Disentangling Perceptual and Response Bias in Unilateral Neglect: A Methodological Proposal

A. Toraldo,[1] R. D. McIntosh,[2] H. C. Dijkerman,[3] and A. D. Milner[2]

University of St. Andrews, Scotland, United Kingdom

Bisiach and his colleagues have developed a new version of Milner's ''landmark task'' for the purpose of separating ''perceptual'' and ''response'' biases in neglect patients. Subjects are required to decide which is the longer (or the shorter) of the two portions of a pre-bisected horizontal line. The authors proposed two indices to measure perceptual and response bias respectively. However, these indices are not mathematically independent of one another. Moreover, they do not exploit all of the information available in the data, since they do not consider the effect of the different transection locations across trials. We now propose an alternative means of analyzing data from the revised Landmark task, which generates independent estimates of perceptual and response biases. The method and its theoretical foundation are summarized, and illustrative data obtained from brain damaged patients and control subjects are presented.

INTRODUCTION

In their revised version of Milner's ''landmark task'' (Harvey, Milner, & Roberts, 1995), Bisiach, Ricci, Lualdi, and Colombo (1998) asked left neglect patients to indicate the shorter (in one condition) or longer (in a second condition) of two collinear segments comprising a 180 mm long horizontal line. The transecting boundary between the two segments (landmark) was varied between nine positions symmetrically distributed around the objective center of the line. In Bisiach et al.'s view, a patient with predominantly perceptual neglect (Input-Related Neglect, IRN) should tend to choose the left part as the shorter (or the right as the longer). In contrast, a patient dominated by a rightward response bias (Output-Related Neglect, ORN) should tend to choose the rightmost line portion, regardless of whether he is asked to indicate the longer or the shorter segment. The proposed indices for IRN and ORN are the following:

$$\text{IRN index: } PB = (LS + RL)/2$$

$$\text{ORN index } RB = (RS + RL)/2,$$

where *PB* is ''perceptual bias,'' *RB* is ''response bias,'' and *LS*, *RL*, and *RS* are, respectively, the percentages of left responses in the shorter condition, of right responses in the longer condition and of right responses in the shorter condition.

Bisiach et al.'s revised Landmark Task constitutes a valuable methodological advance but has two major limitations. (1) *PB* is computed by summing all of the responses made by the subject, regardless of landmark location. This averaging procedure wastes a valuable opportunity which Bisiach's method provides. Since equal numbers of trials are given at a wide range of landmark positions, information about the variations of perception across such locations is potentially available in the data. (2) The ranges of variation of *PB* and *RB* are not mutually independent, which can lead to diagnostic errors. The more extreme the *RB*, the narrower the possible range

[1] Current address: Cognitive Neuroscience Sector, SISSA-ISAS, Trieste, Italy.
[2] Current address: Department of Psychology, University of Durham, UK.
[3] Current address: Psychological Laboratory, University of Utrecht, Netherlands.

of *PB*, and vice-versa. As a consequence, a patient with an extreme *RB* will be assigned an artificially central *PB* score, probably well within the normal range, when, in fact, the properties of his perception are simply unknown.

The intention of the present work is to develop a different mathematical model that allows the above limitations to be overcome.

METHOD

Input-Related Neglect (IRN)

The landmark task was originally devised for the purpose of determining the point of subjective equality (*PSE*), i.e., the position of the transection that would produce two subjectively equal segments (Milner, Brechmann, & Pagliarini, 1992). IRN can be defined as a *pathological shift of the PSE toward one side.* The amplitude of such a shift would thus be a direct measure of IRN.

Output-Related Neglect (ORN)

ORN can be conceived as a trend to ''ignore'' the perceptual experience of the line, and to emit instead a ''default'' response toward one side (typically the ipsilesional). The *proportion of trials* in which this occurs can be considered as a measure of the intensity of ORN.

Derivation of the PSE (index of IRN)

In accordance with standard psychophysical models, the *PSE* is estimated by plotting the transection position against the probability *p* of *perceiving* the right part of the line as the shorter. Approximating the experimentally estimated points by means of a cumulative normal curve, the *PSE* is the point on the line where the function assumes the value 0.5 (i.e., the abscissa of the inflection point: see Fig. 2). The ''perceptual'' probability *p* therefore needs to be estimated at each transection position. Of course, a left neglect subject might sometimes respond 'right' not as a reflection of his visual experience, but because he is dominated by rightward ORN. Nevertheless, a combination of the data from the two response conditions for each specific transection position, allows us to compute *p* partialling out such ORN effects (see Fig. 1).

Derivation of the Index M (measure of ORN)

The caption of Fig. 1 gives the formula for the estimated proportion *m* of trials in which the subject made a default response, as a result of the same logical steps leading to the estimation of *p*. As previously suggested, *m* is a measure of ORN. Nevertheless, *m* values refer to single transection positions, and, taken alone, are not reliable enough to provide a diagnostic measure of ORN. To increase the reliability of the estimation of ORN, *m* values are averaged across all transection positions, thus yielding an overall index, *M*. This value is simply a linear transformation of Bisiach et al.'s *RB.*

Subjects

We administered the Landmark-V (Bisiach et al., 1998) to 18 brain-damaged patients and 12 normal controls. Seventeen patients had suffered a right hemisphere

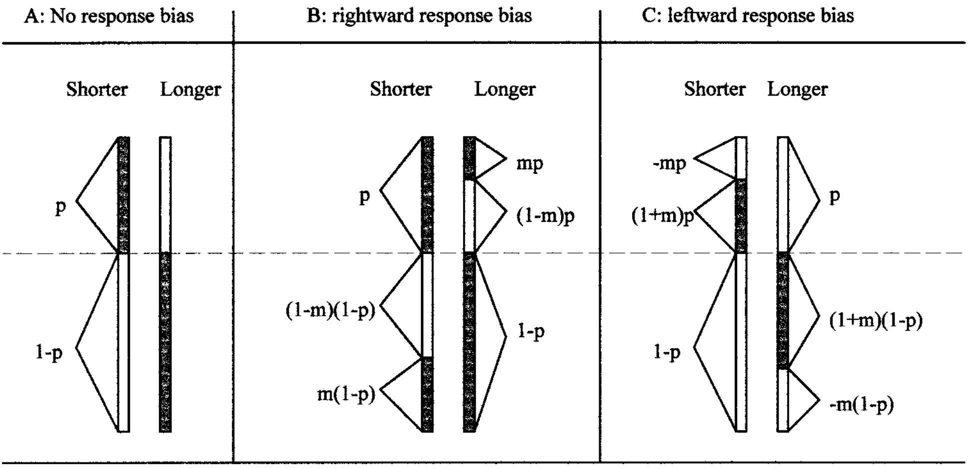

FIG. 1. Proportions of responses at a given transection position are shown as vertical extensions; shaded areas = rightward responses, open areas = leftward responses; the whole vertical extension = the overall set of responses in each specific condition. (A) A subject with no ORN who *perceives* the right segment as shorter in a proportion p of trials will choose the left in the ''shorter'' condition and the right in the ''longer'' condition on an equal proportion p of trials. (B) The same subject, to whom we have ''added'' some rightward ORN, will substitute a proportion m of leftward (open) responses with rightward (shaded) responses in both conditions. Therefore, in the shorter condition an overall proportion $m(1 - p)$ of trials will be substituted by the default response. The same thing will happen in a proportion mp of the trials in the longer condition. We know from observation the proportions of leftward responses made in the shorter condition (LS), and of rightward responses in the longer condition (RL) made by our hypothetical patient. As one can see in part B of the figure, these will be computable as $LS = (1 - m)(1 - p)$ and $RL = mp + (1 - p)$. These latter equations constitute a simple first-degree system, with two unknowns (m and p), LS and RL being known from the experiment. It is thus easy to calculate the values of m and p as follows: $p = (1 - RL)/(1 + LS - RL)$ and $m = RL - LS$. (C) Similar reasoning can be applied in the case of a leftward ORN. Here a proportion m of cases in which the perceptually based response would have been rightward (shaded) will be substituted by leftward (open) responses because of the ORN. Since this time the direction of such pathological responses is towards the left, we will conventionally add a negative sign in front of parameter m, which therefore becomes $-m$. So, $LS = -mp + (1 - p)$ and $RL = (1 + m)(1 - p)$. The solution of these equations is: $p = (1 - LS)/(1 + RL - LS)$ and $m = RL - LS$. Overall, the estimation of m is $RL - LS$. Positive values of this parameter indicate rightward ORN, negative values leftward ORN. As far as p is concerned, a compact formula is: $p = [1 - \max(LS, RL)]/[1 + \min(LS, RL) - \max(LS, RL)]$.

stroke (six of whom showed left neglect on star cancellation); one patient (DL) showed persistent right neglect following a left hemisphere stroke.

RESULTS

Figure 2 illustrates the estimation of the *PSE* for two right hemisphere damaged patients, CS (left neglect on cancellation) and MMcP (no neglect).

Output-Related Neglect (M values)

Since M values had large standard errors, an ORN was diagnosed when the *confidence interval* of a given patient's M value did not overlap the control range. According to this criterion, one only patient (CS) was classified as having a reliable ORN. On the contrary, nine patients did so when Bisiach et al.'s (1998) cut-offs were applied (four towards the contralesional and five towards the ipsilesional side). The reason for this divergence lies simply in the use of confidence intervals instead of

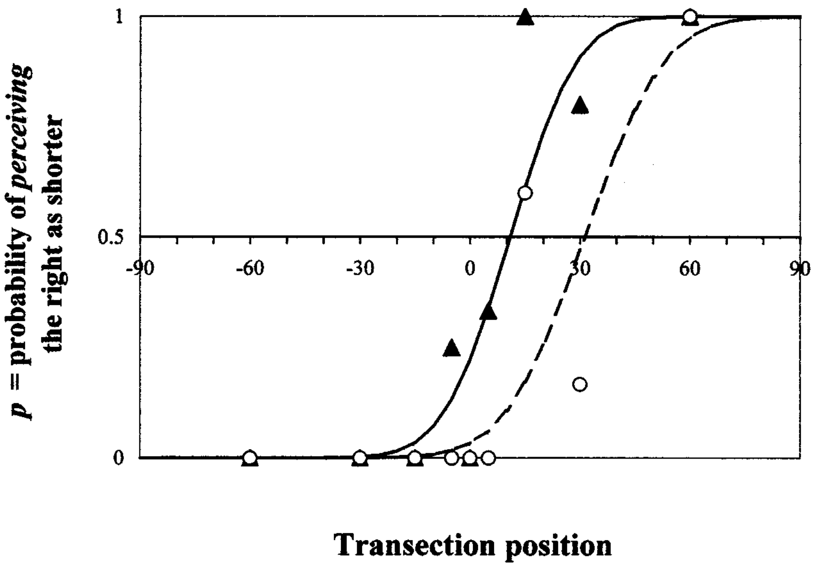

FIG. 2. Patient CS (filled triangles) showed a rightward displaced *PSE* (the solid curve intersects the horizontal axis at +10.9 mm). Patient MMcP, on the other hand, showed a gross rightward perceptual bias (dashed curve fitting open circles, *PSE* = +31.3 mm). The best fit was obtained by means of a Probit Analysis.

local estimates (*M* and *RB* are mathematically equivalent measures). Yet, the use of confidence intervals is not justified only by prudence, but also by the consideration that the instability of perceptual judgements will increase the probability of eccentric *M* values even in absence of true ORN. This problem is particularly severe with brain damaged patients, whose perceptual instability can be quite high. The use of confidence intervals, instead of local estimates, eliminates the risk of misinterpreting an extreme *M* value as a sign of ORN, when its eccentricity is just due to the overall instability of perceptual judgements. By this logic, the eight subjects classified as having ORN according to Bisiach et al.'s criterion, but not according to ours, are likely to be false positives.

Input-Related Neglect (PSE)

As previously outlined, *PSE*s were estimated by means of a Probit Analysis. IRN was diagnosed when a patient's confidence interval for the *PSE* and the normal range did not overlap.

To compare our diagnostic criterion to that of Bisiach *et al.* (1998) *ceteris paribus*, confidence intervals for *PB* were also computed and the same ''nonoverlap'' logic was applied. Our *PSE* index turned out to be a slightly more sensitive measure of perceptual bias than *PB* (four more patients classified as having a perceptual bias, 11/18 vs 7/18).

DISCUSSION

The proposed measures (*PSE* and *M*) are independent in principle, so that any reliable correlation between the two indices in a group of patients will be interpretable exclusively in neuropsychological terms and not as the result of a mathematical artifact.

The indices *PSE* and *M* have mathematical properties that fully reflect the theoretical premises which generated them. This allows us to avoid the diagnostic errors and logical problems of the *RB* and *PB* indices as previously outlined. For instance, a patient who responds rightwards in all of the trials will be correctly given the maximum score for ORN but his IRN will now be correctly defined as *unknown* (all *p* values will be 0/0). In contrast, Bisiach et al.'s *PB* classifies such a patient as having no IRN at all, which is possible but not necessarily the case. Similarly, a patient who simply guesses in many or all of the trials will, by our method, be distinguished by a dramatically shallow normal cumulative function, rather than being judged to have no IRN.

The other main advantages of the present method are:

(i) its parameters take into account a higher quantity of the information contained in the Landmark data;

(ii) even with a constant *PB* score, the parameter *PSE* (point of subjective equality) allows one to distinguish different degrees of perceptual neglect;

(iii) the proposed geometrical interpretation of *PSE* facilitates further investigation of the nature of perceptual neglect itself.

In the present sample of patients, our *PSE* proved to be slightly more sensitive measure of IRN than Bisiach *et al.*'s *PB.* This advantage is expected to increase considerably for patients with severe ORN (absent in our sample) because, whilst the *PB* score of such patients will be prevented from reaching extreme values, the measure of *PSE* will not.

REFERENCES

Bisiach, E., Ricci, R., Lualdi, M., & Colombo, M. R. (1998). Perceptual and response bias in unilateral neglect: Two modified versions of the Milner landmark task. *Brain and Cognition,* **37,** 369–386.

Harvey, M., Milner, A. D., & Roberts, R. C. (1995). An investigation of hemispatial neglect using the landmark task. *Brain and Cognition,* **27,** 59–78.

Milner, A. D., Brechmann, M., & Pagliarini, L. (1992). To halve and to halve not: An analysis of line bisection judgements in normal subjects. *Neuropsychologia,* **30,** 515–526.

This is doi:10.1006/brcg.2001.1423.

Stroke Direction Asymmetry in Figure Drawing: Influence of Handedness and Reading/Writing Habits

Jyotsna Vaid,* Maharaj Singh,† Tripti Sakhuja,‡ and G. C. Gupta‡

**Texas A & M University; †Ch. Charan Singh University; and ‡Delhi University*

Recent studies suggest that asymmetries noted in certain nonlinguistic tasks used in laterality research (e.g., facial affect judgment, line bisection) may in part be influenced by prior reading/writing habits. The present study examined the relative influence of reading/writing direction and handedness on the direction of stroke movement in free-hand figure drawing. One hundred twenty right and left handed brain-intact adult readers of scripts with opposing directionality (Hindi vs Urdu) and illiterate controls were observed while drawing a tree, a hand, a house, an arrow, a pencil, and a fish. Right-handers (including right-handed illiterates) and left-to-right readers drew most figures in a left-to-right direction, whereas left handers (including left handed illiterates) and right-to-left readers more often drew the figures from right to left. These results extend previous findings and contribute to a growing body of evidence demonstrating reading scan biases in nonlinguistic perception and production tasks. It would appear that reading/writing habits cannot be ignored as a potential artifact in studies of hemisphere functional asymmetry employing nonlinguistic stimuli.

INTRODUCTION

Some languages are written and read from top to bottom, others from left to right, and still others from right to left. Exposure to a particular writing system influences the order of recall of visually presented linear arrays of stimuli (see Nachshon, 1985, for a review of studies with Hebrew or Arabic readers). Interestingly, reading/writing biases are not restricted to recall effects or found only for linguistic stimuli, but rather ''invade'' perception or production of nonlinguistic stimuli as well. The presence of such biases has been noted for tasks as varied as aesthetic perception (Chokron & de Agostini, 2000; Nachshon, Argaman, & Luria, 1999), line length estimation (Singh, Vaid, & Sakhuja, 2000), chimeric facial affect judgment (Eviatar, 1997; Sakhuja, Gupta, Singh, & Vaid, 1996; Vaid, 1995), speeded linear dot production (Vaid, 1998), the perception of apparent motion (Morikawa & McBeath, 1992), naming and recall of linearly arrayed pictured objects (Padakannaya, Devi, Zaveria, Chengappa, & Vaid, in press), and covert spatial attention (Eviatar, 1995).

The present research considered the potential influence of reading/writing habits and handedness-linked movement biases on the direction of stroke movement in drawing simple objects, such as an arrow or a flag. It has been noted in other tasks that hand movements directed away from the body are performed more easily than those directed toward the body (van Sommers, 1984; see also Vaid, 1998). In the task of figure drawing, then, one would expect stroke direction asymmetries to vary as a function of the hand used to draw, with drawings executed with the right hand proceeding more easily in a left-to-right direction whereas those done with the left hand proceeding more easily in a right-to-left direction. Whether this hand movement bias might interact with reading/writing habit-induced biases in stroke direction was of additional interest.

In previous related research (Vaid, 1995) we examined the direction in which line drawings of directional objects were oriented by readers of scripts with opposing directionality. Participants were readers of Hindi (a language that is read and written from left to right) and Urdu (a language that is read and written from right to left). From a methodological perspective what is particularly compelling about this language pair is that, on the spoken level, Hindi and Urdu are virtually identical. They differ primarily on the written level: Hindi uses the Devanagari script, a writing system originally developed for Sanskrit, while Urdu uses a Perso-Arabic script. In two previous studies, we examined right-handed adult readers of Hindi and Urdu and Arabic (Vaid, 1995, Study 1) and right- vs left-handed Hindi and Urdu children (Vaid, 1995, Study 2) on a figure drawing task. The results showed that, among right handers, objects tended to be drawn facing left by Hindi readers while Urdu and Arabic readers tended to orient the same objects facing right. Among left-handers, however, the majority of items by Hindi and Urdu readers alike were drawn facing left. These results were taken as support for an influence of reading/writing habit on drawing orientation, but one that was limited to right handers, given that left handers appeared to be more influenced by a hand movement asymmetry favoring extensional movements.

While the present study also compared Hindi and Urdu readers on a figure drawing task, the dependent measure of interest was stroke movement direction, rather than object facing. Moreover, whereas in previous work (e.g., Vaid, 1995) handedness effects had been examined in a child sample, in the present study an adult sample of right vs left handed Hindi and Urdu readers was tested. In addition, the present study tested a sample of (right vs left-handed) illiterate speakers of Hindi/Urdu. Inclusion of the illiterate subgroups should provide an additional way of teasing apart the relative influence of reading/writing habits from handedness effects.

RATIONALE

If figure drawing is uninfluenced by biomechanical or reading/writing habits, then the percent of drawings executed in a leftward direction should be about equal to those executed in a rightward direction. If figure drawing is influenced primarily by biomechanical factors, specifically, innate directional tendencies in muscular movement such that movements directed away from the body are executed more smoothly than those directed toward the body, one would expect more drawings with the right hand to have a left-to-right stroke direction and more drawings with the left hand to have a right-to-left stroke direction, in all three linguistic groups. If, however, movement asymmetries in figure drawing primarily reflect reading/writing habits, then one would expect that, regardless of handedness, readers of Hindi would show a left-to-right movement preference in figure drawing while readers of Urdu would show a leftward movement bias; illiterates would show no consistent directional preference.

Whereas both the biomechanical and the reading/writing habit explanations would predict a left-to-right directional preference in right handed Hindi readers and a right-to-left directional preference in left handed Urdu readers, the two variables make opposite predictions for right handed Urdu and left handed Hindi readers. Thus, the outcome of these latter subgroups (together with that of the illiterate right vs left handers) will be of particular interest in determining the relative strength of the biomechanical versus reading/writing habit explanation.

METHOD

Participants

Subjects were 120 brain-intact subjects between 15 and 20 years of age residing in New Delhi. They were subdivided into three groups in terms of language experience: 40 were native readers of Hindi (a language read and written from left to right), 40 were native readers of Urdu (a language read and written from right to left) and 40 were illiterate age-matched controls speakers of Hindi/Urdu. The Hindi readers (Group 1) were enrolled in an English medium of instruction school. The Urdu readers (Group 2) were enrolled in an Urdu medium of instruction and the illiterates (Group 3) had no formal schooling. Each group was further subdivided by handedness and sex, with 10 subjects per cell. Handedness was assessed by subjects' stated preference on the Annett (1970) inventory and by their observed performance on the inventory items. Only strongly right- or left-handed subjects (defined as those who reported a preference and showed a performance favoring a particular hand for 75% of the activities) were included.

Materials and Procedure

Subjects were tested individually in a quiet laboratory setting with the experimenter seated next to the subject on either the left or the right side, randomly. They were given a response sheet with six marked spaces each containing the name of an item at the top center and were instructed simply to produce a quick sketch of each of the items in the space provided, using their preferred hand. The figures to be drawn

TABLE 1
Right-Handers' Frequency of Figures Drawn from Left-to-Right vs from Right-to-Left as a Function of Reading/Writing Habit (n = 20 per Language Group)

	Illiterate	Hindi	Urdu
Left-to-right	79	85	33
Right-to-left	21	15	67

were a tree, a hand, a house, an arrow, a pencil, and a fish and subjects were told they could draw the figures in any order. For each figure, with the exception of the tree (included as a filler item), the predominant stroke direction employed—right-to-left or left-to-right—was noted.

RESULTS

Per subject, the total number of figures drawn in a leftward vs rightward direction was entered into a series of chi square analyses. (A correction for continuity was computed for all chi square analyses involving a single degree of freedom.) An initial analysis revealed no significant sex difference in direction of movement [$\chi^2 = 2.19$, df = 1, $p = .14$]; males drew 47% and females drew 40.7% of the figures toward the right side. A chi square comparing right and left handers was highly significant [$\chi^2 = 111.6$, df = 1, $p < .0001$]; left handers drew a majority of the figures to the left side (78%) while right handers drew a majority to the right (66%). The chi square analysis by Language group (Hindi readers, Urdu readers, and Illiterates) was also significant [$\chi^2 = 28.53$, df = 2, $p < .01$]; while Hindi readers and illiterates showed no particular directional bias, Urdu readers drew the majority of the drawings toward the left side (71%), in accordance with their right-to-left reading/writing habit.

The data per Language group were analyzed by Handedness in three separate chi square analyses. The results showed a significant effect of Handedness in the Illiterate sample [$\chi^2 = 74.54$, df = 1, $p < .0001$] and in the Hindi sample [$\chi^2 = 72.59$, df = 1, $p < .0001$] but no effect of Handedness in the Urdu sample [$\chi^2 = 1.19$, df = 1, $p = .27$]. For both the Illiterate and Hindi samples, right handers showed a predominant left-to-right stroke movement while left handers showed a right-to-left movement. For the Urdu sample, both right and left handers showed a right-to-left stroke movement bias.

Finally, two additional chi square analyses were performed, comparing the values per Handedness group by Language group. For right-handers (see Table 1), there was a significant effect of Language group [$\chi^2 = 71.85$, df = 2, $p < .0001$] indicating

TABLE 2
Left-Handers' Frequency of Figures Drawn from Left-to-Right vs from Right-to-Left as a Function of Reading/Writing Habit (n = 20 per Language Group)

	Illiterate	Hindi	Urdu
Left-to-right	17	24	25
Right-to-left	83	76	75

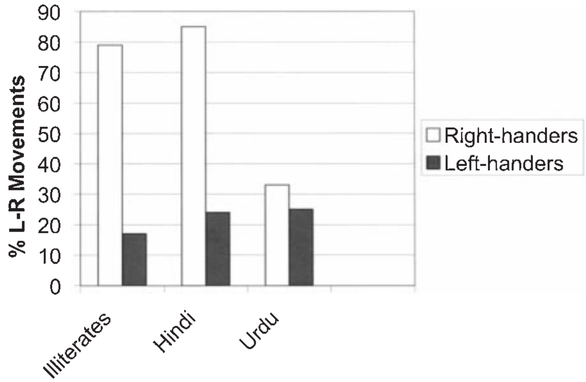

FIG. 1. Incidence of left-to-right movement bias in stroke production as a function of reading habit.

that a propensity to draw figures toward the right side characterized the Hindi readers (who drew 85% of the drawings in a rightward direction) and the illiterate right handers (79% rightward), but not the Urdu right handers, who instead showed a right-to-left directional preference (67% leftward).

In the analysis for left-handers (see Table 2), there was no significant effect of Language group [$\chi^2 = 2.21$, df $= 2$, $p = .33$], with all groups showing a right-to-left movement preference (76, 83, and 75%, for Hindi, Illiterate, and Urdu groups, respectively). A summary figure of the results as a whole is presented in Fig. 1.

DISCUSSION

The present findings on stroke direction biases in figure drawing corroborate those summarized in Vaid (1995) on the facing of directional figures and demonstrate that reading/writing habits clearly influence this ostensibly nonlinguistic task. As such, these results add to the growing body of evidence documenting the influence of reading/writing habits on nonlinguistic tasks. However, reading/writing habits were not the sole influence on asymmetries in figure drawing, as the performance of our illiterate sample demonstrated. While this group showed no overall movement bias when considered as a group, their performance clearly differed when the variable of handedness was considered: right handed illiterates showed a left-to-right stroke bias while left handers showed a right-to-left bias. Thus, a biomechanical influence appears to be operative in figure drawing direction when reading/writing habits are not present.

What happens when reading/writing habits predict one outcome and biomechanical principles predict the opposite outcome? Such was the case for left-handed Hindi readers and for right-handed Urdu readers. In the former case, the results favored a biomechanical influence, as left-handed Hindi readers showed a right-to-left bias, contrary to the expectation based on their (left-to-right) reading/writing habits. However, for right-handed Urdu readers, a reading/writing habit effect appeared to be a stronger influence, in that these individuals showed a right-to-left bias, contrary to the left-to-right expectation for extensional movements with the right hand (see Fig. 1). In summary, our results suggest that reading/writing habits do exert a strong influence that can supersede that of biomechanical principles, but only when the writing hand is the right hand. For left-handers, a reading/writing bias was not sufficient to offset biomechanical principles influencing hand movement asymmetries. These outcomes obtained from a large, adult sample corroborate what had previously been reported for a smaller child sample (Vaid, 1995, Study 2).

Tasks involving asymmetries in perception or production are often used in studies of hemispheric specialization of function. Our findings, taken together with those of several other recent demonstrations of cultural influences on laterality (e.g., Chokron & DeAgostini, 2000; Eviatar, 2000) have certain implications. They suggest that explanations invoking hemisphere-specific mechanisms as the favored account of observed asymmetries in verbal or nonverbal performance cannot afford to ignore either hand movement-related directional biases or directional scanning tendencies arising from reading/writing experience as additional sources of influence to be controlled or systematically explored.

REFERENCES

Annett, M. (1970). A classification of hand preference by association analysis. *British Journal of Psychology,* **61,** 303–321.

Chokron, S., & De Agostini, M. (2000). Reading habits influence aesthetic preference. *Cognitive Brain Research,* **10**(1–2), 45–49.

Eviatar, Z. (1995). Reading direction and attention: Effects on lateralized ignoring. *Brain and Cognition,* **29,** 137–150.

Eviatar, Z. (1997). Language experience and right hemisphere tasks: The effects of scanning habits and multilingualism. *Brain and Language,* **58**(1), 157–173.

Eviatar, Z. (2000). Culture and brain organization. *Brain and Cognition,* **42**(1), 50–52.

Morikawa, K., & McBeath, M. K. (1992). Lateral motion bias associated with reading direction. *Vision Research,* **32,** 1137–1141.

Nachshon, I., Argaman, E., & Luria, A. (1999). Effects of directional habits and handedness on aesthetic preference for left and right profiles. *Journal of Cross Cultural Psychology,* **30**(1), 106–114.

Nachshon, I. (1985). Directional preferences in perception of visual stimuli. *International Journal of Neuroscience,* **25,** 165–174.

Padakannaya, P., Devi, M. L., Zaveria, B., Chengappa, S., & Vaid, J. (2001). Directional scanning effect and strength of reading habit in picture naming and recall. *Brain and Cognition,* doi: 10.1006/brcg.2001.1403.

Sakhuja, T., Gupta, G. C., Singh, M., & Vaid, J. (1996). Reading habits affect asymmetries in facial affect judgments: A replication. *Brain and Cognition,* **32**(2), 162–165.

Singh, M., Vaid, J., & Sakhuja, T. (2000). Reading/writing vs handedness influences on line length estimation. *Brain and Cognition,* **43**(1–3), 398–402.

Vaid, J. (1995). Script directionality affects nonlinguistic performance: Evidence from Hindi and Urdu. In I. Taylor & D. Olson (Eds.), *Scripts and literacy* (pp. 295–310). Dordrecht: Kluwer Academic.

Vaid, J. (1998). Cultural versus biomechanical influences on a graphic production task. *Brain and Cognition,* **37,** 75–78.

Van Sommers, P. (1984). *Drawing and cognition: Descriptive and experimental studies of graphic production processes.* New York: Cambridge University Press.

This is doi:10.1006/brcg.2001.1424.

Reliability of Laterality Effects in a Dichotic Listening Task with Nonverbal Material

Daniel Voyer* and Marguerite A. Rodgers†

**University of New Brunswick Fredericton; and †Saint Francis Xavier University*

The present study examined the reliability of a dichotic listening task using nonverbal stimuli. Twenty undergraduate students (all right-handed native English speakers) had to report whether they had heard a target emotion. The task used English words (bower, dower, power, tower) pronounced with an angry, happy, neutral, or sad emotional tone. Results showed a relatively high level of test–retest reliability for the laterality effect. In addition, a significant gender by ear of presentation interaction was obtained. The interaction reflected the fact that a strong left ear advantage was found in females but not in males. The findings indicate that the task used here should be considered a reliable means to assess the lateralization of emotions. Issues concerning the relation between gender and laterality are addressed in the discussion.

INTRODUCTION

The existence of auditory perceptual asymmetries is one of the clearly established phenomena in the study of human perception (Voyer, 1996). It is thus commonly reported that verbal tasks are performed faster and more accurately when presented to the right ear, whereas nonverbal stimuli are recognized faster and more accurately when presented to the left ear. Based on the pattern of connection in the auditory pathways (see Kimura, 1967), these perceptual asymmetries are typically interpreted as representing the functioning of each cerebral hemisphere. Specifically, the right ear advantage is interpreted as reflecting a left hemisphere advantage, whereas the left ear advantage is believed to reflect a right hemisphere advantage. This type of asymmetrical effects is typically referred to as laterality (Bryden, 1982).

Although the phenomenon of laterality is commonly accepted and many theories are in existence to account for it, very few researchers seem to consider the reliability of the measures they use to assess it (Voyer, 1998). The meta-analysis published by Voyer (1998) actually emphasized the scarcity of research in that area. This author found only 88 effect sizes examining the reliability of laterality effects in either the visual or auditory modality, with 77.3% of those coming from auditory studies with verbal material. In fact, when both modalities are considered, only 6 of the 88 effect sizes (6.8%) were drawn from studies using nonverbal material. It is therefore not surprising that it was one of the conclusions reached by the author that more studies on the reliability of laterality effects using nonverbal stimuli are required. It was therefore the purpose of the present study to investigate the reliability of an auditory task assessing the lateralization of emotions.

METHOD

Participants

Twenty undergraduate students (10 females, 10 males) were recruited from courses in psychology. All participants were right-handed according to their score on the Waterloo Handedness Questionnaire (Steenhuis & Bryden, 1989) and they all reported normal hearing. Participants were paid for their collaboration.

Materials

The words ‘‘bower,’’ ‘‘dower,’’ ‘‘power,’’ and ‘‘tower’’ were pronounced by a male speaker with an angry, happy, neutral, or sad emotional tone, following the procedure developed by Bryden and McRae (1989). The words were recorded on a microcomputer equipped with a 16-bit soundcard. They were all edited to a length of 350 ms and an intensity of 65 dB. All possible combinations of words and emotions in each ear resulted in 144 possibilities. A program written in Mel Professional v2.0 (Schneider, 1995) controlled the timing of each trial, and recorded the accuracy of each response. Koss TD-60 headphones were used for stimulus presentation.

Finally, hand preference was measured with the Waterloo Handedness Questionnaire (Steenhuis & Bryden, 1989). This questionnaire consists of 32 statements concerning hand preference for specific activities. It results in a score ranging from −64 (completely left-handed) to +64 (completely right-handed).

Procedure

Participants were seated in front of a computer in a quiet room. At the onset of each trial, instructions appeared on the computer screen telling participants to press the spacebar when they were ready for the trial. This resulted in the presentation of a pair of dichotic stimuli. Participants were required to listen for a target emotion (anger, happiness, neutrality, or sadness) and to report whether that emotion was presented in the dichotic pair. The target emotion used in the task was counterbalanced across participants. Half the participants indicated their answer by pressing ''0'' on the keyboard when the target emotion was present in the dichotic pair and ''1'' when the target emotion was not present. The remaining participants had the opposite response arrangement.

Participants completed the task twice while monitoring the same emotion each time. For the purpose of clarity, the first 144 trials will be designated as the first testing session in what follows, whereas the second set of 144 trials will be referred to as the second testing session. Performance was scored in terms of the percentage of correct responses for each ear in each of the testing sessions. This design allowed the calculation of a test–retest reliability estimate. To avoid confounding effects due to potential variation in sound intensity between the left and right headphones, half the participants of each gender completed both blocks of the auditory task while wearing the right headphone on the right ear, whereas the other half had the reverse arrangement.

The Waterloo Handedness Questionnaire was completed between the two sessions. Data on that questionnaire indicated that the participants were truly right-handed, with a range from +23 to +64.

RESULTS

Test–retest reliability was calculated separately for each ear on the percentage of correct responses. The results of this analysis showed a reliability estimate of 0.54 for the left ear and 0.86 for the right ear. A simple measure of laterality was computed by subtracting the right ear score from the left ear score. This measure showed a reliability estimate of 0.79 in the whole group. When the reliability estimate was calculated separately for each gender, males and females showed the exact same estimate of 0.79.

Data were also analyzed using a mixed design ANOVA with gender (male, female) as between-subjects independent variable, and testing session and ear of presentation (left, right) as within-subject independent variables. The percentage of correct responses was the dependent variable.

Results showed a main effect of testing session, $F(1, 18) = 13.53$, $p < .01$. This effect indicated that the percentage of correct responses was greater in the second testing session ($M = 93.5\%$) than in the first one ($M = 86.5\%$). A significant gender by ear of presentation interaction was also obtained, $F(1, 18) = 7.61$, $p < .02$ (see Table 1). Simple effect analyses revealed a significant left ear advantage in females, $F(1, 9) = 7.96$, $p < .03$, but not in males, $F(1, 9) = 1.01$. No other main effects or interaction achieved statistical significance (all p's $> .20$).

TABLE 1
Mean Percent of Correct Responses and Standard Deviation (*SD*) as a Function of Gender and Ear of Presentation

	Ear of presentation			
Gender	Left	Right	Difference[a]	Effect size (d)
Males	91.6	95.7	−4.1	−0.34
SD	10.9	6.9		
Females	92.9	80.4	12.5	1.02
SD	7.7	17.4		

[a] The difference score represents left ear performance minus right ear performance.

DISCUSSION

The purpose of the present study was to examine the reliability of a dichotic listening task using nonverbal stimuli. The results showed a satisfactory level of reliability in the task. In fact, the reliability estimate compares favorably with the values obtained with some of the most reliable verbal tasks (see Voyer, 1998). It is worth noting that this relatively high level of reliability was observed in spite of the practice effect observed between sessions, which was likely due in part to the absence of practice trials in the task. This suggests that the task developed by Bryden and MacRae (1989) provides a reliable means of assessing laterality for non-verbal material. It can therefore be added to a literature that is severely lacking in terms of studies examining the reliability of the laterality effects obtained in this type of tasks. However, the addition of familiarization trials to the procedure used here might further improve its reliability. In addition, the present study only examined one of many procedures that might be used in response production. It is likely that the results would be affected by this dimension (Bryden, 1978; Voyer & Flight, in press). Future work in this area should investigate the role of reporting strategy by comparing the results obtained here with what would be obtained with an ABX procedure (see Voyer & Flight, 2000) or a free recall procedure, for example.

The finding that the left ear advantage was significant in females but not in males is somewhat puzzling in view of the finding that the opposite pattern is typically reported (Voyer, 1996). However, it was also reported by Voyer (1996) that gender differences in laterality tend to be small. It is therefore not surprising that results that are contrary to the typical trend should be found. A more important issue concerns the fact that males as a group did not show a significant ear advantage. This could be interpreted as indicating that the reliability estimate obtained for the whole group is not a good reflection of males' laterality. However, data analysis showed that the reliability estimate is the same for both males and females. The results should therefore be interpreted as indicating that the task is reliable for both males and females, but that direction of laterality is more variable in males than in females. It is likely that strategy effects are involved here (Bryden, 1978). For example, it is possible that some of the males found it difficult to dissociate the emotional aspect of the stimuli from their verbal content. This interpretation is supported by the trend for a right ear advantage in males (see Table 1). Similarly, the words themselves might have an emotional load that may interact the emotional tone in uncontrolled way. Future research with this paradigm might benefit from the use of words in an unknown language or pseudowords. However, it is important to remember that the sample size used here is small. This suggests that the sex difference observed in the present study might be sample specific.

To conclude, the present study showed that nonverbal laterality effects can be measured with a reasonable level of reliability in both males and females. This suggests that the task used here should be considered a useful tool in the assessment of right hemisphere functions. More work is required to determine the influence of reporting strategy on the magnitude and reliability of the laterality effects observed in the present study.

ACKNOWLEDGMENTS

The study reported here was supported by a grant from the Natural Science and Engineering Research Council of Canada (NSERC) to D. Voyer and an NSERC Undergraduate Students Research Award to M. Rodgers.

REFERENCES

Bryden, M. P. (1982). *Laterality: Functional asymmetry in the intact brain.* New York: Academic Press.

Bryden, M. P. (1978). Strategy effects in the assessment of hemispheric asymmetry. In G. Underwood (Ed.), *Strategies of information processing* (pp. 117–149). London: Academic Press.

Bryden, M. P., & MacRae, L. (1989). Dichotic laterality effects obtained with emotional words. *Neuropsychiatry, Neuropsychology and Behavioral Neurology,* **3,** 171–176.

Kimura, D. (1967). Functional asymmetry of the brain in dichotic listening. *Cortex,* **22,** 163–178.

Schneider, W. (1995). *MEL Professional V2.0 user's guide.* Pittsburgh: Psychology Software Tools.

Steenhuis, R. E., & Bryden, M. P. (1989). Different dimensions of hand preference that relate to skilled and unskilled activities. *Cortex,* **25,** 289–304.

Voyer, D., & Flight, J. (in press). Gender differences in laterality on a dichotic task: The influence of report strategies. *Cortex.*

Voyer, D., & Flight, J. (2000). Reliability of a dichotic consonant-vowel pairs task using an ABX procedure. *Brain and Cognition,* **43,** 417–421.

Voyer, D (1998). On the reliability and validity of noninvasive laterality measures. *Brain and Cognition,* **36,** 209–236.

Voyer, D. (1996). On the magnitude of laterality effects and sex differences in functional lateralities. *Laterality,* **1,** 51–83.

This is doi:10.1006/brcg.2001.1425.

Colored Photisms Prevent Object-Substitution Masking in Digit-Color Synesthesia

B. M. Wagar, M. J. Dixon, D. Smilek, and C. Cudahy

Department of Psychology, University of Waterloo

For C, a digit-color synesthete, viewing a digit elicits a photism (an experience of a highly specific color), which is perceived as externally projected onto the digit. We used an object substitution paradigm to demonstrate the influence of C's photisms on her perception of digits. Object substitution refers to a form of masking that depends upon two critical factors: attention must be taxed (e.g., distributed among numerous distracters), and the mask must remain onscreen for a sufficient duration after the target has been removed. Results showed that for C, even under conditions that produced maximal object substitution in control participants, her endogenous coloring of target digits prevented object substitution. Thus, the present study clearly showed that C’s photisms influence her ability to detect digits.

INTRODUCTION

Digit-color synesthesia is characterized by a phenomenal experience: when shown a black digit, a digit-color synesthete will perceive the digit in color. For some digit-color synesthetes, the colors are experienced in external space. For example, when the digit-color synesthete C is shown a black 2, she perceives the digit as having a red overlay that adheres to, but completely covers, the shape of the 2. This type of subjective report suggests that under certain conditions the synesthetic colors (called photisms) might influence the ability to perceive black digits. Concretely, if black twos are seen as red, then they should readily stand out against a blue background, but should be hard to detect against a red background. Smilek, Dixon, Cudahy, & Merikle (submitted) required C to localize target digits in a visual search task and identify digits in a masking task. In both tasks the color of the background was manipulated. Results showed that C experienced more difficulty both localizing and identifying target digits when the background color was congruent with C’s photism color for the target (e.g., a black 2 against a red background), relative to when the background color was incongruent with C’s photism color for the target (e.g., a black 2 against a blue background). These results provide strong evidence that the photisms elicited by digits in digit-color synesthesia can influence the perception of digits.

Smilek et al. proposed that when a digit elicits a photism, the endogenously activated color information influences the perception of digits through reentrant pathways in the visual system. Research in the area of visual awareness suggests that what we perceive is a combination of what we *see* and what we *know*. Perceptual processes driven by external sensory signals send information concerning the nature of the stimulus to higher level cortical areas. This information activates concepts stored in memory, which form top-down ‘‘hypotheses’’ of what the stimulus might be. These hypotheses are sent back to the perceptual system via reentrant pathways. The perceptual system compares the features of these hypotheses with those present in the incoming bottom-up sensory information, and informs the conceptual system of the likelihood of each hypothesis—supporting some while discrediting others. Over a series of iterations, these interactions reach a stable state with a single representation reaching perceptual awareness.

The purpose of the present study was to corroborate Smilek et al.’s finding that C’s photisms could influence her perception of black digits. To do so, we used a masking paradigm designed to measure top-down influences on perception. In our adaptation of this paradigm, participants searched for a target digit embedded in a set of homogeneous digit distracters. The target was defined by four dots surrounding one of the digits in the display (see Fig. 1). Under certain conditions, the four dots also served to mask the perception of the target digit (i.e., the perception of the digit was overpowered by the perception of the four dots). Previous studies have shown that this conceptual level masking, called object substitution, occurs when attentional resources are distributed (e.g., over numerous distracters), and when the mask remains on-screen for a sufficient duration after the target has been removed (Di Lollo, Enns, & Rensink, in press; Enns & Di Lollo, 1997).

It is believed that object substitution, like C’s synesthetic perception, is the result of iterative reentrant processing in the visual system. When the necessary conditions for object substitution are met (e.g., long mask duration and numerous distracters),

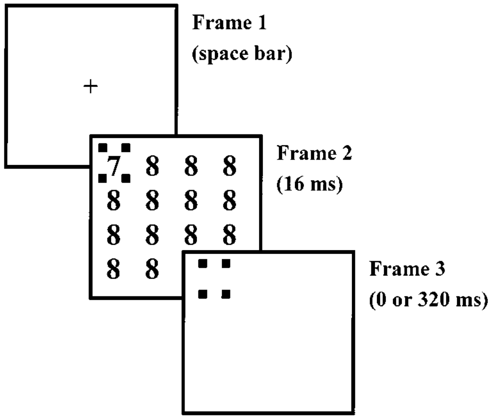

FIGURE 1

most people only see the mask, and not the target contained within it. Object substitution occurs because there is a mismatch between the initial reentrant representation of target-plus-mask (e.g., a 7 surrounded by four dots), and the ongoing lower level activity produced by the four dot mask alone. Basically, the brief presentation of the first display leads to the development of a conceptual level hypothesis that a compound element (target-plus-mask) is present in the display. When the target is removed and continued by the mask alone, a mismatch occurs between the top-down hypothesis of target-plus-mask and the bottom-up sensory information suggesting mask alone. Because the top-down hypothesis degrades rapidly, while the sensory input from the mask remains constant, as the perceptual system gradually accrues information to develop a percept the mask essentially ‘‘overwrites’’ the initial hypothesis, leading to the perception of the mask alone.

Our version of the object substitution paradigm involved black targets (e.g., a black 7) and black distracters (e.g., black 8’s). For nonsynesthetes we predicted that maximal object substitution would occur on trials such as that shown in Fig. 1, where attention is distributed over many distracters and the mask remains on screen for a substantial duration after the digits are removed. By contrast, when C is presented with this same display, she should perceive a yellow 7 embedded in a series of black 8s (the color of C’s photism for 8 is black). Because the target is endogenously colored and the distracters are all black, target–distracter similarity should be reduced for C, allowing the rapid development of a percept for the target digit that negates object substitution masking. Concretely, we predict that even in the condition that produces maximal object substitution in nonsynesthetes (many distracters and long mask durations), C will not show object substitution effects.

METHOD

Participants

Participants were C, a digit-color synesthete, and eight nonsynesthete students.

Stimulus Displays

The display sequence consisted of three frames. The first frame was a fixation cross (0.64°), presented in the center of the screen. The second frame contained the target, the mask, and the distracters, if any. The third frame contained only the mask (see Fig. 1 for an example). Targets and distracters were all digits (0.64° tall). For each trial there was a single target (2, 4, 5, or 7). Distracters were always 8's. The mask consisted of four dots (0.2° on each side), centered on the corners of a notional square (1° on each side). All stimulus elements were presented in black against a gray background. On any given trial, either a single digit (target only), or 16 digits (target plus 15 distracters), were displayed in the cells of a notional 4° by 4° matrix located in the center of the screen. The target was delineated by the four dots, which also served as the mask.

C's associated photism colors for the digit-color pairings are as follows: 2-red, 4-blue, 5-dark green, 7-yellow, 8-black. Because of these pairings, for trials with 15 distracters, C will synesthetically perceive a colored target embedded in black 8s. Controls, by contrast, will perceive a black target embedded in black 8s, hence making them more prone to object substitution.

Procedure

Participants were instructed that reaction time was not a factor and to respond as accurately as possible. C and the eight nonsynesthetes were each tested individually, seated 57 cm from the monitor. Each trial began with a fixation cross presented in the center of the screen. The participant initiated the onset of the second frame, which contained all stimulus elements, by pressing the space bar. After a fixed duration (16 ms), all elements except the mask were turned off, and the third frame containing the mask alone continued for a variable duration (either 0 or 320 ms). The participant then identified the target by pressing the appropriate key, guessing if unsure.

The design of the experiment was a 2 (set size) $\times$ 2 (mask duration) factorial. Each of the four targets appeared in each of the 16 matrix locations once per cell, yielding a total of 256 trials. The four cells of the 2 $\times$ 2 factorial were blocked, while target identity and location were randomized within each cell. The order [set size, mask duration (ms)] of the four blocks was fixed as follows: [1, 0], [1, 320], [16, 0], [16, 320].

RESULTS

It should be noted that in order for C to be considered significantly different from controls, her scores had to fall at least 3 standard deviation units below the control mean (i.e., z scores less than -3.0)—a conservative cutoff value used in conventional outlier analyses.

C vs Control Group

C's error data and the corresponding group means for each cell are shown in Fig. 2. A repeated measures analysis of variance (ANOVA) showed a substantial set size by mask duration interaction for the control group [$F(1, 7) = 28.937$. $MS_e = .002$, $p < .001$]. Controls made many more errors (mean = 19) in the key condition [i.e., large set size (16) and long mask duration (320 ms)], relative to all other conditions (mean error ranged from 2 to 6). This interaction indicated that the combination of large set size and long mask duration in the key condition provided sufficient conditions for object substitution among the nonsynesthetes. C was compared to the control group using z scores. As shown in Fig. 2, C's z scores were within the normal range on the conditions with no distracters and/or with short mask durations ($-.773 \geq z \geq -1.0378$, all within 3 z units of the control means). However, as predicted, in the key the condition where maximal object substitution occurred for nonsynesthetes (an

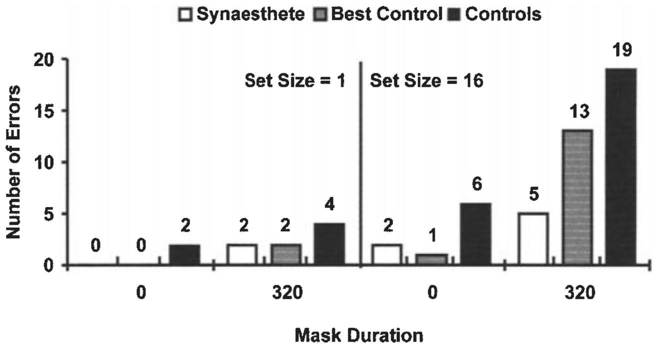

FIGURE 2

average of 19 errors), C was outside the normal range, making only 5 errors ($z = -3.3163$, $p < .004$). A subsequent independent samples ANOVA did not show a set size by mask duration interaction for C [$F(1, 252) = .116$. $MS_e = .034$, ns], providing further evidence that C's photisms prevented object-substitution masking.

C vs Best Control

The error data for the best control participant is also shown in Fig. 2. C and the best control did not differ in terms of the total number of errors made ($\chi^2(1) = 2.00$, ns). An independent samples ANOVA showed a significant set size by mask duration interaction for the best control participant [$F(1, 252) = 7.412$. $MS_e = .053$, $p < .007$], demonstrating a strong object substitution effect. C was compared to the best control using chi-squared analyses. As with the group data, C was no different than the best control on the conditions with no distracters and/or with short mask durations ($\chi^2(1) \leq .34$, all ns). However, C made significantly fewer errors than the best control in the key the condition where maximal object substitution occurred for the best control ($\chi^2(1) = 4.14$, $p < .05$).

DISCUSSION

The above results confirm that C's photisms influenced her perception of black digits. When presented with a display sequence containing a single target (2, 4, 5, or 7), embedded in a set of homogeneous distracters (8's), C's photism for the target prevented object substitution masking, even for long mask durations. When presented with the same display sequence, nonsynesthete controls exhibited the typical object substitution effect. These results clearly show that C's photisms influence her perception of digits.

Previous studies have shown that object substitution depends upon two critical factors: attention must be taxed (e.g., distributed among numerous distracters), and the mask must remain on-screen for a sufficient duration after the target has been removed (Di Lollo, Enns, & Rensink, in press; Enns & Di Lollo, 1997). In our adaptation of this paradigm, participants searched for a target digit embedded in a set of homogeneous digit distracters. Even though homogeneous distracters provide high distracter–distracter similarity, target–distracter similarity was also high (both targets and distracters were digits), thereby placing substantial demands upon the attention system of nonsynesthetes when searching through a 16 element display (see Duncan & Humphreys, 1989). When combined with a long mask duration, this demand on attentional resources allowed object substitution to occur. However, for C, her endogenous coloring of target digits drastically reduced target–distracter similarity, negating any influence of set size on attention, thereby precluding object substitution.

In summary, these results conclusively indicate that C's photisms influence her ability to detect digits. As such, this study contributes to a growing body of evidence from our laboratory showing that photisms elicited in digit-color synesthesia are able to affect the visual perception of digits.

REFERENCES

Di Lollo, V., Enns, J., & Rensink, R. (in press). Competition for consciousness among visual events: The psychophysice of reentrant visual pathways. *Journal of Experimental Psychology.*

Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review,* **96,** 433–458.

Enns, J., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science,* **8,** 135–139.

Smilek, D., Dixon, M., Cudahy, C., & Merikle, P. (submitted). Synaesthetic photisms influence visual perception. *Cognitive Neuropsychology.*

This is doi:10.1006/brcg.2001.1426.

Distinct Neurolinguistic Symptom Clusters in Alzheimer's-Type Dementia and Primary Progressive Aphasia

Chris Westbury,* Dan Bub,† and Howie Chertkow‡

**Department of Psychology, University of Alberta; †Department of Psychology, University of Victoria; and ‡Department of Clinical Neuroscience & Division of Geriatric Medicine, Sir Mortimer B. Davis Jewish General Hospital, McGill University*

Mesulam's (1982) report describing six patients with a slowly progressive aphasia without accompanying signs of dementia led to the recognition of a syndrome now known as Primary Progressive Aphasia (PPA). Many more patients have been described since Mesulam's description was published (see Westbury & Bub, 1997, for a review). However, the published literature is both unsystematic and incomplete, making it difficult to place the findings into a coherent theoretical framework. In addition, little previous work (see Mesulam, 1987) has specifically attempted to specify the difference between PPA and dementia of Alzheimer's type (DAT), although the two disorders are easily confused since many language deficits can masquerade on early presentation as memory or cognitive deficits. In this paper, the linguistic deficits of 11 PPA patients are analyzed, and contrasted with the linguistic deficits of a group of 11 DAT patients. Patients in both groups were tested using an extensive battery of language tests, the Psycholinguistic Assessment of Language Battery (Caplan & Bub, 1990; Caplan, 1992). We consider seven linguistic symptom clusters that differentiate the two groups.

INTRODUCTION

Mesulam (1982) published a report describing six patients manifesting a slowly progressive aphasia without accompanying dementia. Although such cases had been previously reported in the literature (see Poeck & Luzzatti, 1988), it was the publication of Mesulam's cases that led to the recognition of a syndrome characterized by pure aphasia without dementia, now known either as Mesulam's Syndrome or (more commonly) Primary Progressive Aphasia (PPA). The existence of PPA has been widely documented in the years since Mesulam's initial publication. Unfortunately, the existing published literature is unsystematic, consisting largely of single case reports (Westbury & Bub, 1997) that have relied upon a wide variety of different neuropsychological instruments, many of which were developed for other purposes (Caplan, 1992). These limitations have made it difficult to put the published results into a coherent theoretical psycholinguistic framework. It is perhaps due to this lack of coherence that little work (but see Mesulam, 1987) has been done that explicitly attempts to discover if the nature of the language deficits in PPA allow it to be easily distinguished from the most common dementing disorder, and the most important differential diagnosis, dementia of Alzheimer's type (DAT). The distinction between DAT and PPA was never made before 1981. Even today only a few medical practitioners who have a particular interest in language disorders routinely differentiate between the two disorders. Nevertheless, the distinction is an important one. DAT has a different cause and prognosis than PPA. Like PPA, DAT almost invariably affects language function. Since many language problems can masquerade, to both the patient and the psychologically-naive observer, as memory problems, a proper understanding of the role of language deficits in both PPA and DAT is vital to making the diagnostic distinction. This paper is an attempt to add to this understanding by examining the nature of language deficits in both disorders in systematic detail.

METHOD

Participants

Twenty-two subjects who were diagnosed with either probable PPA or probable DAT participated in this study. All eleven of the DAT and 10 of the eleven PPA patients were diagnosed by one of five referring neurologists who specialized in cognitive neurology and were familiar with the literature on both DAT and PPA. One PPA patient was referred with a tentative diagnosis of PPA by a neurologist who did not specialize in cognitive neurology. In that case the diagnosis was confirmed by the authors on the basis of both a neuropsychological assessment and the subject's performance on the test battery.

Some of the patients were seen as part of a clinical consultation. Others participated voluntarily for research purposes. All patients were informed as to the purpose of the study and understood that they would be given a set of computerized tests that had been designed to assess their language functioning.

Demographic data are presented in Table 1. The groups were about equally balanced on sex (PPA: Six males, five females; DAT: Five males, six females). There was no significant difference in years of education ($t(17.1) = 1.81$ $p > .05$). The PPA group was significantly younger than the DAT group ($t(20) = -2.26$ $p < .05$). There was also a significant difference in years since symptom onset ($t(20) = 2.6$; $p < .05$).

Materials

We used a computerized version of the Psycholinguistic Assessment of Language (PAL) battery (Caplan & Bub, 1990; Caplan, 1992; Westbury, 1995). This battery consists of 24 subtests (see Table 2). These tests had been normed on between 20 and 70 elderly normals. Each of the tests in PAL was presented in a consistent manner, using PsychLab software. The tests were run either a Macintosh desktop computer attached to a 14 inch Apple monitor, or a Macintosh Powerbook portable computer. Stimuli size did not vary with screen size. All visual stimuli were presented

TABLE 1
Demographic Data for the 22 Patients in this Study

PPA					DAT				
Name	Age	Sex	Edu.	Years since onset	Name	Age	Sex	Edu.	Years since onset
AB	72	F	10	4	NB	86	M	11	2
DM	57	M	16	6	AB	81	F	14	2
JD	69	M	8	1	DD	69	M	18	2
OD	79	F	14	1/12	EF	76	F	7	2
JH	59	M	16	7	IK	60	M	15	2
BH	74	F	11	1	DO	78	M	5	2
BL	70	F	16	3	MR	63	F	12	2
JL	70	M	16	8	DS	79	F	11	2
CM	50	M	16	3	JS	80	M	13	2
ES	79	F	12	5	RS	86	F	16	1
MW	66	M	12	3	YS	82	F	4	1
Average	67.7		13.4	3.7	Average	76.4		11.5	1.8
SD	9.1		2.9	2.7	*SD*	8.7		4.5	0.4

TABLE 2
The 24 Subtests of the PAL Battery

(A) Tests of word access
- (i) Nonsemantically mediated
  - (1) Phoneme discrimination (PD)
  - (2) Auditory lexical decision (ALD)
  - (3) Written lexical decision (WLD)
  - (4) Reading (READ)
  - (5) Repetition (REP)
- (ii) Semantically mediated
- Word Comprehension
  - (6) Auditory word picture matching (AWP)
  - (7) Written word picture matching (WWP)
  - (8) Semantic access (SEMANTIC)
- Word Production
  - (9) Oral Naming (O-NAME)
  - (10) Written Naming (W-NAME)

(B) Tests of word category access
- (i) Abstract words
  - (11) Auditory comprehension of abstract words (A-ABS)
  - (12) Written comprehension of abstract words (W-ABS)
- (ii) Affixed words
  - (13) Oral production of affixed words (O-PROD-AFF)
  - (14) Written production of affixed words (W-PROD-AFF)
  - (15) Auditory lexical decision of affixed words (ALD-AFF)
  - (16) Written lexical decision of affixed words (WLD-AFF)
  - (17) Auditory synonym judgment of affixed words (A-SYN)
  - (18) Written synonym judgment of affixed words (W-SYN)
  - (19) Auditory word picture matching of affixed words (AWP-AFF)
  - (20) Written word picture matching of affixed words (WWP-AFF)

(C) Tests of sentence-level processing
- (i) Sentence comprehension
  - (21) Auditory comprehension of sentences (A-COMP-SEN)
  - (22) Written comprehension of sentences (W-COMP-SEN)
- (ii) Sentence production
  - (23) Oral production of sentences (O-PROD-SEN)
  - (24) Written production of sentences (W-PROD-SEN)

in black and white. Auditory stimuli were recorded monophonically by a female speech therapist using 8 bit sound and played to the subject over external speakers.

Procedure

Subjects were screened informally with a simple test designed to ensure that they could hear the auditory stimuli and see the visual stimuli. No subjects were eliminated with this screening.

The tests were presented in a roughly standard order, although deviations from that order were sometimes required for practical reasons. Within most tests, the order of stimulus presentation was automatically randomized for each subject. The only exceptions were the sentence production tests, which were presented in a consistent order because they include two unscored ‘‘practice items’’ that must be given first.

Responses were either given verbally, by pointing, or (where required) by writing, and were typed in to the computer by the tester. Since competence rather than performance was the focus of the testing, subjects were informed that they could ask for any test item to be repeated as many times as they liked. They were also allowed to change answers immediately after giving them.

Testing required from 2 to 10 sessions that varied in length from 30 min to about 3 h. The total time spent testing a single patient varied between 5 and 15 h, due to wide variations in patient stamina and speed.

RESULTS

In order to facilitate comparisons between tests, each test result was transformed into a standard score (i.e., expressed in terms of the number of normal standard deviations below the normal average score for that test). A score was considered significantly low if it was two or more standard deviations below the normal average.

For the present analysis, tests were ranked according to ‘‘difficulty’’ within each group. The number of patients in each group who scored significantly low was used as a measure of that test’s difficulty. We ranked not only the global scores from every test, but also the scores from levels within each test, reflecting the fact that the definition of what constitutes a single test is a definition based largely on practical convenience rather than theoretical considerations. For example, we are as interested in a patient’s reading of low frequency words as high frequency words, even though participants read both as part of a single test. There are 129 such subtest items in the PAL battery.

For ease of presentation, in this paper we present the results in terms of seven identified symptom clusters. Each cluster reflects the results from several of the 129 items considered.

Cluster 1: Auditory Comprehension Deficits

Six of the eleven PPA patients had specific difficulty with parsing auditory input. Only one of the eight DAT patients showed evidence of a similar deficit. This difference in proportions is not significant (Fisher’s Exact Probability: .19). No significant differences were seen for repetition of either words ($t(17.0) = -0.88$, $p > .05$) or nonwords ($t(17.0) = -0.03$, $p > .05$). There was a significant group difference, in favor of the DAT group, on one of the tests that is most diagnostic of an auditory input deficit: phoneme discrimination ($t(12.8) = -2.18$, $p < .05$).

The evidence for a significant difference between PPA and DAT patients in auditory input processing is thus weak, but suggestive.

Cluster 2: Reading Deficits

Six PPA patients had clear reading deficits, and an additional four had more ambiguous deficits. Only one of the DAT patients showed a deficit in tests of reading, and that deficit was mild. The difference in proportions is not significant (Fisher's Exact Probability = .06). However, the trend is strong enough to be noteworthy, especially when we consider that the PPA patients were much higher functioning generally than the DAT patients. Since some evidence of reading deficits was seen in almost none of the DAT patients and in almost all of the PPA patients, there is again weak but suggestive evidence of a difference between PPA and DAT patients in reading.

Cluster 3: Deficits in Affixed Word Processing

Every one of the PPA patients had difficulty with comprehension of affixed words, with results from other comprehension tests suggesting that this difficulty was not secondary to a more general difficulty with word comprehension. In contrast, not one of the DAT patients showed an unambiguous problem with affixed word comprehension. This observation of differences at the individual level is confirmed by statistical differences at the group level, which showed that the PPA performed significantly worse on all of the six relevant tests (auditory lexical decision of affixed words: $t(10.9) = 2.63$, $p < .05$; written lexical decision of derived words: $t(13.7) = 2.65$, $p < .05$; auditory word-picture matching of derived words: $t(11.5) = 3.03$, $p < .05$; written word–picture matching of derived words: $t(15.2) = 2.13$, $p = .05$; written synonym judgment of derived words: $t(11.5) = 3.15$, $p < .01$).

No clear analogous dichotomy was seen at the individual level among the affixed word production tests. Six of the PPA patients had a production problem that could be specifically tied to word affixation, but at least three (and perhaps as many as five) DAT patients had the same problem (Fisher's Exact Probability = .71). However, at the group level there were significant differences in favor of the DAT group on both of the most directly relevant tests (oral production of affixed words: $t(10.8) = 3.02$, $p < .05$; written production of derived words: $t(12.0) = 3.18$, $p < .01$).

The evidence suggesting that affixation-specific deficits (especially in comprehension) are much more likely in PPA than DAT is quite strong.

Cluster 5: Naming Deficits

Eight of the eleven PPA patients (73%) had naming deficits. Two of these were clearly anomic (unable to name, but able to access semantic information), and the remainder were agnosic (unable to access semantic information). Six of the eight DAT patients (75%) had naming deficits, split evenly between those who looked anomic and those who looked agnosic. This close accordance (Fisher's Exact Probability = .65) does not suggest that there are obvious systematic differences in naming between the two groups. Further evidence against such a difference is found in the fact that there are no significant group differences in any of three most directly relevant tests or test sets: oral naming, written naming, and a set of tests assessing semantic access.

However, at least one difference does distinguish significantly between the two groups: the proportion of semantic errors to total errors made in naming. The difference between the two groups was analyzed by calculating the contribution made to the total errors within a group by each individual's semantic naming errors. The DAT patients made a larger proportion of semantic errors in both modalities than the PPA patients. (written naming: $t(7.4) = -2.6$, $p < .05$; oral naming: $t(7.6) = -3.9$, $p < .01$).

Cluster 6: Sentence Comprehension Deficits

All but two of the seven PPA patients who completed the auditory and written sentence comprehension tests scored more than two standard scores below normal globally on those tests. In contrast, only two of eight DAT patients who completed the task scored significantly low on auditory sentence comprehension, and only three of eight DAT patients scored globally low at written sentence comprehension.

The global scores for the two groups in oral sentence comprehension are significantly different ($t(12.5) = 2.19$, $p < .05$), but the score for written comprehension do not differ significantly ($t(11.2) = 1.0$, $p > .05$). None of the group differences on sentence types are significant (oral comprehension of semantically constrained sentences: $t(7.0) = 1.0$, $p > .05$; oral comprehension of semantically reversible sentences: $t(7.2) = 1.7$, $p > .05$; written comprehension of semantically constrained sentences: $t(13.7) = -0.03$, $p > .05$).

In sum, the evidence for a specific impairment in PPA of sentence comprehension is equivocal. Differences at the individual level are suggestive of a particular deficit, but the small number of patients in each group who completed these tests precludes the drawing of any definite conclusions.

Cluster 7: Deficits in Abstract Word Access

The differences between the two diagnostic groups in abstract word comprehension are the most striking of any differences documented in this study. Ten of eleven PPA patients scored below norms on the written version of the abstract word comprehension test. All eight PPA patients who completed the auditory version of the test scored significantly low. In contrast, not one of the seven DAT patients who completed the written version scored low, and only two of the eight patients who completed the auditory version scored low. These differences in proportion are statistically significant in the written modality only (written abstract word comprehension: Fisher's Exact Probability = .03; auditory abstract word comprehension: Fisher's Exact Probability = .13). The group differences are significantly different (auditory abstract word comprehension: $t(12.6) = 3.07$, $p < .01$; written abstract word comprehension: $t(13.2) = 5.16$, $p < .005$).

The large differences at both the individual and group level on scores of abstract word comprehension, together with the consistency across modalities within both diagnostic groups, strongly suggest that PPA affects abstract word comprehension in a way that DAT does not.

CONCLUSION

It is possible on the basis of the information presented in this study to design a very simple rule that will distinguish any PPA patient in this study from any DAT patient. Just knowing whether a patient scored significantly low on the test of written comprehension of abstract words would allow one to correctly classify all but one of the patients. It is equally clear that it would be a trivial matter to add another single condition to the rule that allowed one to classify all patients correctly, since that condition would only need to classify the single aberrant patient. Moreover, one could, with little effort, make up any number of similar simple rules, with not more than two or three conditions each, which would have an equal ability to classify every patient in this study correctly.

The problem of taxonomy, of course, is not merely to be able to classify all the patients in a single study, but to classify all possible patients. Moreover, it is preferable to do so using theoretically motivated criteria, rather than relying on post hoc descriptions of empirical observations. Only one neurologically buttressed functional difference was found: the difference in the number of semantic errors made in naming. That difference is grounded in a large body of work showing the role played by the prefrontal cortex (known to be damaged in DAT) in semantic access (Steinmetz & Seitz, 1991). However, many dissociations found—between abstract and concrete words, between morphologically simple and morphologically complex words, and between sentence level versus word level comprehension—are rooted in current neurolinguistic theory (Caplan, 1992), and were built in to the PAL Battery for precisely that reason.

Future work focusing on three functions—abstract word comprehension, affixed word processing in both modalities, and semantic access—is most likely to yield the information needed to reliably distinguish PPA from DAT patients as early as possible.

REFERENCES

Caplan, D. (1992). *Language: Structure, processing, and disorders.* Cambridge, MA.: MIT Press.

Caplan, D., & Bub, D. (1990). *Psycholinguistic assessment of aphasia.* Presented at American Speech and Hearing Association Conference, Seattle, WA.

Mesulam, M-M. (1982). Slowly progressive aphasia without any generalized dementia. *Annals of Neurology,* **11,** 592–598.

Mesulam, M-M. (1987). Primary progressive aphasia—Differentiation from Alzheimer's Disease. *Annals of Neurology,* **22,** 533–534.

Steinmetz, H., & Seitz, R. (1991). Functional anatomy of language processing: Neuroimaging and the problem of individual variability. *Neuropsychologia,* **29,** 1149–1161.

Westbury, C., & Bub, D. (1997). Primary progressive aphasia: A review of 112 cases. *Brain and Language*, **60,** 381–406.

Westbury, C. (1995). *A variation of forms: The cognitive neuropsychology of primary progressive aphasia.* Ph.D. dissertation, McGill University.

This is doi:10.1006/brcg.2001.1427.

Effects of Acoustic Degradation on Syntactic Processing: Implications for the Nature of the Resource System Used in Language Processing

Sasha Yampolsky,* Gloria Waters,* David Caplan,†
Melanie Matthies,* and Peter Chiu‡

**Boston University; †Massachusetts General Hospital; and ‡University of Cincinnati*

This study investigated the effect of noise masking on on-line syntactic processing. Ninety college students were tested on measures of working memory and on-line sentence comprehension. Subjects were divided equally into three listening conditions: no noise masking, −3 dB signal-to-noise ratio (S:N), −4.5 dB S:N. The auditory moving windows (AMW) paradigm was used to measure on-line sentence processing. In the AMW paradigm, subjects pressed a button for the successive presentation of each phrase in two types of sentences (syntactically simple and complex), and listening times were recorded for each phrase. Previous studies have shown that the verb in the more complex sentence type is the most capacity demanding portion of the sentence. Listening times were longer overall with increased noise masking, and listening times were longer overall at the verb of the harder sentence type. However, the increase at the verb was not larger with increased noise masking. All three groups showed similar effects of syntactic structure in the on-line data. The on-line syntactic effects were not due to problems in word recognition. Correlational analyses did not indicate a relationship between the increase in processing time at the capacity demanding region of the harder sentence types and any of the measures of working memory capacity in any of the three listening conditions. Results indicate that on-line sentence processing is not affected by noise masking if lexical access (e.g., word recognition) remains intact.

INTRODUCTION

There are numerous experimental studies indicating that sentence comprehension requires the allocation of processing resources, or working memory capacity (Wanner & Maratsos, 1978; Ford, 1983; Frauenfelder, Sequi, & Mehler, 1980; Holmes & O'Reagan, 1981). However, there is considerable controversy concerning the nature of processing resources used in sentence comprehension. Some researchers claim that there is a general verbal working memory capacity used for sentence comprehension as well as for many other verbal tasks (King & Just, 1991; Just & Carpenter, 1992). Others argue that determining the meaning of a sentence based on the syntactic structure requires a specialized working memory system, one that is separate from that used in other verbal tasks (Caplan & Waters, 1999).

Previous studies have investigated the effects of a memory load, such as remembering a sequence of digits for immediate recall, on sentence comprehension to simulate a reduction in processing capacity (King & Just, 1991; Blackwell & Bates, 1995; Caplan & Waters, 1999; Waters, Caplan, & Hilderbrandt, 1987). If there is only one mental capacity for all verbal tasks, then a memory load should interfere with syntactic processing ability. Thus far, results do not suggest that this is the case.

Background noise masking has been shown to cause reductions in working memory span (Pichora-Fuller, Schneider, & Daneman, 1995). Therefore, taxing working memory via noise masking provides another interesting test of the two views. Previous studies of sentence comprehension in noise masking utilized off-line tasks only, such as gathering accuracy data or end-of-sentence reaction times (Kilborn, 1991; Pichora-Fuller, Lloyd, Dillon, & Kirson, 1998; Dick, Bates, Wulfeck, & Dronkers, 1999). Kilborn and Dick et al. claimed that noise masking affected the syntactic analysis of sentences. However, their studies did not rule out the possibility that interference may have occurred at the lexical level. Pichora-Fuller et al. also found that sentence comprehension accuracy decreased with increased noise masking levels. However, the pattern of errors was the same across noise masking conditions, with no differential effects on syntactically more complex sentences.

On-line experiments measure a person's processing time phrase-by-phrase, throughout sentence presentation (Wanner & Maratsos, 1978; Ford, 1983; Frauenfelder et al., 1980; Holmes & O'Reagan, 1981; King & Just, 1991; Waters & Caplan, 1999). On-line studies have shown that listening or reading times are longer at certain regions of a sentence (e.g., the verb in syntactically more complex cleft-object sentences, when compared with the verb in syntactically simpler cleft-subject sentences), providing evidence that these regions require more processing resources. An important innovation of the present study is the use of an on-line technique to test sentence processing under noise masking. Additionally, previous studies did not include direct measures of word recognition. It is important to investigate whether impairments in word recognition may account for impairments in sentence comprehension that are seen.

In summary, the goals of this study were: (1) to determine whether noise masking negatively affects on-line syntactic processing, (2) to determine if effects of noise masking seen in sentences can be accounted for by interference at the lexical level, and (3) to investigate the relationship between working memory capacity and sentence processing efficiency in the presence of noise masking.

METHODS

Subjects

Ninety university students, age 18–26, participated in the study. Participants were divided into three groups, with each group receiving one level of Signal-to-noise Ratio (SNR): No noise masking ($n = 30$), -3 dB ($n = 30$) and -4.5 dB ($n = 30$). The presentation of the stimuli was counterbalanced in a Latin Square design.

On-line Measure of Sentence Processing Efficiency

The Auditory Moving Windows paradigm (Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996) was used as the on-line sentence processing measure. Participants heard sentences that were segmented into phrases, and they pressed a button on a box interfaced with the computer to hear each successive phrase. Reaction times for each phrase were recorded. Half of the sentences were semantically acceptable and half were unacceptable. Participants made acceptability judgments (by pressing a 'yes' or 'no' button) as soon as they heard the last phrase. The importance of making acceptability judgments is twofold. The acceptability judgement requirement ensured that participants were paying attention to and performing the task as accurately as possible, rather than just pacing their way through the sentences. Also, accuracy of acceptability judgments was used as the off-line measure of sentence comprehension.

The stimuli consisted of 52 semantically acceptable and 52 semantically unacceptable sentences divided equally among two sentence types. Sentences were developed in pairs (the more complex cleft-object, and the simpler cleft-subject), and contained the same words in different order. Examples of each sentence type (with each phrase that participants heard separated by /) are found in Table 1.

The average content word sound-pressure-level (SPL) presentation was 68 dB SPL for the signal-only stimuli and 57 dB SPL for the signal + noise stimuli. All stimuli were presented at suprathreshold levels. At these levels, which are typical conversational speech levels, variations between 68 and 57 dB SPL will have no significant effect on intelligibility.

TABLE 1
Examples of Sentence Stimuli

Acceptable	CS: /It was/[intro] /the fire/[NP1] /that/[pro] /injured/[V] /the policeman/[NP2] /on the highway/[AdjP]
	CO: /It was/[intro] /the policeman/[NP1] /that/[pro] /the fire/[NP2] /injured/[V] /on the highway/[AdjP]
Unacceptable	CS: /It was/[intro] /the man/[NP1] /that/[pro] /delighted/[V] /the camera/[NP2] /in the film/[AdjP]
	CO: /It was/[intro] /the camera/[NP1] /that/[pro] /the man/[NP2] /delighted/[V] /in the film/[AdjP]

Word Recognition

Participants were readministered the Auditory Moving Windows task at least 2 weeks after the initial administration (in the same Noise condition as the initial presentation). Participants self-paced through the sentences and repeated each phrase after it was presented. Word recognition was determined by the number of correctly repeated content phrases.

Noise Masking

The noise masking was a digitally generated and uniformly distributed white noise that was convolved with the long-term average frequency spectrum of all sentences used in the speech corpus. This speech-shaped noise thus had significant energy only in the frequency region of the target sentences. The masking waveform used for each target sentence was the speech-shaped noise modulated with the envelope of the target sentence waveform.[1] The target sentence and masking waveform were checked for peak clipping and then scaled to a specified signal-to-noise-ratio. The signal-to-noise (S:N) levels in the present study were chosen on the basis of a pilot study such that recognition of the content words was below ceiling but still acceptable. Results of the pilot study revealed that 80% recognition accuracy (of each content word) was achieved at −4.5 dB S:N and 91% at −3 dB S:N. These signal-to-noise ratios were chosen for the study.

Measurement of Verbal Working Memory

Participants were administered the working memory tasks below. There were five trials at each span size, and the highest span size was 8. All subjects began each test at span 3. If the subject did not achieve 100% recall accuracy at span 3, then he/she was administered the previous span list in its entirety. Therefore, a basal span, which is the lowest span list recalled with 100% accuracy, was established for each subject. Higher spans were then administered, and the test was discontinued when the participants achieved 2/5 (or less) correct trials at a particular span. The score was the highest span at which the participant achieved 3/5 trials correct (regardless of performance on lower spans). If the participant achieved 2/5 on the last span administered,.5 was added to the span score.

1. *Alphabet Span:* Subjects repeated a series of words after rearranging them in alphabetical order.
2. *Subtract-2 Span:* Subjects repeated a random sequence of digits after subtracting 2 from each digit.
3. *Listening Span:* An adaptation of the Daneman and Carpenter task (Waters & Caplan, 1996) was used. Subjects were presented with sets of acceptable and unacceptable cleft-subject sentences auditorally. Subjects heard a sentence and made an acceptability judgment by pushing a button on the button box. The next sentence in the set was presented immediately after the accuracy judgment. An asterisk appeared on the screen after the last sentence in the series to indicate that the subject must recall each sentence-final word in the correct serial order.

[1] Using a flat envelope noise potentially undermasks higher amplitude phonemes and overmasks phonemes with lower amplitude. The modulated noise keeps the SNR more constant from instance to instance through the presentation of the stimulus.

Correlational analyses between performance on the working memory tasks and on measures of sentence processing were carried out.

RESULTS

On-line Measure of Sentence Processing Efficiency

Listening times for each phrase in acceptable cleft-subject (CS) and cleft-object (CO) sentences for which participants made a correct acceptability judgment were analyzed. Figure 1 shows the mean listening time for the three noise conditions at each phrase for cleft-subject and cleft-object sentences. Overall listening times for the −3 dB S:N and −4.5 dB S:N conditions were longer than those for the no-noise condition. There was no significant difference in listening times between the two noise masking levels.

Unexpectedly, listening times were longer at NP1 in cleft-subject than in cleft-object sentences. One possible source of this effect is that the mean word frequency was higher for NP1 in cleft-object sentences than in cleft-subject sentences. However, since the difference in frequency between the two sentence types was not significant, it is not clear whether this is the source of this effect. More importantly, as expected, listening times were longer at the capacity demanding portion of the harder sentence type (V in CO sentences) in all listening conditions. However, the increase seen at the verb in the harder (CO) sentence type was not larger in the noise masking conditions.

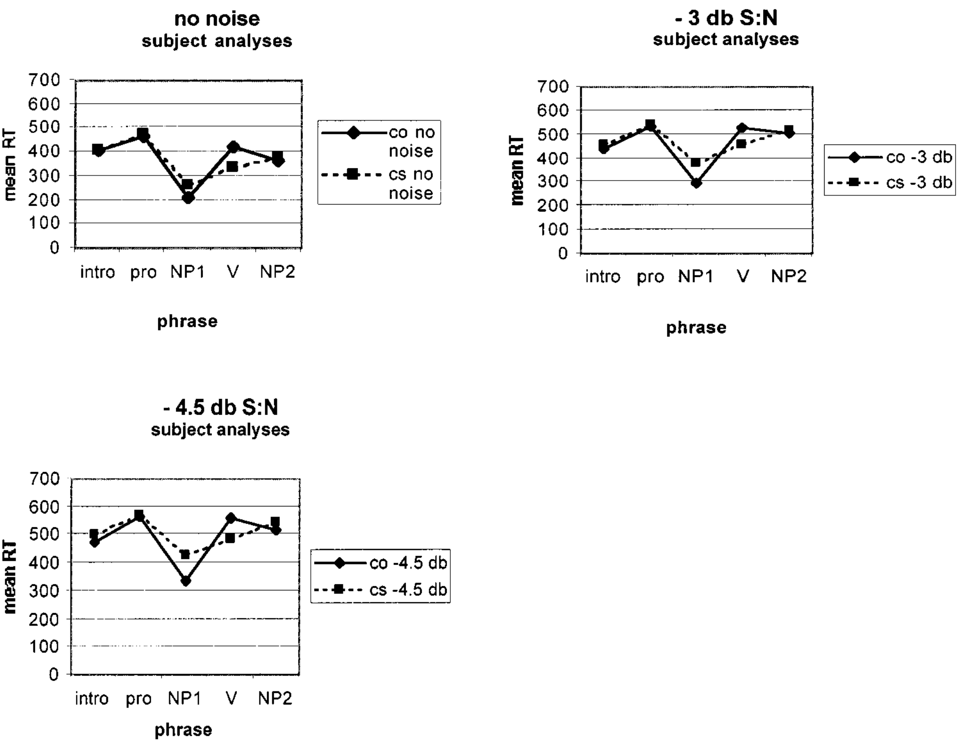

FIG. 1. Mean Listening time (ms) for the three noise conditions, at each phrase for CS and CO sentences.

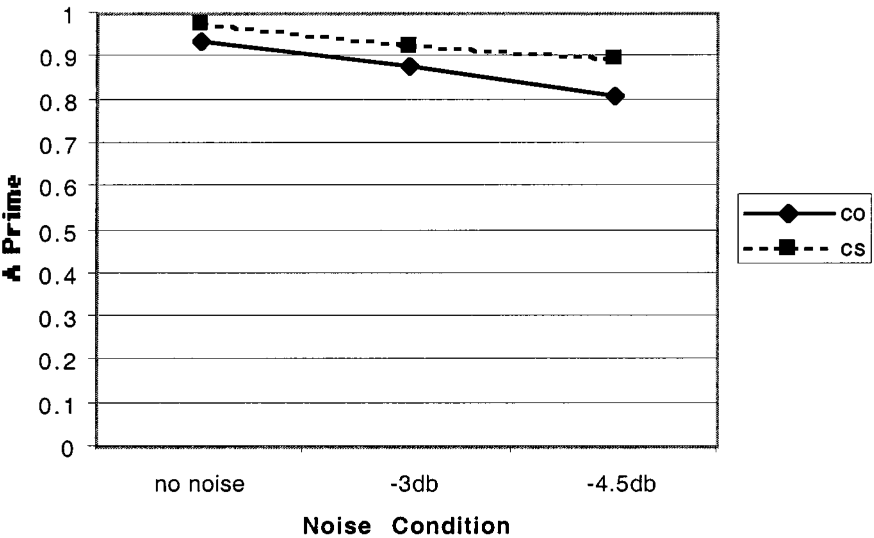

FIG. 2. Mean A′ scores for CS and CO sentences in the three noise conditions.

Off-line Measure of Sentence Processing

Accuracy in making sentence acceptability judgements at the end of each sentence was examined. Judgment accuracy was measured by A-prime (A′), which considers the number of times the subject judged ''correct/yes'' for plausible sentences and ''incorrect/no'' for implausible sentences.[2] This is considered to be a more sensitive measure of accuracy than percent correct, controlling for a strategy of just answering 'yes' (Linebarger, Schwartz, & Saffran, 1983).

Figure 2 shows the mean A′s for cleft-subject and cleft-object sentences in the three noise conditions. Mean A′s in all listening conditions were above .8 for both sentence types. This shows that participants in the noise masking conditions could reliably discriminate acceptable from unacceptable sentences. A′s were higher for the no-noise condition than for the −3 dB S:N and the −4.5 dB S:N conditions, and those for the −3 dB S:N condition were higher than those for the −4.5 dB S:N condition. A′s were higher for the syntactically simpler CS sentences than for the more complex CO sentences in all listening conditions. For the CS sentences, there was a significant difference in A′s between the no-noise and the −3dB S:N conditions but not between the −3dB S:N and −4.5 dB S:N conditions. For the CO sentences, there were significant differences in A′s between all three listening conditions. The results suggest that the accuracy of making plausibility judgments decreases disproportionally for the harder sentence types at −4.5 dB S:N.

Relationship between Sentence Processing and Noise and Working Memory

Difference scores were calculated that reflected listening times at the most capacity demanding portion of each sentence (this was the verb in CO minus the verb in CS). Difference scores that provided an index of increased difficulty in making plausibility judgments for the more complex sentence type (CO) were also calculated. This was the A′ in CO minus the A′ in CS sentences. There was no relationship between working memory and the increase in listening time at the verb in CO sentences in

[2] $A' = .5 + \{[((y - x) * (1 + (y - x))]/[4 * (y * (1 - x)))]\}$, where y = correct yes/total number of plausible sentences; x = incorrect yes/total number of implausible sentences.

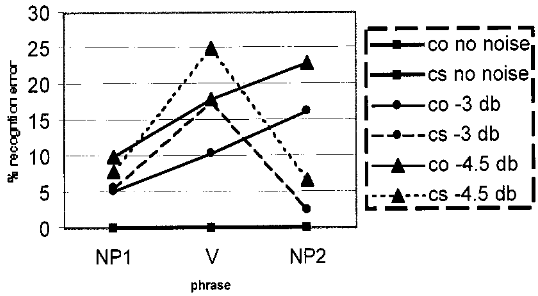

FIG. 3. Mean percent recognition errors for the three noise conditions at each content phrase for CS and CO sentences.

any listening condition. Correlations between the working memory span measures and plausibility judgments also showed no systematic relationship.

Word Recognition Analyses

Word recognition errors were calculated for each content phrase (NP1, V, NP2) for cleft-subject and cleft-object sentences in the three noise conditions. Figure 3 shows the mean recognition errors for the three noise conditions at each content phrase for cleft-subject and cleft-object sentences. Participants did not make any word recognition errors in the no-noise condition. Word recognition errors in the −4.5 dB S:N condition were higher than in the −3 dB S:N condition. For CS sentences, the percent recognition errors was higher on the verb than on any other content word. For CO sentences, percent recognition errors was higher on NP2 than on any other content word. The words on which the subjects made the most recognition errors corresponded to the middle word in each sentence. It is important to note that the middle content word is *not* the capacity-demanding region of the harder sentence type. The findings may be explained by sentential/context effects upon word recognition. In the present stimuli, the phrase immediately following ''It was the NP1 that __ '' can either be a noun or a verb, depending on the sentence type being presented. This is the middle content word, and maximum uncertainty to word identity occurs here. This may cause recognition difficulty. This phenomenon has been encountered in the audiological literature (Bilger, Nuetzel, Rabinowitz, & Rzeczkowski, 1984) and has been demonstrated in a widely used clinical test of word recognition under noise masking (Speech Perception In Noise (SPIN) test).

Relationship between Sentence Processing in Noise and Word Recognition

There was no relationship between percent recognition error and listening time (at the content phrases, NP1, V, NP2) in either noise masking condition. These results indicate that longer listening times are not due to word recognition impairments. Taken together, the word recognition and correlational data do not indicate a relationship between word recognition and on-line sentence processing in conditions of increased noise masking.

DISCUSSION

The main finding of the study was that listening times under increased noise masking were not disproportionately longer at the capacity demanding region of the harder sentence type. Self-paced listening showed the same pattern of local increases in syntactic processing load in all listening conditions. Results also indicated that measures of working memory span were not related to on-line measures of sentence processing efficiency in adverse listening conditions. There was a relationship between noise masking and syntactic complexity in end-of-sentence plausibility judgment data, such that judgment accuracy decreased disproportionally for the more complex cleft-object sentences with increased noise masking. This may reflect participants reviewing their memory of the sentence in order to be confident of their decision about each thematic role (who did what to whom). Results of word recognition analyses indicate that the pattern of increased processing time at the capacity demanding region of the more complex sentence type is not due to problems in word recognition. In summary, noise masking has been previously shown to be a form of extrinsic load that can cause reductions in verbal working memory capacity. The results therefore provide evidence that on-line syntactic operations are at least in part separate from the general verbal working memory system.

ACKNOWLEDGMENTS

The authors acknowledge support from the National Institute of Aging and the Dudley Allen Sargent Research Fund.

REFERENCES

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research*, **27**(1), 32–48.

Blackwell, A., & Bates, E. (1995). Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience,* **7**(2), 228–257.

Caplan, D., & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences,* **22**(1), 77–126.

Dick, F., Bates, E., Wulfeck, B., & Dronkers, N. (1998). Stimulating deficits in the interpretation of complex sentences in normals under adverse processing conditions. *Brain and Language,* **65,** 57–59.

Ferreira, F., Henderson, J., Anes, M., Weeks, P., & McFarlane, D. (1996). Effects of lexical frequency and syntactic complexity in spoken language comprehension: Evidence form the auditory moving-window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **22**(2), 324–335.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior,* **22,** 203–218.

Frauenfelder, U., Segui, J., & Mehler, J. (1980). Monitoring around the relative clause. *Journal of Verbal Learning and Verbal Behavior,* **19,** 328–337.

Holmes, V. M., & O'Reagan, J. K. (1981). Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior,* **20,** 417–430.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review,* **99**(1), 122–149.

Kilborn, K. (1991). Selective impairment of grammatical morphology due to induced stress in normal listeners: Implications for aphasia. *Brain and Language,* **41,** 275–288.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing. The role of working memory. *Journal of Memory and Language,* **30,** 580–602.

Linebarger, M., Schwartz, M., & Saffran, E. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition,* **13,** 215–225.

Miyake, A., Emerson, M., & Friedman, N. (1999). Good interactions are hard to find (commentary). *Behavioral and Brain Sciences,* **22**(1), 108–109.

Pichora-Fuller, K., Lloyd, L., Dillon, L., & Kirson, R. (1998). *Effects of sentence type and background* *noise on comprehension by listeners with low- or high-span working memory*. Poster Presented at Cognitive Aging Conference.

Pichora-Fuller, K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America,* **97**(1).

Wanner, E., & Maratsos, M. (1978). An ATN approach to comprehension. In M. M. Halle, G. Miller, & J. Bresnan (Eds.), *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press.

Waters, G., Caplan, D., DeDe., & Kemtes, K. (submitted). The measurement of working memory capacity in young and elderly subjects.

Waters, G., Caplan, D., & Hilderbrandt, N. (1987). Working memory and written sentence comprehension. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 531–555). London: Erlbaum.

This is doi:10.1006/brcg.2001.1428.

Diurnal Time Courses in Psychomotor Performance and Waking EEG Frequencies

C. Lafrance,* J. Paquet,* and M. Dumont*,†

**Laboratoire de chronobiologie, Hôpital du Sacré-Coeur de Montréal, Montréal, Québec, Canada; and †Département de psychiatrie, Université de Montréal, Montréal, Québec, Canada*

The present study examined the similarities between the diurnal time courses in the waking EEG activity and the psychomotor performance. The aim was to verify if some ongoing changes in the excitability of the cortical nerve cells across the daytime could facilitate the sensorimotor processing. EEG recordings and performance results for the Four Choice Reaction Time Test (FCRTT) were obtained every two hours, from morning to late evening period, in 8 young normal subjects (21.3 ± 0.5 years). ANOVAs were used to verify the presence of diurnal variations in the two measures. Nonparametric correlations were obtained to test the similarity between the changes across the day in the two measures. Three EEG frequency bands (delta, sigma, and beta1) and the reaction time measures varied across the daytime. The changes in the sigma (12.00–13.75 Hz) and the beta1 (14.00–19.75 Hz) bands were similar to the diurnal variations in the reaction time measures. It is suggested that the changes in the sigma and the beta1 bands may facilitate the processing of the sensorimotor treatment.

INTRODUCTION

Alertness refers to the capacity to maintain states of cortical and behavioral wakefulness (Goldenberg, 1988). Perceptual and cognitive tests have been developed to evaluate behavioral wakefulness. These tests generally consider speed and accuracy of performance (% of errors, gaps) as reflecting the ability of the individuals to process information. They presuppose that the better the performance efficiency, the higher the alertness level.

Physiologically, electroencephalography (EEG) activity reflects the synchrony between the discharges of the cortical neurons (Klimesch, 1999). Cortical wakefulness is characterized by high excitability of the nervous cells and assures processing of cortical information. The quantitative analysis of the waking EEG activity, by measuring the changes in the synchrony of the cortical neurons, may reflect the physiological processes underlying the processing of information.

In the absence of sleep deprivation, EEG activity and performance efficiency vary across the daytime. A few studies have examined the time course of daytime waking EEG frequencies in normal individuals (Cacot et al., 1995; Lorenzo et al., 1995;

Cummings et al., 2000; Lafrance et al., 2000). Diurnal variations were found throughout the full EEG spectrum, with maximum activity occurring at a later time for higher frequencies. Variations among times of day were examined for both perceptual and cognitive performance in recent studies (Lorenzo et al., 1995; Lenne et al., 1998; Owens et al., 1998; Owens et al., 2000). A time of day effect for the reaction times were found, with an improvement across the daytime hours (Owens et al., 1998; Lenne et al., 1998; Owens et al., 2000). The steady daytime decrease in measures of reaction time was also observed in a study which used counterbalanced testing times, i.e. in which data sampling hours were scheduled among different times of day for different subjects (Lenne et al., 1997). This suggests the presence of true diurnal variations and not a practice effect on reaction time measures. However, the time course of performance accuracy (% of errors, gaps) differed in function of the performance test used, or simply revealed no variation over the times of day.

Several studies have examined the waking EEG oscillations during cognitive performance in normal individuals. Most of these reports claimed that good performance at memory, attention and reaction time tasks is associated with an increase in the theta activity and a decrease in the alpha frequency band (Klimesch et al., 1996; Klimesch et al., 1998; Doppelmayr et al., 2000; for a review, see Klimesch et al., 1999). Other studies have demonstrated a link between the fast cortical activity (beta and gamma bands) and the cognitive processing, such as in sensorimotor processing (Classen et al., 1998; Haig et al., 2000). However, these protocols did not include repeated sampling of EEG activity and performance across the day. The similarities between the daytime courses of the two measures have not been investigated.

In the present study, the waking EEG activity and reaction time (RT) performance were recorded from morning to late evening hours. Three questions were asked:

1. In the absence of sleep deprivation, what are the time courses in EEG bands and performance measures during the daytime?
2. Are there EEG frequency bands which follow temporal patterns typical of psychomotor performance over the times of day?
3. At the individual level, is there an association between the diurnal changes in the waking EEG frequency bands and the diurnal changes in the performance?

MATERIALS AND METHODS

Subjects

Eight subjects (2 men, 6 women) aged 20 to 24 years old (21.3 years $\pm$ 0.5) were enrolled in the study. They had no sleep or vigilance complaints, no extreme scores of morningness/eveningness (Horne & Ostberg, 1976), and no history of psychiatric/neurological disorders. Subjects did not report using drugs and medications, were nonsmokers, and refrained from alcohol and caffeine intake. They signed a consent form and were paid for their participation.

Procedures

Prior to the laboratory part of the study, subjects maintained a regular sleep schedule (4 days, bedtime/waketime: 00:00–08:00 h). Alertness and performance measures were administered during the daytime following a night of polysomnography. Throughout the testing day, the subjects remained in their room in dim illumination ($<$ 30 lux) and pursued quiet activities.

Alertness and performance measures. Two measures were used to evaluate levels of alertness and performance: waking EEG activity recordings and the Four Choice Reaction Time Test (FCRTT) (Wilkinson & Houghton, 1975). The measures were conducted seven times a day, every two hours, on even hours (10:00 to 22:00) for the waking EEG and on odd hours (11:00 to 23:00) for the FCRTT. The standard procedures used to record the waking EEG activity (continuous focusing on a target during 2 min, avoiding body movements) may induce some kind of relaxation. Conversely, the execution of the FCRTT in itself causes an increase in the levels of alertness. The measures were separated by a one-hour interval to avoid overlapping effects.

The EEG was recorded from four derivations (C3/A2, C4/A1, O1/A2, O2/A1) in addition to the electromyogram (EMG) and the electroocculogram (EOG). Signals were recorded on paper with an 8-channel Grass polygraph (EEG sensitivity 7.5 μV/mm, bandpass 0.3–90 Hz, speed 15 mm/s) relayed to a personal computer, digitized at 256 Hz and filtered with a cutoff frequency at 64 Hz (digital filter). One out of every two points were stored on disk (128 Hz). Each recording lasted for 2 min during which the subjects had to keep their eyes open and fixed on a target. Epochs contaminated by artifacts (e.g., eye movements, muscular, and cardiac activity) were rejected. For each recording session, 9- to 15 4-s epochs (32–60 s) were selected (C3/A2 derivation), then subjected to amplitude spectral analysis (commercial software, RHYTHM, Stellate Systems, Montréal, Canada) with Fast-Fourier Transform (resolution of 0.25 Hz, cosine window smoothing). Absolute amplitudes (μV/Hz) were added within six bands: delta (0.75–3.75 Hz), theta (4.00–7.75 Hz), alpha (8.00–11.75 Hz), sigma (12.00–13.75 Hz), beta1 (14.00–19.75 Hz), and beta2 (20.00–31.00 Hz). The values of 4-s epochs were averaged across each recording session.

The FCRTT is a reaction time test where the subjects are instructed to respond as quickly and accurately as possible, that lasts 10 min, and provides no feedback on performance level. Results include mean reaction times (RT), percentage of errors (%err), and total number of gaps (reaction times $>$ 1 s). Three practice trials were given on the evening of admission in the laboratory.

Statistical Analyses

The first set of analyses tested temporal variations in EEG activity and performance (RT, %err, gaps) among the times of day. For each subject, the results of EEG amplitude, RT, %err and number of gaps obtained at each time of day were transformed into a *z* score calculated over the 7 sessions. For the performance measure, a global performance *z* score (Lafrance et al., 1998) was obtained by summing the *z* scores for the RT and the %err for each subject at each time of day. This composite *z* score reflects the general ability of the subjects to cope with all the demands of the task and is independent of the strategy favored (e.g., increased speed in the cost of more errors or decreased speed with less errors). One-way ANOVAs for repeated measures (Huynh–Feldt corrections) were conducted on the *z* scores for EEG and performance measures. Significant ANOVAs were followed by post-hoc comparisons (Tukey LSD).

Correlations were used to compare the temporal patterns between the EEG and the performance (Markowitz et al., 1988). The analyses were conducted on the results which displayed significant temporal changes. The subjects scores (transformed in *z* scores) were first averaged for each time of testing. Spearman rank-order correlations were used to compare the 7 averaged values representing temporal changes in specific EEG frequency band and in performance efficiency.

The last statistical analysis tested for a possible similarity between the two measures at the individual level. For this purpose, a sign test was applied on individual coefficients of correlation (Spearman rho) obtained for each pair of variables (Monk et al., 1997; Lafrance et al., 2000). Coefficients of correlation were calculated on individual *z* scores for each variable. The sign test is useful to test whether two variables tend to show either a positive or a negative correlation in a significant number of subjects.

RESULTS

Diurnal Variations

The delta and the beta1 frequency bands of the waking EEG showed significant diurnal variations ([$F(6, 42) = 2.49$ $p < .04$]; [$F(6, 42) = 2.88$, $p < .02$], respectively), and the changes in the sigma band approached significance [$F(6, 42) = 2.17$, $p < .08$]. The delta band showed high activity values at 12:00 and 16:00 h and low activity values at 10:00 and 22:00 h ($p < .03$). For the beta1 band, the activity values obtained at 14:00, 16:00, 20:00, and 22:00 h were higher than those obtained during the first recording (10:00 h), and activity values obtained at 16:00 h were higher than that of 12:00 h ($p < .02$). The sigma band showed lower values at 10:00 h when compared to 16:00, 18:00, and 20:00 h and the minimum obtained at 12:00 h was

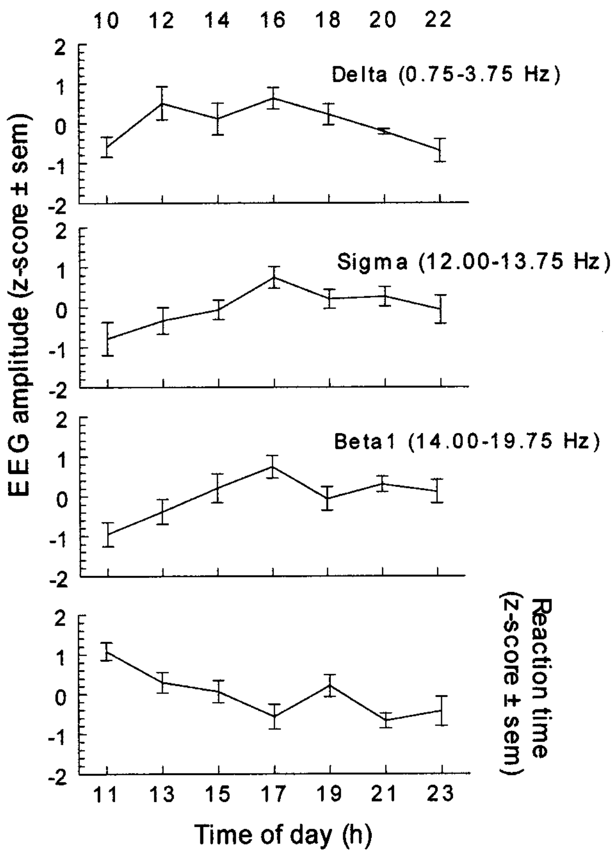

FIG. 1. Time course of the EEG frequency bands and the reaction times from the morning to the late evening. For each subject, data were transformed into z-scores calculated over the 7 times of day. Each point represents the mean of 8 subjects (± sem).

different from the value reached at 16:00 h ($p < .04$ for all comparisons). Variations in the theta, alpha and beta2 bands were not significant ($p > .25$).

The reaction time was the only performance measure to show changes across the day ($F(6, 42) = 4.21$, $p < .002$). The percentage of errors showed a slight increase across the day, but the changes were not significant ($p = .19$). The number of gaps and the global performance scores did not vary over the daytime ($p > .60$). The mean z score value for the reaction time measures obtained at 11:00 h was higher than those obtained from 15:00 to 23:00 h. Reaction time value reached at 13:00 h was longer than the values obtained at 17:00 and 21:00 h. Finally, the increase in the measure of reaction time at 19:00 h differed from the subsequent lower values obtained at 21:00 h ($p < .05$ for all comparisons). The changes across the day in EEG amplitude and reaction time measures are illustrated in Fig. 1.

Group Correlations between Temporal Patterns

There was a negative correlation between variations in the sigma band and the beta1 band and the reaction time measures ($R = -0.86$, $p < .02$; $R = -0.93$, $p < .003$, respectively). However, the correlation between the changes in the amplitude of the delta band and reaction time measures was not significant ($R = -0.04$, $p = .94$). Figure 2 illustrates the diurnal changes of the paired measures that showed significant correlations.

Intra-individual Correlations between Temporal Patterns

In 6 of the 8 subjects, there was a negative association between the sigma and the beta1 bands and the RT over the daytime. However, the sign test did not reach significance ($p = 0.14$ for the three associations).

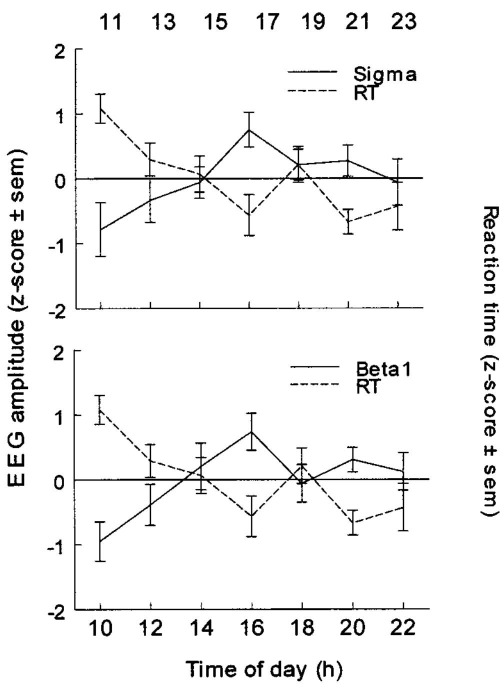

FIG. 2. Representations of the Spearman rank-order correlations obtained between the EEG bands and the reaction times. Correlations were calculated over the 7 mean points (± sem).

DISCUSSION

The waking EEG activity showed significant changes across the daytime but in comparison to previous reports (Cacot et al., 1995; Lorenzo et al., 1995; Cummings et al., 2000; Lafrance et al., 2000), less frequency bands displayed significant variations. This may possibly be due to the small number of participants in the present study. However, the frequency bands which displayed changes in amplitude were consistent with earlier conclusions: the more rapid the frequency band, the later its peak of activity.

Reaction times were longer during the morning hours in comparison to the evening hours. These changes were consistent with those reported in other protocols (Lenne et al., 1998; Owens et al., 1998; Owens et al., 2000), including those which controlled for the practice effect (Lenne et al., 1997). This suggests the presence of a true diurnal improvement in the time to respond and not only a practice effect. In addition, there was no significant decrease in the percentage of errors. To the contrary, the errors slightly increased from the morning to the late evening, as shown by group data. If there was a strong practice effect over the day, it would probably have improved all aspects of performance efficiency.

It is unexpected that, even with the significant decrease in reaction times, global performance (RT + %err) did not improve across the daytime hours. It is possible that the increase in the %err across the day was too small to be statistically significant, while being sufficiently important to counteract the improvement in the measures of reaction time.

The diurnal temporal patterns of the sigma and the beta1 bands were similar to the changes seen in the speed of performance across the day. These similarities were also observed at the individual level, in 6 of the 8 subjects. Recently, it was shown that the coherence of brain electrical activity in the beta range (13–21 Hz) was enhanced between the visual and motor cortex in subjects who performed a visuomotor task (Classens et al., 1998). It was suggested that the synchronization in the 13–21 Hz activity over different regions may have a functional significance, in favor of a better processing of information between large and multiple neuron networks. In our study, it was not possible to make direct comparisons between the EEG activity and the performance efficiency, because there was a lag of 1 h between the taking of the two measures and because only the central derivation was examined. However, it is not excluded that the ongoing increase in the 12.00–19.75 Hz activity over the daytime may represent an increase in the excitability of the cortical nerve cells which have favored or helped the subjects to process the information at a higher speed.

In conclusion, three EEG frequency bands, as well as the measure of reaction time, showed diurnal variations. The time course of the frequencies ranging from 12.00 to 19.75 Hz was similar to the time course of speed of execution. It seems possible that the changes in the 12.00–19.75 Hz range across the day facilitate the sensorimotor processing.

REFERENCES

Cacot, P., Tesolin, B., Sebban, C. (1995). Diurnal variations of EEG power in healthy adults. *Electroencephalographic Clinical Neurophysiology,* **94,** 305–312.

Classen, J., Gerloff, C., Honda, M., & Hallett, M. (1998). Integrative visuomotor behavior is associated with interregionally coherent oscillations in the human brain. *Journal of Neurophysiology,* **79**(3), 1567–1573.

Cummings, L., Dane, A., Rhodes, J., et al. (2000). Diurnal variation in the quantitative EEG in healthy adult volunteers. *British Journal of Clinical Pharmacology,* **50**(1), 21–26.

Doppelmayr, M., Klimesch, W., Schwaiger, J., et al. (2000). The time locked theta response reflects interindividual differences in human memory performance. *Neuroscience Letters,* **278**(3), 141–144.

Goldenberg, F., Weil, J. S., & Von Frenckell, R. (1988). Utilisation du rapport des puissances spectrales thêta/alpha comme indice de vigilance. *Neurophysiol. Clin.,* **18,** 433–445.

Haig, A. R., Gordon, E., Wright, J. J., et al. (2000). Synchronous cortical gamma-band in task-relevant cognition. *Neuroreport,* **11**(4), 669–675.

Horne, J. A., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology,* **4,** 97–110.

Klimesch, W., Doppelmayr, M., Russeger, H., & Pachinger, T. (1996b). Alpha frequency, reaction time, and the speed of processing information. *Journal of Clinical Neurophysiology,* **13**(6), 511–518.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Review,* **29,** 169–195.

Lafrance, C., & Dumont, M. (1998). Daytime vigilance after morning bright light exposure in volunteers subjected to sleep restriction. *Physiology and Behavior,* **63**(5), 803–810.

Lafrance, C., & Dumont, M. (2000). Diurnal variations in the waking EEG: comparisons with sleep latencies and subjective alertness. *Journal of Sleep Research,* **9,** 243–248.

Lenne, M. G., Triggs, T. J., & Redman, J. R. (1997). Time of day variations in driving performance. *ACCID Anal. Prev.,* **29,** 431–437.

Lenne, M. G., Triggs, T. J., & Redman, J. R. (1998). Intereactive effects of sleep deprivation, time of day, and driving experience on a driving task. *Sleep,* **21**(1), 38–44.

Lorenzo, I., Ramos, J., Arce, C., et al. (1995). Effect of total sleep deprivation on reaction time and waking EEG activity in man. *Sleep,* **18,** 346–354.

Markowitz, M. E., Arnaud, S., Rosen, J. F., et al. (1988). Temporal interrelationships between the circadian rhythms of serum parathyroid hormone and calcium concentrations. *Journal of Clinical Endocrinological Metab.,* **67,** 1068–1073.

Monk, T. H., Buysse, D. J., Reynolds, C. F., Berga, S. L., Jarrett, D. B., Begley, A. E., & Kupfer, D. J. (1997). Circadian rhythms in human performance and mood under constant conditions. *Journal of Sleep Research,* **6,** 9–18.

Owens, D. S., Macdonald, I., Tucker, P., et al. (1998). Diurnal trends in mood and performance do not all parallel alertness. *Scandinavian Journal of Workplace Environmental Health,* **24**(3), 109–114.

Owens, D. S., Macdonald, I., Tucker, P., et al. (2000). Diurnal variations in the mood and performance of highly practised young women living under strictly controlled conditions. *British Journal of Psychology,* **91**(1), 41–60.

Wilkinson, R. T., & Houghton, D. (1975). Portable four-choice reaction time test with magnetic tape memory. *Behavior Research Methods Instrumentation,* **7,** 441–446.

This is doi:10.1006/brcg.2001.1449.