 minutes. Primed and unprimed trials were randomized across trials and participants. The different categories were also randomized across trials and participants. The protocol was programmed using the PsyScope software.

After the priming experiment, participants were asked if they had heard anything particular in the right ear (unattended sentences) and then, they were specifically asked if they had noticed the prime.

3. EXPERIMENT 1

The purpose of Experiment 1 was to investigate the lexical processing of unattended speech using continuous and non degraded speech in the unattended message.

3.1. Participants

43 university students were paid for their participation. They were English native speakers with no hearing, language or attentional impairment. They were not selected for handedness.

3.2. Design and procedure

We used the general method described above. The attended messages were always presented to the left ear and the unattended messages to the right ear. The pitch of the attended messages was changed with the PSOLA algorithm to a constant fundamental frequency of 180Hz. The unattended messages were constructed from the original voice and were not further

processed. Half of the trials were primed and half were unprimed (an unrelated word to the target in the unattended message preceded the target word). We counterbalanced the primed and unprimed trials in 2 lists of items.

3.3. Results and discussion

Fig. 2 shows the average data for the primed and unprimed trials. We found that participants react faster when the target is preceded by a prime than when it is not. This priming effect was 22 ms and it was significant ($F(1,41)=13.06$; $p<.0008$; ($t(60)=4,53$; $p<0.00005$). The percentage of correct detections was equivalent in both conditions of priming (73% of correct detection). None of the participants were conscious of the prime.

This priming effect shows that unattended speech processing is possible using non-degraded, continuous speech in the unattended message and presenting the unattended message to the right ear.

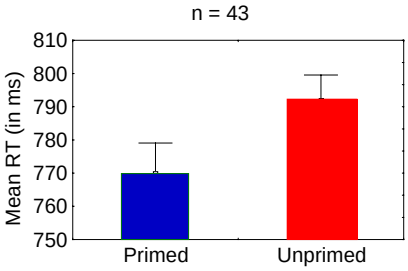

Figure 2: *Mean reaction times (in ms) for the primed (blue column) and unprimed (red column) trials averages across the 60 trials and the 43 participants (Experiment 1). Errors bars shows the standard error across participants.*

4. EXPERIMENT 2

The purpose of experiment 2 was to evaluate the effect of the ear in which the unattended message was presented.

4.1. Participants

48 university students were paid for their participation. In Experiment 2a, 39 were right handed and 9 were left handed, whereas in Experiment 2b, the 48 participants were right handed. They all had a normal audiogram for both ears. They were English native speakers with no hearing, language or attentional impairment.

4.2. Design and procedure

We used the same procedure as in Experiment 1 except that for half of the trials, the attended message was presented to the left ear and the unattended message to the right and for the other half, the attended message was presented to the right ear and the unattended message to the left. The primed and unprimed trials were counterbalanced in two lists. The right and left trials were also counterbalanced in two lists. Priming and ear were within-subject factors.

In Experiment 2a, the participants were not selected for handedness, as in the Dupoux et al. (2003) study. In Experiment 2b, we used only right handed participants in order to assess more accurately possible hemispheric asymmetries.

4.3. Results

4.3.1. Experiment 2a

Fig. 3 shows the average data for the primed and unprimed trials when the prime is presented to the right or to the left ear and when participants are right and left handed (Panel a). When the prime is presented to right ear (as in Experiment 1), we found a significant priming effect of 24 ms ($F_1(1,40)=12.01$; $p<.001$). When the prime is presented to left ear, we also found a significant priming effect of 22 ms ($F_1(1,40)=7.32$; $p<.01$). There was no significant interaction between the priming factor and the ear in which the prime was presented ($F_1<1$; $F_2<1$).

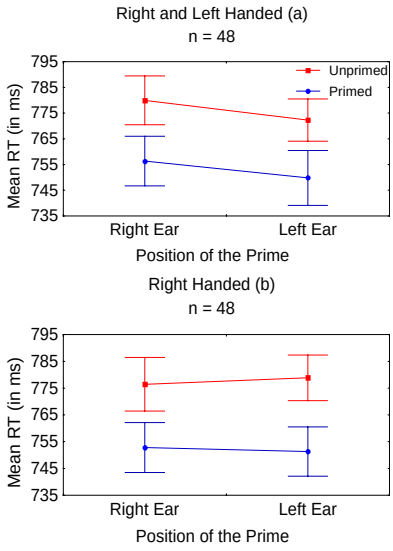

Figure 3: *Panel (a) shows the mean reaction times (in ms) for the primed (blue circle) and unprimed (red square) trials when the prime is presented to the right or to the left ear (Experiment 2a). Panel (b) shows the same data for only right handed participants (Experiment 2b).Errors bars shows the standard error across participants.*

4.3.2. Experiment 2b

Figure 3 (panel b) shows the average data for the primed and unprimed trials when the prime is presented to the right or to the left ear and when participants are exclusively right handed. The results are not different from Experiment 2a. The same significant priming effect occurs whether the prime is presented to the right (24 ms) or to the left ear (27 ms).

4.4. Discussion

These data show that unattended speech processing does not depend on the ear in which it is presented. Thus the ear in which the prime was presented cannot explain the different results in our study and in Dupoux et al. (2003).

5. EXPERIMENT 3

The last difference remaining between Dupoux's paradigm and ours was the pitch range difference between the attended and unattended messages. In Experiments 1 and 2, we used different pitch range voices for the two messages, whereas they were the same in Dupoux's study. The purpose of Experiment 3 was to assess the effect of the pitch range difference between the attended and the unattended message on the lexical processing of the unattended message.

5.1. Participants

51 university students were paid for their participation. They were English native speakers with no hearing, language or attentional impairment.

5.2. Design and procedure

We used the same procedure as in Experiment 1 except that for half of the trials, the voice of the attended message had an average fundamental frequency of 180Hz and for the other half, the pitch of the voice was kept at the original fundamental frequency (140Hz). In both cases the pitch was constant. The unattended message was always presented at the original pitch (average f_0 = 140Hz). The primed and unprimed trials were counterbalanced in two groups. The same and different pitch range trials were also counterbalanced in two lists. Priming and pitch range difference were within-subject factors.

5.3. Results and discussion

The results are shown on figure 4. As in the previous experiments, we found a significant priming effect of 29 ms when the attended and unattended messages have a different pitch range ($F_1(1,48)=12.90$; $p<.005$); $F_2(1,59)=15.56$; $p<.0005$). However no priming effect was found when they have the same pitch range ($F_1<1$; $F_2<1$).

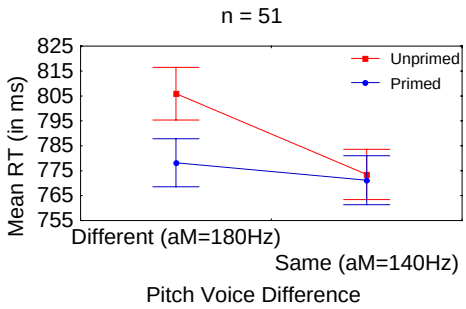

Figure 4: *Mean reaction times (in ms) for the primed (blue circle) and unprimed (red square) trials when the attended and unattended messages have a different or same pitch range (Experiment 3). Errors bars shows the standard error across participants*

These data may explain the discrepancy of results between our previous experiments in which we always used a different pitch range voice for the attended and unattended messages and Dupoux et al. (2003) experiments in which he always used the same pitch range for the two messages. Thus it appears that unattended speech can be lexically processed only if it has a different pitch range voice than an attended speech which is simultaneously presented.

6. EXPERIMENT 4

The purpose of Experiment 4 was to test the generalization of the results of Experiment 3 to different values of pitch difference for the unattended message. We also wanted to get more observations for each participant and each experimental condition, thus the pitch range condition were a between factor.

6.1. Participants

80 university students were paid for their participation. They were English native speakers with no hearing, language or attentional impairment.

6.2. Design and procedure

We used the same procedure as in Experiment 1, but we added a new set of stimuli. We had the two conditions of pitch range difference as in Experiment 3: attended and unattended message could be in the same or different pitch range condition. What was new is that the unattended message could be the original voice (average f_0 = 140Hz) or increased by 40Hz (average f_0 = 180Hz; synthesised with the PSOLA algorithm). The attended message could have a fundamental frequency of 140Hz or 180Hz depending on the pitch range condition. See table 1 for a summary of the different experimental conditions.

Pitch	Different	Same
aM	180Hz	140Hz
UaM	**140Hz**	**140Hz**

Pitch	Different	Same
aM	140Hz	180Hz
UaM	**180Hz**	**180Hz**

Table 1: *Different pitch range condition used in Experiment 4. aM refers to the attended message and UaM refers to unattended message.*

Each participant was assigned to one of the 4 groups of the pitch range difference factor. The primed and unprimed trials were counterbalanced in two groups. Priming was a within-subject factor whereas the pitch range difference was a between-subjects factor.

6.3. Results and discussion

The results are shown in figure 5. When the unattended message is presented at its original fundamental frequency (as in Experiment 3), we found a significant priming effect of 34 ms ($F_1(1,72)=10.48$; $p_1<.005$; $F_2(1,59)=13.27$; $p_2<.001$) only when the attended and unattended message were in a different pitch range. The same results were observed when the unattended message is synthesized with an average fundamental frequency of 180Hz: there was a significant priming effect of 30 ms when the two messages were in a different pitch range ($F_1(1,72)=8.17$; $p_1<.01$; $F_2(1,59)=29.64$; $p_2<.0001$) but no priming effect when they were in the same pitch range ($F_1<1$; $F_2<1$).

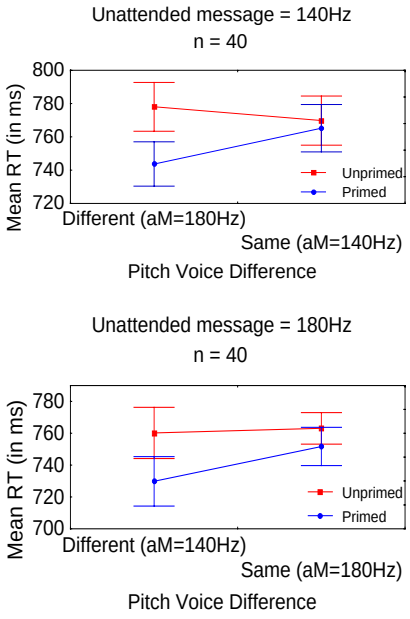

Figure 5: *Mean reaction times (in ms) for the primed (blue circle) and unprimed (red square) trials when the attended and unattended messages have a different or same pitch range (Experiment 4). The top panel shows the data when the unattended message has an average fundamental frequency of 140Hz (original voice) and the bottom panel shows the data when the unattended message has an average fundamental frequency of 180Hz (original voice synthesised). Errors bars shows the standard error across participants.*

These data confirm the results obtained in Experiment 3 and show that unattended speech can be lexically processed if it is not degraded and has a different pitch range from the other simultaneously-presented speech sounds.

7. EXPERIMENT 5

The purpose of Experiment 5 was to check whether priming could still be obtained when attention was more explicitly controlled. We wanted to be sure that participants' attention was not directed to the prime.

7.1. Design and procedure

We used the same procedure as in Experiment 1 except that we asked participants to perform a second task. As before, they had to detect the target in the attended message but they also had to remember a word in the attended message during each trial. The word that they had to remember was marked by a loud and low pitch tone presented in the attended ear. This word could be presented at a neutral position (far from the target word) or just before the target word. When the to-be-remembered word was presented just before the target, it occurred at the same time as the prime (because the end of the prime was always synchronised to the start of the target). Thus half of the targets were preceded by a to-be-remembered word and half were presented alone (as in the previous experiments). At the end of each trial, participants had to write down the to-be-remembered word.

Our hypothesis was that if participants were able to recall correctly the to-be-remembered word they could not have been actually attending to the (nominally unattended) prime. Furthermore, if a priming effect were still observed in the recall condition, then priming could not be due to an attentional switch toward the unattended message. Again, the primed and unprimed trials were counterbalanced in two groups. The recall conditions (target alone vs target preceded by a to-be-remembered word) were also counterbalanced in two lists. These two factors were within-subject.

7.2. Results and discussion

In a large majority of the trials, participants recalled the word which was presented in same time as the tone, instead of the word presented after the tone. This means that they recalled the word preceding the prime and not the one which was in the same time as the prime. However, we considered that it is very unlikely to argue that participants could have 1- recall the word preceding the prime, 2- switch their attention to the unattended ear to listen to the prime and 3- switch back to the attended ear to detect the target, all this in less than 1 second! Thus we analyzed the data considering that if participants recalled correctly the word presented in the same time as the tone, they did not switched their attention to the prime. In order to estimate the priming effect in the recall condition, we only kept the trials in which participants correctly recalled the word which was in the same time as the tone. The results are shown in figure 6. When the target is presented alone (as in the previous experiments), we found a significant priming effect of 28 ms ($F_1(1,30)=5.57$; $p_1<.05$; $F_2(1,59)=27.50$; $p_2<.0001$). In the recall condition, we also found a significant priming effect of 30 ms ($F_1(1,30)=9.43$; $p_1<.005$; $F_2(1,59)=13.58$; $p_2<.0005$). These two values did not differ ($F_1<1$; $F_2<1$).

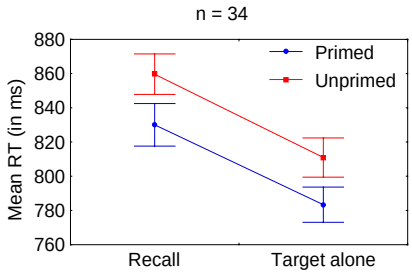

Figure 6: *Mean reaction times (in ms) for the primed (blue circle) and unprimed (red square) trials when the target is preceded by a to-be-remembered word and when it is presented alone (Experiment 5). Errors bars shows the standard error across participants*

These data show that the priming observed in the previous experiments still remains on trials when we have independent evidence that participants are actually attending to the nominally attended message.

8. DISCUSSION

In this study, we have been able to show a priming effect when the prime was unattended, indicating that it is possible to process unattended speech. In Experiment 1, we showed a priming effect of 22 ms with an unattended prime using

continuous speech in the unattended message. In Experiment 2, we showed that this priming effect did not depend on the ear in which the prime is presented; we found a priming effect of 24 ms when the prime was presented in the right ear and 22 ms when it was presented to the left ear. In Experiment 3, we found that the priming effect depended on the pitch range difference between the attended and unattended messages: we found a priming effect of 29 ms when the pitch range was different but this priming disappeared when there was no pitch range difference between attended and unattended messages. This result was confirmed in Experiment 4 in which we tested different values of pitch range difference and different groups of subjects for each pitch range condition. Finally in Experiment 5 we were able to prove that participants do not process the prime with attentional switches; we found a priming effect of 30 ms when participants had to remember a word in the attended message which was presented at the same time as the prime.

Our data are not consistent with those of Dupoux et al. (2003). They showed that it was not possible to get a priming effect with an unattended prime while using continuous speech in the unattended message. The reason for this discrepancy of results may be that the authors degraded the prime by temporally compressing it. Furthermore the authors used the same pitch range in the attended and unattended messages, whereas we have shown that in order to the get a priming effect, it was necessary to use different pitch ranges.

Our study presents the first evidence of unattended speech processing using continuous speech in the unattended message. The major criticism made of past studies showing a lexical processing of unattended messages lay on the inadequate control of attentional switches [5][6]. However, we have been able to show that even if we compel participants to focus on the attended message by asking them to remember some specific words, the priming effect remained (Experiment 5).

The reason why priming disappeared when there is no pitch range difference between the attended and unattended message may lie in the perceptual organization of the unattended message which may be harder in such conditions than when the pitch range is different. The unattended message needs to be perceptually well-segregated from the louder attended message, and a similarity in pitch may cause perceptual degradation of the unattended message. The louder, attended message would not be affected.

The nature of the processing of the unattended message may be lexical. In our experiments, we used repetition priming, thus simple phonological processing of the prime would be sufficient to get a priming effect without necessitating lexical processing. The reason why we think lexical processing may be involved is that Dupoux et al. (2003), using the same repetition priming paradigm, showed that priming was not found with non-words, which are not able to activate a lexical representation. One possible way of directly testing whether lexical processing is in fact involved would be to test unattended speech processing using semantic priming.

9. REFERENCES

[1] Cheery, E. C. (1953). Some experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America, 25*, 975-79.

[2] Broadbent, D. E. (1958). *Perception and Communication*. New York: Pergamon Press.

[3] Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin, 20,* 12-16.

[4] Deutsch, J. A., & Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological Review, 70,* 80-90.

[5] Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *The behavioral and Brain Sciences, 9,* 1-66.

[6] Dupoux, E., Kouider, S., & Mehler, J. (2003). Lexical access without attention? Exploration using dichotic priming. *Journal of Experiment Psychology: Humain Perception and Performance, 29.*

[7] Kucera, H., & Francis, W. N. (1967). Computational Analysis of present day American English. *Providence,* RI: Brown University Press.

[8] Hampton, J. A., & Gardiner, M. M. (1983). Measure of internal category structure: A correlational analysis of normative data. *British Journal of Psychology, 74,* 491-516.