ON READING HANDWRITING

H.J. van Jaarsveld

ON READING HANDWRITING

ON READING HANDWRITING

PROEFSCHRIFT
ter verkrijging van de graad van doctor
in de sociale wetenschappen
aan de Katholieke Universiteit te Nijmegen
op gezag van de Rector Magnificus Prof Dr J H G I Giesbers
volgens besluit van het College van Dekanen
in het openbaar te verdedigen
op donderdag 30 juni 1983
des namiddags te 2 00 uur

door

Hendricus Johannes van Jaarsveld
geboren te Boskoop

Sneldruk Boulevard Enschede
1983

Promotor: Prof. Dr. W.J.M. Levelt

ACKNOWLEDGEMENTS

For assistance in statistical analyses I am indebted to drs. M. Berger at the Katholieke Hogeschool in Tilburg (Experiments 5 and 6), to the Unit for Mathematical-Statistical Service (M.S.A.) at the Katholieke Universiteit in Nijmegen (Experiment 8) and to drs. P. Sanders of Cito in Arnhem (selection experiment). Thanks are due to F. Maurer for taking care of many randomizations and to A. Kerkhoff, who as a student-assistant, helped me in running experiments. I also want to express my gratitude to the many people who gave a hand in writing or reading.

CONTENTS

CHAPTER 1

INTRODUCTION

The topic of this thesis is the reading of handwriting by adults. 'Handwriting' will be understood to refer to samples of linguistic material (letters, words, or texts) that are produced by means of some writing utensil and involve movements of the hand or arm of a producer.

The most salient aspect of handwriting is undoubtedly its variability. The variation is so large that every single handwriting can rightly be regarded as a highly individual way of expression, that may even (unwantedly) betray some carefully hidden aspects of the personality of the producer. Because of this large variability, the reading of 'handwriting' provides an example *par excellence* of pattern recognition. Figure 1.1 on page 2 and 3 shows only some of the variations that a reader may have to adapt to when reading handwriting. The samples, displayed in Figure 1.1, are part of a larger collection of handwritings, put together during the research reported in this thesis. All displayed samples were written by aduls.

Apart from aspects that do not even belong to the handwriting itself like inter-line spacing or orientation on the page, handwritings may differ in overall characterisicts like size (samples *A* and *B* in Figure 1.1), slope (*C* and *D*), regularity (*E* and *F*). Less obvious dimensions like 'horizontal/vertical' can also be distinguished (*G* and *H*). Such characteristics need not be consistent within a handwriting as can be seen in sample *I* in which letters are both upright and slope to left or right.

Variations in handwriting will also involve more local features. A variation dealing with word shape is presented in samples *J* and *K*. While in sample *J* the word shape is very pronounced, ascending or descending line-segments in sample *K* do not provide reliable cues for letter identity. As will have been noted in considering the samples in Figure 1.1, handwritten letters come in myriads of forms. The isolation of single letters is not always self-evident as can be seen in sample *L* in which some line-segments appear to belong to two letters at the same time. In other handwritings, however, large inter-letter spaces separate the letters from each other (sample *M*). Such differences in segmentation may not only involve the letters within words, but even the words themselves (sample *N*).

A remarkable form variation in letters appears in sample *O* in which handwriting upper and lower case are combined. Confusing form varite grote hoeveelheden vers groen

A gras hadden zien drijven, en hy

in zyn versleten pandjesjas en te wyde halsboord

B kon hy er uit zien als een vogelverschrikker

rieten mat op de vloer. Daar

C slapen ze ook 's Avonds worden

seling bij de kant van de weg een kleine

D kudde giraffen aan de kruinen van

in groten getale verkocht maar

E die verkoop is in de laatste

De sponningen waren de [illegible] die van een snel [illegible] konden zeggen. In het begin van

F deze eeuw toen er bijna nog geen auto was, werd er reeds veel gebruik gemaakt van

meeste bewoners in armoedige

G omstandigheden leefden. Bovendien

was geweest, dat, wat de dichtheid betrof,

H kon worden vergeleken met de wouden langs

inderdaad zo is, kan men slechts ontdekken

I door een proefboring. Men richt dan een

Figure 1.1

het platteland naar de stad gekomen en de
hoop wat meer geld te verdienen. Maar dat … J

waren meestal arme lieden, die met
het schrijven van brieven voor rijke kooplui K

sloopten zijn gezondheid. Ernstig ziek werd
hij opgenomen in een zenuwinrichting. L
Zijn werkkracht was echter niet gedoofd.

saus, vlees, melk, noten, kaas,
zoetigheden en tal van andere M

zaad, hoeveel zorg de oude vrouw voor die vogeltjes N
had. hoe zij ze koesterde en verpleegde. In de angst

ogen, is het bijna niet meer te geloven dat er in de
eerste wereldoorlog van de radio bijna geen sprake ge- O

Een radarscherm doet denken aan een oud televisie. P

en waarmee zelfs paarden vervoerd konden
worden. In onze landstreek hebben de Q

Tot nu het water in [illegible] waren hier en daar
plekken met drijfzand, zodat en het hek R

Variations in handwritings

ations in which a letter takes the disguise of the standard form of another letter also occur (sample *L* in which the *z* is a pretty good *r*) It is hardly surprising that letters take different forms within the same handwriting (the *d* in sample *P*)

Apart from all kinds of emotions that handwritings may give rise to, it is obvious that aspects like the ones mentioned above will create large differences in legibility (samples *Q* and *R*) In view of the interesting variations that occur in handwriting, it is surprising that reading research uses dull material like standard typefaces on an IBM electric type-writer at all

Experimental research of reading handwriting as reported in the next chapters may serve two different functions in reading research as a whole

First, it can be argued that any valid theory about reading should not be limited to print, but should also be able to explain how handwriting is read Although it cannot be ruled out that reading handwriting will be a special process with mechanisms of its own, it seems more likely that reading print and handwriting will have much in common, especially with respect to non-visual processes Differences between reading print and handwriting may, however, very well be found for visual processing Handwritings often confront the reader with problems that are not, or are in a quite different way, present in typed material A well-known example is letter segmentation which is probably more complex in cursive handwriting (Neisser, 1967) Data on human recognition of handwriting might therefore provide indications whether theories about reading, based on experimental findings with printed material, should be modified in order to be generalizable to reading of handwriting

Second, research with handwritten material may prove to be a valuable addition to usual experimental techniques As Figure 1 1 shows, handwritings possess in ample measure the reduced legibility and form variation that are often artificially introduced into typed materials by researchers Examples of these manipulations are the use of visually degraded stimuli (e g , Meyer, Schvaneveldt, and Ruddy, 1975) or stimuli that consist of different type-faces (e g , Adams, 1979) Experimental results based on these manipulations have contributed significantly to theorizing about the reading process Reading research that uses handwritings may do likewise, with more natural stimulus materials

This study limits itself to the human recognition of handwriting Research dealing with the machine recognition of handwriting will deliberately be left out of consideration Although it may ultimately turn out that human and machine recognition of handwriting have much in common, it seems premature to compare machine and human recognition of handwriting A large discrepancy exists in the amount of research that has been carried out in the two areas While the machine recognition of handwriting is being investigated rather intensively, experimental research of human recognition is virtually non-existent as is evidenced by the very few available studies that are described in the next section

Neisser and Weene (1960) studied the recognition of single handwritten letters and numbers The material consisted of manuscript letters which were taken from written names and addresses The mean accuracy for correct recognition was found to be 94 9 percent Most erroneous readings (3 2 %) involved only 21 examplars No errors were found for the letters *a*, *h*, *k*, *m*, *r* and *z*, while *t*, *u*, *v*, and *y* resulted in the largest proportions of incorrect classifications
In a study conducted by Corcoran and Rouse (1970), typed and handwritten words were presented tachistoscopically In their Experiment I, two conditions were used In an unmixed condition, all stimulus materials were of the same kind (typed or handwritten), in a mixed condition, typed and handwritten words were randomly mixed Handwritten words were found to result in fewer correct recognitions than printed words and the mixed condition resulted in poorer performance than the unmixed condition for both typed and handwritten words Mixed and unmixed presentation was used for two different handwritings in Experiment II and for upper and lower case in Experiment III In these experiments, however, no effect of mixing was obtained Corcoran and Rouse suggested that possibly different sub-routines are used for print and handwriting Part of the recognition process consists in a decision whether the stimulus is typed or handwritten and to switch to the appropriate sub-routine Because the reading of handwriting is generally more time-consuming than reading print, the sub-routine for print might be shorter, contain a different sequence of operations, and would have no need for letter segmentation procedures Corcoran and Rouse also proposed that for handwriting and print, a number of different routines may be required that would deal with a certain spread of specimens

Ford and Banks (1977) used a memory search task (Experiment 1 and 2) and a word naming task (Experiment 3) to study differences between reading handwritten and printed words In the memory search task (Sternberg, 1969, 1975), the difference between typed and handwritten words appeared to affect only the intercept and not the slope of the RT function The intercept is assumed to represent, among other things, the duration of the stimulus encoding stage, the slope represents the amount of time required for the comparison of the probe with items held in memory. In the word naming task, reaction times for handwritten words were found to be longer for printed words by about the same amount of time as the difference between the intercepts for printed and handwritten words in the memory search task

Because the memory search task showed slopes for printed and handwritten not to be different, it may be deduced that perceptual processes in reading handwriting do not use information in active memory According to Ford and Banks, the results therefore cast doubt on a model for (handwritten) word recognition in which conceptually-driven processes contribute to recognition To explain the mean overall difference in latencies between print and handwriting, they suggested that, following Corcoran and Rouse, a handwriting sub-routine might take longer than a print sub-routine They also considered that switching time for the correct sub-routine may be responsible for the overall difference.

A study that did not use handwriting, but nevertheless suggested some general processing characteristics of handwriting was carried out by Bryden and Allard (1976) They presented ten different typefaces (capital letters) in right and left visual field Although most typefaces resulted in a right visual field superiority, a left visual field was found for three typefaces Left visual field superiority was found to correlate with position on a dimension scriptlike-printlike' Bryden and Allard related the left visual field superiority for more script-like typefaces to the greater ability of the right hemisphere for global preprocessing. Script-like typefaces and handwriting generally require more cleaning-up of the initial representation and filtering out of irrelevant detail, which have been described as properties of visual processing by the right hemisphere

Thomassen and Hudson (1982) tried to extend the Bryden and Allard results to small letters and to words They presented small and capital single letters and words in four different typefaces which differed in their script-like characteristics Although the more script-like typefaces were generally less legible, no differences were observed for recognition in right and left visual field for single letters For the word stimuli, a right visual field superiority was found This effect was somewhat reduced for the more script-like typefaces

In the present study two aspects of visual processing of handwritings were selected for experimental investigation, which would seem, on first view, rather promising for establishing specific characteristics of the reading of handwriting

One aspect dealt with the segmentation of the handwritten word into its constituent letters (Chapters 2 and 3) Cursive writing is conspicuously different from printed material in that letters are connected, while in print letters are separated by inter-letter spaces It seems likely that segmentation of the word into letters will be less easily achieved in cursive handwriting and will require some additional mechanisms that are not necessary for print

A second topic of investigation was the perceptual learning of handwriting, i e , the perceptual adaptation of a reader to the characteristics of particular handwritings (Chapters 4 and 5) It is a common experience that one may get used to handwriting which is evidenced by improvements in reading speed This suggests that the (initial) reading of a handwriting is characterized by concurrent perceptual learning Because of the extensive experience with common typefaces, this process is almost certainly absent in the reading of print by adults It should be noted that experiments about practice effects also serve a methodological purpose It is clear that experiments with handwriting will become more difficult to interpret if indications are found that perceptual processes in reading handwriting change considerably in a short amount of time

The research of these two aspects displayed two different methodological approaches to handwriting research These two approaches will, for convenience s sake, be referred to as a depth and a breadth approach They differ primarily in intended generalizability of experimental findings across different handwritings

A depth approach involves the detailed investigation of some specific aspect of handwritten material that does not have to occur in each and every handwriting (it may even be very rare) Selected aspects will preferably be unique for the reading of handwriting but might also be artificially created in printed material (see, for instance, Brooks (1977) for imaginative examples of form variations in printed material) Consider, for instance, the fact that in some handwritings letters take different forms dependent on their relative position in the word investigation of this so-called 'position-specific allographic variation' may try to establish whether recognition is facilitated or hampered by this variation and may search for related variables or factors that affect performance Attempted generalizations would not primarily involve different handwritings but, for instance, relations between letter-form and position information.

For the investigation of segmentation, a 'depth' approach was adopted with the aim of establishing some general aspects of segmentation procedures in handwriting, independent of particular handwritings.

A 'breadth' approach tries to establish whether particular processes are common to all handwritings or to certain 'types' of handwritings. This research strategy figures, for instance, in attempts to establish overall differences between 'handwriting on the one hand, and 'print' on the other. 'Types' of handwriting can be distinguished on the basis of very divergent dimensions like esthetic quality, production antecedents (age, left- or right-handedness of the writer, relative speed with which it was produced), graphic (physical) characteristics, relative legibility, or even graphological aspects. Obviously, the choice of a particular dimension will depend on kinds of problems one wants to investigate. In a 'breadth' approach, representative sampling procedures will be required to warrant generalizability of experimental findings along a certain dimension. It should be pointed out that in reading research that uses printed material, an analogous problem exists regarding the generalizability of experimental findings across different type-faces

For the experiments dealing with perceptual learning, a 'breadth' approach was adopted Handwritings were sampled that differed in legibility and it was tested whether perceptual learning displayed a systematic relation with handwriting legibility. This procedure implied that handwritings were classified on the basis of legibility and that generalizability of experimental findings across handwritings was sought along this dimension

The choice of the legibility dimension was inspired by the fact that it resembles common experimental manipulations in research of the reading of printed material The legibility of a handwriting can be defined as the ease with which the handwriting can be read and which will be apparent in the reading speed or in the accuracy Legibility is not a property of the handwriting alone, independent of what has been written On the contrary, legibility will be jointly determined by form and content, reflecting the general principle that visual recognition makes use of both data-driven and conceptually-driven processing (Lesgold and Perfetti, 1982, Norman and Bobrow, 1975) Recent research (Meyer *et*

al., 1975) shows that poor stimulus quality can be compensated for by semantic context.
Establishing the relation between certain physical aspects of the typed word and its legibility has been the topic of an extensive research program (Tinker, 1965). Similar research may be carried out for handwriting It might, for instance, very well be that handwritings which slope to the left are generally less legible than handwritings that are upright. Research, that has been carried out in the context of the development of handwriting scales, has suggested some physical determinants of handwriting legibility (although none of these have been tested experimentally). Handwriting scales are used for the evaluation of the handwriting quality of children of different ages. Three major scales were developed at the beginning of this century (Thorndike, 1910, Ayres, 1912; Freeman, 1915). In Thorndike's scales handwritings are judged for 'general merit', a qualification that involves esthetic quality and clearness in line and form. In Ayres' scale, the judged characteristic is legibility which is measured by reading speed In Freeman's scale the quality of the handwriting is judged as the sum of a number of factors: uniformity of slope and direction of letters, line quality, letter form, and spaces between letters and words. In a study of Anderson (1969), the legibility judgment was found to be related to the size, slope and uniformity of slope in the handwriting.
A rather different approach to the study of legibility can be found in the study by Pressey and Pressey (1927) On the basis of their analysis of three thousand illegible segments in different handwritings they concluded that relative illegibility is mainly caused by certain letters (especially *r*) or letter combinations Quant (1946) also found that the quality of the single letter-forms was one of the main determinants of the legibility of the handwriting, while more general characteristics like slope, line-quality, or direction of letters were found to be negligible. In Chapter 4 of this thesis, it is suggested that the legibility of a handwriting is determined by its overall resemblance to print. Implicit in this suggestion is the idea that the physical determinants of legibility will not only be based on general aspects like the ones mentioned above, but also on the resemblance to particular forms which have been encountered very often and which have become rather easy to process

As noted above, a researcher who wants to use handwriting(s) is confronted with a generalizability and related sampling problem. Due to the large variability in handwritings, the possibility exists that certain experimental findings will be valid for the selected handwritings only and cannot be replicated with other handwritings Sampling procedures for handwritings need therefore some special considerations. The importance of sampling procedures may be illustrated with the study of Ford and Banks mentioned above In their experiments, handwritten words were reported to require longer recognition times than typed words. They used 19 different handwritings for the handwritten stimuli and upper case for their printed stimuli Apart from this asymmetrical selection procedure (which was probably advantageous for the printed material), no information was presented about the number of handwritings the above mentioned finding was valid for Such information is indispensable for evaluating the statement that handwritten words generally require longer visual recognition times than typed words. It will be obvious that results heavily depend on the characteristics of the selected handwritings. Selection of handwritings with good legibility might have resulted in insignificant differences or perhaps even in an opposite finding. Besides the omission of necessary statistical information to support their general conclusion, Ford and Banks also did not provide any description of the sampling procedures they used, which makes their experiments rather difficult to replicate.

Sampling procedures will differ for a depth approach (the detailed investigation of some specific aspect) and a 'breadth approach (establishing common processes for certain types of handwriting)

In a 'depth' approach, the sampling of particular handwritings will be based on the aspect that is being studied A 'depth approach may lead to the selection of very unusual or unique handwritings It may even require the use of some manipulations that would not normally occur in handwritten material (see, for instance, Experiment 2 of the present study) Clark (1973) discussed experimental manipulations for which it is very difficult or even impossible to use systematic sampling procedures. For such cases Clark proposed investigators using such intuitive procedures should be as explicit as possible about the constraints they were trying to stick to so that other investigators can construct similar samples (ibid , p 352) The adoption of such guidelines may ensure that experiments will remain replicable

It was pointed out above that a 'breadth approach requires representative samples along particular dimensions Intuitively graphic characteristics and legibility seem to be the most relevant dimensions for reading research

A selection on the basis of graphic characteristics would primarily pay attention to the physical attributes of the writing like relative size, slope, or overall regularity A parallel in typed material is the distinction between upper and lower case, upper case misses the graphic aspect of different word shapes The importance of graphic aspects for handwriting recognition is stressed in the suggestion of Corcoran and Rouse (1970) that possibly different sub-routines exist for the recognition of very dissimilar handwritings

In a 'breadth' approach, representative samples, based exclusively on external characteristics, will be difficult to obtain for the following reasons First, it is not known which graphic aspects of handwriting ought to be distinguished Apart from the characteristics mentioned above, less obvious ones like 'variations in line-width or internal letter similarity also deserve consideration Moreover, physical distinctions do not necessarily coincide with perceptual distinctions, i e , it may be that certain physical variations are irrelevant perceptually A second problem consists in the fact that in each handwriting different graphic features are indissolubly connected For instance, it is impossible to vary the slope of the handwriting without simultaneously affecting the single letter forms It will therefore not be feasible to select handwritings on the basis of one graphic aspect, while keeping others constant. To obtain representative samples of different combinations of characteristics, the number of handwritings may become unmanageably large These observations are, of course, not intended to suggest that one might not be able to select handwritings on the basis of certain physical characteristics, they point to difficulties in obtaining representative samples of handwritings based on these characteristics

Representative sampling on the basis of legibility in a breadth' approach has the inherent disadvantage that handwritings that are equal for legibility may at the same time be very dissimilar in physical respects, i e , the cause of the relative (il)legibility may be very different

Sampling procedures for legibility may use preliminary experimental testing, which would establish the legibility for a whole range of handwritings beforehand An important limitation connected with this procedure is discussed below in connection with the predictive validity of legibility judgments A second procedure involves intuitive judgments about legibility An example is described in Appendix H of this study If such judgments are fairly reliable, this procedure will have the advantage that the sampling of handwritings can be achieved rather easily

It should be pointed out that the use of intuitive judgments for sampling procedures has some problematic aspects In the use of handwriting scales uncertainty exists about the exact property that is being judged (Herrick and Erlebacher, 1963) It is not self-evident that the 'quality of a handwriting can be equated with its legibility Conversely, this suggests that in the legibility judgment esthetic considerations may play a role The construct validity of the notion relative legibility is therefore doubtful It certainly seems likely that within a particular level of legibility additional esthetic qualifications can be assigned

Legibility should be regarded as a continuum on which handwritings can be ordered In a handwriting scale, legibility has to assume a discrete number of values In such a case a second problem consists in the number of levels that ought to be distinguished in legibility Thorndike distinguished 17 levels while Herrick and Erlebacher identified 24 The reliability of the classification of handwritings will increase with smaller numbers of levels, but may leave legibility distinctions within categories Herrick and Erlebacher (ibid) indicated that the number of levels to be used will, in general, depend on the ability of judges to discriminate reliably between different levels On the basis of this general criterion, they found five scale units to be sufficient for judging handwritings

Another important aspect of obtaining legibility judgments has to do with the range of legibility, i e whether the variation in the random sample can be regarded as representative of the population. The representativeness of any particular sample may, of course, be doubted for the extremities of relative legibility Thorndike solved this problem by including an articifially produced, illegible handwriting in his sample Although this problem may be relevant for handwriting scales that attempt to cover the whole range of legibility, it is not of prime importance for reading research Initially, a reasonable variation in legibility seems sufficient for investigating the reading of handwriting

For reading research, the most important aspect of legibility judgments concerns their predictive validity In testing this predictive validity, the following considerations (which also apply to sampling procedures based on preliminary testing) need to be taken into account

It was noted that legibility can be measured in different ways The accuracy or the speed of identification may be dependent variables Each of these can be measured in a variety of experimental tasks In reading research, it is not always clear whether or how different experimental procedures reflect normal reading (e g , Jackson and McClelland, 1979) In particular, it cannot be ruled out that handwritings which show legibility differences in one task (for instance, normal reading) may turn out not to be significantly different in others Such inconsistencies need not be a nuisance, they can be used to determine more exactly the cause (or the level) of the relative legibility and thus lead to more precise assessment of legibility.

The general outline of this thesis is as follows. In Chapters 2 and 3, four experiments are reported that dealt with segmentation. In the general discussion at the end of Chapter 3, the results of the four experiments are assessed. Experiments 4-8, that dealt with perceptual learning, are reported in Chapters 4 and 5. Conclusions drawn from these experiments are presented in the general discussion at the end of Chapter 5.

CHAPTER 2

SPACES AND CONFIGURATIONS

It was pointed out in Chapter 1 that differences between reading print and handwriting are likely to involve aspects of visual processing Virtually every model of the reading process assumes that visual processing in reading consists in an initial extraction of elementary features, which are mostly identified with curved or straight line-segments in different positions and orientations Subsequent letter (or word) recognition is thought to be based on the activation of central memory representations by specific combinations of these features

Assuming that reading handwriting, like print, will involve an initial feature-extraction, research of reading handwriting may try to establish whether features are different for handwriting and print More promising, however, for investigating differences between the recognition of print and handwriting will be segmentation, i e., the segregation of the units to be recognized

A remarkable difference exists between print and handwriting with respect to the physical segregation of letters within the word· in print, letters are separated by inter-letter spaces, while the handwritten word often confronts the reader with a continuous, complex line in which letter-boundaries are not self-evident It cannot be ruled out that these differences cause the adoption of quite different segmentation procedures for print and handwriting, leaving open the theoretical possibility that letter identification in handwriting is changed by the fact that letters are connected

Some informal observations support this suggestion Writing connected letters is such a remarkable phenomenon that a separate name -cursive writing- has found acceptance in order to distinguish it from handwriting in which letters are separate (manuscript) At some schools in the U S , cursive writing is introduced in the curriculum only in third grade The first two years the child reads and writes manuscript which is thought to be more legible (Herrick, 1963) The apparent ease with which people can read cursive handwriting is impressive in view of the many failed attempts to get machines to read it Although progress has been made in the categorization of separate characters, segmentation of cursive script still remains a notorious problem (Suen, Berthod and Mori, 1978).

In models of the reading process, based on research with printed materials, the role of individual letters in visual word recognition constitutes one of the major problems (Bradshaw, 1975) Some models view the whole word as the primary recognition unit and accordingly assign individual letters to a subordinate role (Johnson, 1975,1977, Smith, 1971) In letter-mediated models, however, the individual letters function as the fundamental psychological units in the reading process (Estes, 1975a,1977, Gough, 1972, Massaro et al , 1980) Letter models can differ in postulating serial or parallel letter identification A serial model (Gough, 1972) supposes that letters are identified successively from left to right, parallel models assume that all constituent letters can be analyzed simultaneously (Estes, 1975a,1977, Johnston and McClelland, 1980, Massaro et al , 1980)

A letter model encounters an additional problem that does not arise in a whole-word model It will have to explain how the reader arrives at the individual letters as sub-units in the word as a whole Before letters can be recognized they must first be isolated The view that identification must be preceded by segmentation is put forward, among others, by Neisser (1967), who also claims that the two processes are qualitatively different in nature The segmentation process is genuinely "global" and "wholistic" (ibid p 89), while identification is a more analytic process and requires focal attention Pattern recognition is a partly sequential process attentive acts are carried out in the context of the more global properties already established at the preattentive level (ibid , p 90)

Generally, the segmentation of letters in print is thought to be based on the inter-letter spaces, which function as unambiguous signals for letter-boundaries In connection with the distinction between serial and parallel processing of letters, inter-letter spaces differ in function In a serial model, inter-letter spaces indicate which features must be combined as input for letter-detectors In a parallel model, it must be prevented that features of different letters are combined, which would result in uninterpretable chaos or in intolerably high levels of errors Parallel models may therefore propose separate input-channels, each of which corresponds with a letter (or position) Johnston and McClelland (1980) provide an example of such a model We assume that a word target is first preprocessed so that each letter in it is allocated to a position-specific letter-processing channel Within each of these channels information is analyzed for the presence of different letter properties or features (ibid , p 505) In a parallel model, the inter-letter spaces play the important role of physically demarcating the input-channels, ensuring that each letter-detector is supplied with a correct, interpretable combination of features. It should be noted that not in every parallel model input-channels are assumed to coincide with letter-positions. In Estes' model (Estes, 1975b), for instance, each input-channel corresponds with a certain segment of the total visual field. The density of these input-channels decreases in the periphery.

The kind of models outlined above do not readily explain the segmentation of cursive handwriting. In a serial letter identification model, it is no longer self-evident which features must be combined to activate a letter-code. A possible extension of this model could consist in a recursive procedure, which would cumulatively analyse more features simultaneously if, on the basis of a certain sampling, no letter recognition occurred. Apart from the laboriousness of this procedure, its application would result in many errors due to the overlap in features of many letters (e.g., *m-n*, *o-c*, *w-v*). In a parallel model in which input-channels correspond with letter-positions, no cue for segregating features is present in cursive handwriting because the input-channels are not physically demarcated. In parallel models in which input-channels do not coincide with letter-positions, extracted feature-information is often assumed to be coded for position (e.g., Estes, 1975a). However, 'most theorists have not considered the problem of parsing features into groups and by default have assumed that features are automatically placed in appropriate groups' (Wolford, 1975, p.192). Inter-letter spaces likely play an important role in an automatic grouping of features. In general, the theoretical considerations presented above suggest that segmentation of letters will be more difficult in cursive handwriting than in print.

From a functional point of view, cursive writing is composed of two kinds of line-segments: letter-segments that are part of letters and contribute to letter identification, and connecting segments, which function as inter-letter spaces and are of no consequence for the identification of letters. It is conceivable that segmentation of handwriting proceeds by means of an initial classification of line-segments as letter-segments or as connecting segments (a computer program by Eden, described in Neisser (1967), works in exactly that way). Because connecting segments look like letter-segments such a distinction will, however, not be an easy one to make.

In the first experiment of this study, the effect of removing connecting segments was investigated. It was expected that in a comparison between handwritten words in which letters were connected or separate, explicitly segmented words would result in faster recognition. In the latter case, the reader can make use of familiar inter-letter spaces and does not need to distinguish letter-segments from connecting segments. The removal of connecting segments aimed to facilitate the segmentation process by providing clear, physical boundaries between letters as in printed material. If the prediction is borne out, evidence that segmentation in handwriting is made more difficult by the presence of connecting segments will have been obtained

The effect of letter spacing in print may be similar to the removal of connecting segments in handwriting. Letter spacing, the introduction of empty letter positions between letters, will make segmentation of print even more easily to perform. Much of the research that has made use of letter spacing has not been concerned with the segregation of perceptual units like letters but with effects of lateral inhibition between letters (e g., Estes, 1972; Krumhansl and Thomas, 1977). This research shows that letter recognition generally improves with increased distance between letters. Although the effects may be similar, it will be noticed that important differences exists between letter spacing in print and the removal of connecting segments in handwriting The removal of connecting segments has presumably little or no consequences for the amount of lateral inhibition. In the stimulus materials used in the first experiment of this study the average distance between the letters was held about equal for segmented and connected forms and no systematic manipulation of the inter-letter distances was applied.

A few studies have investigated the effect of letter spacing on the recognition of other perceptual units than letters Mewhort (1966) presented 8-letter pseudo-words with and without spacing. The pseudo-words were 0-order or 4-order approximations to English or hybrid combinations of 0-order and 4-order. For the recognition of letters in the 0-order approximations spacing had no effect, but the recognition of letters in the 4-order was impaired in spaced condition. According to Mewhort, letter spacing slowed the scanning process which for the 4-order approximations results in a fading of the material before it could be chunked' Gibson and Levin (1975) pointed out that the letter spacing in the 4-order approximations may have caused disruption of the higher-order visual units which are present in those arrays.

Schindler, Well, and Pollatsek (1974) and Terry, Samuels and LaBerge (1976) used letter spacing in words to disrupt the word as a familiar visual unit In the simultaneous matching task of Schindler et al , the letter spaced words resulted in longer latencies, but the effect was dependent on the expectations of the subject If the subject expected both words and nonwords to be presented, letter spacing had an equal effect for words and nonwords; when the subject expected only nonwords, spacing had an effect primarily for the words. In the study by Terry et al., irregular letter spacing had no negative effect on recognition compared with regular spacing. A clear effect of explicit segmentation of psychological units by letter spacing was obtained by Taft (1979). His study of the representation of poly-syllabic words is very similar in design to the first experiment of this study. In his lexical decision experiment (Experiment I), words were presented in segments which were separated by a letter space. The position of the letter space coincided with an important constituent-boundary, according to an orthographic or pronunciation analysis of the word. The aim of the letter space was 'to guide the analysis imposed by subjects in the retrieval process' (ibid., p. 24). If the segmentation of the word coincides with the stored representation, shorter latencies will be observed for those stimuli compared with words for which the segmentation does not coincide. Taft's results showed that orthographically segmented words were accessed faster than words segmented according to their pronunciation. His results indicate that the explicit physical segmentation of psychological units can facilitate their recognition.

Similar to Taft's design, two different segmentations were applied in the first experiment. In a comparison between handwritten words in which letters are connected or separate, faster recognition of the segmented forms will not provide conclusive evidence that facilitation is due to segmentation. The segmented forms simultaneously contain less stimulus information and require less feature-analysis. To show that the facilitation is due to segmentation, a control condition was introduced in the experiment in which the stimuli also contained less information but where connecting segments were left intact. This control condition involved the removal of line-segments within letters. The removal of these line-segments may impair letter identification. It was attempted to minimize this effect by removing segments which may be deemed inessential for the identification of the letter. If there is no difference between segmented and control stimuli and both are identified faster than the connected stimuli, facilitation for the segmented forms will have to be attributed to less stimulus information. A larger facilitative effect for the segmented forms in comparison with the control condition will suggest that reading handwriting is selectively hampered by the presence of connecting segments.

The removal of line-segments within letters does not only result in incomplete letter information, but provides, at the same time, misleading information for a segmentation procedure that relies on physical interruptions. In combination with facilitation for the segmented condition, longer latencies for the control condition will provide additional support for the importance of physical separation of letters for segmentation.

EXPERIMENT 1

METHOD

Stimulus materials

Words

The 60 stimuli were randomly selected, 5-letter Dutch nouns. They were all mono-syllabic to control for effects of syllable length on naming time (Eriksen, Pollack and Montague, 1970) and had a frequency of less than 25 (per 720.000) according to the Uit den Boogaart (1975) count. Each of the 60 words was randomly assigned to one of three arbitrary classes of 20 words each. These three classes corresponded with the three form conditions of the experiment: one class contained words in which letters were connected (connected condition); the second class consisted of words in which letters were separate (segmented forms); the third class contained words in which line-segments were removed within letters (control condition). Appendix A contains a list of the individual items.

Handwritings

In the course of this research, about a hundred handwriting samples were collected. From this larger set, 20 handwritings were selected that were roughly equal in size to ensure about equal visual angles for different handwritings. Additionally, it had to be possible to remove connecting segments without damaging the letters. Some handwritings in the larger set could not be used because letters were written so closely together that connecting segments were missing. One word from each of the three form classes mentioned above was written in each handwriting. Each individual writer was contacted to write these three words. Assignment of handwritings to an arbitrary set of three words was random. Examples of the stimuli used are presented in Figure 2.1. The preparation of the handwritten stimuli required that connecting segments had sometimes to be added to words in the connected and control conditions. These additions could be made without any apparent

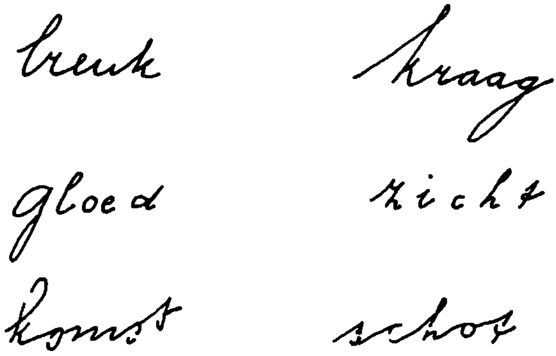


Figure 2.1 Examples of stimuli presented in Experiment 1. The upper row are instances of connected forms, the middle row of segmented forms, the bottom row of the control condition.

damage to the 'natural appearance' of the words. With a special pen, line-segments were erased in words for the segmented and control condition. No traces remained of the removed segments. In words for the segmented condition, all four connecting segments were removed. In words for the control condition, the removal of segments satisfied several conditions. Most importantly, no other letter was created as a result (cf. *c* and *o*) and distinctive characteristics of the letter were left complete (e g., the crossbar in the *t*). Removals had to result in perceptible interruptions within the letter. Only one removal per letter was applied. With these restrictions the position of the disruptions was random. Representative examples of removals of line-segments within letters were leaving out a line-segment in the up-going or down-going line-segments of ascenders and descenders or in the left-most part of small letters like *a* or *o*. The length of the removed segments was about 2-4 mm and, in general, somewhat larger for the connecting segments.

Procedure and Subjects

Words were projected on a translucent screen by means of a Kodak Caroussel slide-projector The stimuli were presented within a visual angle of about 2°. Task for the subjects was word naming. Latencies were measured from stimulus-onset till initiation of the vocal response.

For order of presentation, the 60 stimuli were divided in 20 classes of three words each, one out of each of the three form conditions. Within

these classes, order of form condition itself was random. As an additional constraint, two words from the same handwriting could not follow each other. Within these general restrictions, order of presentation was random for each subject.
To familiarize the subject with the procedure, 10 printed words were presented prior to the experimental items. Experimental sessions lasted about half an hour.
The 20 subjects were all students at the University of Nijmegen. They were paid f 7,- for their participation in the experiment.

RESULTS

Latencies
Some correct responses had an very long RT. These values presumably reflect idiosyncratic difficulties with a particular word. Two parallel analyses were carried out: one, including all correct responses, and a second in which all correct responses exceeding three standard deviation units above the general mean of 1067 ms were left out. On the basis of this criterion, 2.2 % of all valid observations were discarded in the analysis. Results differed only in minor points for both analyses. The analysis reported below was based on the 'cleaned' data.
Separate analyses were carried out for subjects and items (Clark, 1973). The subject analysis was based on the means per form condition per subject. The item analysis was based on the means for single items. Form was significant in the subject analysis [F_s (2,38) = 14.64, $p < .01$] and in the item analysis [F_i (2,57) = 4.14, $p < .05$; *min* F' (2,84) =3.23, $p < .05$]. Means for the connected, segmented and control condition were 982, 994 and 1114 ms, respectively. The *t*-tests for comparisons among means showed the control condition to be significantly different from the segmented and connected conditions.
It is not implausible to assume that explicit segmentation will a larger effect for handwritings with reduced legibility. Generally, the letter-forms will be more deviant in these handwritings and particularly in those cases a clear demarcation of the letters might facilitate recognition.
On the basis of their means, the 20 handwritings were divided into two classes. The 10 handwritings with the largest means were considered to be poorly legible; for the other 10 handwritings good legibility was as-

sumed. Within these two classes, individual handwritings were considered random replications.

In the analysis of variance, the interaction between Legibility and Form appeared not to be significant ($F < 1$), although both main effects Legibility [F (1,19) = 50.19, $p < .01$] and Form [F (2,38) = 14.02, $p < .01$] were significant. Means for the six conditions are presented in Table 2.1.

Table 2.1
Mean Naming Latencies (in milliseconds) for
Three Form Conditions in Handwritings with
Good and Poor Legibility in Experiment 1.

	Legibility of Handwriting	
Form condition	Good	Poor
Connected	910	1058
Segmented	936	1056
Control	1056	1172

Errors

Different kinds of errors -experimental error (.6 %), erroneous readings (2.7 %), and 'illegible' responses (2.7 %)- amounted to 6 % of all data. The number of erroneous readings for the three form conditions -connected, segmented, and control- was 4, 13, and 16 respectively. For the 'illegible' responses these numbers were 6, 10, and 16.

The erroneous readings involved 12 different words. The corresponding figure for the 'illegible' responses was 17, but the words were partly the same.

DISCUSSION

The experiment compared two segmentation procedures for handwriting. One procedure involves a distinction between letter-segments and connecting segments. A second is based on empty spaces between letters, as has been supposed for printed material. It was expected that the latter procedure will result in faster recognition because it uses familiar physical cues in the signal itself for combining features. Segmentation by means of an initial classification of line-segments will be more complex because, generally, letter and connecting segments look very similar. No significant difference, however, was found between the segmented and the connected condition. This negative finding indicates that the presence of connecting segments does not necessarily make segmentation more difficult. The result can be explained by assuming that the segmentation of handwriting is carried out on the basis of contours, configurations of spatially adjoining features. If contours provide sufficient cues for segmentation, the presence of connecting segments may be of no consequence for recognition.

The control condition, in which line-segments were removed within letters, resulted in longer latencies than both the segmented and the connected condition. Different interpretations may be considered for this finding. Longer latencies for the control condition can be attributed to the misleading information which is inherent to interruptions within the letters. The disruptions caused incorrect combinations of features, which were uninterpretable for the letter detectors. This interpretation is supported by considerations relating to the insensitivity of the experimental design like the following. In the task for the subject -naming of words- segmentation constitutes only a minor component of the total process. In the composition of the RT, which will also contain components like accessing a phonological representation and articulatory program, the variance connected with segmentation may be considered negligible. Effects of this process will only become apparent under abnormal conditions like misleading information. According to this interpretation, the longer RT for the control condition points to the preferential strategy of the reader to segment at interruptions.

A second interpretation attributes the longer latencies for the control condition to incomplete letter information. Unfortunately, most experiments (for a review, see Krueger, 1975) in which the effects of mutilated letters were studied, are so different from this experiment that their results can hardly be used as support for this interpretation. In a study of Massaro (1980), letters were mutilated in a similar way as in Experiment 1. The length of the horizontal line-segment in the letter *e* was varied, so that the *e* gradually became a *c*. The same operation was applied to the ascending line-segment of the *h*, which gradually became a *n*. In both cases, the discrimination of relative length of the manipulated line-segment improved under longer viewing conditions. The 'mutilation' of the letters in this study involved line-segments that can be regarded as essential for the discrimination between letters. In Experiment 1, however, those segments were left untouched

The equal RT for segmented and connected forms suggests that segmentation of handwriting is carried out on the basis of contours. This interpretation may also explain the longer RT for the control condition. The interruptions in this condition caused quite different contours, which were no longer sufficient for correctly assigning letter-boundaries The disruptions within letters broke up the contours of letters which constituted important information for segmentation. Because contours may also have an effect on letter identification, an interpretation of the longer latencies for the control condition as due to incomplete letter information seems more valid It should be noted that one aspect of the results argues against the 'incomplete letter information' hypothesis It was pointed out above that the legibility of a handwriting may primarily be determined by the legibility of its letters Incomplete letter information will carry extra weight for the less legible handwritings, in which letter information is already deficient In the analysis of variance, however, no significant interaction between handwriting legibility and form condition was found To provide empirical support for the interpretations of the results obtained in this experiment, the role of contours in segmentation was investigated in the next experiment.

EXPERIMENT 2

In studying the relevance of contours for segmentation, two different kinds of contours need to be considered The handwritten stimulus as the total configuration of features has a contour of its own This contour can loosely be described as the overall size, measured in length and height Within this total configuration, 'sub-contours' will be discriminated that are relevant for imposing letter-boundaries The discrimination of these latter contours may not be independent of certain aspects of the configuration as a whole Segmentation procedures on the basis of contours may therefore involve wholistic aspects of the stimulus These wholistic properties are referred to as global configurational aspects because they deal with the spatial distribution of features beyond the level of constituent letters The contribution of these configurational aspects to segmentation can be illustrated with the handwritten words in Figure 2.2.

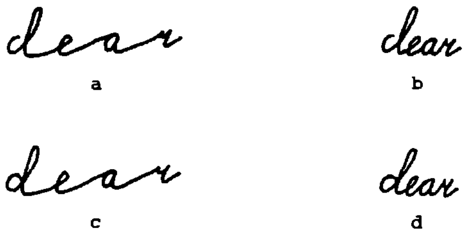


Figure 2.2 Handwritten words that demonstrate that segmentation is affected by global and local configurational aspects.

Comparing Figures 2.2 (a) and (b), in which the first segment is identical in both figures, the likelihood of perceiving one (*d*) or two (*c*, *l*) letters will be seen to be partly dependent on overall length. In Figure 2.2 (a), the first segment is more likely to be perceived as one letter because this interpretation will make the relative distances between letters more uniform. For the same reason, the likelihood of perceiving two letters in the first segment in Figure 2.2 (b) seems to be greater than in Figure 2.2 (a).

Segmentation might also be determined by more local aspects of configurations of features as can be seen by comparing Figures 2.2 (a) and (b) with Figures 2.2 (c) and (d). Closing the gap in the first segment resolves the potential ambiguity of the first segment and will make it a more prototypical *d*.

Figures 2.2 (a) -(d) demonstrate two quite different ways in which configurational aspects of the handwritten stimulus can be manipulated. Changes in the configuration might leave the contours for letters untouched. These manipulations will involve the connecting segments between letters as was done in Figures 2.2 (a) and (b). On the other hand, changes in the configuration might simultaneously affect the contours of the letters (cf. Figures 2.2 (a) and (b) versus Figures 2.2 (c) and (d)). For the study of segmentation processes, the latter manipulations have the disadvantage that they might also have an effect on letter identification. For instance, it can be argued that differences between Figures 2.2 (a) and (b) on the one hand and Figures 2.2 (c) and (d) on the other are related to the identification of the *d* only and do not involve segmentation. In Experiment 2 the effect of global configurational aspects on segmentation is investigated. Experiments 3 and 4, reported in Chapter 3, deal with the effects of local configurational aspects.

The discrimination of letter contours may be due to the operation of principles for grouping features. Principles for grouping (or segmentation) were first enunciated by Gestalt psychologists (Wertheimer, 1923; Koffka, 1935) and have also inspired more recent research (Hochberg, 1974; Sutherland, 1973). Among these principles, the following seem intuitively relevant for segmenting handwriting:

1. *closed boundaries*. Features will be grouped when they constitute closed regions. This principle is, of course, prominent for closed letters like *a* and *o* or for parts of letters like *b, e, d,* or *l*.

2. *similarity*. Handwriting is built out of curved and straight line-segments. Transitions between these two types of line-segments may provide cues for letter-boundaries. Ascending or descending lines may be regarded as another instance of the operation of the similarity principle.

3. *spatial contiguity*. Features that are close together are likely to be grouped for the activation of letter representations

For segmenting handwriting, a partial hierarchy between these principles might be supposed that is based on the degree of contextual dependency. Closed boundaries should be considered primary: if a segment is closed, no alternative segmentation will be possible. Moreover, the closedness of a letter is independent of the relative distance or similarity between constituent components (e g , a stretched, partly flat *o* remains an *o*). As can be seen in Figure 2.2 and below in Figure 2.3, spatial contiguity is, of course, only determined with respect to the whole configuration and should therefore be regarded as basically context-dependent Similarity seems relatively less dependent on context, but the distinction between curved and straight line-segments will be an unreliable cue for segmentation, because transitions between the two types of line-segments may occur between letters as well as within letters.

Letters will differ in the degree in which their contour discrimination is self-evident. Such differences will be related to the operation of the grouping principles mentioned above. Letters with closed boundaries will readily be isolated by segmentation procedures, while segments with 'open' contours will be more difficult to process. This would also mean that the segmentation of letters with open contours will be more dependent on configurational aspects of the stimulus than letters with closed contours.

To demonstrate that segmentation is (partly) determined by global configurational aspects of the stimulus, both spatial contiguity and similarity of adjoining features were manipulated in the experiment described below.

Before discussing the experimental manipulations, it should be mentioned that in Experiment 2, subjects were presented with a *n*-detection task. The stimuli were all illegal Dutch pseudo-words. Such stimulus materials have the advantage that they will limit the possible assistance of different kinds of linguistic knowledge (orthographic, lexical) in segmentation and will make Experiment 2 more sensitive than Experiment 1 for detecting differences in initial visual processing.

The manipulation of spatial contiguity involved the violation of the principle that neighbouring elements (features) belong to the same pattern (letter). To facilitate discussion, some terminology must be introduced. The distribution of letters across the stimulus as a whole is the inter-letter spacing (inter-spacing). This kind of spacing has to do with the relative distance between letters; i.e., with the length of the connecting line-segments. The relative distance between features within letters is the intra-letter spacing (intra-spacing). This kind of spacing is essentially the letter-width (cf. *i* versus *m*). Across different handwritings, both inter-spacing and intra-spacing vary considerably in absolute ways. More important for segmentation, however, are the interrelations between inter-spacing and intra-spacing, because disruptions in the regularity of these relations may cause the break-down of the grouping principle of spatial contiguity. The manipulation of spatial contiguity is conveniently discussed with assistance of Figure 2.3 which provides prototypical examples of the stimuli used.

In Figure 2 3 (a), both inter-spacing and intra-spacing are regular. The constituent letters have a constant width and the connecting segments do not vary in length. Intuitively, this configuration is the one most easily identified.

In Figure 2 3 (b), the inter-spacing is irregular (the connecting segments vary in length), but the letters themselves have a constant width Although this irregularity is a little disturbing, the effect on recognition is presumably not nearly as large as in Figure 2 3 (c), where the letter width itself is irregular and the connecting segments are regular. The presumed difference in the recognition of Figure 2.3 (b) and (c) can be explained as follows. In Figure 2 3 (b), the relative distance between features within letters is almost equal to the relative

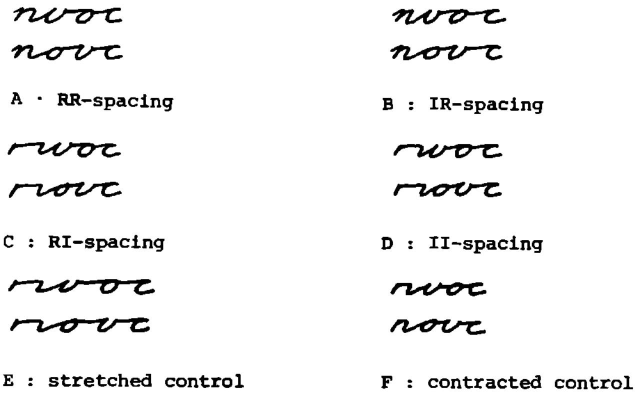


Figure 2 3 The six Spacing Conditions used in Experiment 2 Figure (a) is Regular for Inter-spacing and Intra-spacing, Figure (b) is Irregular for Inter-spacing and Regular for Intra-spacing, Figure (c) is Regular for Inter-spacing and and Irregular for Intra-spacing, Figure (d) is Irregular for Inter-spacing and Intra-spacing Figures (e) and (f) display the control conditions (see text) The upper display in each figure is an instance of the similar condition, the lower display of the dissimilar condition

distance between features of adjoining letters (the exact metrics are supplied below). In Figure 2 3 (c), a gross violation of the principle of spatial contiguity occurs The distance between the two legs of the *n* is much larger than the distance between the right leg of the *n* and the left-most features of the adjoining letter On the basis of this spatial contiguity, part of the *n* and the adjoining letter will initially be combined as input for letter detectors But this input will not activate a letter detector (for half the stimuli, see below) and will in all cases leave the first segment uninterpreted Therefore, a new segmentation will have to be attempted to arrive at a consistent interpretation of all features

Figure 2 3 (d) is a combination of Figures 2 3 (b) and (c). both inter-spacing and intra-spacing are irregular It is difficult to make precise predictions for the relative difficulty of identification for this spacing condition, but one would expect that Figures 2 3 (d) will be more difficult than Figures 2.3 (c), due to the combined irregularity of inter-spacing and intra-spacing

The conditions, displayed in Figures 2 3 (a) to (d), will be referred to as the 'spacing quadrant' The four conditions reflect a continously increasing probability that features, on the basis of relative spatial contiguity, will be combined in the wrong way Virtually equal overall length was common to all stimuli in the quadrant. It is clear that overall length needs to be constant if relations between inter-spacing and intra-spacing are to be irregular

Although the interrelations between inter-spacing and intra-spacing seem to be a global configurational aspect that does not involve any single letter, it should be noted that the conditions in the spacing quadrant also differed with respect to the contour of the *n* In the irregular intra-spaced conditions, the *n* has a rather deviant height-width ratio, while the *n* in the regular intra-spaced conditions may be considered 'normal' in this respect Longer latencies for the irregular intra-spaced conditions might therefore also reflect difficulties in recognizing the deviant *n* An appropriate control condition added to the experiment is discussed below

To manipulate similarity of adjoining features, the *n* was positioned in half of the stimuli next to a *m, u, v,* or *w* (the distractors). In the other half, the *n* was separated from these letters by one or more other letters The letters *m, u, v, w* were selected because they contain straight, vertical line-segments like the *n*. Conditions in which distractors were adjacent to the target, are referred to as similar conditions Conditions in which target and distractor were not adjacent, are dissimilar conditions As can be seen from the two instances of each spacing condition in Figures 2 3 (a) to (d), the adjacency of the *n* and the distractor causes additional problems for segmentation procedures For the similar conditions, the probability that part of the configuration will erroneously activate a letter seems larger In the similar, irregular intra-spaced conditions, the similarity and contiguity of features create configurations that provide strong evidence for a particular letter (w), while, in fact, this letter consists of parts of adjacent letters A re-organization of the features will be more difficult for these conditions than for the dissimilar conditions, in which only the principle of spatial contiguity is violated As is suggested by Figures 2 3 (a)-(d), the similarity principle might not by itself be basic for segmentation but only in conjunction with spatial contiguity A significant interaction between spacing and similarity was expected The effect of similarity will be larger for irregular intra-spaced conditions than for regular intra-spaced conditions

Figures 2 3 (e) and (f) are examples of the control conditions added to the experiment. It might be thought that the larger difficulty of recognizing irregular intra-spaced forms is caused by the unfamiliar appearance of the *n* only and has nothing to do with the irregular distribution of features across the stimulus as a whole Figure 2 3 (e), the 'stretched' control condition, was meant to test the validity of this interpretation In the stretched control condition, the distance between features within letters (except for the *n*) was smaller than the distance between features of the *n* and those of adjoining letters If the spatial distribution of features across the stimulus is an important factor in segmentation, this control condition should be as easily identifiable as the RR-condition in the spacing quadrant and should certainly be less difficult than the irregular intra-spaced conditions If, however, the stretched control condition results in latencies in the same order of magnitude as the irregular intra-spaced conditions, support for the 'unfamiliarity' explanation is provided

The 'contracted' control condition, displayed in Figure 2 3 (f), was designed as a control on an interpretation in terms of differential lateral inhibition effects As is well known, the amount of lateral inhibition increases with smaller inter-letter distance (Eriksen and Eriksen, 1974) or decreases with blank spaces between letters (Estes and Wolford, 1971) In the irregular spaced conditions, the amount of lateral inhibition may be larger because (some) letters are closer together In the contracted control condition, all connecting segments were reduced to the shorter length used in the irregular spaced condition If differential lateral inhibition effects are the main cause of increased difficulty of recognizing irregular spaced forms, no difference should be found between these conditions and the contracted control condition

METHOD

Stimulus materials

Linguistic properties

a *General* All 144 stimuli satisfied the following general conditions They were 4-letter arrays, that were illegal with respect to Dutch orthography (illegal pseudo-words) and contained at least one vowel

Stimuli differed in relative pronouncability. The stimuli were made exclusively with the letters *n* (the target), *m*, *u*, *v*, *w* (the distractors) and the small letters *a*, *c*, *e*, *i*, *o*, *r*, *s*, and *z*.

All target stimuli also contained one of the four distractors. The nontarget stimuli contained two distractors in various combinations. Depending on whether the distractor was adjacent to the *n*, target stimuli were classified as similar or dissimilar A comparable variation was applied to the nontarget items. These stimuli were considered similar if the two distractors were next to each other. In dissimilar nontarget items, the two distractors were separated by one or two other letters.

As described in the introduction, the spacing quadrant consisted in variations in the regularity of inter-spacing and intra-spacing. Of the 144 stimuli, 96 were written according to these spacing conditions. The remaining 48 were control stimuli; the stretched and contracted condition contained 24 stimuli each

The ratio of target to nontarget stimuli was 2 to 1 for both the quadrant and the control conditions, resulting in 96 targets and 48 nontargets for the whole set of 144 stimuli

As the constraints imposed on the stimuli were different in each case, a more detailed description of target and nontarget stimuli in the quadrant and of the control conditions is provided below Consultation of the list of stimuli in Appendix B will facilitate reading of the next sections.

b. *target stimuli* For the 64 target stimuli in the quadrant, the position of the *n* was systematically varied In each of the four letter positions, the *n* occurred 16 times In each position the *n* was paired four times with each of the four distractors (*m*, *u*, *v*, *w*) Of the four stimili with *n* in position *x* (1-4) with a particular distractor, two were similar stimuli and the other two dissimilar ones

In both similar and dissimilar stimuli, the *n* always preceded the distractor in the letter-array (except, of course, for stimuli with *n* in fourth position). In the dissimilar stimuli with *n* in first (fourth) position, every distractor occurred once in third (first) position and once in fourth (second) position Across stimuli, this assignment procedure caused every distractor to occur three times in positions 1 and 2, and five times in positions 3 and 4.

A set of eight stimuli (the *n* in fixed position for similar and dissimilar stimuli separately) was the basis for the assignment of the remaining letters (*a*, *c*, *e*, *i*, *o*, *r*, *s*, and *z*) In each of these sets, each remaining letter occurred twice. Across all stimuli, each of the remaining letters occurred 16 times The assignment of the remaining letters to combinations of 'target-position and distractor' was random

The 64 stimuli were assigned to the four spacing conditions according to interlocking latin squares for target-position and distractor In each spacing condition, the target occurred four times in a particular position, each time with a different distractor and each combination of target and distractor occurred four times (but in different positions for each of the spacing conditions) These latin squares were evenly distributed across similar and dissimilar stimuli The spacing conditions were equal with respect to the position of the target and combinations of target and distractors The spacing conditions differed, however, with respect to the position of combinations of targets and distractors, i e , target-distractor combinations and their position in the array were completely confounded with spacing condition

c *nontarget stimuli* Combinations of distractors were equally distributed over the 32 stimuli Each combination (*mu*, *mv*, *uw*, *vw*) occurred in eight stimuli, four of them similar (with adjoining distractors), the other four dissimilar In each block of four stimuli, the order of the distractors in the letter-array was counterbalanced The position of the distractors in the stimulus was systematically varied In each block of four similar items, the distractors occurred twice in positions 2 and 3 (permutated for relative order), once in positions 1 and 2, and once in positions 3 and 4 Similar variations were applied for blocks of dissimilar items In these items, the distractors occurred twice in positions 1 and 4, once in positions 2 and 4, and once in positions 1 and 3 In each block of four stimuli, each of the eight remaining letters occurred once

To limit the use of a simple decision-rule like a long, straight horizontal line at the top is a yes' by the subject, all nontargets with a *m* were assigned to the *irregular intra-spaced conditions* The *m is a letter* which can be stretched in a similar way as the *n* With this general exception, the assignment of stimuli to spacing conditions was counterbalanced for combinations of distractors and positions

d *control conditions* For similar and dissimilar target stimuli, basic sets of four stimuli were created In each of these sets, the position of the *n* was systematically varied in combination with a particular distractor The *n* always preceded the distractor in the array, except for *n* in last position Each of the remaining letters occurred once in the sets of four stimuli Assignment of items to stretched or contracted condition was made according to interlocking latin squares for position of target and combination of target and distractor Control conditions were equal to spacing quadrant conditions with respect to the position of target and target-distractor combinations, but not for the position of target-distractor combinations in the array (cf. conditions in the quadrant). Controls that were applied to the nontarget stimuli in the quadrant were also in effect for nontarget stimuli in the control conditions. Nontargets with *m* were assigned to the stretched control condition only.

Writing conditions

The size of each individual letter was 3 by 3 mm. For the irregular intra-spaced conditions, the *n* was 6 mm long and 3 mm high. The *m* appeared in two different forms in irregular intra-spaced conditions: one in which the distance between the first and second leg was 2 mm and the distance between second and third leg 4 mm. In the other form, these interrelations were reversed. The width of all other letters was not varied. Individual letters differed in slope because of the size requirements (e.g., the *i* is a diagonal line extending 3 mm, but no slope is apparent for the *o*). For some letters, the transition between letter segment and connecting segment is fluent (e.g., for the *e*). Letter-forms were held constant by tracing the chosen letter-forms for every single stimulus. The selected letter-forms appear in Figure 2.4.

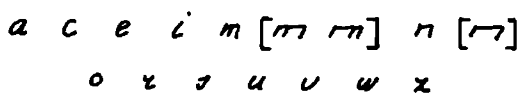


Figure 2.4 Single letter-forms used in Experiment 2. Letters in parentheses were used for irregular intra-spaced conditions.

Connecting segments differed in orientation. For most letters, the outgoing connecting segment was a direct continuation of the last letter-segment. For a few letters (*o*, *v*, *w*), the outgoing segment started at the top of the letter. The connecting segment 'entered' half-way for some letters (e.g., *e*, *s*), but for most at (nearly) the top. In general, connecting segments were drawn as diagonal lines with the top at the right.

In the RR-spaced condition (regular for both inter-spacing and intra-spacing), stimuli were 19.5 mm long. Each connecting segment extended 2.5 mm.

In the IR-spaced condition (irregular inter-spacing, regular intra-spacing), three different forms were used, depending on which of the three connecting segments was lengthened. The lengthened connecting segment was 4 mm, the other two 2 mm each. The total length of the stimuli in this condition was 20 mm Distribution of the three forms over combinations of target-position and distractor was equalized as much as possible. Each form occurred eight times (across target and nontarget stimuli).

Obviously, in the RI-spaced condition (regular inter-spacing, irregular intra-spacing), four different forms had to be used, depending on the position of the *n* or *m*. The irregular letter-form (*n* for targets and *m* for nontargets) extended 6 mm; the remaining three letters 3 mm each. The connecting segments were 1.5 mm each, resulting in a total length of 19.5 mm for stimuli in this condition.

In the II-spaced condition (irregular for both inter-spacing and intra-spacing), one letter was irregular (6 mm) and the connecting segments were of unequal length. In each of these stimuli the lengthened connecting segment was 3 mm, the other two were 1 mm each (the total length amounted to 20 mm). For this condition, the number of possible forms was very large. This diversity was limited in the following way. For stimuli with first or second irregular letter, the third connecting segment (between the third and fourth letter) was lengthened. For stimuli with irregular letters in third or fourth position, the first connecting segment was lengthened.

In the stretched control stimuli, only the *n* (in targets) or *m* (in nontargets) were stretched to 6 mm All other letters were 3 by 3 mm. Every connecting segment in these stimuli extended 4 mm, resulting in an overall length of 27 mm.

In the contracted control stimuli, all connecting segments were reduced to 1.5 mm with constant letter width of 3 mm. The total length of stimuli in this condition amounted to 16.5 mm.

Procedure and Subjects

Stimuli were projected on a translucent screen by means of a Kodak Caroussel slide-projector. Subjects had to perform a *n*-detection task. The subject responded by pressing a *yes* (right hand) or *no* (left hand) button.

The stimuli for the spacing quadrant were presented at a visual angle of 2.02°. The visual angle for the stretched stimuli was 2.48° and 1.41° for the contracted stimuli. Stimuli were projected symmetrically across a fixation point, which was visible between trials. Each stimulus was displayed for 4 seconds maximally, but disappeared when the subject responded.

For order of presentation, the 144 stimuli were divided in 24 groups of six stimuli, one out of each of the spacing quadrant conditions, one stretched, and one contracted stimulus. No more than three target responses succeeded each other. Within these general constraints, order of presentation was random for each subject. Before presenting the experimental items, twenty typed practice items were presented to familiarize the subject with the procedure

After having performed the *n* detection task, subjects were presented the stimuli a second time and were asked to pronounce the individual letters. This second run was intended as a check on the accuracy of the identification process. No latencies were recorded for this pronounciation task. Experimental sessions lasted about one hour.

All 16 subjects were students at the University of Nijmegen. They were paid f 7,- for their participation in the experiment. Data of one subject were discarded because of high error-rates (20 %).

RESULTS

Mean latencies for all conditions are presented in Table 2.2. Results for nontargets proved to be very difficult to interpret. The lack of a clear pattern for these conditions may be due to the fact that different spacing conditions were completely confounded with target-distractor combinations. Combinations with *m* occurred in irregular intra-spaced conditions only. Moreover, as can be seen in Table 2.2, nontargets differed considerably in percentages of error. For these reasons, results with respect to nontargets are left out of consideration.

For the target stimuli in the spacing quadrant, the position of the target was completely confounded with target-distractor combination. In each spacing condition, the *n* occurred four times in each of the four positions (across similar and dissimilar conditions), but each time with a different distractor.

For the latencies in the spacing quadrant, two analyses were carried out, based on the correct responses only. One analysis involved the

Table 2.2
Mean Reaction Times (in milliseconds) and, In Parentheses, Percentages of Errors for Similar and Dissimilar Targets and Non-targets in Six Spacing Conditions.

		Response Type	
Spacing	Similarity	Target	Non-target
RR	Similar	632 (-)	798 (1.6)
	Dissimilar	640 (3.3)	799 (-)
IR	Similar	671 (1.6)	782 (-)
	Dissimilar	695 (5.8)	865 (3.2)
RI	Similar	755 (4.2)	978 (20)
	Dissimilar	691 (4.2)	1157 (21.6)
II	Similar	980 (6.7)	1097 (18.3)
	Dissimilar	747 (-)	907 (10)
Stretched Control	Similar	706 (.8)	925 (5.0)
	Dissimilar	748 (1.6)	886 (16.6)
Contracted Control	Similar	632 (1.6)	726 (-)
	Dissimilar	637 (1.6)	784 (3.3)

position of the target in the letter-array; the other, the target-distractor combinations. For the analyses the two replications per target-position or target-distractor combination were averaged. The *F*-values reported below are based on the analysis for distractor combinations, but results for position effects are also reported. Generally, the two analyses showed *F*-values that differed only slightly for spacing and similarity effects.

Latencies

1. Spacing quadrant.

In the analysis for the target responses (replications were averaged across distractor combinations), the effects of Inter-spacing [F (1,14)

= 32.00, $p < .01$] and Intra-spacing [F (1,14) = 15.97, $p < .01$] were both significant. The main effect of Similarity [F (1,14) = 6.35, $p < .02$] was also significant. For similar responses (distractors adjacent to target), the mean latency was 759 ms; for dissimilar responses (distractors nonadjacent to target), 693 ms.

The factor Distractor-combination was significant [F (3,42) = 4.28, $p < .01$]. Combinations of the target with *m*, *u*, *v*, and w resulted in mean latencies of 755, 754, 711, and 685 ms, respectively.

The two-way interaction between Inter-spacing and Intra-spacing was significant [F (1,14) = 4.55, $p < .05$]. Tests of the simple main effects (Kirk, 1968) showed that Inter-spacing had no effect in regular intra-spaced conditions, but was significant for irregular intra-spaced conditions [F (1,14) = 22.92, $p < .01$]. Intra-spacing was significant both for the regular inter-spaced conditions [F (1,14) = 7.04, $p < .02$] and for the irregular inter-spaced conditions [F (1,14) = 34.89, $p < .01$]. Similarity interacted significantly with Intra-spacing [F (1,14) = 18.28, $p < .01$].

The three-way interaction Inter-spacing × Intra-spacing × Similarity, was also significant [F (1,14) = 7.33, $p < .01$]. The interaction is displayed in Figure 2.5.

The two-way interaction between Similarity and Distractor-combination was significant [F (3,42) = 5.76, $p < .01$]. For the *m*, *u*, *v*, and *w*, the difference between similar and dissimilar versions was 153, 18, 21, and 72 ms, respectively. The three-way interaction Inter-spacing × Similarity × Distractor-combination, was also significant [F (3,42) = 12.51, $p < .01$] as was the interaction between Intra-spacing × Similarity × Distractor-combination [F (3,42) = 8.03, $p < .01$]. Means for the different target-distractor combinations in the spacing quadrant conditions are presented in Table 2.3.

As can be seen from this table, results for the different distractors display a rahter complex pattern. These results deal with intricacies of the segmentation process that are beyond the scope of the present experiment. The experimental materials seem rather complex with respect to configurational properties. Most experiments that have studied configurational effects in visual information processing have used stimuli that seem simpler to describe than the materials used in this experiment (for instance, Palmer, 1977; Pomerantz, 1978; Weisstein and Harris, 1974). This configurational complexity is additionally complicated by the fact that the stimuli can be differentially suggestive for interpretation in meaningful patterns like letters. Interpretation of the effects of

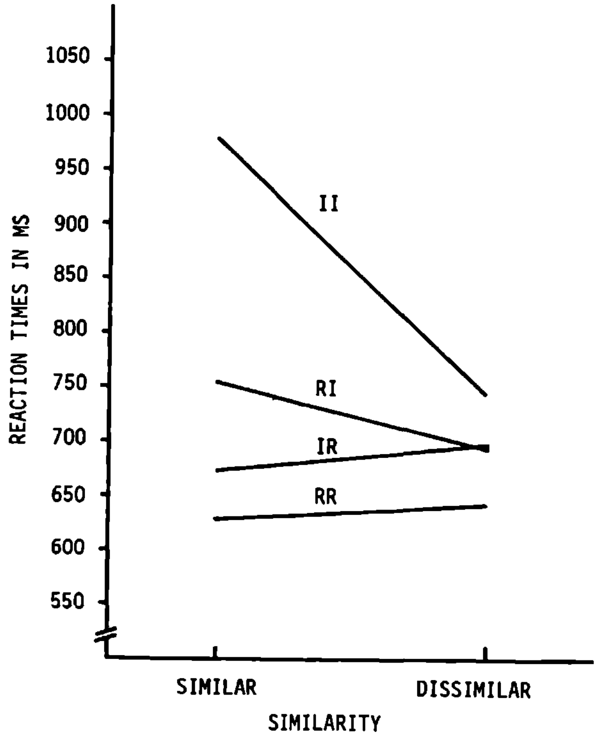


Figure 2.5 Latencies for Spacing Quadrant Conditions as a Function of Similarity (target-responses only).

particular distractors awaits, therefore, more research into relevant configurational aspects of handwritten stimuli for segmentation.

In the analysis for position effects (target-responses only), the main effect of Position was significant [F (3,42) = 25.48, $p < .01$]. The means for positions 1-4 were 651, 694, 751, and 815, respectively. Position interacted significantly only with Inter-spacing [F (3,42) = 3.08, $p < .03$]. The overall larger effects of target-distractor combinations, compared with position effects, have been the main reason to report F-values based on the former analysis.

Table 2.3
Mean Reaction Times (in milliseconds) for Target-Distractor Combinations in Similar and Dissimilar Spacing Conditions (target-responses only) in Experiment 2.

		Distractor			
Spacing	Similarity	M	U	V	W
RR	Similar	575	708	682	561
	Dissimilar	671	614	609	666
IR	Similar	709	644	616	714
	Dissimilar	688	844	671	575
RI	Similar	818	863	656	680
	Dissimilar	660	666	756	685
II	Similar	1223	838	931	927
	Dissimilar	693	856	769	670

Although the *F*-values for distractor and position analyses were generally in close agreement, they differed with respect to the interaction between Similarity and Inter-spacing. This interaction was significant in the position analysis only [$F\ (3,42) = 4.69,\ p < .05$].

2. Control conditions

To analyze differences between the spacing quadrant and control conditions, subject means for target conditions were calculated. Significance of differences were determined by means of a *t*-test for dependent samples summed across subjects. Means for the control conditions are presented in Table 2.2. The stretched control condition was clearly more difficult than the RR-spaced condition, both for similar and dissimilar conditions [$t(14) = -2.75$ and $t(14) = -4.17$, $p < .01$] The other relevant comparison involves the RI-spaced condition which had, like the stretched condition, one stretched letter and regular inter-letter spacing. The stretched control condition did not differ significantly from the RI-spaced condition [$t(14) = 1.79$, $p = .09$ for similar condition, and $t(14) = -1.58$, $p = .14$ for dissimilar condition].

The most relevant comparison for the contracted control condition is the II-spaced condition. In the latter, two of the connecting segments were

1 mm, causing the adjoining letters to be very close together. In the contracted stimuli, all connecting segments were 1.5 mm. Differences between the II-spaced and contracted control condition were significant [$t(14) = 4.62$, and $t(14) = 4.44$ for similar and dissimilar responses]. As an additional check, the contracted control condition was also compared with the RR-spaced condition. No significant differences were found in this analysis

Errors

As can be seen in Table 2 2, error-rates were unevenly distributed across different conditions. Especially high error-rates were obtained for nontargets in irregular intra-spaced conditions Statistical tests for these differences are not provided The results for the pronounciation task showed 4 2 % errors (based on the stimuli as a whole) Most of these errors involved confusions between *u* and *v* and between *r* and *z* For the irregular intra-spaced nontarget stimuli, segmentation errors occurred the stretched *m* was sometimes identified as *n* and *r*

DISCUSSION

In the introduction to this experiment, it was proposed that Gestalt principles for segmentation are operative in the discrimination of letter contours Excluding closed boundaries, the experiment studied the effect of spatial contiguity by introducing irregular interrelations between inter-spacing and intra-spacing The operation of the similarity principle was studied by means of adjacency of letters with similar features Before discussing the effects of these two manipulations, the effects of target-position in the array will be assessed As this factor was completely confounded with Distractor-combination, interpretations of these effects are, of course, subject to strong limitations Nevertheless, it seems that from the patterns of interactions some inferences can be drawn about the processing of the stimuli presented in the experiment. Firstly, it should be noted that the factor Distractor-combination interacted with various other factors that appeared to affect RT. Distractor-combinations interacted with both Inter-spacing and Intra-spacing, and with Similarity Target-position, on the other hand, interacted only with Inter-spacing (at a 3 % level of significance) It is proposed that this latter interaction should be interpreted as an accidental effect of target-distractor combinations This interpretation may not be implausible, considering the overall patterns of interactions and the fact that effects of Inter-spacing appeared to depend on intra-spaced conditions.

The significant main effect of Position does not force one to assume serial visual processing. After the stimulus is coded in letters, response execution requires a read-out of this representation. Because latencies increased with the position of the target in the array (about 60 ms per position), this read-out process can be assumed to be serial and self-terminating (Sternberg, 1966). The absence of interactions between target-position and other factors (with the one exception mentioned above) in combination with the about equal differences between mean latencies for different positions, seems to exclude the possibility that attention was selectively directed to only parts of the stimulus and that response execution was carried out on the basis of this partial information (Shiffrin and Geisler, 1973). It appears more likely that the subject coded the whole stimulus in letters before initiating his response. This assumption of complete visual coding is important for the interpretation of other effects. Differences between conditions do not primarily reflect the speed with which the *n* was detected, but, more generally, the speed with which the stimulus as a whole was coded in letters.

In the *spacing quadrant* conditions, Inter-spacing had no effect in regular intra-spaced conditions, but was significant for irregular intra-spaced conditions. Intra-spacing was significant in both regular and irregular inter-spacing. These results show that the spatial contiguity principle is indeed important for segmentation. This interpretation is supported by a detailed consideration of the metrics of the stimuli. (A different explanation will be considered in connection with the results for the stretched control condition). In the RR-spaced condition, the width of the letters was 3 mm and each of the connecting segments was 2.5 mm. In this condition, the distance between features within letters was virtually equal to the distance between features of adjoining letters. In the IR-spaced condition, the lengthened connecting segment (4 mm) resulted in a distance between features of some adjoining letters of only 2 mm, which was smaller than the relative distance between features within letters. Similarly, in the RI-spaced condition, the stretching of the *n* reduced each of the connecting segments to 1.5 mm. This length was only half the distance between features within letters. Following this line of reasoning, it will be clear that the II-spaced condition meant a relatively dense concentration of features in

certain segments which were difficult to interpret Additional support for the relevance of spatial contiguity in segmentation is provided by the effects of similarity discussed below.
The correct responses for irregular spaced conditions suggest that the system must be flexible in being able to carry out segmentations that violate the principle of spatial contiguity. On the basis of relative spatial contiguity alone, features in the irregular spaced conditions will be combined incorrectly Nevertheless, as can be seen in Table 2 2, the target responses did have rather low error-rates. This means that a re-combination of features must have been carried out, otherwise the irregular spaced forms would have resulted in extremely high error-rates or in 'illegible' responses Grouping by spatial contiguity should therefore be regarded as a foremost heuristic principle for segmentation, which will be successful in virtually all cases
The low error rates for the irregular spaced conditions also suggest that segmentations need to satisfy certain sufficiency criteria. The special (and natural) status of letters for segmentation procedures consists in the fact that they are the sole criterion for successful segmentations; presumably a set of letter candidates is checked against the features in the stimulus to see whether they (the letters) are necessary and sufficient When a certain set does not satisfy the sufficiency criteria, new segmentations will have to be carried out.
A complicating factor resides in the fact that sufficiency criteria themselves may be flexible. It is well-known that subjects neglect or even distort local details in order to arrive at a coherent semantic interpretation (e g., Palmer, 1975). In reading, overlooking spelling errors has been regarded as a reflection of the inaccuracy of a post-access check (O'Connor and Forster, 1981) In the results of this experiment, a similar phenomenon may have been present in the high error-rates for the nontargets in irregular intra-spaced conditions The asymmetrical *m* in these conditions strongly suggested an *n* and may have caused the subject to neglect the remaining part of the *m* as a strange, irrelevant squiggle. At the same time, the normal symmetry of the *m* was absent, making an interpretation as *m* even less likely Results for the pronounciation task revealed, on the other hand, that the stretched *m* may also have been interpreted as an *r* and *n*

Effects of *similarity* of adjoining features on segmentation were found to depend on spatial contiguity conditions In the RR-spaced and IR-spaced conditions, the effect of similarity was small and slightly in the opposite direction. Clear effects of similarity occurred in the

RI-spaced and II-spaced conditions. These results indicate that similarity and spatial contiguity of features create contours within the configuration as a whole that are able to activate letter codes. Due to irregularities in spatial contiguity, adjoining features of adjacent letters may themselves activate letter representations. This activation will, in general, depend on the configurational properties of a certain segment. The effects of similarity therefore suggest that segmentation is jointly determined by global configurational aspects and configurational aspects of single letters Within the whole configuration, spatial contiguity and similarity of features may create misleading letter contours, but this also means that letter recognition is automatically 'triggered' by specific configurations of features.
The effect of similarity raises the same problems with respect to the occurrence of correct readings as the irregular spaced conditions. Analogous to the interpretation considered for the irregular spaced conditions, it is assumed that effects of similarity have to do with sufficiency demands If a certain contour activates a letter representation directly, strong evidence is provided for the presence of that letter. The system will initially try out different segmentations for remaining segments, leaving the recognized 'letter' intact Only when these attempts fail will the 'letter' be broken up in segments to allow for more segmentation possibilities. The similar items were more difficult to recognize because they forced the system to re-interpret strong evidence for a 'good' segmentation in order to arrive at a veridical representation of the stimulus in letters

The *stretched control condition* was introduced as a check on the interpretation of spacing effects as due the unfamiliar appearance of the stretched *n* The stretched condition did not differ significantly from the RI-spaced condition but was more difficult than the RR-spaced condition. Because the planned comparison between the stretched and RI-spaced condition showed no significant difference, the unfamiliarity explanation cannot be ruled out as an alternative interpretation for the differences found in the spacing quadrant conditions It should be noted, however, that this interpretation does not explain the significant difference between the RI-spaced and II-spaced condition
The unfamiliarity of the stretched *n* had to do with the violation of a 'normal' height-width ratio This ratio determines the relative distance between features within the letter and may contribute to a rather characteristic contour The results for the stretched control condition suggest that the speed of letter recognition is affected by the

distinctiveness of the letter, the extent to which the configuration of its features approaches a standard form.
Template-matching models for visual pattern recognition include initial preprocessing procedures to 'clean up' the original input. Standard examples from the text-books are operations on size or orientation The longer latencies for the stretched *n* condition suggest another preprocessing mechanism: the normalization to a standard height-width ratio.

The results for the *contracted control condition* make it unlikely that the effects of different spacing conditions should be attributed to differences in lateral inhibition Otherwise the contracted control condition should have resulted in latencies comparable to those obtained for the RI-spaced and II-spaced conditions The latencies for the contracted control condition were, however, of the same order of magnitude as observed for the RR-spaced condition It will be noted that in both conditions, relations between Inter-spacing and Intra-spacing were regular and that the *n* had a normal height-width ratio

In summary, the results provided evidence that global configurational aspects of the stimulus having to do with spatial contiguity and similarity of adjoining features play a role in the discrimination of letter contours. Indications were found that letter recognition is affected by the distinctiveness of the letter In the experiments, reported in the next chapter, it is investigated whether this distinctiveness also contributes to segmentation

CHAPTER 3

GRAPHONYMY

As was noted in the introduction to Experiment 2, segmentation may not only be determined by global configurational aspects of the stimulus, but also by aspects of configurations of features at the local letter level Changes in these latter configurations may, however, simultaneously affect letter recognition, as was shown in Experiment 2. Stretched *n*'s were generally less easily identified than *n*'s with a normal height-width ratio. This suggested that the speed of the letter recognition is affected by the degree in which a letter approximates a standard form. This property of configurations of features at the letter level was referred to as the distinctiveness of the letter

To demonstrate that segmentation is affected by local configurational aspects, the distinctiveness of letters and the 'segmentability' of configurational context were varied orthogonally in Experiments 3 and 4. If segmentation is also affected by more local configurational aspects, effects of distinctiveness of letters will depend on the configurational context. Larger effects of distinctiveness of letters may be expected for configurational contexts that are more difficult to segment.

It wil be noted that the segmentation process is manipulated by variations in the segmentability of configurations, and letter recognition by variations in the distinctivenss of the letter An interaction between configurational context and distinctiveness of letters will indicate that segmentation and letter recognition cannot be assigned to different stages in processing (Sternberg, 1969)

In normal, everyday handwriting, 'sloppy writing' will decrease the distinctiveness of single letters. Such writing will impair letter recognition, but may also have an effect on segmentation, as is evidenced by graphonyms (Sommer, 1977). Graphonyms can be defined as configurations of features that can plausibly be segmented in different ways; i e , they are ambiguous with respect to segmentation For this ambiguity to occur, the contour of the graphonym must be compatible with different sets of letters An example of a graphonym, reproduced in Figure 3.1, is given by Rumelhart

The ambiguous segment at the beginning of the word can be read either as *ev* or as *w*. 'Graphonymic segment' and 'graphonym' are used interchangeably. Strictly speaking, a graphonym is a whole word that can be read in two (or more) ways. In the usage adopted here, graphonym-

Figure 3.1. Example of a graphonym (after Rumelhart, 1977)

ic stimuli also denote items which contain (potentially) graphonymic segments. Familiar graphonyms are *eu* (w) or *ui* (*m*). Graphonymic segments are not limited to cases in which one letter can be read as two. Combinations of two letters which can be read as two different letters also occur (*nu-mi*) It is even conceivable that in longer sequences (*muvnum*) many alternative readings are possible. Graphonyms are also not limited to certain letters, like *m*, *n*, *u* or *v*. Less obvious letters like *k* (*l* and *i*), *d* (*c* and *l*) or even *a* (*c* and *i*) can also be graphonyms. In general, the occurrence of graphonyms is heavily dependent on the particular handwriting, i e., on the way particular letters are written. In all cases, however, the same contour can be interpreted as different sets of consecutive letters.

As Sommer (ibid.) remarks, graphonyms, being ambiguous figures, require some fudging This fudging implies that the letters that constitute the graphonym cannot be prototypically clear but must be rather indistinctive. The graphonym in Figure 3 1 shows this fudging in the relative closedness of the *e*. If the closed loop of the *e* had been more pronounced, different interpretations would not have been possible. On the other hand, the relative closedness of the *e* may only be important in contexts where ambiguity might arise but irrelevant to segmentation or identification in others. It will be clear that graphonyms (non-experimentally) demonstrate that indistinctiveness of letters may have an effect on segmentation.

It should be noted that graphonyms cannot generally be explained as part-whole relationships between letters Many letters show these relations (like *o*, *c*, and *e*) but sequences of these letters are not likely to cause segmentation problems (although there might be some uncertainty about the identity of the letters) For the same reason, graphonymy cannot be dealt with as a complex instance of lateral inhibition.

For the experiments described below, stimuli were selected which contained segments that were potentially graphonymic These segments consisted in all cases of an *i*, preceded or followed by one or more *n*'s, *m*'s, or *u*'s in various combinations It was assumed that connected, consecutive *n* s, *m* s, *u* s, and (undotted) *i*'s constitute configurations that would be more difficult to segment because of an increased likelihood that features would be combined in a wrong way Nongraphonymic stimuli also contained an *i*, but in these stimuli the *i* was surrounded by letters, which had mostly round contours or by ascenders and descenders It will be noticed that the graphonymic stimuli contained configurations in which similar features were spatially contiguous. The results of Experiment 2 indicated that such handwritten segments are difficult to process Longer latencies for the graphonymic stimuli will therefore provide confirming evidence for the operation of Gestalt principles in creating letter-boundaries

The effect of the distinctiveness of letters on segmentation was studied by means of the dot on the *i* The dot has the advantage of being an isolable, salient feature which presumably makes the *i* more distinctive. Leaving out the dot will slow the identification of the letter *i*, but if the dot also affects segmentation, the effect of omitting the dot will be different for graphonymic stimuli and nongraphonymic stimuli For the graphonymic stimuli, which are more difficult to segment, the dot on the *i* will not only facilitate the perception of the *i*, but might also resolve the ambiguity of the graphonymic segment or limit the number of alternative segmentations that need to be considered. The effect of dotting the *i* was, therefore, expected to be larger for the graphonymic stimuli.

An aspect of segmentation procedures not investigated thus far is the effect of linguistic context In Experiment 1, words were used for a naming task The use of such meaningful materials might have made the experiment less sensitive for detecting differences in initial visual processing In Experiment 2, linguistic context was therefore kept to a minimum by using illegal pseudo-words Pattern recognition has often been described as being both bottom-up and top-down (Lindsay and Norman, 1972; Palmer, 1977) Lexical context has been shown to facilitate letter recognition (Reicher, 1969) and might likewise have an effect on segmentation Linguistic context may particularly exert its influence by inducing guessing letters for ambiguous graphonymic segments. On the basis of letters that are identified relatively fast, words may be activated that suggest letters for the ambiguous segment Such top-down processing can easily be demonstrated with Rumelhart's graphonym in Figure 3 1 Placing the first segment in the context '--idence', where interpretation as w would make a nonword, shows that lexical context may bias the perception of constituent letters and hence segmentation. It is even conceivable that the ambiguity of graphonymic segments might not be resolved in nonwords To assess the effect of meaningful linguistic context on segmentation, a lexical-decision task was used that allowed presentation of both words and nonwords If lexical context influences segmentation, it may be expected that differences between words and nonwords will be larger for the graphonymic stimuli On the basis of segments that are identified rather fast, lexical candidates will be activated that may guide the segmentation of the graphonymic segment. Reading graphonymic nonwords will be especially difficult because processing of the graphonymic segment will have to proceed in a completely bottom-up fashion The contribution of linguistic context to segmentation will be less for the nongraphonymic stimuli in which contours were more indicative of letter-boundaries

EXPERIMENT 3

METHOD

Handwriting and stimulus materials

In this experiment, legibility of handwriting was not varied (cf Experiment 1). Pilot work had shown that even in fairly legible handwritings, considerable effort is required to decipher sequences of adjoining *n*'s, *m*'s, or *u*'s, due to the indistinctiveness of single letters. Such handwritten materials would leave open the possibility that differences between the graphonymic and nongraphonymic conditions would primarily be due to the relative distinctiveness of letters, other than the *i*, in the two conditions To control for this confounding factor, the decision was made to write all stimuli in the standard handwriting taught at Dutch primary schools, for which a good legibility of individual letters may be assumed Even in such handwriting the graphonymic segments will be more difficult to segment than nongraphonymic segments Examples of the presented stimuli are presented in Figure 3 2

In the graphonymic stimuli, the letter *i* was adjacent in all cases to at least one *m*, *n*, or *u* Stimuli in this condition were, however, rather heterogeneous with respect to the length of the graphonymic segment The number of consecutive 'legs varied from three (as in knie' ('knee')) to eight (as in 'onmin (discord)) The position of the graphonymic segment in the stimulus was not systematically varied

All nongraphonymic items also contained an *i*, but in no item was this letter next to an *n*, *m*, or *u*. With a few exceptions, these letters did

Figure 3.2 Examples of stimuli presented in Experiment 3.

not occur in these stimuli. The selection of nongraphonymic items was primarily determined by contour characteristics of adjacent letters of the *i*. By preference, these contours should be maximally constrastive with the *i*. Therefore, items in which adjacent letters were ascenders or descenders (e.g., 'spil' ('pivot')) or had round contours (e.g., 'spion' ('spy')) were selected. Not all items satisfied this criterion, resulting in a larger variety of contours for adjacent letters (mainly *a*, *e*, *r*, *s*). The mean frequency of the graphonymic and nongraphonymic words was 38 and 13, respectively (Uit den Boogaart, 1975) The direction of this difference was contrary to the hypothesis.

For both the graphonymic and the nongraphonymic stimuli, most nonwords were formed as anagrams of the words In this way words and nonwords were roughly equated for constituent letters. All nonwords were legal pseudo-words according to Dutch orthography.

To introduce uncertainty about the number of letters, stimuli consisted of either four or five letters To avoid additional cues for segmentation, the physical length of the stimuli was held constant at about 1.5 cm.

Twenty items were selected for each combination of levels of factors (except the dot), the total set amounted to 160 items

Two versions of each item were made that were exactly the same, except for the presence of the dot on the *i* First a slide was made of the item without a dot, then the dot was added (above the *i*) and a new slide was made. Each item (type) was presented twice, but in two different versions (tokens), with and without the dot on the *i*. Appendix C contains a list of the individual items.

Procedure and Subjects

Stimuli were projected on a translucent screen by means of a Kodak Caroussel slide-projector Stimuli were presented within a visual angle of 2° and centered around a fixation point, which was visible between trials. The experimental task for the subject was vocal lexical decision, responding 'yes' for words and 'no' for nonwords. The 320 stimuli were divided into two sets of 160 each. Each subject was first presented all 160 items (types) and then their 160 counterparts. Within this general constraint, the order of presentation was random for each subject. Prior to the experimental items, twenty handwritten stimuli were presented to familiarize the subject with the procedure. Experimental sessions lasted about one hour.

The twenty subjects were all students at the University of Nijmegen. They were paid f 7,- for their participation in the experiment. Data of one subject were discarded because of extremely long latencies.

RESULTS

Latencies

Means for all conditions are presented in Table 3.1. Separate analyses were carried out for subjects and items (Clark, 1973). The subject analysis was based on the mean RT per condition per subject (every mean was based on 20 observations). The item analysis was based on the means for the individual items. Errors were excluded from both analyses.

The main effect of Graphonymy was significant only in the subject analysis [F_s (1,18) = 5.80, $p < .05$]. Latencies for nongraphonymic stimuli were shorter (829 ms) than for graphonymic stimuli (844 ms).

The dot on the *i* had a significant effect on lexical decision times [F_s (1,18) = 14.23; F_i (1,152) = 32.30; *min* F' (1,59) = 8.49, $p < .01$]. Stimuli without a dot resulted in longer latencies (850 ms) than stimuli with a dot (823 ms).

The lexical status of the letter sequences had a marked effect on latencies, both in the subject analysis [F_s (1,18) = 82.35, $p < .01$] and in the item analysis [F_i (1,152) = 92.71, $p < .01$; *min* F' (1,59) = 43.61, $p < .01$] Words were responded to faster (784 ms) than nonwords (890 ms).

Stimuli with four letters were identified faster (824 ms) than stimuli with five letters (850 ms). This difference was significant in the subject analysis [F_s (1,18) = 14.29, $p <$ 01] and in the item analysis [F_i (1,152) = 5.77, $p < .05$; *min* F' (1,126) = 4.11, $p < .05$].

The interaction between Graphonymy and Dot was significant in the subject analysis only [F_s (1,18) = 5 10, $p < .05$] The pattern of this interaction was, however, contrary to expectations, the dot had a larger

Table 3.1
Mean Latencies (in milliseconds) and, In Parentheses, Percentages of Errors for Conditions in Experiment 3.

Graphonymy	Dot on 'I'	Lexical Status	Stimulus Length in Letters 4	5
Graphonyms	Dotted	Words	746 (1.3)	799 (2.4)
		Nonwords	899 (1.8)	893 (.8)
	Undotted	Words	780 (2.6)	812 (5.8)
		Nonwords	909 (3.9)	917 (.3)
Nongraphonyms	Dotted	Words	749 (2.1)	769 (2.6)
		Nonwords	858 (2.6)	874 (2.9)
	Undotted	Words	780 (3.4)	838 (4.2)
		Nonwords	868 (2.9)	899 (2.4)

effect for the nongraphonymic stimuli. The mean RT for dotted graphonyms was 834 ms and 855 ms for undotted graphonyms. For the nongraphonyms the corresponding figures were 814 ms and 845 ms.
The presence of the dot showed differential effects for words and nonwords in the item analysis [F_i (1,152) = 4.44, $p < .05$]. For dotted words the mean RT was 766 ms and for undotted words 803 ms. Corresponding means for the nonwords were 881 ms and 898 ms.
The interaction between Graphonymy and Word was significant in the subject analysis [F_s (1,18) = 21.52, $p < .01$]. The mean RT of the graphonymic and nongraphonymic words was 784 ms. The mean RT was 905 ms for the graphonymic nonwords and 875 ms for the nongraphonymic nonwords. Separate analyses for order effects (each stimulus type was presented twice) are presented in Appendix D.

Errors
The total number of errors amounted to 2.5 % of all data. Wrong decisions ('yes' for nonwords and 'no' for words) accounted for 1.8 %. As can be seen from Table 3.1, slightly more errors occurred for the undotted stimuli than for their dotted counterparts and more errors were made for words than for nonwords.

DISCUSSION

Although only the main effects (except Graphonymy) appeared to be significant in both the subject and the item analysis, the interactions that were significant in either the subject or the item analysis, are also discussed. Segmentation as part of the initial processing may in general be completed very fast, which will make reliable experimental findings more difficult to obtain. The significant interactions found in this experiment can be regarded as potentially revealing trends in the data for which more convincing support may be found in subsequent research. The interpretation of these interactions can, of course, only be tentative.
The factor Graphonymy reflected conditions under which the segmentation process was assumed to differ in relative speed. Because the graphonymic sequences were potentially ambiguous with respect to letter-boundaries, longer latencies were expected for these stimuli than for nongraphonymic stimuli. The effect of Graphonymy was only significant in the subject analysis at the 5 % level.
Several factors may have contributed to the relative weakness of the effect. The graphonymic segments may not have been ambiguous for segmentation. As was mentioned above, stimuli were written in such a way that good legibility for single letters could be assumed. These writing conditions will reduce the potential ambiguity of the graphonymic segments. Another important cause of the weakness of the effect may have been the contribution of other letters, outside the graphonymic segment, to the segmentation. When these letters activate a word, the graphonymic segment 'falls into letters'. The observed interaction between Graphonymy and Word clearly supports this suggestion. Graphonymic nonwords were the most difficult to process because letters, outside the graphonymic segment, did not offer (correct) lexical candidates to guide the letter perception in graphonymic segments.

The presence or absence of the dot on the *i* was significant in both the subject and the item analysis The distinctiveness of the *i* was manipulated by means of the dot. The generally shorter latencies for the dotted stimuli indicate that letter recognition is facilitated by an increased distinctiveness of the letter (cf Experiment 2)

Effects of increased distinctiveness of letters on segmentation should be apparent in a significant interaction between Graphonymy and Dot Contrary to expectations, the effect of the dot appeared to be larger for the nongraphonymic stimuli (the interaction was significant in the subject analysis only) A plausible explanation of this result is not easy to find and will have to invoke configurational aspects of the nongraphonymic stimuli (cf Experiment 2) As the contours of the neighbouring letters of the *i* in these stimuli were quite diverse, the absence of the dot on the *i* may have resulted in unexpected complex configurations in which different letters were difficult to discriminate Consistent with this observation is the weak effect of graphonymy in this experiment Stimulus materials in Experiment 2, in which similar features were also spatially contiguous, resulted in generally longer latencies It was predicted that larger effects of letter distinctiveness would be observed in contexts that were more difficult to segment. Because the graphonymic and nongraphonymic were not sufficiently different in this respect, the expected interaction was not observed

To investigate the interaction more closely, a post-hoc analysis was made of the graphonymic stimuli As noted above, this set of stimuli was rather heterogeneous with respect to the number of consecutive legs, i e , the number of adjacent *m* s, *n*'s, or *u*'s Longer graphonymic segments will be more difficult to segment Mean latencies for undotted versions were therefore expected to increase with the length of the graphonymic segment Moreover, the dot was expected to have a larger effect as the number of legs increased, i e , the mean difference-scores between the dotted and undotted tokens of types should increase with the number of legs In Table 3 2, the results of this post-hoc analysis are presented As can be seen in this table, the supposed relationships held rather well for the words, but were quite irregular for the nonwords

A notable aspect of the mean RT for the nonwords is that they are centered at about 900 ms An explanation of the irregular relations may therefore invoke the notion of a dead-line (Coltheart et al , 1977) A dead-line processing assumes that a no -response is made if after *t* ms no word has been recognized Applying a dead-line decision rule for nonwords would mean that the subject stopped trying different segmen-

Table 3.2

Mean Reaction Times (in milliseconds) for Undotted Versions and Mean Difference-scores (in milliseconds) between Undotted for Undotted Versions Based on the Number of Consecutive Legs in the Stimulus.

		Mean RT Undotted Versions		Mean Difference-scores Undotted-dotted Versions	
Number of Consecutive Legs		Words	Nonwords	Words	Nonwords
Nongraphonymic	1 (40,40)[a]	809	884	50	18
	3 (17,16)	779	926	9	5
	4 (8, 7)	791	883	13	-11
Graphonymic	5 (5, 6)	742	934	24	48
	6 (9,11)	832	900	63	25
	8 (1,-)	1078	-	-3	-

[a] The first number in parentheses is the number of words in each condition; the second the number of nonwords.

tations. Such a strategy will make the RT independent of the number of alternative segmentations, if this number exceeds a certain critical value.

One might also consider that word-like properties of the nonwords have contributed to the irregular relationships. Anomalous findings like a negative difference-score for nonwords with four legs could be explained by assuming that stimuli with a dot erroneously activated word-detectors. Such initially false recognitions as words would be detected in a post-access check (Forster, 1976) and would force the subject to a new analysis of the stimulus. The versions without a dot activated less word-detectors and were classified as nonwords comparatively faster.

The finding that words were classified faster than nonwords agrees with the results of numerous studies that have used the lexical decision task. Different models have been proposed to explain this basic finding (Forster, 1976; Morton, 1969, 1979). The significant interaction between Dot and Word (in the item analysis) can easily be explained in the framework of Morton's logogen model The visual input logogen functions as a counting device that accepts sensory evidence The presence of the dot, being clear bottom-up evidence for the letter *i*, meant that the threshold for the logogen was reached earlier.

EXPERIMENT 4

Variations in the distinctiveness of letters will not only affect letter identification processes but may also have an effect on segmentation, as is demonstrated by graphonyms Dependent on the configurational context, the indistinctiveness of letters may lead to segments that are ambiguous for segmentation In Experiment 3, it was attempted to create such configurations by leaving out the dot on the *i* in sequences of adjoining *m*'s, *n* s, and *u*'s The effect of adding the dot was expected to be larger for these graphonymic stimuli than for nongraphonymic stimuli *in which the i was surrounded by letters with quite different contours* than the *i* Results showed the effect of the dot, however, to be larger for the nongraphonymic stimuli This negative result was attributed to the small effect of graphonymy itself Experiment 3 was designed to demonstrate that effects of the distinctiveness of letters vary with the segmentability of the configurational context It was assumed that graphonymic and nongraphonymic stimuli represented configurations that differed in the speed with which the segmentation process is accomplished Because all stimuli were written in a handwriting with good legibility for single letters, the intended segmentation ambiguity of the graphonymic stimuli might have been lost However, because graphonymic and nongraphonymic stimuli were about equally difficult to segment, the predicted larger effect of the dot for the graphonymic stimuli was not observed

Experiment 4 was a replication of Experiment 3 with the same experimental manipulations Graphonomy, Dot, Word, and Letter-length An additional factor Connectedness was added to the experimental design: stimuli were segmented or connected forms Connecting line-segments between letters were removed in segmented forms, but were present in connected forms Both Connectedness and Graphonymy were factors that intended to manipulate the speed of the segmentation process It was expected that Connectedness would be more effective in creating configurations that differ in speed of segmentation than Graphonymy in Experiment 3. The same manipulation was applied in Experiment 1. Words were presented in which letters were connected or physically separated by inter-letter spaces. No significant differences were found between the two conditions. Experiment 4 provide data comparable to those obtained in Experiment 1 and will demonstrate whether the results obtained in Experiment 1 are replicable.

To explain the results of Experiment 3, it was assumed that graphonymic and nongraphonymic stimuli were about equal with respect to segmentation. If this assumption is correct, no interaction will be observed between Connectedness and Graphonymy When graphonymic stimuli are indeed more difficult to segment, the effect of explicit segmentation will, however, be larger for the graphonymic stimuli.

Of special interest in this experiment is the interaction between Connectedness and Dot. Leaving out connecting segments will create configurations in which segmentation is carried out relatively fast and the omission of the dot will primarily slow the recognition of the *i*. In the connected forms, for which segmentation is assumed to be more difficult, the distinctiveness of the *i* may also contribute to segmentation. The effect of dotting the *i* was, therefore, expected to be larger for the connected stimuli.

It will be seen that Experiment 4 closely resembled Experiment 3. The expected larger effect of the dot for the graphonymic stimuli in Experiment 3 is equivalent to the expected larger effect of the dot for the connected forms in this experiment In both conditions, segmentation was assumed to be more difficult and the effect of the distinctiveness of letters was expected to be larger for such configurations.

Predictions for other main effects and interactions were generally the same as formulated in connection with Experiment 3. Significant effects were expected for Graphonymy, Dot, Word, and Letter-length Dotted stimuli will be recognized faster than undotted stimuli and words will be classified faster than nonwords. A replication of the result found in Experiment 3 requires that 4-letter stimuli will result in shorter latencies than 5-letter stimuli

Contrary to expectations, the effect of the dot was found to be larger for the nongraphonymic stimuli in Experiment 3 It will be of interest to observe whether this result will be replicated in this experiment. Different effects of the dot for graphonymic and nongraphonymic stimuli should be especially pronounced for the connected forms For graphonymic and nongraphonymic segmented stimuli, the effect of omitting the dot should be considerably less. A similar prediction can be made for the interaction between Connectedness, Graphonymy and Word Graphonymic nonwords will only be more difficult in connected forms; in segmented forms the potential ambiguity of the graphonymic stimuli is certainly absent.

METHOD

Stimulus materials

The 160 stimuli (types) that were used in Experiment 3 were again used in this experiment. In Experiment 3, twenty stimuli were used for each combination of levels of factors. Because of the introduction of the extra factor Connectedness, this number had to be reduced by half. For the assignment of individual items to the connected or segmented condition, the twenty stimuli in each condition in Experiment 3 were rank ordered for mean RT. Individual items were alternately assigned to the segmented and connected condition to equate stimuli in both conditions for overall speed of identification. Difference-scores between dotted and undotted versions were also taken into account in the assignment. If discrepancies in these difference-scores turned out to be rather large for corresponding classes, re-assignment was made for some of the items. The mean RT and mean difference-scores for the segmented and connected conditions, as based on the results of Experiment 3, are presented in Table 3.3

As can be seen in this table, the balancing of the items was fairly close with respect to the mean RT, but was less good for the mean difference-scores It should be noted, however, that the direction of these differences was contrary to the hypothesis. For the segmented stimuli, smaller differences between dotted and undotted versions were expected than for the connected stimuli. An unavoidable consequence of the assignment procedure was that words and nonwords in a particular condition were no longer equated for constituent letters.

By tracing the original exemplars, the stimuli were also physically identical to the ones used in Experiment 3 In the segmented condition, all connecting segments between letters were left out Tokens with and without a dot were produced in the same way as in Experiment 2.

Procedure and Subjects

Procedure and task were the same as in Experiment 3. Order of presentation was completely counterbalanced across subjects and items.

Table 3.3
Mean Reaction Times (in milliseconds) and Mean Difference-scores between Undotted and Dotted Versions for Segmented and Connected Stimuli as Based on the Results of Experiment 3.

			Mean RT		Mean Difference-scores Undotted-dotted Versions	
Graphonymy	Lexical Status	Letter Length	Connected	Segmented	Connected	Segmented
Graphonyms	Words	4	760	766	24	44
		5	815	796	2	26
	Nonwords	4	901	908	25	-6
		5	908	902	28	18
Nongraphonyms	Words	4	762	768	25	36
		5	807	799	60	80
	Nonwords	4	864	862	3	16
		5	882	893	2	43

Twenty subjects were paid f 7,- for their participation. They were all students at the Institute for General Linguistics at the University of Nijmegen. Again data of one subject had to be discarded (30 % of his latencies were larger than three standard deviation units from the overall mean).

RESULTS

Latencies

As in Experiment 3, separate analyses were carried out on subject means and item means. Means for all conditions are presented in Table 3.4.

The main effect of Connectedness was significant, both in the subject [F_s (1,18) = 37.97, p < .01] and in the item analysis [F_i (1,144) = 6 86, p < .01, *min F'* (1,159) = 5 81, p < 05] Segmented forms (741 msec) were recognized faster than connected forms (766 ms)

The effect for Graphonymy was again rather weak It was significant in the subject analysis only [F_s (1,18) = 9 91, p < 01]. Graphonymic stimuli resulted in longer latencies (759 ms) than nongraphonymic stimuli (748 ms).

Stimuli with a dot were recognized faster (743 ms) than their undotted equivalents (764 ms) This effect was significant both in the subject [F_s (1,18) = 17.44, p < 01] and in the item analysis [F_i (1,144) = 21.98, p < .01; *min F'* (1,54) = 9.72, p < 01].

The effect of the lexical status of letter sequences was also significant in the subject analysis [F_s (1,18) = 31 88, p < .01] and in the item analysis [F_i (1,144) = 48.83, p < 01; *min F'* (1,47) = 19 29, p < 01]. Words were classified faster (721 ms) than nonwords (786 ms)

The effect of Length (four or five letters) was not significant in both analyses

A significant interaction between Connectedness and Dot was obtained. The *interaction* was *significant* both *in* the *subject* [F_s (1,18) = 11.97, p < 01] and in the item analysis [F_i (1,144) = 14 32, p < .01; *min F'* (1,56) = 6.52, p < .05]. The mean RT for dotted segmented forms was 739 ms and 743 ms for the undotted segmented forms. For the connected forms, these figures were 747 ms and 784 ms

In the subject analysis, Connectedness interacted significantly with the word status of the letter sequences [F_s (1,18) = 6 30, p < 05] For segmented and connected words, mean latencies were 705 ms and 737 ms, respectively For the nonwords, the corresponding means were 778 ms and 795 ms.

The significant interaction between Graphonymy and Word [F_s (1,18) = 7.77, p < 05] was a replication of the result found in Experiment 3 Again, the effect was significant only in the subject analysis. The mean RT for graphonymic words was 721 ms and 722 ms for nongraphonymic words. Corresponding latencies for the nonwords were 798 ms and 775 ms

The three-way interaction Connectedness × Dot × Length was significant both in the subject [F_s (1,18) = 5 49, p < 05] and the item analysis [F_i (1,144) = 6.11, p < .05; *min F'* (1,59) = 2 89, n s] For the undotted stimuli, the difference between segmented and connected forms was 50 ms for the 4-letter stimuli and 32 ms for the 5-letter stimuli. For the dotted stimuli, the differences between segmented and con-

Table 3 4
Mean Reaction Times (in milliseconds) and, In Parentheses, Percentages of Error for Conditions in Experiment 4

Graphonymy	Dot on 'I'	Lexical Status	Connected: 4 letters	Connected: 5 letters	Segmented: 4 letters	Segmented: 5 letters
Graphonyms	Dotted	Words	691 (.5)	737 (7.4)	681 (3.2)	705 (4.2)
		Nonwords	791 (2.1)	795 (2.1)	792 (2.6)	791 (1.6)
	Undotted	Words	771 (6.8)	779 (13.7)	692 (3.7)	710 (2.1)
		Nonwords	820 (1.1)	813 (-)	791 (2.1)	790 (.5)
Nongraphonyms	Dotted	Words	700 (.5)	721 (2.6)	723 (2.6)	701 (1.6)
		Nonwords	761 (2.6)	783 (.5)	768 (1.1)	756 (1.6)
	Undotted	Words	732 (1.6)	767 (7.4)	709 (1.1)	721 (2.6)
		Nonwords	812 (2.1)	782 (-)	746 (1.1)	790 (1.6)

nected forms was much smaller the difference was negative for the 4-letter stimuli (-5 ms) and was 21 ms for the 5-letter stimuli The interaction is displayed in Figure 3 3
Analyses for effects of order of presentation are presented in Appendix F

Errors
Experimental error and wrong decisions amounted to 3 5 % of all data The experimental error constituted 9 % Connected, graphonymic words caused a relatively large number of errors Inspection of the raw data showed that these errors were mainly due to a few items

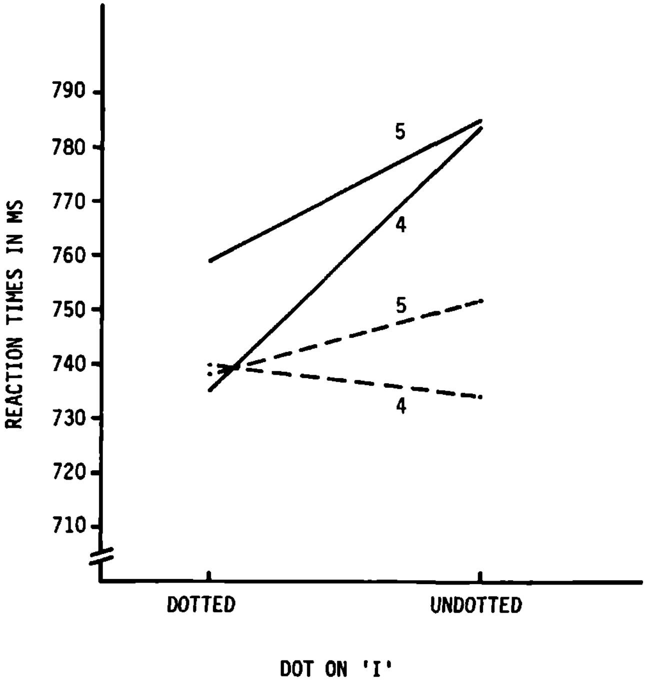


Figure 3.3 Effects of dotting the i for segmented (— — —) and connected (———) stimuli of four and five letters

DISCUSSION

For reasons stated in the discussion of the results for Experiment 3, effects that were only significant in either the subject or the item analysis are also discussed In the first part of this discussion, the effects of the added factor Connectedness are assessed In the second part, the results of Experiments 4 are compared with the results obtained in Experiment 3.

The main effect of Connectedness was found to be significant The facilitation of physically isolating the letters by removing connecting line-segments did not involve configurational aspects of single letters The observed effect must therefore be due to general configurational aspects of the stimulus (cf Experiment 2) It was proposed earlier that letter contours are due to the operation of Gestalt principles for grouping features It may not be implausible to regard explicit segmentation as an instance of the principle of closed boundaries, the inter-letter spaces provide clear boundaries for configurations of features that constitute a letter Explicit segmentation did not result in faster recognition in Experiment 1 The inconsistency in the results is not easily explained and may be related to the different tasks (naming and lexical decision) in both experiments

The interaction between Connectedness and Graphonymy was not significant This result supports the suggestion that graphonymic and nongraphonymic stimuli did not differ with respect to segmentation If the graphonymic stimuli had been more difficult to segment, larger effects of Connectedness for these stimuli had likely been found The graphonymic stimuli did tend to result in higher latencies than nongraphonymic stimuli (the main effect of Graphonymy was significant in the subject analysis) The difference between the two kinds of stimuli might be attributed to other factors than segmentation, e g , differences in the speed with which letters in the two conditions were recognized In the graphonymic stimuli, *m*'s, *n* s, and *u*'s were very frequent and may be more difficult to perceive than letters with ascenders/descenders or closed contours that were used in the nongraphonymic stimuli

The significant interaction between Connectedness and Dot provides evidence that the distinctiveness of letters does facilitate segmentation In segmented forms, the effect of omission of the dot was found to be much smaller than in connected forms The size of the effect was found to be partly dependent on the letter-length of the stimulus The connected forms were configurations of features that were more difficult to segment and the larger effect of the dot in these stimuli suggests that the dot also facilitated segmentation

The interaction between Connectedness and Word was found to be significant in the subject analysis The effect of Connectedness was larger for words than for nonwords It is suggested that this result is comparable with the significant interaction between Dot and Word found in Experiment 3 The effect of adding the dot was found to be larger for the words in Experiment 3 The significant interaction between Connectedness and Dot indicated a virtual equivalence between the dotted and undotted *i* in segmented condition Apparently, the *i* can be made to stand out' in two different ways by adding the dot or by leaving out its connecting line-segments The interaction between Connectedness and Word reflects faster letter recognition, made possible by the explicit segmentation The letter information is fed directly into word-detectors (logogens) that reach threshold earlier

As in Experiment 3, significant main effects of Graphonymy, Dot, and Word were obtained The effect of Letter-length, that was significant in Experiment 3, was not significant in this experiment
In Experiment 3, a significant interaction between Graphonymy and Dot was obtained in the subject analysis A larger effect of adding the dot was observed for the nongraphonymic stimuli This interaction between Graphonymy and Dot was not replicated in this experiment The related three-way interaction between Connectedness, Graphonymy, and Dot was also not significant
As in Experiment 3, a post hoc analysis of the graphonymic stimuli was carried out with respect to the number of consecutive legs (the number of adjoining *n*'s, *m* s, or *u* s) According to the number of consecutive legs, graphonymic stimuli were analyzed for mean RT for undotted versions and mean difference-scores between dotted and undotted versions The results of the analysis are presented in Table 3 5 As can be seen in this table, the overall effect of the dot was larger for the connected forms than for segmented forms, in both words and nonwords This effect reflects the significant interaction between Connectedness and Dot For the connected, graphonymic words, the effect of the dot actually decreased with the number of legs This result is opposite to the one obtained in Experiment 3, in which difference-scores were found to increase with the number of legs In Experiment 3, the mean RT for the undotted words also tended to increase with greater number of legs, but the data in Table 3 5 do not display this trend
A minor, but interesting detail of the data is apparent in the mean RT for undotted nonwords In Experiment 3, it was suggested that possibly a dead-line decision rule was applied for these stimuli with the criterion-value set at about 900 ms In Table 3 5 it can be seen that in Experiment 4, a criterion was set at a much lower level of about 800 ms This criterion-shift may have been caused by the introduction of the segmented forms, which were more easily identified, i e , the criterion-value may be influenced by the overall level of identification-speed Intuitively, a similar phenomenon occurs in reading handwritings with

Table 3.5
Mean Reaction Times (in milliseconds) for Undotted Versions and Mean Difference-scores (in milliseconds) between Undotted and Dotted Versions as a Function of the Number of Consecutive Legs.

	Number of Consecutive Legs	Mean RT Undotted Versions		Mean Difference-scores Undotted-dotted Versions	
		Words	Nonwords	Words	Nonwords
Connected Forms					
Nongraphonymic	1 (20,20)	750	797	39	25
Graphonymic	3 (10, 8)	785	825	89	-4
	4 (4, 3)	772	816	62	28
	5 (3, 1)	724	812	41	65
	6 (2, 8)	786	808	17	43
	8 (1, -)	820	-	-57	-
Segmented Forms					
Nongraphonymic	1 (20,20)	715	768	3	6
Graphonymic	3 (7, 8)	695	805	-3	5
	4 (4, 4)	722	788	24	-13
	5 (2, 5)	714	784	24	-6
	6 (7, 3)	690	768	5	6

poor legibility. The reader knows that something meaningful has been written and accordingly increases his tolerance for deciphering

The significant interaction between Graphonymy and Word in the subject analysis was a replication of the result found in Experiment 3 The effect of graphonymy was larger for the nonwords. In Experiment 3, it was proposed that the lexical context facilitated the segmentation of the graphonymic segment by inducing guessing of certain letters It was predicted that the interaction beween Graphonymy and Word would be especially pronounced for the connected forms (in the segmented forms, the potential ambiguity of the graphonymic stimuli is certainly lost).

The three-way interaction between Connectedness, Graphonymy, and Dot was, however, not significant

The unreliability of the effects of dotting the *i* in graphonymic and nongraphonymic stimuli and the absence of an interaction between Connectedness and Graphonymy support the interpretation of Experiment 3 Graphonymic stimuli did not represent configurations that were clearly more difficult to segment than nongraphonymic stimuli The removal of connecting segments was more effective in this respect connected forms were generally recognized slower than segmented forms In configurations that are more difficult to segment, distinctiveness of letters will facilitate segmentation as was shown by the significant interaction between Connectedness and Dot This result indicates that segmentation is not only determined by global configurational aspects of the stimulus, but also by local configurational aspects at the letter level

GENERAL DISCUSSION

The results of Experiments 1-4 indicate that at least the following three factors have an effect on letter segmentation in cursive handwriting

The configuration of features in the stimulus as a whole will partly determine letter-boundaries Gestalt principles for grouping features are operative in the discrimination of sub-contours within the configuration as a whole that are likely to coincide with letters Configurational aspects that have to do with closed boundaries, spatial contiguity and similarity of adjoining features are of special relevance for segmenting handwriting Although no direct empirical evidence for the relevance of closed boundaries configurations was obtained, inter-letter spaces can be regarded as a rather specific manifestation of this principle, which serves to isolate features that belong to the same letter The facilitation for segmented forms, found in Experiment 4, may accordingly be considered as evidence for the relevance of closed boundaries Effects of violating Gestalt principles of spatial contiguity and similarity of features was found in Experiment 2 Configurations in which letter-boundaries were inconsistent with these principles resulted in significantly longer latencies than configurations in which the application of these principles lead to correct interpretation in letters

The apparent operation of Gestalt principles in the segmentation of handwriting does not exclude the possibility that certain letters are identified without reference to the whole configuration (see below) and are processed virtually simultaneously with more global characteristics of the stimulus (Navon, 1977) Such very fast letter recognition could provide cues about the average letter-size that is to be expected and would constitute valuable additional information for segmentation procedures.

It should be kept in mind that configurational aspects of the stimulus as a whole are inextricable connected with aspects of the the configuration of features at the local letter level (Garner, 1979) This consideration is also relevant for the evaluation of the effect of distinctiveness of letters on segmentation (see below)

The configuration as a whole will play a role in assessing the sufficiency of particular segmentations that have been carried out, i e , it will be used to determine whether a particular set of letters is sufficient and necessary to 'explain all features It seems likely that the strictness of the sufficiency criteria will vary, particularly if performed segmentations result in meaningful words

Indirect evidence for this function of the whole configuration was obtained in Experiment 2, where irregular configurations that violated Gestalt principles for grouping nevertheless resulted in rather low error rates The same may have happened in Experiments 3 and 4 in which initially wrong segmentations for the graphonymic stimuli were probable The assessment of the sufficiency of segmentations may easily be incorporated into models that postulate some kind of verification procedure as part of the word recognition process (cf Becker, 1976, 1979, Forster, 1976)

A second factor that contributes to segmentation is the distinctiveness of letters Letters may become indistinctive by lacking feature information or by deviant spatial relations between features (like the stretched *n* s that were used in Experiment 2) Support for the importance of clear letter forms for segmentation was found in Experiments 3 and 4, in which the effect of letter-distinctiveness (leaving out the dot on the *i*) appeared to depend on configurational contexts which differed in segmentability The most convincing demonstration was found in Experiment 4, in which the effect of leaving out the dot was found to be much larger in connected forms than in segmented forms

The observed interaction between letter-distinctiveness and segmentability of the configurational context has implications for the view that letter recognition is preceded by segmentation In Neisser s analysis of segmentation, for instance, preattentive processes, which are global and wholistic, serve to separate figures from each other, as a framework for subsequent and more detailed analyses which require focal attention Thus, segmentation and letter recognition are assumed to be partly sequential (Neisser, 1967). According to Sternberg's additive factor method (Sternberg, 1965,1967), effects of experimental variables will be independent and additive if they influence different stages An interaction is likely to be found when two factors affect the same stage Applying this logic to the experimental manipulations in Experiments 3 and 4, the following reasoning suggests itself The segmentability of the configuration (manipulated by graphonymy or inter-letter spaces) affects the segmentation, and the letter-distinctiveness (leaving out the dot) will influence subsequent letter recognition Because letter-distinctiveness interacted significantly with configurational context, letter recognition and segmentation cannot be assigned to different stages It was argued above that segmentation is determined by configurational aspects of the stimulus as a whole that have to do with spatial relations between features A common stage for segmentation and letter identification will, therefore, have to do with extracting features and the perception of their spatial relations

Because both segmentation and letter-recognition involve features and their spatial relations, the possibility must be left open that segmentation can be immediate, i.e., based on specific configurations of features that can activate letter representations directly. Letter perception can, therefore, be simultaneous with perception of the configuration as a whole. This interpretation implies that segmentation is not exclusively based on global stimulus characteristics, as some stage theories would have it, but that local letter features can also contribute to segmentation at an early stage of the processing

In the introduction to Experiment 1, it was pointed out that models of the reading process may explain segmentation by postulating separate input channels, segments of the total visual field that coincide with letter positions. These models do not readily handle the segmentation of cursive handwriting, because obvious physical cues like inter-letter spaces that are used for the segregation of input channels are absent in handwriting The segregation of input channels might be based on other cues than inter-letter spaces Banks and Prinzmetal (1976) have suggested that Gestalt principles of organization may be operative in the formation of visual channels They suggested that a perceptual 'parser' is responsible for organizing information in parallel from all over the field into groups, and the groups it formed would be determined by such factors as the Gestalt principles (ibid , p 367). The precise workings of such a parser, however, are as yet completely unspecified. If such input channels would explain the effects of global stimulus characteristics on segmentation, the effect of local features like the dot on segmentation is still not easily accounted for. Effects of global and local configurational aspects are both incorporated in the interpretation presented above

A third factor that influences segmentation is the linguistic context, the fact that a string of letters forms a meaningful word. This factor explains the interaction between lexical status of letter sequences and graphonymy in Experiments 3 and 4
Evidence for this factor is still limited. This is partly due to the availability of alternative explanations (like differences in confusability between graphonymic and nongraphonymic stmuli), but may also be a consequence of the way this factor operates in segmentation. Linguistic context may induce the guessing of letters for segments that are particularly difficult to segment Thus, the effect of the word context on segmentation would only be indirect, i.e , mediated by letter-recognition If this hypothesis is correct, it would imply that a subset of the word's letters can be sufficient to activate the word representation (Rubinstein, Garfield, and Millikan, 1970), and that this representation may, in its turn, facilitate the perception of letters that are identified more slowly It should therefore be noted that linguistic context can only mediate in the process of segmentation, if a substantial difference exists in the identification speed for the different letters of a word If all letters are recognized at about the same time, the segmentation will be complete before the word is recognized The graphonymic stimuli of Experiments 3 and 4, for which linguistic context effects were found, presumably were stimuli in which letters were identified at different speed the graphonymic part of the stimulus was probably slow as far as letter identification was concerned, whereas the surrounding letters were recognized relatively fast

CHAPTER 4

INITIAL PRACTICE EFFECTS IN READING HANDWRITING

The experiments reported in this chapter are related to a set of common observations with respect to reading handwriting. Everyone will have come across handwritings which initially require considerable effort to decipher. After a while, however, the reading of such 'difficult' handwritings tends to become easier, although some seem to resist any practice effects. It is as if one gets accustomed to the specific graphic, physical characteristics of the particular handwriting, especially once one has discovered that 'that odd scribble' stands for the letter *x*. Fortunately, there also exist handwritings which do not require such laborious deciphering and which from the start seem to be read as easily as print. The experiments in this chapter aim to provide some insights in the nature and generality of such initial practice effects in reading handwriting. Initial practice effects will be understood as (gradual) improvements in reading speed that occur during the first and rather limited acquaintance with particular handwritings. Research of these practice effects will indicate how the reader adapts to the characteristics of individual handwritings.

A plausible explanation of initial practice involves perceptual learning of the handwriting. The general theory about perceptual learning developed by LaBerge (LaBerge and Samuels, 1974; LaBerge, 1976) has been worked out in considerable detail for reading print and will be adopted to derive some general notion as to what may be involved in the perceptual learning of handwriting. Because it is unclear whether the theory was intended to cover perceptual learning of handwriting, the results of the experiments reported below cannot directly be taken as relevant empirical evidence for the model. A similar caveat should be made with respect to specific predictions of experimental results. These should be regarded as plausible inferences from LaBerge's theory. On the other hand, it should be noticed that predictions or results may be compatible with similar theories about perceptual learning, for instance Gibson's discrimination theory (Gibson, 1969, 1971; Gibson and Levin,1975). Because the experiments are a first exploration into perceptual learning of handwriting, it can hardly be expected that results can be used to decide between alternative theories. Moreover, theories that have not been explicitly concerned with handwriting may be extended or modified to accomodate particular findings.

LaBerge's theory about perceptual learning encompasses three stages· the feature discovery, the extraction of relevant features; the coding stage, the efficient grouping of features into higher-order codes, and the automatic coding, in which a code can be activated without attention. Feature discovery involves sensitivity to aspects of visual patterns which distinguish one pattern from another. For written language these features are identified with, for instance, lines, angles, and intersections.

According to LaBerge (1976) 'the process by which perceptual elements are selected and grouped into unit codes through experience is considered here to be the crux of perceptual learning' (ibid., p. 247). A code 'represents an analytic operation on the array of features activated by sensory input by resonating to a particular selected set, and also represents a synthesis of this set of features when they are integrated into a single output for further cognitive processing (ibid., p. 245).

In the automatic coding stage, the features are combined into a code automatically, i.e., without the assistance of attention In the coding stage, considerable amounts of attention are still required for the synthesis of the features into a code but due to repeated exposure to the pattern this synthesis becomes automatic. For a perceptual learning interpretation of initial practice in reading handwriting, the first two stages may be especially relevant. Automatic coding may occur only after a considerable amount of practice

From LaBerge's theory it can be deduced that initial perceptual learning of handwriting may involve feature discovery. It cannot be ruled out that different handwritings (or type-faces) require different sets of distinctive features or draw selectively from a universal set of features. For instance, the distinctive feature-chart in Gibson (1969) was intended to describe Roman capital letters. Other proposals for feature-lists for printed language have been made (Massaro and Schmuller, 1975), some with the limited intent to describe a particular type-face (Rumelhart and Siple, 1974) This leaves the possibility that for each single handwriting (slightly) different sets of distinctive features needs to be extracted

Perceptual learning of handwriting may also consist in the combination of features into higher-order codes According to LaBerge, this aspect deals with the perception of combination information' A likely candidate for this kind of information is the spatial relations between the features. The results of Experiment 2 indicated that stretched *n*'s with an a-typical height-width ratio were more difficult to perceive. Perceptual learning may consist in the pickup of such characteristic spatial relations between features. Apprehending that the writer intends his idiosyncratic scribble to stand for the letter *x* may be regarded as a proto-typical example of this kind of learning that may also take place (although unconsciously) for less deviant cases In LaBerge s theory, the higher-order codes which are formed may be letters, spelling-patterns or words These possibilities may also be assumed for handwriting, although it seems likely that at first these codes will primarily be letter codes

Perceptual learning will result in more efficient processing which would explain the improvements in reading speed This efficient processing may especially be brought about by the formation of higher-order codes 'The code which groups features of the input pattern also acts as a selective filter Irrelevant features are screened out because they do not feed into the code which incorporates the grouping of the particular combination of relevant features' (LaBerge, 1976, p 241)

Apparent practice effects differ considerably among handwritings For some handwritings the effects are hardly noticeable while for others obvious improvements in reading speed are observed In Chapter 1, it was argued that sampling procedures for handwritings may consider physical aspects like size and slope or relative legibility Dependent on the physical characteristics of the handwriting, certain distinctive features may be more easily extracted or some kinds of noise may be filtered out more readily than others Because nothing is known about these processes, it is at present not possible to formulate predictions about practice effects based on specific physical characteristics of handwritings

A very general prediction can be made if it is assumed that the initial legibility of a handwriting is determined by its overall resemblance to print In Chapter 1 attention was called to various physical aspects of handwritings that may contribute to their legibility Configurations of these characteristics in single handwritings will result in an overall resemblance to standard type-faces This resemblance will, of course, be difficult to describe formally because of the complex interplay between different physical aspects Because perceptual learning will have taken place for print, small practice effects will be observed for handwritings that physically resemble print rather closely For those handwritings, large amounts of transfer from printed material seem likely Thus, the observation of practice effects can be related to the initial legibility of the handwriting Due to different amounts of transfer, practice effects will increase with reduced legibility of the handwriting

For the experiment described below, nine handwritings were selected that differed markedly in legibility For each of these handwritings, practice material of about 5200 words was provided in the form of seven prose-passages to be read consecutively The experiment also contained a control condition in which the prose-passages were presented in print instead of handwriting The seven selected texts were semantically unrelated, because otherwise decreasing reading times for consecutive texts would be ambiguous they might reflect perceptual learning or ongoing familiarity with the contents of the reading materials

This might be the right place to notice that another, less obvious explanation exists for practice effects observed in everyday reading of handwriting This interpretation of early practice effects deals with semantic/conceptual aspects of the reading process It was pointed out before that reading handwriting will involve both bottom-up and top-down processing The gradual improvement in reading speed observed for some handwritings may reflect more efficient top-down processing To see how such an interpretation might explain initial practice effects, the following kind of considerations are examplary Some recent theories about reading comprehension view text understanding as the construction of an underlying model or schema which organizes and augments the surface structure in the text (Spiro, Bruce, and Brewer, 1980) An example of such a theory is the progressive refinement theory of Collins, Brown, and Larkin (1980) In this theory, it is assumed that text understanding proceeds by progressive refinement of an initial, partial model that is triggered by the beginning elements of the text As the reader proceeds in the text, models become complete and the search for relevant information becomes more constrained The overall process of text understanding is referred to as constraint satisfaction and may be regarded as an instance of increasingly efficient top-down processing It is not implausible to assume that processes like constraint satisfaction may be operative in, for instance, interpreting words that are difficult to read As more information is obtained about the text, better candidates will be selected which will result in increased reading speed

Apart from gradually more efficient top-down processing in the course of a text made possible by the specific contents of the reading material, the reader might also develop (or apply more effectively) more general top-down reading strategies A familiar example of such a stategy is the reading of remaining parts of sentences (or even paragraphs) first before trying to decipher illegibilities

A prediction for the observation of practice effects according to a 'top-down' interpretation is based on the assumption that the contribution of bottom-up and top-down processes to overall reading will differ with the legibility of the handwriting In handwritings with good legibility, the bottom-up processing is fast and the overall reading speed will be determined by the speed of higher-order processes that deal with the comprehension of the reading material In less legible handwritings, upper limits on reading speed may, however, be set by slow bottom-up processes. In such handwritings, top-down processes may lower the limits on reading speed

In a top-down' interpretation, practice effects are understood as changes in the relative speed of bottom-up and top-down processes More specifically, practice effects reflect an increased speed of top-down processes relative to the speed of bottom-up processes In less legible handwritings, the perceptual, bottom-up processing is (and will remain) rather slow Initially this is also true for the top-down processing, but gradually the speed of these processes increases and will become significantly faster than bottom-up processing In handwritings with good legibility, the bottom-up processes will, however, always be faster than the higher-order processes that deal with the comprehension of the material and no practice effects will be apparent Thus, a top-down interpretation of practice effects also makes it likely that these effects will only be observed for less legible handwritings

If perceptual learning for handwriting occurs, reading times for the seven consecutive texts should decrease The size of the decrease will depend on the relative legibility of the handwriting Clear effects should be observed for handwritings with poor legibility but much smaller effects, if any, should be obtained for handwritings with good legibility Practice effects for less legible handwritings in the text materials might also reflect, however, the more efficient application of certain general top-down strategies mentioned above

As a check on the interpretation of decreasing reading times in the text materials, the practice materials were preceded and followed by a task involving the recognition of single words For this pretest/posttest a word naming task was selected Nonwords in less legible handwritings are likely to cause large numbers of illegible responses and therefore the lexical decision task, used in Experiments 3 and 4, is less suited Differences between the pretest and posttest latencies (the pretest/posttest gain scores) will be used as a context-free' measure of perceptual learning The pretest establishes the initial level of legibility of the handwriting, uncontaminated by context-effects beyond the

word; the posttest provides an indication of increased legibility after having read a substantial amount of material, also without the assistance of larger context. If decreasing reading times for the practice materials reflect perceptual learning, results for the pretest/posttest will parallel the results for the practice materials: the size of the gain-scores will vary with relative legibility If the decreasing reading times for the practice materials reflect the more efficient application of top-down reading strategies, no decrease in reading times for single words should be found in the pretest/posttest. Gain scores should be zero for all handwritings, independent of initial legibility
A problem was posed by the task that subjects should perform when reading the practice materials Time to read aloud is too insensitive a measure of speed of reading and would cause too much fatigue when large amounts of materials have to be read Subsequent tests for comprehension would probably induce the subject to perform extensive (unwanted and unnecessary) conceptual or memory processing Salient semantic anomalies were inserted in the texts (Tinker, 1965), which the subject had to detect while silently reading the materials This task provides the necessary check that the subject is indeed reading for comprehension.

EXPERIMENT 5

METHOD

Stimulusmaterials

For the pretest/posttest single word naming task, two series of 10 5-letter nouns were used Words in the two series were matched for frequency None of the words used was a replacement word in the experimental texts (see below) A list of the stimuli appears in Appendix G.

The reading materials consisted of 11 prose-passages of approximately 700 words each Every text contained 10 salient semantic anomalies (e g , At the desk a man was busy putting notes and rain-drops in his suit-case'), which the subject had to mark with a pencil Because a pilot study had shown practice effects on the anomaly detection task itself, 4 of the 11 texts were used as (printed) practice texts, the remaining 7 being the (handwritten) experimental texts Passages were taken from primary (upper grades) and secondary (lower grades) school books. They were descriptive, narrative stories on a variety of topics, ranging from a report of a bank-robbery to a simple outline of the development of the radio. The stories had to fulfil the requirement that they would make fairly interesting reading materials and should not be too redundant. On the other hand, texts should not be so complex that reading times would primarily be a reflection of comprehension difficulty Stories of the above mentioned level should satisfy these two general requirements, and, upon questioning, subjects showed this intuition to be largely correct

An important assumption underlying a valid interpretation of consecutive reading times is that the texts are about equal for ease of comprehension. A check on this assumption is provided for in the data analysis

Apart from differences in overall conceptual difficulty, inequality between texts may also be caused by differences in the saliency of the inserted semantic anomalies The pilot study had shown some anomalies to be more readily detectable than others. To control for these differences, each anomaly was obtained by replacing a word by another that very obviously did not fit the context Further standardization was achieved by distributing the anomalies uniformly over different parts of speech, with the additional advantage of ruling out a possible strategy of the subject in which he would pay attention to certain parts of speech only. Of the 10 anomalies in each text, four involved nouns, three verbs, one an adjective and the remaining two, other parts of speech Anomalies were inserted in the passages at random intervals The passages were read by two colleagues to check for possible metaphorical readings of the replacements

Handwritings

In a separate selection experiment, described in Appendix H, ratings of the relative legibility of 32 handwritings were obtained (12 subjects were used). Each handwriting was presented in the form of a short 50-word prose-passage under which a 5-point scale, running from 'very easy' to 'very difficult was printed for the rating For each handwriting the mean number of points ('very easy' = 1, 'very difficult' = 5) was calculated On the basis of these scores the handwritings were rank-ordered for relative legibility In accord with analyses of the results of the selection experiment, three main categories of legibility -high, medium, and low- were decided upon. The first category includes handwritings which give the impression of being very legible,

they are clear cases of 'easy' handwritings. The second category is formed by handwritings which are of average difficulty. They constitute an 'intermediate' class, neither being very difficult nor particularly easy. The third category includes cases of 'difficult handwriting'; i.e., handwritings with a clearly reduced legibility. To ensure that results would not be too dependent on idiosyncratic characteristics of a single handwriting, three handwritings were used in each of these categories. Mean legibility scores were 1, 1.1, and 1.1 for handwritings in the high legibility condition; 2.6, 2.8, and 3.0 for handwritings in the medium legibility condition and 3.6, 4.2, and 4.7 for handwritings in the low legibility condition. The rank-orders for the three handwritings in the high legibility condition were 1, 2.5, and 2.5; for the 'intermediate' handwritings 15.5, 17, and 19 and for the three handwritings in the low legibility condition 26, 31, and 32.
The nine writers (selected from the original 32) were paid f 50,- for writing the seven experimental texts. They were asked to write consistently in the same style they had written the prose-passages used for the rating. They were instructed not to use indentation, to write proper names in capitals and to start a new page with a new sentence. Due to differences in the size of the handwritings, writers needed two to six pages per story. The original exemplars were reproduced by off-set lithography.

Design and Procedure

Five subjects were randomly assigned to each of the nine handwritings. A control condition was added in which the seven experimental texts were typed. Of the 10 subjects assigned to this control condition, 5 read the experimental texts in Courier typeface, the other 5 in Script typeface, each produced on an IBM electric type-writer. The four practice texts were set in Gothic typeface. Every experimental session began and ended with the presentation of 10 single words (the pretest/posttest). The two series were alternated for the pretest and posttest. Latencies for this single word naming task were recorded by voice-key. The stimuli were projected on a screen, 1.5 m in front of the subject. The visual angle at which the words were presented varied considerably, due to the varying size of the handwritings. There was, however, no systematic relation between legibility and size of the handwritings used. Order of presentation was random for each subject. Before being presented with the words that were typed or handwritten in the same typeface or handwriting in which the subject was going to (or had) read the experimental texts, the subject was given five practice items (in Gothic typeface) to become familiar with the procedure Subjects were instructed to respond as quickly as possible and to say 'illegible' when they could not read a particular word After the pretest, the subject was given the four practice texts and the seven experimental texts to read Order of presentation of texts was random for each subject For each group of five subjects, a particular experimental text was the first or last presented text only once Subjects were told that in each story words had been inserted that did not fit the context They were told to read the stories silently while marking those words The instruction contained an example of the kind of anomaly they could expect Although subjects were told to read fast, the instruction also stressed accuracy It was pointed out that they were not to look for subtle logical errors or faulty reasonings, neither were they expected to mark stylistic infelicities For the reading of the practice texts, feedback (if necessary) was given between texts If the subject made more than one miss during the experimental texts he was told to read more carefully Reading times were measured with a stop-watch Each text was given a number which had to be read aloud as a signal to start timing After having read each text, the subject was to say 'stop' Experimental sessions lasted between an hour and an hour and a quarter for the control and high legibility conditions, up to two hours in the low legibility condition

Subjects

All 55 subjects were students at the University of Nijmegen They were recruited through advertisement in the local university newspaper and were paid f 7,- an hour for their services Two subjects had to be replaced due to poor understanding of task requirements

RESULTS

The experimental design is a multiple time-series design with four groups of subjects (control, high, medium, and low legibility conditions) Pretest and posttest were added (Campbell and Stanley, 1963) No results for single handwritings are reported because generalization involves different degrees of legibility (see Chapter 1)

Word recognition

Latencies

Mean latencies for pretest and posttest word naming for the four groups of subjects are presented in Table 4.1. Due to experimental failure, one subject is missing in the control group.

The analysis of variance, based on the subject means, showed significant differences between the four groups in the pretest [F (3,50) = 12.39, $p < .01$], and in the posttest [F (3,50) = 7.82, $p < .01$].

Table 4 1

Mean Reaction Times (in milliseconds) and, In Parentheses, Numbers of Different Errors for Pretest and Posttest Word Recognition.

	Legibility Condition	Pretest	Posttest	Pretest/posttest Gain Scores
1	Control (n = 9)[a]	659 (-,-)[b]	652 (-,-)	7 (-,-)
2	High (n = 15)	830 (1,-)	797 (2,-)	33 (-1,-)
3	Medium (n = 15)	1101 (19,4)	931 (18,3)	170 (11,1)
4	Low (n = 15)	1093 (16,8)	1049 (13,6)	43 (3,2)

[a] n is the number of subjects.

[b] The first number in parentheses is the number of erroneous readings, the second is the number of 'illegible' responses.

Scheffe's post-hoc test showed in pretest and posttest, differences between group 1 (control) and group 2 (high legibility condition) not to be significant. Group 3 (medium legibility condition) did not differ significantly from group 4 (low legibility condition) in the pretest and posttest Post-hoc tests for the posttest latencies showed the difference between groups 2 and 3 not to be significant ($F < 1$). All other comparisons were significant

A test for different amounts of perceptual learning is provided by a one-way analysis of variance on the pretest/posttest gain scores. The analysis revealed no significant differences between the four groups [F (3,50) = 2 12, $p =$ 11] Table 4 1 clearly shows a trend, however, suggesting that the lack of significance may be due to the large within-group variances (the within-group variances were 8,019, 14 110, 34,907 and 62,187 for groups 1, 2, 3, and 4 respectively) These variances are very heterogeneous [Bartlett s $\chi^2(3) = 28\ 18$, $p < 01$] It was therefore decided to use a non-parametric test to test differences in perceptual learning for the four groups The Kruskall-Wallis test (Siegel, 1956) for the rank ordering of the gain scores showed nonsignificant differences between the four groups [$H = 7\ 00$, $p = 07$] Testing whether the gain scores were significantly different from zero for the four groups separately showed significant results for the medium legibility condition only [$t(14) = 3\ 53$, $p < 01$]

Errors
Missing observations due to experimental error constituted 4 5 % of all data The percentage of erroneous readings (words read incorrectly) was 8 8 for the medium and low legibility condition together, about equally distributed over pretest and posttest Of the words presented in the medium and low legibility conditions, 3 5 % was reported illegible They were also about equally distributed over pretest and posttest

Text materials

a. Reading times for handwritten experimental texts.
It is well known that time-dependent measures may be expected to have large correlates between adjacent responses and will have decreasing correlates as measurements are made further apart in time For this reason an analysis of variance may not be the most adequate technique and an analysis taking into account this serial correlation structure was applied Reading times for the seven consecutive experimental texts were analyzed by growth-curve analysis (e g , Timm, 1975) The relevant statistic Θ, mentioned in the analyses reported below, is the largest root statistic, s, m and n are the corresponding largest root distribution parameters
Two aspects of the curves are of special relevance the intercepts, indicating whether the conditions differ significantly with respect to legibility and trends, indicating whether the conditions differ with respect to decreases in reading times for consecutive texts By means of growth-curve analysis two relevant hypotheses were tested the equality of regressions (whether the curves coincide) and parallism of regressions (whether the curves run parallel) The analysis for the four groups simultaneously showed significant differences with respect to coincidence [$\Theta = 60$], and parallelism [$\Theta = 49$, the critical value $\Theta_{01} = 42$ with distribution parameters $s = 3$, $m = 1\ 5$, and $n = 21$] To control for the overall probability of making Type-I errors, simultane-

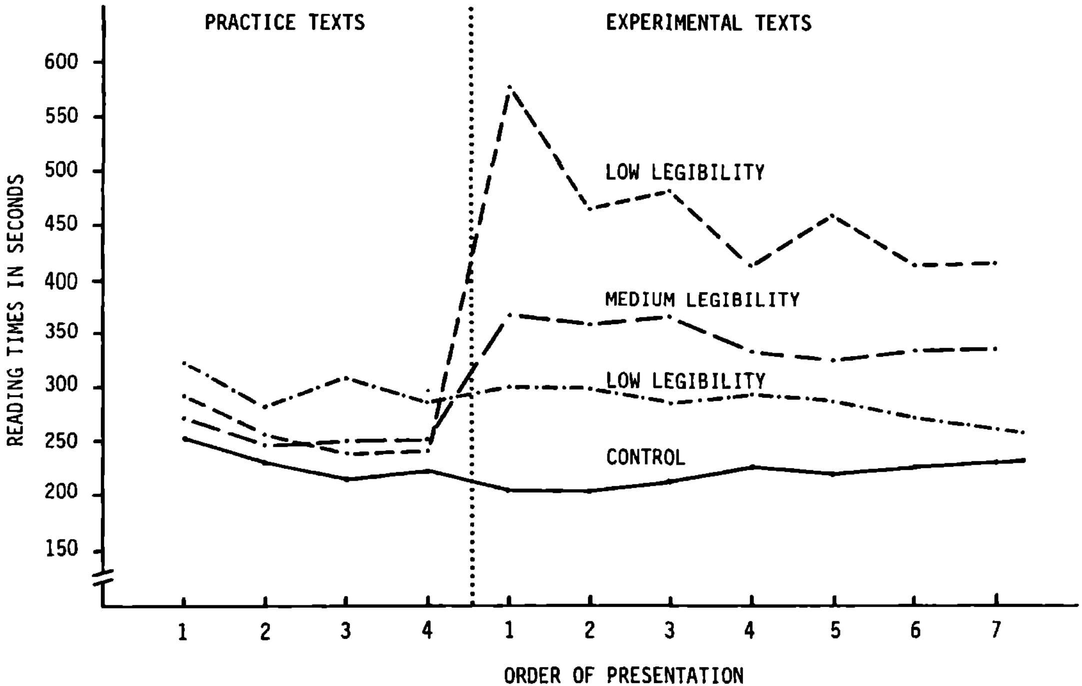


Figure 4.1. Mean Reading Times for Consecutive Practice and Experimental Texts for Four Groups of Subjects.

ous test procedures were applied Tests for different combinations of groups showed that all combinations involving group 4 (low legibility) were significant The other three groups did not differ significantly Mean reading times for the four groups are displayed in Figure 4 1

A separate growth-curve analysis was performed on groups 1, 2, and 3 This analysis showed significant differences in height for the three groups, $\theta = 54$, group 1 (control) was significantly different from groups 2 and 3 which coincided The groups did not differ significantly for parallelism, $\theta = 32$ For these tests, $\theta_{05} = 40$ with distribution parameters s = 2, m = 2, and n = 15 A profile analysis (Timm, 1975) was used to test whether the curves for the three groups can be regarded as horizontal The analysis showed that the reading times for consecutive texts in the medium legibility, the high legibility and the control condition may considered to be constant [$\theta = 41$, $\theta_{05} = 42$ with distribution parameters s = 3, m = 1, and n = 14 5]

b. Printed practice texts.

As is shown by the analyses presented in this section, data for the printed practice texts indicated that no effects of ongoing task-profiency were present during experimental texts and that different groups of subjects may be regarded as equivalent in reading ability

Data for the practice texts can be used to determine whether task-proficiency is asymptotic when subjects are presented with the experimental texts, which will rule out an alternative explanation of decreasing reading times If practice effects on the anomaly-detection task would still be present during experimental texts, reading times would steadily decrease, leading to an overall difference in height of both curves Testing for the coincidence of practice and experimental curves will therefore provide an indication of asymptotic task-proficiency during practice texts In order to compare the practice and experimental curves, the experimental curves must first be shown to be of degree 3 or less, because the practice curves consist of only four measurements which can maximally be described by a third-order polynomial This can be done by testing whether higher-order parameters for the experimental curves are zero This reduction of the number of relevant parameters describing the experimental curves is allowed for groups 1, 2, and 3 [$0 = 11$, $0_{05} = 31$, s = 3, m = - 5, n = 16 5], but not for group 4 (low legibility condition) The test for coincidence of practice and experimental curves showed no significant differences for groups 1 and 2 There was, however, a significant rffect for group 3, the medium legibility condition, [

Table 4 2
Numbers of Different Errors in Consecutive Practice and Experimental Texts for Four Groups of Subjects

	Practice Texts				Experimental Texts						
Legibility Condition	1	2	3	4	1	2	3	4	5	6	7
(a) Missed anomalies											
Control	8	14	15	13	11	6	3	6	7	4	7
High	22	10	13	14	6	7	3	5	2	3	4
Medium	22	19	21	16	3	8	8	4	8	4	9
Low	16	14	19	19	7	9	12	5	7	13	8
(b) Incorrect markings											
Control	21	5	8	9	6	9	8	7	7	3	3
High	32	26	23	20	18	21	7	8	10	18	10
Medium	16	12	13	18	9	9	18	6	7	9	10
Low	21	13	11	17	20	4	11	10	3	7	5
(c) Illegible markings											
Control	-	-	-	-	-	-	-	-	-	-	-
High	-	-	-	-	-	-	-	1	-	-	-
Medium	-	-	-	-	3	4	7	4	4	-	3
Low	-	-	-	-	15	11	8	4	10	7	9

F (4,11) = 10.38, $p <$ 01] This analysis indicates that no further practice effects for the anomaly detection itself occur when subjects are presented with experimental texts The profile analysis for the experimental texts, reported above, in which the curves for groups 1, 2, and 3 were shown to be horizontal, provided another indication that task-proficiency is asymptotic during experimental texts The comparison between practice and experimental texts showed that for group 2 the handwritings are as legible as print For group 3, the handwritings are less legible than print, but there appears to be no practice effect, while group 4 showed a significant effect for legibility as well as a significant practice effect in the text materials

Differences between groups of subjects may be due to relative legibility alone or to relative legibility in combination with reading ability. Differences in reading ability between groups can be determined by testing the coincidence of the four practice curves The analysis showed that the four curves coincide [$\theta = .15$, $\theta_{.05} = .27$, $s = 3$, $m = 0$, $n = 23$]. The four groups may therefore be regarded as equivalent in reading ability.

c *Equal difficulty of experimental texts.*

Interpretation of results will be more valid if the texts may be regarded as equivalent with respect to ease of comprehension. Means for different texts are displayed in Figure 4 2.

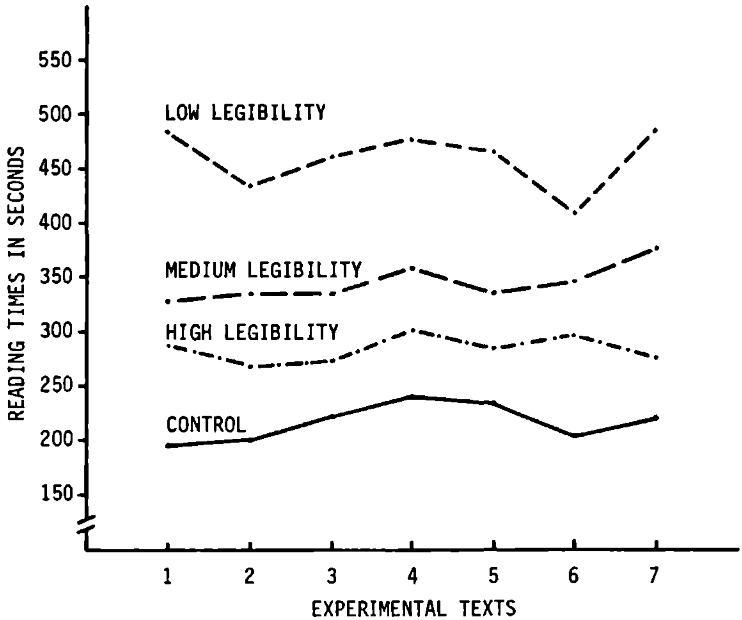


Figure 4 2 Mean Reading Times per Experimental Text for Four Groups of Subjects

Because order of presentation could not be counterbalanced, the mean of a particular text may be influenced disproportionately by the fact that it was presentrd first or second in low legibility condition An adequate test for equal difficulty of texts was provided by the text means for groups 1 (control condition) and 2 (high legibility condition), for which no difference between practice and experimental texts was found. The profile analysis for these two groups showed a significant difference between texts [$\theta = .67$, $\theta_{.05} = .55$; s = 2, m = 1.5, n = 8]. Inspection of the curves in Figure 4.2 and the table of errors (Table 4.3) indicates that text no. 4 is slightly more difficult than the others. The profile analysis for groups 1 and 2, excluding this text, showed non-significant text differences [$\theta = .49$; $\theta_{.05} = .51$; s = 2, m = 1, n = 8.5].

Text differences appeared to depend on legibility conditions as was shown by the test for parallelism of the four profiles [$\theta = .34$; $\theta_{.05} = .33$; s = 3, m =1, n =22]. This result does not necessarily imply that texts differ in conceptual difficulty, because accidental effects of order of presentation or random variations in the relative sloppiness (even for low legibility condition) of the writing over texts may also have contributed to differences between texts.

Errors.

To avoid very long reading times, subjects were given the special option, to be used with restraint, of marking words they could not read. The three types of errors, misses (undetected anomalies), incorrect markings ('normal' words marked as anomalies), and illegible markings, were analyzed separately because each represented a distinct error category. Error data are presented in Tables 4.2 and 4.3. An analysis of error data was carried out for the experimental texts only.

There were significant differences in error frequency between the four groups, both for misses [$\chi^2(3) = 15.64$, $p < .01$], and incorrect markings [$\chi^2(3) = 8.52$, $p < .05$]. Group 3 and 4 differed significantly in the number of illegible markings [$\chi^2(1) = 17.08$, $p < .01$]. Friedmans two-way analysis of variance by ranks (Siegel, 1956) was used to test whether errors were equally distributed over texts and order of presentation. Differences for order of presentation were not significant, neither for misses nor incorrect markings. Texts differed in misses [$\chi^2(6) = 14.23$] and incorrect markings [$\chi^2(6) = 14.72$, $p < .05$]. Data in Table 4.3 show that in text no. 2 subjects made only few misses and incorrect markings, while text no. 4 brought out a large number of incorrect markings.

Table 4.3
Numbers of Different Errors per Experimental
Text for Four Groups of Subjects.

Legibility Condition	Experimental Texts 1	2	3	4	5	6	7
	(a) Missed anomalies						
Control	6	4	5	11	3	8	7
High	2	2	1	3	6	9	7
Medium	7	3	6	6	3	9	10
Low	9	4	7	12	9	9	11
	(b) Incorrect markings						
Control	2	2	3	22	3	4	7
High	13	4	11	30	11	10	13
Medium	6	2	10	30	4	10	6
Low	9	3	4	24	11	6	3
	(c) Illegible markings						
Control	-	-	-	-	-	-	-
High	-	-	1	-	-	-	-
Medium	1	5	2	4	3	8	2
Low	12	8	12	6	14	5	7

DISCUSSION

Although practice effects are the main interest in this experiment, differences between the four legibility conditions also deserve consideration A minor aspect of these data is the fact that reading times for text materials and mean latencies for the word naming task generally agreed with the obtained legibility ratings This suggests that the procedure adopted in the selection experiment (Appendix H) may be regarded as a valid instrument to ensure systematic variation in the relative legibility of sampled handwritings

Of greater importance are legibility differences between the handwriting conditions and the printed control condition Both Corcoran and Rouse (1970) and Ford and Banks (1977) found in their experiments that handwritten words were less easily identified than printed words This experiment did not replicate their results The analysis of differences in height between curves for the four groups showed that only the low legibility condition was significantly different from the other conditions For the subjects in the high legibility condition, the handwritings were as legible as print as was shown by the comparison between printed practice and handwritten experimental texts In the pretest/posttest word naming task, differences between high legibility and printed control condition were not significant These results are testimony to the often noticed impressive tolerance for form variation in human pattern recognition (Kolers, 1975, Neisser, 1967) and indicate that reading handwriting is not necessarily more difficult than reading print Corcoran and Rouse (1970) have suggested that different recognition routines may exist for print and handwriting The data obtained in this experiment indicate that important differences in processing are related to the legibility of the handwriting Differences between reading print and handwriting may be only one aspect of more general differences in legibility

Turning to practice effects, the results showed that these effects depend on the relative legibility of the handwriting The high legibility condition did not show any practice effects Neither significant gain scores for the pretest/posttest nor decreasing times for the text materials were observed in this condition For the medium and low legibility conditions practice effects were apparent For the medium legibility condition practice effects consisted in significant gain scores for the pretest/posttest and for the low legibility condition in decreasing reading times for consecutive texts

Although these findings undoubtedly reflect floor-effects in reading speed, they nevertheless are an example of the kind of qualified generalisation that may turn out to be commonplace for handwriting research Practice effects will not be observed for any particular handwriting as is suggested by Gibson and Levin (1975), but are limited to handwritings with reduced legibility

Results are less clear with respect to the nature of initial practice effects When these effects reflect perceptual learning, reading times for consecutive texts may be expected to decrease The size of the decrease was expected to be related to the initial legibility of the handwriting, larger decreases will be observed for handwritings with reduced legibility Because decreasing reading times for the text materials may, however, also reflect more efficient top-down processing strategies, a pretest/posttest involving single word recognition was introduced in the design Significant gain-scores for this test serve as unambiguous evidence of perceptual learning

The results support a perceptual learning interpretation only partially Decreasing reading times were observed for the low legibility handwritings but no corresponding decrease was found for the gain-scores in the pretest/posttest For the medium legibility condition, significant gain-scores were found but no decrease in reading times for the practice materials

This pattern of results can plausibly interpreted in the following way LaBerge and Samuels (1974) suggested that considerable application of attention is necessary if the reorganization into higher-order units is to take place When a person does not pay attention to what he is practicing, he rules out opportunities for forming higher units because he simply processes through codes that are already laid down' (ibid , p 315) It was suggested in the introduction that in less legible handwritings readers use more extensive top-down processing This way of processing may imply that the reader actually pays relatively little attention to the graphic details of the handwriting itself The reader will, for instance, solve' illegibilities by considering the context in which they occur rather than by extensive perceptual analysis The adoption of this strategy would explain that, when confronted with single words in the posttest, no effects of perceptual learning are apparent Significant pretest/posttest gain scores were obtained for the medium legibility condition Handwritings in this condition, despite their reduced legibility, do not force the subject to rely as heavily on context as the handwritings in the low legibility condition As a consequence, more attention is paid to the graphic aspects of the handwriting with positive results for perceptual learning as indicated by significantly shorter latencies for the posttest Decreasing reading times for texts in the low legibility condition reflect increasingly efficient top-down processing strategies Such strategies are not developed for the medium legibility condition because the handwritings are still rather legible

Different results for skilled and less skilled readers would support this interpretation Skilled readers have been characterized as making optimal use of various kinds of contextual constraint (Goodman, 1976)

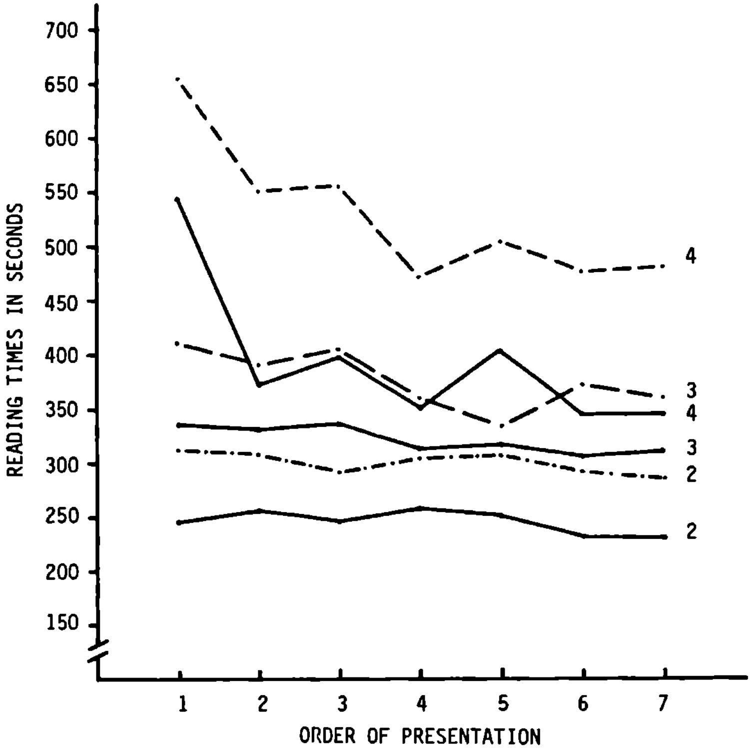


Figure 4.3 Mean Reading Times of Skilled (———) and Less Skilled (— — —) Readers for Consecutive Texts in Three Legibility Conditions. Numbers represent Groups of Subjects (2 = High Legibility, 3 = Medium Legibility, and 4 = Low Legibility).

On the basis of this characterization one would expect skilled readers to develop a 'context strategy' earlier and more efficiently than less skilled readers; i.e., a significant difference for slope in consecutive reading times for skilled and less skilled readers is expected To determine skilled and less skilled readers, the average reading time for the four practice texts for each subject in the handwriting conditions was calculated Subjects with an average above the median of 243 seconds were regarded as skilled readers, readers with a lower average as less skilled ones There were 6, 9, and 7 skilled readers in groups 2, 3, and 4, respectively. Relevant data are presented in Figure 4 3, which shows similar results for groups of skilled and less skilled readers Using simultaneous test procedures, no significant differences were found for height and parallelism between skilled and less skilled readers in any of the three handwriting conditions

EXPERIMENT 6

In Experiment 5 some supportive evidence for a perceptual learning interpretation of initial practice effects was found The analysis of variance for the pretest/posttest gain scores showed nonsignificant differences between the four groups of subjects (control, high, medium, and low legibility conditions) Tests for the four groups separately showed significant gain scores for the medium legibility condition In the low legibility condition, decreasing reading times were found for the seven consecutive texts which were used as practice materials This decrease was, however, not reflected in a corresponding gain-score for the pretest/posttest

The experimental conditions in Experiment 5 might have been less optimal for observing effects of perceptual learning in reading handwriting Texts were presented that were rather easy to comprehend Such reading materials allow extensive top-down processing which will be especially helpful in reading handwritings with low legibility This way of processing may be adverse to perceptual learning because the reader needs to pay less attention to the actual feature-information Perceptual learning might, in other words, depend on the amount of bottom-up processing that the reading materials require

The purpose of Experiment 6 was to find evidence for perceptual learning under conditions in which the visual processing is less assisted by context and where the system is consequently forced to operate in a more bottom-up fashion Experiment 6 was a replication of Experiment 5, but used conceptually difficult texts as practice-materials Reading speed is known to depend heavily on the difficulty of the material to be read (Huey, 1908, 1968) This difficulty is based on the overall predictability or conceptual redundancy of the reading material, which is

especially important in top-down processing. Conceptually difficult texts will curtail the top-down processing and will force the reader to pay more attention to the feature information.
Pretest/posttest gain-scores on a word naming task were again used to assess perceptual learning, free from possible context-effects beyond the word.
Predictions were similar to those formulated in Experiment 5. If practice effects reflect perceptual learning, significant gain-scores for the pretest/posttest should be observed. The observation of decreasing reading times for the text materials will depend on the conceptual demands of the reading materials. Some decrease might occur, but if the reading materials are very difficult, effects of perceptual learning will be masked by ceiling effects due to understanding the materials.
A 'top-down' interpretation of early practice effects predicts no decrease in consecutive reading times. Because the practice materials are inherently difficult, an attempted application of top-down processing strategies will not be successful According to this interpretation, the pretest/posttest gain scores should be zero (cf. Experiment 5).

METHOD

For comparison's sake, Experiment 6 resembled Experiment 5 rather closely. The same handwritings as in Experiment 5 were used (writers received the same instructions). On the basis of the results in Experiment 5 the decision was made to exclude the handwritings with good legibility These handwritings were shown not to differ significantly from print and no practice effects of any kind were observed for these handwritings. As in Experiment 5, the reading of the practice materials was preceded and followed by a word-naming task for single words. The stimuli of Experiment 5 were used for this task.
Prior to the reading of the experimental texts, subjects read two printed practice texts The practice texts allow to differentiate between practice effects for the experimental task (anomaly detection) and practice effects for the handwriting. Only two practice texts of 750 words each were used in Experiment 6, because no further practice effects for the anomaly detection task occurred after the second practice text in Experiment 5 (see Figure 4.1). Two practice texts of Experiment 5, which had resulted in small numbers of misses were selected.

The prose-passages selected for the handwritten experimental texts were taken from an advanced history text-book. The passages were selected from 'Erflaters van onze beschaving' by J. and A. Romein. The book is about well-known personalities in Dutch history. All selected passages were written by J. Romein. Because conceptually difficult materials will require more time to read, the experimental texts were reduced in length to keep the total amount of reading time roughly equal to the reading times in Experiment 5 Texts consisted of about 350 words. Four anomalies were inserted in each text. The anomalies involved the replacement of two verbs and two nouns.

The experimental procedure was the same as in Experiment 5. The 40 subjects, which were randomly assigned to conditions, were paid f 7,- an hour for their participation in the experiment.

RESULTS

Word recognition

Latencies

The mean RT for the three groups of subjects are presented in Table 4 4

The analysis of variance, based on the subject means, showed significant differences between three groups in the pretest [F (2,37) = 27 01] and in the posttest [F (2,37) = 19 93, p < 01 for both]. For both the pretest and the posttest, Scheffe s post-hoc test showed differences between the two handwriting conditions not to be significant (a similar result was obtained in Experiment 5)

The three groups of subjects did not differ significantly with respect to the pretest/posttest gain scores ($F < 1$) As in Experiment 5, a non-parametric test was used because of the heterogeneous within-group variances. The Kruskall-Wallis test for the rank order of the gain scores showed nonsignificant differences between the three groups [H = 2 68, p = .20]. Tests to determine whether the gain scores were significantly different from zero for the three groups separately showed significant results for the low legibility condition only [$t(14)$ = 2.61, p < 01].

Errors

The number of missing observations, due to experimental error, amounted to 3 8 % of all data Erroneous readings amounted to 7 % in

Table 4.4
Mean Reaction Times (in milliseconds) and, In Parentheses, Numbers of Different Errors for Pretest and Posttest Word Recognition for Three Groups of Subjects.

	Legibility Condition	Pretest	Posttest	Pretest/posttest Gain Scores
1	Control (n = 10)[a]	654 (-,-)[b]	641 (-,-)	14 (-,-)
2	Medium (n = 15)	1129 (9, 7)	1010 (6,2)	119 (3,5)
3	Low (n = 15)	1101 (12,12)	1001 (11,5)	100 (1,7)

[a] n is the number of subjects.
[b] The first number in parentheses is the number of erroneous readings, the second is the number of 'illegible' responses.

the pretest for the medium and the low legibility conditions combined; for the posttest this percentage was 5.6. In the pretest, 6.3 % of the words presented in the medium and the low legibility conditions were found to be illegible, the corresponding percentage for the posttest was 2.3.

Text materials.
a. *Reading times for consecutive experimental texts*
The mean reading times for the seven experimental texts are displayed in Figure 4 4.
The growth-curve analysis showed that the curves for the three groups of subjects differed significantly in height [0 = .50], but not in parallelism [0 = .24, the distribution parameters are s = 2, m = 2, and n = 15; the critical value 0_{01} = .48]
Simultaneous test procedures showed that only group 1 (control) and group 3 (low legibility condition) differed significantly with respect to height. The profile analysis showed that the three curves may be regarded as horizontal lines [0 = 28; 0_{05} = .42, s = 3, m =1 and n = 15] indicating that in the (handwriting) conditions no significant decrease in reading times occurred

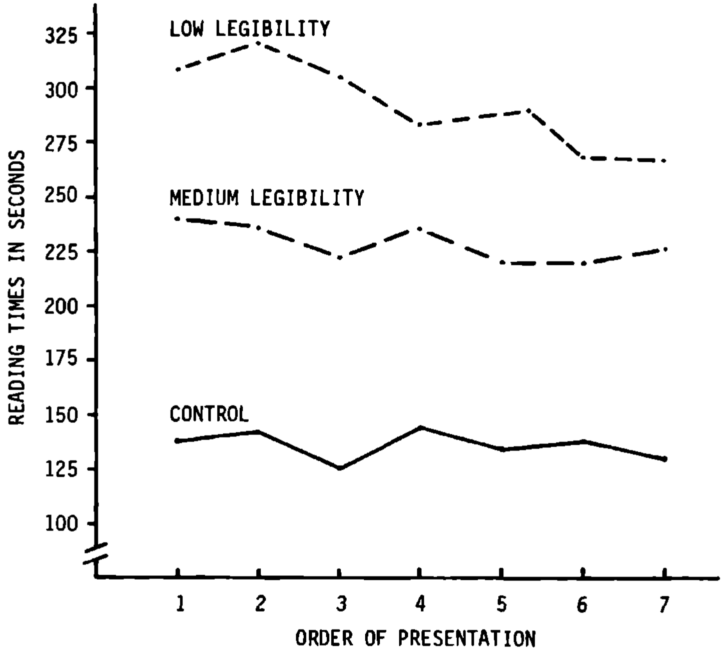


Figure 4.4 Mean Reading Times for Consecutive Experimental Texts for Three Groups of Subjects.

b. *Printed practice texts*

Results for the printed practice texts were used to determine whether differences between the three groups are (partly) due to differences in reading ability. The analysis of variance for the three groups and order of presentation showed differences between the groups not to be significant [F (2,37) = 1.96, p = .15].

c. *Equal difficulty of experimental texts*

Interpretation of results for the handwritten practice materials will be more valid when texts can be regarded as equal for ease of comprehension. Relevant data are presented in Figure 4.5. and Table 4.6.

The analysis for seven consecutive texts showed that the curves for the three groups run parallel and may be regarded as horizontal lines. The equal difficulty of texts was therefore tested by a comparison between

Table 4.5
Numbers of Different Errors in Consecutive Experimental Texts for Three Groups of Subjects.

	Experimental Texts						
Legibility Condition	1	2	3	4	5	6	7
	(a) Missed anomalies						
Control	2	7	6	5	3	3	6
Medium	7	6	7	8	6	10	10
Low	8	7	9	4	9	6	7
	(b) Incorrect markings						
Control	11	10	6	7	8	7	8
Medium	33	20	17	20	16	10	26
Low	21	16	22	12	25	13	13
	(c) Illegible markings						
Control	-	-	-	-	-	-	-
Medium	5	8	9	5	5	3	5
Low	43	39	20	18	19	28	21

the seven experimental texts for the 40 subjects simultaneously. This analysis showed significant differences between texts [F (6,34) = 8.90, $p < .01$]. For the control group separately, however, texts did not differ [F (6,9) = 1.33, $p = .40$]. A test for the text-profiles indicated that the curves in Figure 4.5 cannot be regarded as parallel [$\theta = .41$, $\theta_{.05} = .38$; $s = 2$, $m = 1.5$, and $n = 15$]. As can be seen in Figure 4.5, text no. 6 appears to be more difficult than the others in low and medium legibility conditions.

The tests for equal difficulty of texts are not unequivocal: on the one hand, no significant differences between the texts were found for the control group, but on the other hand, significant differences between texts were obtained across 40 subjects and for the handwriting groups separately. It is tempting to infer that the differences are related to variations in the handwriting instead of being variations in conceptual

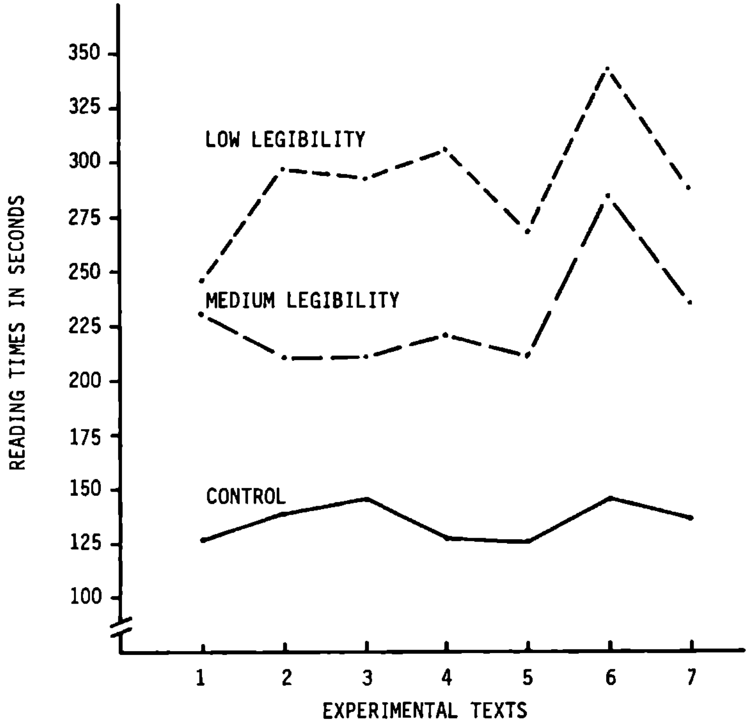


Figure 4 5 Mean Reading Times per Experimental Text for Three Groups of Subjects

difficulty, but the agreement between handwriting conditions makes this interpretation rather unlikely

Errors

The error data are presented in Tables 4 5 and 4 6 The experimental texts were analyzed for the three types of errors -misses, incorrect markings, and illegible markings- separately The three groups of subjects did not differ significantly for the number of misses ($x^2(2) < 1$) The control group made significantly less incorrect markings than the two handwriting groups [$x^2 = 14\ 09$, $p < \ 01$] The two handwriting

Table 4 6
Numbers of Different Errors per Experimental Text for Three Groups of Subjects

Legibility Condition	Experimental Texts 1	2	3	4	5	6	7
	(a) Missed anomalies						
Control	3	4	6	5	3	3	8
Medium	7	3	13	10	2	8	11
Low	9	2	9	10	7	2	11
	(b) Incorrect markings						
Control	7	9	4	6	9	9	13
Medium	23	11	7	17	15	49	20
Low	12	18	9	22	16	33	12
	(c) Illegible markings						
Control	-	-	-	-	-	-	-
Medium	9	-	2	6	8	5	10
Low	15	26	19	29	27	42	30

groups differed significantly for the number of illegible markings [χ^2 = 96 08, $p <$ 01]

Friedman s test was used to determine whether errors were evenly distributed across texts and evenly across order of presentation for the three subject groups All tests showed nonsignificant results except for the number of misses across texts [χ^2 = 13 75, $p <$ 05]

DISCUSSION

Most relevant for a perceptual learning interpretation of initial practice effects are the results for the word naming task As in Experiment 5, the pretest/posttest gain scores for the three groups of subjects did not differ significantly In tests for the three groups separately significant gain-scores were obtained for the low legibility condition It was supposed in the introduction that the use of conceptually difficult materials as practice materials would force the reader to pay more attention to the characteristics of the handwriting and would thus provide more optimal conditions for perceptual learning Induced greater attention for the handwriting is evidenced by results for the low legibility condition In Experiment 5, no significant gain-scores were obtained for this condition, but Experiment 6 showed the opposite result For the medium legibility condition, however, significant gain-scores were obtained in Experiment 5, but this result was not replicated in this experiment An explanation of this negative finding is not easily found

Comparing the results for the text materials obtained in Experiments 5 and 6, it appears that the occurrence of practice effects, as measured by reading times for text materials, also depend on the nature of the reading materials In Experiment 5, the low legibility condition displayed a significant decrease in reading times for consecutive texts In Experiment 6, the low legibility condition showed some decrease in consecutive reading times but this trend was not significant Texts in the two experiments differed with respect to conceptual difficulty and the combination of results suggests that the availability of elaborate top-down processes does contribute to improvements in reading speed for handwritings with reduced legibility

As in Experiment 5, a separate analysis for skilled and less skilled readers was carried out This analysis may provide support for a 'top-down interpretation Skilled readers may develop a 'context-strategy earlier and more extensively than less skilled readers

The relevant test was carried out for the 30 subjects in the two handwriting conditions simultaneously, because the curves for the seven consecutive texts in the two handwriting groups were found to be parallel The 30 subjects were divided into two groups, based on the median reading times for the practice texts Curves for the two groups are presented in Figure 4 6 The test for parallelism showed differences between the two groups of subjects not to be significant [F (8,32) = 1 07, p = 41]

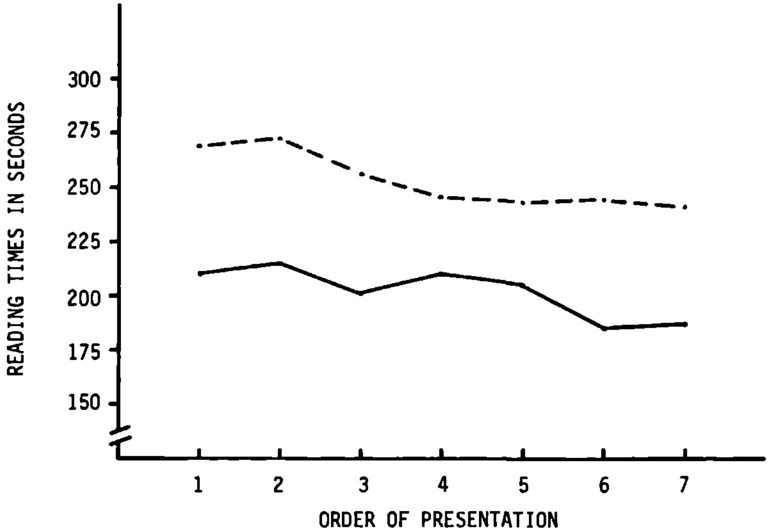


Figure 4.6 Mean Reading Times of Skilled (——) and Less Skilled (— —) Readers for Consecutive Experimental Texts.

EXPERIMENT 7

Experiments 5 and 6 measured improvements in reading speed for handwritings in two different ways. One measurement involved the reading times for consecutive texts, the second dealt with gain-scores for a pretest/posttest of single word recognition. Two different interpretations of initial practice were considered. Improvements in reading speed may reflect perceptual learning or more efficient top-down processing. Results for the pretest/posttest were of special importance. Significant gain-scores would support a perceptual learning interpretation and are contrary to a 'top-down' interpretation. In Experiment 5, significant gain-scores were obtained for the medium legibility condition and in Experiment 6 for the low legibility condition. Thus, the perceptual interpretation did receive some empirical support. It should be noticed, however, that these effects appeared to be rather unreliable. The significant gain-scores for the medium legibility condition in Experiment 5 were not replicated in Experiment 6. Perceptual learning also seemed to depend on the attention that is paid to the handwriting itself, as was shown by the pretest/posttest results for the low legibility condition. The attention directed at the handwriting was manipulated by presenting reading materials that differed in conceptual difficulty. Reading materials that are difficult to comprehend provide the reader with less resources to decipher less legible handwritings Under those conditions the reader will have to perform a more extensive perceptual analysis of the characteristics of the handwriting

The pretest/posttest results for the low legibility condition may be representative for the practice effects that are observed in every-day reading of handwriting Under normal conditions, the reader is concerned with the meaning of what has been written and his attention will be directed at the highest meaningful level This basic characteristic of the reading process will also be present in the reading of less legible handwritings. In keeping with his *main goal -the comprehension of the text-* the reader will preferably make (maximal) use of semantic/conceptual information in reading these handwritings. Paying considerable attention to the graphic aspects of the handwriting will obviously be less integrated in the overall reading process and might even be a hindrance to good comprehension As an aside, this interpretation may be satisfying from a functional point of view. Clearly, learning the idiosyncrasies of every handwriting is not very useful if one might never encounter the handwriting again.

Data for decreasing times in text materials are consistent with this interpretation Decreasing reading times were found to depend on the conceptual difficulty of the reading materials In Experiment 5 in which reading materials were used that were easy to comprehend a significant decrease in reading times was found for the low legibility condition In Experiment 6 in which conceptually difficult materials were used a non-significant trend for decreasing reading times was found.

Although the interpretation presented above is admittedly somewhat speculative, its plausibility will be strengthened if it can be shown that top-down information is indeed more important for reading less legible handwritings. In Experiment 7, it was therefore attempted to demonstrate an interaction between the legibility of the handwriting and the conceptual difficulty of the reading materials Conceptually difficult texts will result in longer reading times than 'easy' texts. If more use is made of top-down processing in reading handwritings with poor legibility, the increase in reading times for difficult texts will be more pronounced for these handwritings than for handwritings with good legibility. Although Experiment 7 did not deal directly with practice effects, the predicted interaction will provide circumstantial evidence for the correctness of a 'top-down' interpretation of practice effects.

For Experiment 7, three handwritings were selected, one out of each of the three legibility categories used in Experiment 5. A difficult and easy text was presented in each of these legibility conditions. Differences in conceptual difficulty between the texts were determined beforehand by measuring reading times for printed versions.

It will be noticed that Experiment 7 is analogous to experiments that have demonstrated an interaction between visual degradation and semantic context (e.g., Becker and Killion (1977), Massaro et al. (1978), Meyer, Schvaneveldt and Ruddy (1975)). Experiment 7 closely resembled an experiment of Becker and Killion (ibid.). In their Experiment I, they used a lexical decision task in a semantic priming paradigm with visually degraded stimuli. The semantic relation between words presented first and second corresponds with reading easy and difficult text materials. The targets in their experiments were presented under three stimulus intensity conditions: low, medium, and high. The variation in the legibility of the handwritings may be considered an analogous manipulation.

For reasons that will become clear below, Experiment 7 was carried out in two different versions. These versions differed with respect to the selected handwritings.

METHOD

Stimulus materials

Texts

Six prose-passages of approximately 250 words were selected. The three easy passages were the initial sections of three experimental texts (nos. 3, 5 and 7) used in Experiment 5. The three difficult passages were taken from the same history text-book that provided the reading materials for Experiment 6. In each of these texts, four semantic anomalies were inserted at random intervals. Two anomalies involved nouns and the other two verbs.

Before reading the handwritten texts, the subject was presented the two printed practice texts used in Experiment 6 to familiarize him with task and procedure

To each of the three legibility conditions -high, medium, and low- one easy and one difficult text was assigned Preliminary testing of the six texts, all set in Courier type-face, with 12 subjects, showed significant differences for Conceptual Difficulty [F (1,11) = 42 73, $p <$ 01] Texts within a particular category may be regarded as equal with respect to conceptual difficulty [F (4,44) = 1 82, $p =$ 14] Mean reading times for easy texts 1, 2, and 3 and difficult texts 4, 5, and 6 were 59, 63, 62, and 93, 86, and 82 seconds respectively Pairing texts 1 and 6 (23 sec), 2 and 4 (30 sec), and 3 and 5 (24 sec) equalized the differences between pairs of easy and difficult texts as much as possible The combination with the smallest difference (1-6) was assigned to the low legibility condition, the combination with the largest difference (2-4) to the high legibility condition The asssignment was contrary to the hypothesis

Handwritings

Legibility ratings for the three selected handwritings (which were different from the ones used in Experiments 5 and 6) were obtained in the selection experiment described in Appendix H Mean legibility scores for the three handwritings were 1 5, 2 9, and 4 2 respectively The corresponding rank orders were 4 5, 15 5, and 30 (out of 32)

Writers were instructed in the same way as in Experiment 5 and 6 To control for differences in legibility due to the relative speed of writing (the easier texts being written faster), they were instructed to write 3-4 words at a time Original exemplars were reproduced by off-set lithography

Procedure and Subjects

Subjects had to read the texts silently while marking the anomalies Each subject was first presented the two practice texts before reading the handwritten passages Order of presentation was counterbalanced within easy and difficult texts separately No two texts in the same handwriting could follow each other Instruction, feedback for errors, and timing procedures were the same as in Experiment 5 and 6 Experimental sessions lasted about half an hour

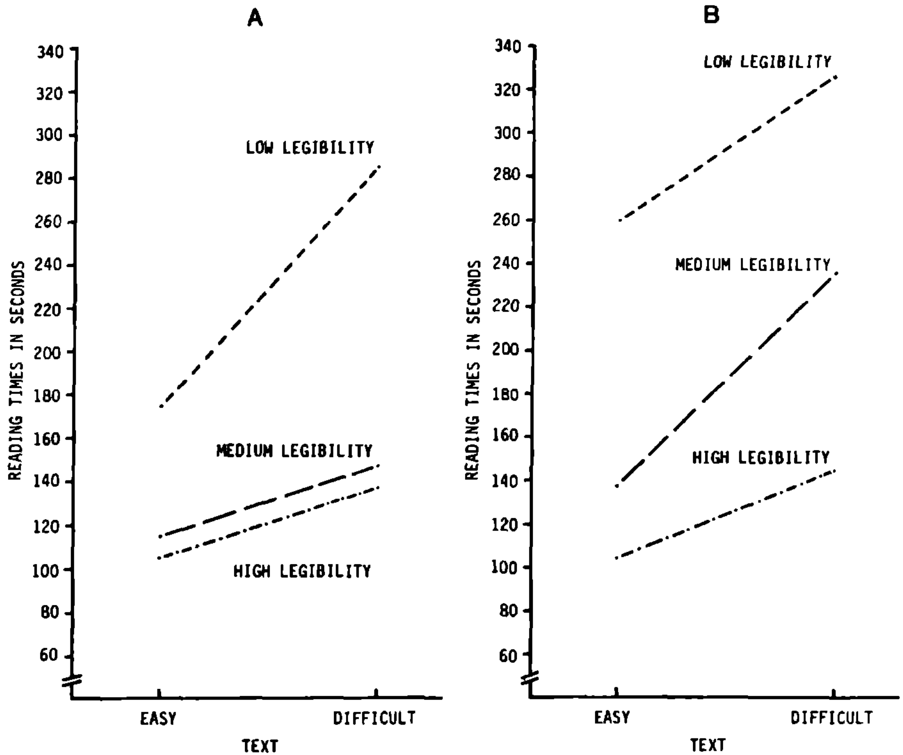


Figure 4.7. Mean Reading Times for Handwritings with Different Legibility As a Function of Conceptual Difficulty of Texts. Figure (a) and (b) correspond with two versions of Experiment 7.

24 subjects were used All subjects were students at the University of Nijmegen. They were not paid for their services

RESULTS

Reading times were entered into an analysis of variance with both Conceptual Difficulty (easy versus difficult texts) and Legibility as within-subjects factors. The main effects of Conceptual Difficulty [F (1,23) = 107 84, p < 01] and Legibility [F (2,46) = 142 36, p < 01] were significant, as was their interaction [F (2,46) = 37 58, p < 01] Mean reading times are displayed in Figure 4 7 (a) All tests for simple main effects were significant at the one percent level
Error data are presented in Table 4 7 (a) Easy and difficult texts did not differ significantly for the number of misses, but significantly more incorrect markings as well as illegible markings were made in difficult texts (Wilcoxon, p < 01)
For the incorrect markings, Friedmans test showed no difference between the handwritings in both easy [$\chi^2(2)$ = 1 19, p = .50] and difficult texts [$\chi^2(2)$ = 77, p = 70] Table 4 7 (a) shows a large increase of illegible markings for difficult text in the low legibility condition, which is consistent with the hypothesis

DISCUSSION

Results confirmed the predicted interaction between the conceptual difficulty of the reading materials and the legibility of handwritings. The interaction appeared in both latencies and errors Particularly the large increase in the number of illegible markings provides evidence that higher-order information is essential for the efficient reading of handwritings with low legibility
When reading handwriting involves a continuous interplay between bottom-up and top-down information, it may be expected that the effect of conceptual difficultty will increase with reduced legibility of the handwriting Evidence for such a continuous increase was found by Becker and Killion (1977) In their experiment the effect of semantic context amounted to 70 ms at low target intensity, 40 ms at medium intensity, and 30 ms at high intensity (Becker and Killion, ibid) The obtained

Table 4 7
Mean Reading Times (in seconds) and Numbers of Different Errors for Conceptually Easy and Difficult Texts in Handwritings of Different Legibility

	Text					
	Easy			Difficult		
Legibility Condition	High	Medium	Low	High	Medium	Low
(a)						
Mean reading times	104	115	174	137	146	284
Missed anomalies	-	10	3	6	6	4
Incorrect markings	10	14	19	43	32	53
Illegible markings	5	1	5	5	10	88
(b)						
Mean reading times	103	137	259	143	234	326
Missed anomalies	6	13	18	13	13	13
Incorrect markings	15	16	23	24	33	36
Illegible markings	-	10	102	2	57	155

results do not display such a continuous increase the effect of conceptual difficulty varied from 33 sec ± 4 for the high legibility condition to 110 sec ± 4 for the low legibility condition, and 31 sec for the medium legibility condition This result may be due to a sampling artefact Despite the fact that the legibility ratings for the handwritings were reasonably apart, reading times did not reflect an even sampling of the legibility continuum

To show a continuous increase in the effect of conceptual difficulty the experiment was replicated with different handwritings This replication was identical to the first version, except for the use of different handwritings and subjects Subjects in this second version were paid f 7,- for their services

Legibility ratings for the three handwritings were 2 8, 3 8, and 4 7 Their corresponding rank order was 19, 29, and 32 (out of 32) The handwritings in the high and low legibility conditions were also used in Experiment 5 From the rating data, it can be inferred that the handwriting in the high legibility condition in fact represents a medium legibility condition, the other two handwritings represent low legibility conditions In the analysis of variance for the reading times, Conceptual Difficulty [F (1,23) = 64 18], Legibility [F (2,46) = 101 43] and their interaction [F (2,46) = 8 05, p < 01] were significant Mean reading times are displayed in Figure 4 7 (b) All test for simple main effects were significant at the one percent level

Error data are presented in Table 4 7 (b) No significant differences appeared between easy and difficult texts with respect to misses or incorrect markings, but difficult texts caused more illegible markings (Wilcoxon, p < 01) in low legibility conditions

The legibility conditions did not differ for misses and incorrect markings, but differences for illegibility markings proved to be significant for easy [Friedman s $\chi^2(2)$ = 29 89] and difficult texts [$\chi^2(2)$ = 34 57, for both p < 01]

The reading times for easy texts in the high and the low legibility conditions were further apart than in Experiment 7 (a) The handwritings used in this replication represent therefore a more evenly spread sampling of the legibility continuum

The expected result, an increasing effect of conceptual difficulty with reduced legibility, was clearly not obtained In fact, as can be seen in Figure 4 7 (b), the effect of conceptual difficulty is larger for the medium legibility condition than for the low legibility condition This negative finding may be contributed to the extreme illegibility of the handwriting in the low legibility condition A salient feature of Table 4 7 (b) is that even in the easy text a large number of illegible markings was obtained for the low legibility condition A continuously increasing effect of conceptual difficulty will therefore not be observed for the whole spectrum of relative legibility, due to ceiling-effects in reading speed

The results of Experiment 7 have clearly shown that in reading less legible handwritings more extensive use is made of top-down information Experiments 5 and 6 provided inconclusive evidence with respect to a perceptual learning or top-down interpretation of early practice effects Although Experiment 7 did not deal directly with prac-

tice-efects, the observed interaction between legibility of the handwriting and conceptual difficulty of the reading materials provides support for a 'top-down' interpretation.
A rather conservative conclusion that may be drawn from the results of Experiments 5-7 is that in the initial reading of a handwriting, perceptual adaptation to its characteristics is rather limited. The improvements in reading speed that are observed in the first acquaintance with 'difficult' handwritings may primarily be due to changes in processes that deal with the conceptual aspects of reading materials.

CHAPTER 5

PERCEPTUAL LEARNING OF HANDWRITING

Experiments 5 and 6 were concerned with initial practice effects in reading handwriting. These initial practice effects were defined as improvements in reading speed that occur during the first and limited acquaintance with a particular handwriting. Two interpretations of these practice effects were considered. According to a perceptual learning interpretation, the improvements in reading speed are due to changes in perceptual processes. The extraction of distinctive features or the formation of higher-order codes were considered likely components of perceptual learning. A 'top-down' interpretation attributes improvements in reading speed to increasingly efficient top-down processing. Although the obtained empirical evidence was not clearly in favor of one of the interpretations, results tended to support a 'top-down' interpretation of initial practice effects observed in every-day reading. The results of Experiments 5 and 6 do not, of course, constitute evidence that perceptual learning does not occur for handwritings; they indicate that perceptual learning is limited in the initial reading of a particular handwriting.

Perceptual learning may be a rather slow process that requires larger amounts of practice than was provided in Experiments 5 and 6. This hypothesis may be investigated, analogous to Experiments 5 and 6, by presenting larger amounts of handwritten material over longer periods of time. Handwriting, however, provides a unique and more convenient way of investigating effects of long-term familiarity on reading speed. The writer of a handwriting is in a privileged position for the perceptual learning of the characteristics of a particular handwriting: he has read everything that has ever been produced in that handwriting. Long-term practice effects may therefore be investigated by studying the reading of a person's own handwriting. If perceptual learning occurs for handwritings, a distinctive facilitation for 'own handwriting' may be expected.

For Experiment 8 which assessed the effect of 'own handwriting', six handwritings were selected from the sample of handwritings used in the selection experiment (see Appendix H). The six handwritings were at about equal distances in the obtained rankorder for the 32 handwritings and differed in legibility. A Latin Square design was used in which each subject was presented with samples of each of the six handwritings, one of which was his own. The mean reading time for a handwriting, across six subjects, was used as a measure of legibility; the mean reading time of a subject, across six handwritings, was used as a measure of reading ability.

The effect of 'own handwriting' for a particular subject can be measured as the difference between the mean reading time for the 'own handwriting' and the mean reading time for the other five handwritings, taking into account differences in legibility between handwritings. If facilitation for 'own handwriting' exists, the mean reading time of the writer for his own handwriting will, on the average, be lower than his mean reaction time for the other five handwritings. Thus, it was predicted that the mean 'own handwriting' effect (as the mean facilitation effect across six subjects) would be significantly different from zero. The size of 'own handwriting' effect will not be equal for all six subjects. In connection with differences in relative legibility of the handwritings, the size of the 'own handwriting' effect may be expected to increase with reduced legibility of the handwriting. For the providers of handwritings with good legibility, the facilitation for the 'own handwriting' will be smaller because of floor effects in reading speed.

It is conceivable that a particular subject reads his own handwriting much faster than the other five handwritings while at the same time it appears that his reading times for the other five handwritings are also generally shorter than the mean reading times of the other five subjects. The facilitation for 'own handwriting' may therefore be only an aspect of more general differences in reading ability. An important aspect of the 'own handwriting' effect is its specificity; i.e., there should be no systematic differences between subjects except for their ability to read the 'own handwriting'.

As noted above, the largest own handwriting effects may be expected for the writers of less legibible handwritings. To test the specificity of the 'own handwriting' effect, the three providers of the more legible handwritings were compared with the three providers of less legible handwritings for their ability to read less legible handwritings. If writers of less legible handwritings are found to be better readers of these handwritings in general, the 'own handwriting' effect may be part of a more general ability to read less legible handwritings.

Apart from the perceptual learning of a particular handwriting, which may be investigated by studying the effects of 'own handwriting', general perceptual learning might exist for reading handwriting. Frequent reading of (less legible) handwritings may turn one into a 'good handwriting reader' and this general ability may facilitate the reading of every individual handwriting. To investigate this possibility, professional librarians were included as a separate group of subjects in this experiment. Librarians at desk-service come across many different handwritings, some of which will undoubtedly be less legible. Moreover, reading materials (titles of books, names of authors, street-names) often provides minimal semantic cues for deciphering Such reading materials may be considered optimal for the perceptual learning of handwriting because the reader will have to pay considerable attention to the physical characteristics of handwritings (cf. Experiment 6). If general perceptual learning for reading handwritings exists, the librarians as a group will be faster readers of handwriting than a randomly selected group of university students, which presumably read fewer handwritings. Because of floor-effects in reading speed, their greater ability in reading handwriting will be especially pronounced for the less legible handwritings.

EXPERIMENT 8

METHOD

Stimulus materials

Words

The 360 words used were 180 nouns and 180 verbs Of these 180 nouns and verbs, 90 were high frequency words (> 70, according to the Uit den Boogaart (1975) count) and 90 were low frequency words (< 20) Thus, four main *word-categories* of 90 words each resulted. high and low frequency nouns and high and low frequency verbs.

The selected words varied in length between five to eight letters In general, verbs tended to be somewhat longer than nouns The four main word-categories were equalized as much as possible for length. To each of six arbitrary *word-blocks*, fifteen members of each of the four main word categories were randomly assigned These six word-blocks of 60 words each were matched for word-length and frequency (for this purpose, the 180 high-frequency words were sub-divided into classes with interval-size 100).

Handwritings

The six selected handwritings fulfilled two conditions. First, they were positioned at about equal distances in the obtained rankorder for the 32 handwritings used in the selection experiment. The rank-order position for the six handwritings *A, B, C, D, E, F* was 2.5, 7, 11.5, 19, 24, and 28 respectively. The corresponding mean legibility scores were 1.1, 1.8, 2.5, 2.8, 3.3, and 3.8. The six handwritings had not been used in Experiments 5-7. Second, the handwritings were physically rather dissimilar. Samples of the six handwritings apppear in Figure 5.1.

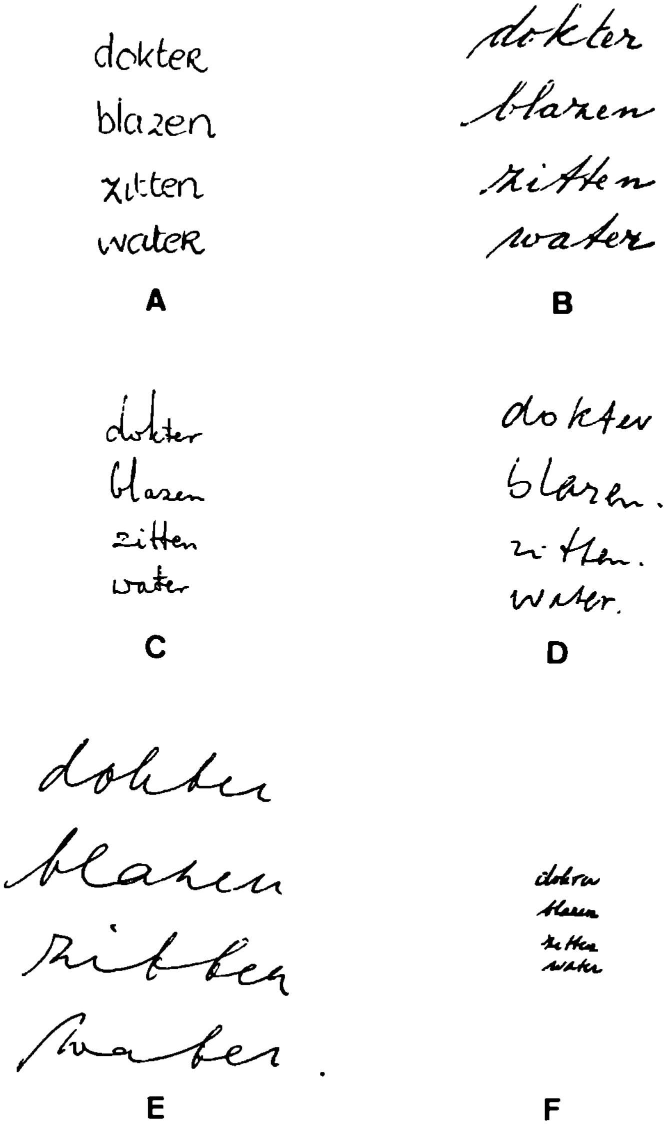
Each of the six subjects wrote the 360 words referred to above. The words were dictated by the experimenter in a different random order for each subject. Each word was written without capitals on a small, white card (12 × 7.5 cm). Subjects were allowed to use their preferred writing-utensil. All subjects wrote with a ball-pen, except subject 3 who preferred her own fountain-pen.

Immediately after having written the 360 words, subjects were asked to recall them. Except for subject 1, who remembered some 30 items, no subject correctly recalled more than 10 words.

Subjects were instructed to write in their 'normal, usual' handwriting. Upon questioning, none of the subjects appeared to use different styles of handwriting, although five of them noted that when writing in a hurry, their handwriting tended to become 'sloppy'. For all subjects, the writing session took approximately one hour.

Design

The design for the reading session of the experiment, which took place a month after the writing session, was a 6 × 6 Latin Square with the word-blocks entering the cells of the square. Columns 1-6 of the Latin Square are the handwritings *A* (most legible) through *F* (least legible), with subjects randomly assigned to rows. These assignments marked the cells which were to contain the subjects' own handwritings. The construction of a suitable Latin Square involved two separate randomization procedures. First, each of the six word-blocks, with nos. 1-6, was randomly assigned to one of the own handwriting cells. In this way, it was ensured that all six word-blocks were read once in 'own handwriting' and that each subject read a different word-block in 'own handwriting'. A second, incomplete randomization, also using tables of random numbers, filled the remaining cells. Thus, each of the 2160 to-



Samples of the Six Handwritings used in Experiment 8.

kens (6 × 360 types) was used once in the reading session The used Latin Square is presented in Figure 5.2.

		Handwritings					
		A	B	C	D	E	F
Subjects	1	5	1	6*	3	4	2
	2	1	6	5	2	3	4*
	3	2*	5	4	1	6	3
	4	4	3*	2	6	5	1
	5	6	2	3	4	1*	5
	6	3	4	1	5*	2	6

Figure 5.2 Latin Square used in Experiment 8

The adopted procedures have the advantage of minimizing the contribution of a memory factor which would favor the reading of 'own handwriting'. Every subject read all 360 types he had written a month earlier, but in six different handwritings, leaving to the own handwriting only the small advantage of having seen 60 tokens a month before The adopted Latin Square was used for two groups of subjects One group was the 'own handwriting group', for which one of the handwritings was their own The second group were six librarians, which were a 'replication' of the own handwriting group in the sense that for each subject in the 'own handwriting' group there was a corresponding subject in the librarian group which was assigned to the same treatment, i e , the same pairings of word-block and handwriting

Procedure
Stimuli were presented by means of a video-disk recording system (Ampex MD-400) and a PDP 11/34 computer Words were recorded on disk in the order they would be presented to the subject.
Order of presentation of the words (types) was random for each subject (in a particular group) Handwritings were presented in 60 blocks of six words each, with one token of each handwriting in each block Order of handwritings within these blocks was random No two instances of the same handwriting could follow each other. Due to practical cir-

cumstances every order of presentation had to be used twice: once for a subject from the 'own handwriting group, a second time, for one of the librarians.

The subject was seated in front of a monitor at a viewing distance of about 75-100 cm. The length of the presented words varied from 2.5 cm (5-letter words in Handwriting *F*) to 20 cm (8-letter words in Handwriting *E*). Most words did not exceed 8 cm Specification of the visual angle is rather difficult because subjects differed in preferred viewing distance.

The computer recorded the naming latency for each single word, measuring from onset of the display till initiation of the vocal response by the subject The subject spoke into a microphone which was attached to headphones he was wearing. Each word remained on the screen 500 ms after the subject had responded. By pressing a key the subject could present the next stimulus, allowing for a built-in delay of one second.

The instructions for the subjects stressed accuracy over speed. Subjects were told that every display would be a familiar Dutch word If the subject could not read a particular word, he was to respond 'illegible'. When faced with two plausible alternative readings, he was instructed to say the word that had occurred to him first. Erroneous readings were written down by the experimenter, who also kept a record of the 'illegible' responses.

Prior to the presentation of the handwritten materials, a series of 40 typed words was given as practice to familiarize the subject with the procedure Experimental sessions took about 45 minutes.

Subjects

Of the six subjects in the 'own handwriting group, five were students, one a faculty member at the University of Nijmegen Four of these subjects were men, two of them (nos. 3 and 5) were women. These subjects were told beforehand that the experiment would consist of a writing and reading session Subjects suspected that the experiment had something to do with the perception of own handwriting', but had no knowledge of the design

The group of librarians were all women, except for subject no 12 On the average, they had worked for about eight years at desk-service in libraries at the University of Nijmegen This group was told that the experiment was to investigate whether they were better readers of handwriting than students Subjects were not paid for their services

The within-cells variances appeared to be very heterogeneous. Because the number of observations in the cells of the square are reasonably large and approximately equal, this heterogeneity is of little consequence for the validity of the statistical analyses performed. To reduce the heterogeneity of the within-cells variances, all responses with latencies longer than 4000 ms were set to a maximum value of 4000 ms. This data-cleaning involved only half a percent of all correct responses, but was unevenly distributed across handwritings (only the handwritings *D, E,* and *F*) and subjects.
Although the legibility of a handwriting is primarily reflected in the latencies for the correct responses, some valuable additional information might be obtained by including latencies for erroneous readings and 'illegible' responses in the analysis. In connection with the rather large number of stimuli used in this experiment (cf. Experiment 5 and 6), it seemed worthwile to carry out two separate analyses: one (A-analysis) involving only the correct responses, and a second (B-analysis) involving all responses including the different kinds of errors. Both analyses carried a maximum value of 4000 ms. The two analyses showed the same overall pattern of results. The congruence of the two analyses adds to the reliability of the analysis for the correct responses only (A-analysis), which is reported below (unless specifically noted otherwise).

Latencies

1. *Word-blocks*

For the estimation of the effects of the handwritings, subjects and their interaction, the word-blocks constitute a variable that is in itself of no interest in this experiment. As described above, the six word-blocks were matched for grammatical category, letter-length, and frequency. Although it may be assumed that this matching procedure suffices for establishing the homogeneity of the linguistic materials in the six word-blocks, a statistical test of this assumption is appropriate. In examining the effect of the word-blocks, an analysis of variance for an incomplete three-way lay-out (the Latin Square) with fixed effects was carried out.
For the estimation of the overall mean and the effects of the main factors Handwritings, Subjects, and Word-blocks 16 degrees of freedom are needed. Because there are only 36 degrees of freedom for the complete square, only 20 are left for the estimation of the interaction ef-

fects. It was assumed that the second and third order interactions involving the factor Word-blocks itself were neglible. It was also assumed that that some of the handwriting-subject interactions were zero. Two tests with respect to the effects of the word-blocks were carried out one (I) with the assumption that there was no interaction between the more legible handwritings *A*, *B* and *C* and the two fastest subjects (subject 1 and 5 for the 'own handwriting' group and subjects 11 and 12 for the librarian group) A second test (II) was carried out with the assumption that there was no interaction between handwritings *A*, *B*, and *C* and subjects 1 and 2 in the own handwriting' group and between these handwritings and subjects 8 and 11 in the librarian group. Subjects 2 and 8 also appeared to be fast readers For both assumptions and for both groups, an A-analysis (correct responses only) and B-analysis (errors included) was carried out As can be seen in Table 5.1, all these tests resulted in p-values greater than 40 On the average, the absolute values of the word-blocks effects were about 50 ms

Table 5 1
F-values and associated p-values for different analyses (see text) of the effects of Word-blocks.

	Type of Analysis			
	A-analysis		B-analysis	
	I	II	I	II
Subjects 1-6	F = .59	F = .34	F = 1.02	F = .20
	p = .71	p = .89	p = .40	p = .96
Subjects 7-12	F = .41	F = .02	F = .33	F = .11
	p = .84	p = 1.0	p = .89	p = .99

The two alternative analyses -I and II- showed completely different estimations of the effects of the word-blocks This is probably due to the confounding of the word-block effects with effects of handwriting-subject interactions In both analyses, the hypothesis of no handwriting-subject interaction had to be rejected These analyses clearly support the assumption of no systematic differences between the word-blocks. The analysis of variance was therefore simplified to a two-way scheme with factors Handwriting and Subjects

2. *Own handwriting*

The analysis of variance for subjects 1-6 (own handwriting group) showed significant results for Handwritings [F (5,1996) = 65.97, $p < .01$] This result shows that the handwritings differed significantly in relative legibility. Subjects [F (5,1996) = 53.40, $p < .01$] and the interaction between Handwritings and Subjects [F (25,1996) = 4.18, $p < .01$] were also significant. Mean latencies are presented in Table 5.2.

Table 5.2
Mean Latencies (in milliseconds) of Twelve Subjects
for Six Handwritings in Experiment 8.

Subjects	Handwritings						
	A	B	C	D	E	F	Mean
1	581	655	602*	790	984	954	761
2	640	605	661	788	1089	870*	776
3	681*	861	824	1198	1543	1503	1102
4	748	740*	843	1053	1265	1350	1000
5	578	561	563	665	724*	696	631
6	788	757	848	785*	1015	1025	870
Mean	670	696	724	880	1103	1066	857
7	740	753	794	1009	1210	1066	929
8	723	717	750	877	927	1116	852
9	692	718	704	1096	1175	933	886
10	682	681	759	1027	1129	1424	950
11	547	556	570	602	840	898	669
12	732	689	729	776	868	805	767
Mean	686	686	718	898	1025	1040	842

Note Subjects 1-6 are the 'own handwriting' group (starred means are 'own handwriting'); subjects 7-12 are the librarians.

The estimates of the 'own handwriting' effects for subjects 1-6 were -31, -138, -280, -120, -185, and -130 ms, respectively. The mean 'own handwriting' effect was -147 ms (SE = 26 ms), which was significantly different from zero [t (1996) = -5.8, p < .01] The size of the 'own handwriting' effect did not increase regularly with reduced legibility of the handwriting; the largest effect (-280 ms) was observed for a handwriting which may be assumed to be fairly legible.

It was pointed out in the introduction that the facilitation for 'own handwriting' may not only reflect perceptual learning, but also more general differences in reading ability. As a check on this interpretation, providers of good handwritings (subjects 1, 3 and 4) were compared with providers of poor handwritings (subjects 2, 5 and 6) for their ability to read less legible handwritings. The mean RT of subjects 1, 3 and 4 for the 'good' handwritings (*A, B, C*) was 752 ms and for the 'poor' handwritings (*D, E, F*) 1182 ms. Corresponding figures for subjects 2, 5 and 6 were 667 and 880 ms (in the calculations the data for 'own handwriting' were left out). Scheffe's test showed differences between subgroups of subjects to be significant [*F* (5,1996) = 20.71, p < .01]. The interaction effect between subgroups of subjects and subgroups of handwritings, 217 ms (SE = 43 ms), was significantly different from zero [t (1996) = -5.05, p < .01] Providers of less legible handwritings read other 'difficult' handwritings than their own generally faster than providers of handwritings with good legibility, which suggests that the facilitation for 'own handwriting' may be part of a more general ability to read less legible handwritings.

3. *Group-effects*

The introduction of the librarians in the experiment was based on the consideration that, in their profession, they read many different handwritings, including presumably less legible ones. Their large experience might make them good readers of handwriting compared with randomly selected students. Moreover, the librarians may be expected to be especially good in reading less legible handwritings. In testing the differences between the two groups, data for 'own handwriting' were left out.

The students appeared to be faster readers than the librarians. An overall difference of 39 ms (SE = 14 ms) was found between the mean latencies for the student group and the mean latencies for the librarian group. This difference was significant [*t* (3631) = -2.77, p < .01]

Also contrary to expectations, the 'own handwriting' group read the more difficult handwritings *D, E* and *F* faster than the librarians. The interaction effect, calculated in an analogous way as for the subgroup analysis in the 'own handwriting' group, amounted to 69 ms (SE = 28 ms). This was found to be significantly different from zero [*t* (3631) = -2.45, p < .01]

Errors

The total number of errors -experimental error, erroneous readings, and 'illegible' responses- was 128 (5 9 %) for subjects 1-6 and 150 (6.9 %) for subjects 7-12.

For subjects 1-6, the percentages for each category separately were 2.2 for the experimental error, 2.7 for the erroneous readings, and .9 for the 'illegible' responses. Corresponding figures for subjects 7-12 were 2.3, 4, and 6 %. The experimental error was due to some malfunction of the apparatus or to an insufficiently loud response of the subject.

With respect to the erroneous readings, the following general observations can be made, which also valid for errors made in the other experiments. The number of erroneous readings increased with reduced legibility of the handwriting For both groups of subjects, an average of 88 % of all erroneous readings involved only handwritings *D*, *E* and *F*. In most cases, latencies for the erroneous readings were longer than the mean RT for the correct responses in a particular handwriting. Many erroneous readings involved the faulty recognition of one or two letters (66 % for the 'own handwriting' group, and 58 % for the librarians). It should be noticed that in some cases additional context was needed to decide between two plausible interpretations. This occurred especially if letters were highly confusable and when alternatives made words. The complexity of the ensuing decision processes is likely to be responsible for the notably longer latencies in most cases. The decision to treat these cases nevertheless as errors was based on the consideration that the writer intended a different word to be read.

Handwritings *D*, *E*, and *F* also caused most of the 'illegible' responses. An interesting detail of the 'illegible' responses deals with their latencies. The 'illegible'-RT was sometimes shorter than the longest correct RT. That suggests that in some cases the reader can readily decide that he will not be able to decipher the presented word.

DISCUSSION

The results for the 'own handwriting' group showed facilitation for the 'own handwriting': the mean 'own handwriting' effect was found to be significantly different from zero. This effect indicates that perceptual learning occurs for the characteristics of particular handwritings. In the introduction to Experiment 5, it was suggested that perceptual learning of handwriting likely consists in the discrimination of certain distinctive features or the formation of higher-order codes It may be assumed that both aspects play a role in facilitating the reading of the 'own handwriting' The result, obtained in Experiment 8, is in contrast with results obtained in Experiments 5 and 6 In those experiments, unreliable indications of perceptual learning were found It will be noted that Experiment 8 differs from Experiments 5 and 6 in the amount of practice In Experiment 5, subjects read about 5200 words in a particular handwriting and in Experiment 6 about 2500 words These numbers are, of course, much larger for the own handwriting' Apparently, perceptual learning of handwriting is a rather slow process that requires considerable amounts of practice

Some aspects of the results for the own handwriting' group deserve special consideration First, it will have been noticed that the providers of less legible handwritings read their own handwriting slower than handwritings with good legibility It was proposed earlier that the legibility of a handwriting is determined by its overall resemblance to print Providers of less legible handwritings will read more print than own handwriting which explains that they read handwritings with good legibility faster than their own

Second, the *facilitation for the own handwriting' was found not to be* specific and may therefore be partly due to general differences in reading ability. Providers of less legible handwritings generally read other 'difficult' handwritings than their own faster than providers of handwritings with good legibility In an extensive study on individual differences in reading, Jackson and McClelland (1979) found three independent correlates of individual differences in reading speed general language comprehension skill, processes involved in accessing letter-identity information, and use of complex spelling-to-sound correspondences They state 'it may be surprising that mature college student readers at a major state university have not reached asymptotic levels of letter-code access ability Surprisingly or not, our results do not support the statement of some (.) who have said that beyond the grade school level, individual differences in reading ability are only differences in comprehension difficulty (ibid , p 179/180) Individual differences in accessing letter-identity information may be magnified in the reading of less legible handwritings In Experiment 5, attention was called to large individual variations which were apparent in the large within-group variances Even among university students, which may be considered to be a relatively homogeneous group of readers, reading speed for handwritings with poor legibility may differ by a factor four or more (for the more legible handwritings the discrepancies are much less). In this connection it is also of interest that in Experiments 5 and 6 no evidence was found for more extensive use of context by more skilled readers (West and Stanovich, 1978).

Individual differences in reading ability may partly be determined by more general aspects of cognitive functioning In the context of this experiment the following speculation suggests itself In writing, two factors will be of special importance the speed of production and the legibility of the product. These two factors will often be in conflict· legible handwritings will be produced more slowly. Providers of less legible handwritings will be 'speed-oriented' producers. They may also be 'speed-oriented' in perception, as shown by their overall faster performance It should also be noticed that the generally greater ability of providers of less legible handwritings to read these handwritings may in fact be due to the 'own handwriting' It may not be unreasonable to assume that less legible handwritings show an increased resemblance. Due to the frequent reading of their own handwriting, providers of less legible handwritings will display larger amounts of transfer for other 'difficult' handwritings than providers of handwritings with good legibility.

The librarians were introduced as a separate group of subjects in the experiment to investigate 'general' perceptual learning of handwriting. The results for the librarians showed them in general to be slower readers of handwriting than university students They even appeared to be significantly slower than students in their ability to read less legible handwritings This result suggests that long-term practice in reading many different handwritings does not lead to a 'general' perceptual learning for reading handwritings Perceptual learning of handwriting apparently involves some specific handwriting, as was shown by the facilitation for the 'own handwriting'

GENERAL DISCUSSION

The general conclusion that may be drawn from Experiments 4-8 is that perceptual learning for handwriting is rather slow and limited. Although everyday experience suggests that one readily adapts to the specific characteristics of single handwritings, more controlled experimental conditions indicate that increases in reading speed can as well be due to other factors than perceptual learning Evidence for this limited perceptual learning was obtained in Experiments 5 and 6, in which no reliable gain-scores for the pretest/posttest single word recognition were observed. Aspects of the results of Experiment 8 also supported the general conclusion. Although facilitation for the own handwriting was found, subjects with less legible handwritings still read their own handwriting less easily than handwritings with good legibility they have never seen before. Extensive experience with many different handwritings does not make librarians better readers of handwritings than university students.

The limited perceptual adaptation to characteristics of single handwritings is compatible with the notion that abstract representations involved in pattern recognition remain rather constant over time and are not easily modified. Existing differences in legibility can then be explained as smaller or greater deviations from these constant representations. It seems likely that these representations contain information about the usual appearance of graphemes, i.e , about common typefaces. In this connection it is of interest that evidence for case-specific codings has been found in research that used printed materials (e g , Henderson and Chard, 1978; McClelland, 1976).

Although perceptual learning seems very limited during the initial reading of a handwriting, the facilitation for 'own handwriting' shows that, with more extensive practice, some perceptual learning does occur. The experiments did not provide cues as to what in fact has been learned. In the introduction to Experiment 5, it was suggested that perceptual learning might consist in the extraction of certain distinctive features or in the formation of higher-order perceptual codes This interpretation of perceptual learning effects may be a plausible one but it should be noted that alternative interpretations can be put forward. Experiments 5 and 6 are similar to studies, Kolers (e g , Kolers and Perkins, 1975; Kolers, Palef and Stelmach, 1980) has carried out with geometrically transformed texts in so far as both involve practice effects in reading unfamiliar typography According to Kolers, various transformations require subjects to exercise skills like rotating or ordering letters and words Practice effects are explained as facilitation due to repeated application of particular sets of such pattern-analyzing operations These conceptions can also be applied to the reading of handwriting (an example of a pattern-analyzing operation might be the reduction to standard height-width ratios that was discussed in the context of Experiment 2) and can be used to explain observed practice effects. It will be noted, however, that operations like rotation and ordering are rather easily specified for clearly definiable manipulations like geometrical transformations but analogous pattern-analyzing operations will be much more difficult to specify for the complex form variations that occur in handwriting

Kolers (1975) has argued that, when coding operations are skilled (as in normal typography) a reader is able to attend more to meaning. When coding operations are not automatic, however, the reader will have to attend more to lower-level aspects of the text Masson and Sala (1978) showed, however, that the reading of inverted typography is characterized by slow and elaborate data-driven processes and relies heavily on the use of conceptually-driven processes They even suggested that 'succesful and efficient surface processing is dependent on grasping the meaning of the message (ibid , p 268) The observed practice effects for reading continuous text support this suggestion.

The results of Experiments 5 and 6 make it likely that initial improvements in reading speed, observed in everyday reading of handwriting, should largely be contributed to changes in the interaction between bottom-up and top-down processes As the reader proceeds through the text and more knowledge is accumulated about its contents, top-down processes are able to provide more efficient assistance to the elaborate decipherment of less legible handwritings. Evidence for the correctness of this hypothesis was found in the decrease of consecutive reading times for the practice materials. In Experiment 5, conceptually easy texts were used which allowed for the development of efficient top-down processing strategies and decreasing reading times were observed In Experiment 6, the conceptually difficult texts limited this development and no significant decrease in reading times was found. This combination of findings indicates that the observation of decreasing reading times for handwritten material will depend on the conceptual difficulty of the reading material Experiment 7 provided indirect support for this interpretation of practice effects by showing that reading times for conceptually difficult texts increase markedly for less legible handwritings

The constant interplay between top-down and bottom-up processing may explain why perceptual learning of handwriting is limited Because the attention of the reader is mainly directed at semantic/conceptual information that will further his comprehension, the graphic aspects of the handwriting are hardly attended to In this connection it is of interest that significant gain-scores for the pretest/posttest word recognition were observed for handwritings in the low legibility condition in Experiment 6, while they were absent in Experiment 5 Limiting the amount of top-down processing forces the reader to pay more attention to the feature-information and will provide conditions that are more optimal for perceptual learning This aspect of the results supports the suggestion of LaBerge and Samuels (1973) that attention is needed for perceptual learning (see also Shiffrin and Schneider, 1977)

It will have been noticed that the explanation of the results emphasizes aspects of 'normal' reading This is justified by the fact that experimental conditions were a rather close approximation to normal reading Practice materials consisted of continuous text and the task for the subject required reading for meaning These experimental conditions do not imply that perceptual learning will generally be difficult to achieve It may well be that perceptual learning is rather fast when subjects are instructed to 'study' the handwriting Such conditions can, however, hardly be regarded as representative for normal reading

In Chapter 1 a review of experimental research of handwriting recognition was presented It suggested that reading handwriting is generally more difficult and may need different recognition routines than print In Experiments 4-8, handwritings were systematically sampled for legibility and results showed handwritings with good legibility not be significantly different from the printed control condition Dependent on their legibility, handwritings show different ways of processing. Interactions between top-down and bottom-up processes were found to differ with the legibility of handwriting. top-down information will often be indispensable for the decipherment of less legible handwritings. Moreover, these interactions do not remain constant, but change in the course of the reading process as was shown by practice-effects for reading continuous text The distinction between handwriting and print may therefore be nothing more than a salient aspect of more general legibility differences in visible language

APPENDICES

APPENDIX A

Stimuli used in Experiment 1 are listed below. The consecutive words in each form class are written in 20 different handwritings. Each item is followed by its mean latency in milliseconds.

Connected form

plein (square) 938; graan (corn) 1103; front (front) 936; breuk (crack) 1148; dwang (coercion) 1061; kraag (collar) 884; speld (pin) 940; slang (snake) 778; smart (grief) 896; bocht (bend) 936; klomp (lump) 908; brein (brain) 1137; kwaal (ailment) 1135; vuist (fist) 939; staaf (bar) 984; naald (needle) 921; ernst (seriousness) 1077; kloof (gap) 912; kwast (brush) 1011; klauw (claw) 994.

Segmented form

wraak (revenge) 998; drang (urge) 1098; storm (storm) 821; gloed (glow) 991; stoom (steam) 979; zicht (sight) 1601; nicht (niece) 929; vloot (fleet) 981; sluis (sluice) 1043; kraan (tap) 856; leeuw (lion) 984; gunst (favor) 1128; proef (test) 897; stoep (doorstep) 907; prooi (prey) 1302; kreet (scream) 1107; stank (stench) 1032; laars (boot) 959; bruid (bride) 1049; drank (drink) 964.

Control form

sloot (ditch) 1095; stier (bull) 1000; slaaf (slave) 1151; komst (coming) 1241; taart (cake) 1298; schot (shot) 910; draad (thread) 1130; kruis (cross) 1079; knaap (lad) 995; griep (flu) 917; spier (muscle) 918; fruit (fruit) 1017; strik (knot) 1179; kaars (candle) 1191; klank (sound) 990; schim (shadow) 984; baard (beard) 1308; plank (plank) 1115; oogst (harvest) 1461; poort (gate) 1203.

Appendix B: Stimuli used in Experiment 2. The number behind each item is the mean RT across 16 subjects

Spacing Quadrant Conditions								Control Conditions			
RR		IR		RI		II		Stretched		Contracted	
TARGETS											
Similar											
nmsa	555	nmci	628	nvae	591	nuzs	841	nvic	604	nmoc	646
nwor	542	nurz	658	nwec	713	nvio	848	nwaz	601	nueo	613
anvs	609	enuz	628	cnwi	664	inuc	808	rnui	571	rnvo	585
onwr	580	znve	658	rnmo	792	snma	1105	znmi	678	onws	568
acnu	623	esnv	702	oinv	686	renw	942	aznv	713	ernm	621
canm	595	iznw	665	zrnu	783	sonm	1278	sznu	650	ecnw	618
irvn	755	ocmn	790	coun	943	azvn	1057	asmn	924	sevn	662
zsun	793	sewn	788	rimn	877	eawn	931	riwn	913	caun	766
Dissimilar											
nevz	523	nacv	618	nrsu	722	ncwe	642	nomz	652	narv	575
nsui	645	niow	605	nzam	611	nomr	640	nrue	638	nozw	543
cnem	594	anzw	545	enav	736	onrw	698	ensw	588	cnzu	569
znou	496	insm	650	rncu	610	sniv	844	cniv	619	incm	650
vanc	695	minr	725	menz	719	uzne	795	manr	1012	uona	665
wrni	580	uons	886	wcna	693	vsno	694	wanr	876	vzno	602
iwrn	777	ucin	789	evcn	729	ruon	916	uisn	772	iwcn	733
mosn	781	vsan	764	wazn	689	zmen	747	evsn	861	msen	768
NON-TARGETS											
Similar											
uwoi	863	wvze	715	murs	902	vmca	1193	vmei	840	vwas	729
ewuz	765	ovws	761	iumo	1123	omvr	1111	amvz	951	zwvo	686
iwvr	751	auwc	729	evmz	921	zmue	1107	oumr	994	euwr	691
cavw	809	rswu	923	simv	1054	caum	970	scmu	913	ciwu	796
Dissimilar											
wouc	721	ruiw	996	uiam	1172	mreu	790	mcve	783	iusw	820
uzaw	841	wseu	943	amsu	1148	ucmo	989	ovsm	859	wouz	849
awsv	759	vrwi	744	mzva	1020	merv	830	uimz	1069	vewc	713
vcow	875	wezv	778	vocm	1256	svim	1031	rmau	917	rwav	731

APPENDIX C

The stimuli used in Experiment 3 are listed below The first number following each item is the mean latency for the undotted version, the second number is the mean latency for the dotted equivalent

GRAPHONYMIC ITEMS

Words, 4 letters duim (thumb) 734 676, emir (emir) 948 892, enig (sole) 712 735, etui (case) 856 752, klim (climbing) 780 773, knie (knee) 729 683, luis (louse) 806 738, mier (ant) 771 757, mist (mist) 719 725, muil (muzzle) 882 733, muis (mouse) 743 691, nimf (nymph) 780 828, pink (little finger) 741 712, puin (debris) 765 769, ruim (large) 859 779, slim (smart) 708 688, smid (smith) 756 709, trui (jersey) 772 711, tuin (garden) 698 710, uier (udder) 854 865

Words, 5 letters animo (gusto) 808 792, brein (brain) 747 730, bruin (brown) 738 657, einde (end) 766 741, genie (genius) 773 759, glimp (glimpse) 879 872, kluis (vault) 756 771, luier (diaper) 784 754, manie (mania) 883 895, minst (fewest) 843 819, nieuw (new) 731 766, onmin (discord) 1078 1081, opium (opium) 886 785, pruim (prume) 956 817; schim (shadow) 767 805, snuit (snout) 820 733, sluis (lock) 757 788, thuis (at home) 668 736, uiten (utter) 908 952, uniek (unique) 700 723

Nonwords, 4 letters muig 971 889, mire 996 1028, gine 845 879, tuip 983 975, klin 861 859, keni 907 802, slui 1052 962, kime 863 847, stim 905 906, limu 827 902, uims 912 791, anim 902 897, kinp 873 886, nuip 806 806, imur 908 834, mils 864 875, dims 853 791, prui 976 1064, tinu 908 908, buif 968 1086

Nonwords, 5 letters onima 913 910, ebrin 867 860, rinbu 824 841, ednie 776 846, egine 871 907, plimp 859 876, suilk 1075 886, uiler 999 976, amnie 957 930, nimst 896 874, weinu 982 885, nomni 871 886, pomui 880 843, rumip 868 874, misch 842 937, pluig 885 875, nuist 912 897, wuist 1061 1043, etinu 1036 910, inuke 958 811

Words, 4 letters arıa (arıa) 809 780; blık (can) 753 708, drıl (drıll) 835 975; gıro (gıro) 768 733; gıst (yeast) 753 737; grıl (whım) 789 676, ıdee (ıdea) 702 712; kıes (molar) 709 688; kıst (box) 738 717; kleı (clay) 742 706; klıp (rock) 791 795; lıst (lıst) 709 689, olıe (oil) 743 664; pıon (pawn) 820 787; prık (stab) 787 743; rıte (rıte) 903 906; rıts (zıpp) 739 701; silo (sılo) 885 838, spıl (pıvot) 833 690; trıo (trıo) 789 736.

Words, 5 letters bızon (buffalo) 951 814; dıcht (closed) 763 689; dosıs (dose) 820 724; drıft (temper) 828 733; emaıl (enamel) 882 919; fıets (bıcycle) 708 668, fırma (fırm) 821 758; kıosk (bookstall) 825 766; lakei (lackey) 814 751, legıo (legıon) 814 726; motıe (vote) 826 832; pıano (pıano) 807 693; pıoen (peony) 1066 980; prıor (prıorate) 897 883; radıo (radıo) 748 673; rıant (ample) 825 797; spıon (spy) 840 745; stıer (bull) 802 760, strık (button) 830 704; tıran (tyrant) 886 764.

Nonwords, 4 letters aıra 976 1019; klıg 799 792; lırd 808 842; lırg 824 810; grıo 948 836; stıg 831 847; eıde 926 862; sırk 864 843; ekil 912 831; plik 917 943; esık 861 840; ılst 924 829; elıo 779 863; onpı 840 789; krif 954 978; ıtre 795 841; sırt 824 848; olsı 830 806; ılps 806 831; rıto 945 917.

Nonwords, 5 letters bınzo 869 819; chıto 957 890; osdıs 785 814; frıdt 763 813; alıem 946 864; tıfse 895 810; frıam 901 837, sıoke 931 875; alkie 894 1010, giloe 868 896, adıpo 905 807; nepıo 849 773; meıto 971 858; ripro 853 895, trıan 1057 1170; odrıa 877 807; erıst 1017 896; nosıp 797 834; tirks 925 799; antır 921 1056.

APPENDIX D

The results of the analysis for order of presentation in Experiment 3 are presented below In Experiment 3 each type was presented twice once with a dot on the *i* and once without the dot on the *i* Order of presentation of dotted and undotted versions was not completely counterbalanced across subjects and items The deviations were, however, generally very small The analysis for effects of order was based on the means for first and second presentation for each condition separately

The critical F-values for all effects in the subject analysis were F (05, 1, 18) = 4 41 and F (01, 1, 18) = 8 29 For the item analysis these values were F (05, 1, 152) = 3 91 and F (01, 1, 152) = 6 88 Means are only presented for effects involving Order For the analysis involving the other factors (Graphonomy, Dot, Word Status and Length) only the F-values are reported with the understanding that means display the same pattern as for the analysis of first and second presentation combined

The main factor Order itself was significant both in subject analysis (F (1, 18) = 12 82) and in item analysis (F (1, 152) = 138 98) The overall mean for tokens presented first was 852 ms, for tokens presented second the mean was 799 ms As in the analysis for first and second token combined, the main effect of Graphonymy was significant in the subject analysis only (F (s) = 7 98, F (i) = 1 04, n s) The effects of Dot (F (s) = 19,53, F (i) = 29 23) and Lexical Status (F (s) = 77 64, F (i) = 97 62) were significant in both subject and item analysis The effect of Length (F (s) = 33 02, F (i) = 3 75, n s) was significant in the subject analysis only

In the subject and the item analysis the interaction between Order and Lexical Status was significant (F (s) = 23 90, F (i) = 24 70) Order of presentation had a larger effect for nonwords (85 ms) than for words (25 ms) Order of presentation did not interact significantly with any other factor or combinations of factors in the subject analysis, but in the item analysis the interaction Order × Graphonymy × Lexical Status (F (i) = 4 21) was significant, as was the interaction Order × Graphonymy × Length (F (i) = 7 40) For graphonymic words the difference between first and second presentation was 39 ms, for graphonymic nonwords 66 ms Corresponding differences for the nongraphonymic

stimuli were 23 and 97 ms. For graphonymic 4-letter stimuli the difference between first and second presentation was 37 ms; for the 5-letter stimuli 67 ms. Corresponding figures for the nongraphonymic stimuli were 55 and 41 ms.
In agreement with the analysis for first and second presentation combined, the following interactions were found to be significant. The interaction between Graphonymy and Dot was significant in the subject analysis ($F(s) = 10.94$; $F(i) = 2.94$, n.s.). Graphonymy interacted also significantly with Lexical Status ($F(s) = 14.88$). The interaction between Dot and Lexical Status was significant in the item analysis ($F(i) = 5.15$).

Comparing these results with the ones described for the analysis of the data for first and second presentation combined, the two analysis, with one exception, do not differ with respect to effects of Graphonymy, Dot, Lexical Status, Length, and their interactions. In the analysis for first and second presentation combined, the main effect of Length was significant in both subject and item analysis. In the analysis reported above Length was significant only in the subject analysis.
As an additional check, separate analyses were carried out for first and second presentations. These analyses showed the same pattern of results as observed in the two analyses reported.

APPENDIX E

The stimuli used in Experiment 4 are listed below. The first number following each item is the mean latency for the undotted version, the second number the mean latency for the dotted equivalent.

CONNECTED FORMS

Graphonymic items

Words, 4 letters etui (case) 781 673; knie (knee) 670 646; luis (louse) 784 751; mier (ant) 945 684; mist (mist) 680 700; nimf (nymph) 767 737; pink (little finger) 701 697; trui (jersey) 822 676; tuin (garden) 670 635; uier (udder) 893 710.

Words, 5 letters bruin (brown) 811 723; glimp (glimpse) 731 717; kluis (vault) 787 657; luier (diaper) 831 708; manie (mania) 855 782; minst (fewest) 805 802; nieuw (new) 726 664; onmin (discord) 820 887; schim (shadow) 730 736; uniek (unique) 691 692.

Nonwords, 4 letters dims 725 737; imur 796 686; kinp 809 798; klin 747 723; limu 764 697; muig 910 921; prui 885 932; slui 877 843; stim 923 841; uims 763 736.

Nonwords, 5 letters amnie 832 753; ednie 766 715; egine 766 797; inuke 812 747; nomni 822 788; onima 783 789; plimp 801 788; pomui 794 750; suilk 839 843; wuist 913 982.

Nongraphonymic items

Words, 4 letters aria (aria) 854 781; blik (can) 682 670; giro (giro) 690 715; gist (yeast) 712 653; idee (idea) 683 629; klei (clay) 717 672; list (list) 678 665; pion (pawn) 740 769; rite (rite) 809 760; trio (trio) 753 687.

Words, 5 letters dicht (closed) 728 630; dosis (dose) 744 694; fiets (bicycle) 613 666; motie (vote) 887 704; pioen (peony) 852 794; prior (priorate) 868 811; riant (ample) 753 768; spion (spy) 734 709; stier (bull) 708 769; strik (button) 778 669.

Nonwords, 4 letters aira 960 877; eide 786 786; elio 796 776; ilst 800 708; itre 858 709; klig 728 741; lirg 730 723; rito 894 772; sirk 767 719; sirt 799 797.

Nonwords, 5 letters adipo 718 721; alkie 780 756; antir 835 843; binzo 764 869; friam 818 753; meito 926 829; nosip 703 744; osdis 749 751; ripro 775 814; sioke 753 751.

Graphonymic items

Words, 4 letters duim (thumb) 651 630, emir (emir) 904 781; enig (sole) 659 647; klim (climbing) 670 675, muil (muzzle) 725 694; muis (mouse) 658 645; puin (debris) 676 702; ruim (large) 660 701; slim (smart) 657 628; smid (smith) 655 705.

Words, 5 letters animo (gusto) 676 750, brein (brain) 695 678; einde (end) 689 761; genie (genius) 714 703, opium (opium) 749 692; pruim (prume) 709 683; sluis (lock) 694 672; snuit (snout) 752 678; thuis (at home) 641 622, uiten (utter) 776 807.

Nonwords, 4 letters anim 765 767; buif 952 826; gine 730 816; keni 748 709; kime 697 758, mils 779 759, mire 928 938; nuip 715 713; tinu 765 785, tuip 830 853

Nonwords, 5 letters ebrin 756 762, etinu 902 823; misch 746 747; nimst 754 760; nuist 777 851, pluig 798 816; rinbu 731 753; rumip 785 759; uil-er 896 864; weinu 759 778.

Nongraphonymic items

Words, 4 letters dril (drill) 769 705, gril (whim) 672 718, kies (molar) 675 706; kist (box) 650 645; klip (rock) 732 729; olie (oil) 670 657; prik (stab) 706 719; rits (zipp) 794 854; silo (silo) 734 781; spil (pivot) 684 709.

Words, 5 letters bizon (buffalo) 736 783; drift (temper) 733 711;email (enamel) 745 758; firma (firm) 702 688; kiosk (bookstall) 698 685; lakei (lackey) 808 685, legio (legion) 722 728; piano (piano) 654 672; radio (radio) 635 622; tiran (tyrant) 775 677.

Nonwords, 4 letters ekil 724 721, esik 718 720; grio 831 834; ilps 677 752, krif 757 820 lird 679 677; olsi 754 762; onpi 708 761; plik 799 839; stig 812 792

Nonwords, 5 letters aliem 813 779, chito 720 771, erist 825 792; fridt 710 674; giloe 717 736, nepio 739 667, odria 793 744; tifse 752 682, tirks 744 781, trian 1084 931.

APPENDIX F

The results of the analysis for order of presentation in Experiment 4 are presented below. As in Experiment 3, each 'type' was presented twice: once with the dot on the *I* and once without the dot on the *I*. Order of presentation was completely counterbalanced across subjects and items. Because the capacity of the computer program (BMD-08V) used was exceeded, no subject analysis could be carried out in which all factors, including Order of presentation, were analyzed simultaneously. Therefore, for the subject analysis, results for the first presentation only are presented.

In the item analysis the main factor Order was significant ($F (1, 144) = 343.12, p < .01$). The overall mean for the tokens presented first was 789 ms; for tokens presented second the mean was 714 ms. Connectedness was significant ($F (1, 144) = 18.18, p < .01$) as were the main effects of Dot ($F (1, 144) = 5 92, p < .05$) and Lexical Status ($F (1, 144) = 45.41, p < .01$). The main effects of Graphonymy and Length were not significant
Order of presentation interacted significantly with Connectedness ($F (1, 144) = 9 46, p < 01$). For the segmented forms the difference between first and second presentation was 63 ms; for the connected forms 88 ms. No other interaction involving Order of presentation was significant. The interaction between Connectedness and Dot was significant ($F (1, 144) = 10 02, p < 01$) and the three-way interaction between Connectedness, Dot, and Length ($F (1, 144) = 5.32, p < .05$).

In the subject analysis for the first presentation only (involving different stimuli for different subjects), all main effects were significant The F-values for the factors Connectedness, Graphonymy, Dot, Word, and, Length were 28 04, 5 17, 13 68, 23 07 and 4 56 respectively The critical F-values for this analysis were $F (.05, 1,18) = 4 41$ and $F (01, 1,18) = 8.29$ The interaction between Connectedness and Dot was significant ($F (1, 18) = 10 34$). The three-way interaction Connectedness × Graphonymy × Lexical Status was also significant ($F (1, 18) = 6 59, p < 05$) Means for this analysis showed the same pattern as observed for the analysis of first and second presentation combined

APPENDIX G

The series of words used for the pretest and posttest in Experiments 5 and 6 are listed below. The two series were alternated for pretest and posttest.

Series A: agent (agent); bezit (posssesion); droom (dream); gebed (prayer); klauw (claw); loods (shed); molen (mill); nagel (nail); offer (sacrifice); sport (sport).

Series B; appel (apple); brood (bread); draad (thread); getal (number); klank (sound); laars (boot); motor (motor), neger (negro); oever (bank); storm (storm).

APPENDIX H

This appendix contains a description of the procedure and results of the selection experiment in which legibility ratings were obtained for the handwritings used in Experiments 4-8.

Out of a larger, randomly collected sample of 100 handwritings (see also Chapter 2, Experiment 1), 32 handwritings were selected that differed considerably in legibility The 32 handwritings differed also in graphic characteristics like size, slope, manuscript versus cursive. In the selection of the handwritings no attempt was made to vary legibility while keeping graphic features constant.
Each handwriting was presented by means of a little piece of prose of about 50 words, which was different for each handwriting Research into the quality of handwriting often makes use of one and and the same piece of linguistic material. This procedure seems to be less suited for obtaining judgments about the relative ease with which handwritings can be read. Because of order effects in presentation, the legibility of difficult handwritings may be overestimated To reduce the influence of the conceptual difficulty of the texts on the legibility judgment, the prose passages were all taken from primary school books.
Under each passage a 5-point scale was presented with verbal labels: zeer gemakkelijk (very easy) = 1; tamelijk makkelijk (fairly easy) = 2; niet makkelijk/niet moeilijk (not easy/not difficult) = 3; tamelijk moeilijk (fairly difficult) = 4; zeer moeilijk (very difficult) = 5 Subjects were instructed to mark only one of the labels
For order of presentation the 32 handwritings were split in two groups. All subjects were first presented with seven handwritings which covered about the whole range of legibility in the sample, followed by the remaining 25 handwritings. Within these two groups, order of presentation was random for each subject To avoid having the judgments based on a first impression only, subjects were instructed to read the passage before they passed judgment Subjects were asked explicitly to pay attention to relative legibility only and not to consider esthetic qualities. The experiment was carried out group-wise for 12 subjects, all students at the Faculty of Language and Literature at the University of Nijmegen.

On the basis of their mean judgment the 32 handwritings were ordered for relative legibility. The mean judgment varied from 1 to 4.66, showing that a reasonable variation of relative legibility was present in the sample.

An important aspect of the rank-ordering is the degree of agreement between judges. To determine this agreement, differential weightings were assigned to differences in judgments. For complete agreement (the use of the same scale value) a 1 was assigned, for complete disagreement (differences of four scale values) the weight 0 was used. To differences of one, two and three scales units the weights of .75, .50 and .25 were assigned. The mean sum of the number of observed agreement, multiplied by their corresponding weights, constituted the percentage of observed agreement between two judges. The mean percentage of observed agreement for the 66 pairs of judges was 81; the highest value was 91; the lowest one 71.

Inspection of the judgment for single handwritings showed that seven of the 32 handwritings appeared to be difficult to judge, because the judgment was found to be rather inconsistent (differences of three scale values or more) across judges. A calculation of the mean percentage of observed agreement, leaving out these seven handwritings, was shown to be 84; the highest value was 92; the lowest one 71. Herrick & Erlebacher (1963) noted that handwritings that appeared difficult to classify, fell into three categories: extremely large or small, extreme slope or irregularity in letter or word forms. Inspecting the seven handwritings for these characteristics showed that three of them did display a pronounced slope. On the other hand, however, there were handwritings in the sample that showed the characteristics mentioned above, but which were classified rather consistently.

Although the use of a 5-point scale results in a relatively high percentage of observed agreement between judges, a reduction in the number of scale units was applied to increase this agreement. Analyses were carried out for two different reductions to three scale units. In these analyses weight '1' was assigned to complete agreement; for differences of two scale units a '0' was assigned and .50 was used for differences of one scale unit. In a first reduction, categories 1 and 2 (very easy and fairly easy) and 4 and 5 (fairly and very difficult) were combined. This reduction to three scale units resulted in a *lower* mean percentage (74) of observed agreement. A second reduction to three scale units, however, did result in a higher percentage of observed agreement. This reduction involved combining categories 2, 3 and 4 (fairly easy, not easy/not difficult, fairly difficult). The mean percentage observed agreement in this analysis was 90, the highest value was 97; the lowest one 80. In combination, these two analyses suggest that judges find it rather difficult to distinguish between the middle scale values It seems therefore likely that legibility judgments make use of three distinct categories clear cases of very legible handwritings, clear cases of handwritings with poor legibility and a broad, diffuse class of handwritings with average or medium legibility On the basis of these analyses three categories of legibility -high, medium, and low- were decided upon to represent a 'reasonable' variation in legibility

ADAMS, M.J. Models of Word Recognition *Cognitive Psychology*, 1979, 11, 133-176.

ALLPORT, A. Word recognition in reading (Tutorial paper). In P.A. Kolers, M Wrolstad & H. Bouma (Eds.), *Processing of visible language Volume I*. New York Plenum Press, 1979.

ANDERSON, D.W. What makes Writing Legible? *The Elementary School Journal*, April 1969.

AYRES, L.P., A Scale for Measuring the Quality of Handwriting of School Children. New York. Russell Sage Educational Bulletins, No. 113, 1912

BANKS, W.P., & PRINZMETAL, W. Configurational effects in visual information processing. *Perception & Psychophysics*, 1976, Vol. 19(4), 361-367.

BECKER, C.A Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 556-566.

BECKER, C.A., & KILLION, T.H Interaction of Visual and Cognitive Effects in Word Recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, Vol 3, No 3, 389-401.

BRADSHAW, J L , Three interrelated problems in Reading: *Memory & Cognition* 3, 1975, 123-134.

BROOKS, L. Visual Pattern in Fluent Word Identification. In A.S Reber and D L Scarborough (Eds) *Toward A Psychology of Reading* N J · Lawrence Erlbaum, 1977

BRYDEN, M.P , & ALLARD, F Visual Hemifield Differences Depend on Typeface. *Brain and Language*, Vol 3 (1), 1976, 191-200

CAMPBELL, D.T., & STANLEY, J C *Experimental and Quasi-Experimental Designs for Research* Chicago Rand McNally, 1963

CLARK, H H The language-as-fixed-effect-fallacy A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 335-359

COLLINS, A , BROWN, J S., & LARKIN, K M Inference in Text Understanding In. R J Spiro, B C Bruce, & W F. Brewer (Eds) *Theoretical Issues in Reading Comprehension* N J Lawrence Erlbaum, 1980

COLTHEART, M. Critical notice. *Quarterly Journal of Experimental Psychology*, 1977, 29, 157-167.

COLTHEART, M., DAVELAAR, E., JONASSON, J.T., & BESNER, D. Access to the Internal Lexicon. In. S. Dornic (Ed) *Attention and Performance VI*. London· Academic Press, 1977.

CORCORAN, D.W., & ROUSE, R.O. An aspect of perceptual organization involved in reading typed and handwritten words. *Quarterly Journal of Experimental Psychology*, 1970, 22, 526-530.

ERIKSEN, B.A. & ERIKSEN, C.W. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 1974, Vol. 16, No. 1, 143-149.

ERIKSEN, C.W., POLLACK, M.D., & MONTAGUE, W.E. Implicit speech: Mechanism in perceptual encoding? *Journal of Experimental Psychology*, 1970, *84*, 502-507.

ESTES, W K. Interactions of signal and background variables in visual processing. *Perception & Psychophysics*, 1972, 12, 278-286.

ESTES, W. K. The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General*, 1975, *104*, 122-145 (a).

ESTES, W. K. Memory, perception, and decision in letter identification. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hillsdale, N.J.: Erlbaum, 1975 (b).

ESTES, W K. On the Interaction of Perception and Memory in Reading. In: D. Laberge & S. Samuels (Eds.) *Basic Processes in Reading: Perception and Comprehension*. Hillsdale, N.J.: Erlbaum, 1977.

FORD, B., & BANKS, W. P. Perceptual differences between reading handwritten and typed words. *Memory & Cognition*, 1977, 5, 630-635.

FORSTER, K.I. Accessing the mental lexicon. In R.J. Wales & E. Walker (Eds.) *New approaches to language mechanisms*. Amsterdam: North-Holland, 1976.

FREEMAN, F.N. An Analytical Scale for Judging Handwriting. *Elementary School Journal*, 1915, 15, 432-441

GARDNER, G.T., Evidence for Independent Parallel Channels in Tachistoscopic Perception. *Cognitive Psychology*, 1973, 4, 130-155.

GARNER, W.R. Letter Discrimination and Identification. In: A. Pick (Ed.) *Perception and its Development: a tribute to E.J. Gibson*. Hillsdale, N.J.: Erlbaum, 1979

GIBSON, E.J. *Principles of perceptual learning and development*. New York. Prentice-Hall, 1969.

GIBSON, E.J. Perceptual Learning and the Theory of Word Perception. *Cognitive Psychology*, 1971, 2, 351-368. New York: Basic Books, 1970.

GIBSON, E.J., & LEVIN, H *The psychology of reading*. Cambridge: MIT Press, 1975.

GOODMAN, K.S Reading. A Psycholinguistic Guessing Game In H. Singer & R.B Ruddell (Eds.) *Theoretical Models and Processes of Reading*. (2nd ed.) Newark International Reading Association, 1976.

GOUGH, P.B. One second of reading In J.F Kavanagh & I.G. Mattingly (Eds.) *Language by ear and by eye*. Cambridge, Mass.. MIT Press, 1972.

HENDERSON, L. & CHARD, M J On the Nature of the Facilitation of Visual Comparisons by Lexical Membership *Bulletin of the Psychonomic Society*, 7, 1976, 432-434

HERRICK, V.E. (Ed) *New Horizons for Research in Handwriting*. Madison: The University of Wisconsin Press, 1963

HERRICK, V E. & ERLEBACHER, A The Evaluation of Legibility in Handwriting In V E Herrick (Ed) *New Horizons for Research in Handwriting*. Madison: The University of Wisconsin Press, 1963.

HOCHBERG, J Organization and the Gestalt Tradition In : E.C. Carterette & M.P. Friedman (Eds) *Handbook of perception, Volume 1*. New York· Academic Press, 1974.

HUEY, E B *The psychology and pedagogy of reading*. Cambridge, Mass. MIT Press, 1968. (Originally published, New York: Macmillan, 1908)

JACKSON, M D , & McClelland, J L Processing Determinants of reading speed *Journal of Experimental Psychology: General*, 1979, Vol. 108 (2), 151-181

JOHNSON, N.F. On the Function of Letters in Word Identification. Some Data and a Preliminary Model *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 17-29

JOHNSON, N F. A pattern-unit model of word identification In· D. LaBerge & S J Samuels (Eds) *Basic processes in reading: Perception and comprehension* Hillsdale, N J . Erlbaum, 1977

JOHNSTON, J C., & McCLELLAND, J L Experimental Tests of a Hierarchical Model of Word Identification *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 503-524.

KIRK, R E *Experimental Design: Procedures for the Behavioral Sciences* Belmont, CA Brooks/Cole, 1968

KOFFKA, K *Principles of Gestalt Psychology* New York Harcourt, Brace & World, 1935

KOLERS, P Memorial consequences of automatized encoding *Journal of Experimental Psychology, Human Learning and Memory*, 1975, 1, 689-701

KOLERS, P , & PERKINS, D N Spatial and Ordinal Components of Form Perception and Literacy *Cognitive Psychology*, 1975, 7, 228-267

KOLERS, P A , PALEF, S R , & STELMACH, L B Graphemic analysis underlying literacy *Memory & Cognition*, 1980, Vol 8 (4), 322-328

KRUEGER, L E Familiarity Effects in Visual Information Processing *Psychological Bulletin*, 1975, Vol 82, No 6, 949-974

KRUMHANSL, C L , & THOMAS, E A Effect of level of confusability on reporting letters from briefly presented visual displays *Perception and Psychophysics*, 1977, Vol 21 (3), 269-279

LaBERGE, D Perceptual Learning and Attention In W K Estes (Ed) *Handbook of Learning and Cognitive Processes. Vol. 4 Attention and Memory* Hillsdale, N J Erlbaum, 1976

LaBERGE, D , & SAMUELS, S J Toward a Theory of Automatic Information Processing in Reading *Cognitive Psychology*, 1974, *6*, 293-323

LESGOLD, A M , & PERFETTI, C A (Eds) *Interactive Processes in Reading* N J Lawrence Erlbaum, 1981

LINDSAY, P H , & NORMAN, D A *Human Information Processing: An introduction to psychology* New York Academic Press, 1972

MASSARO, D How does orthographic structure facilitate reading? In J F Kavanagh & R L Venezky (Eds) *Orthography, Reading, and Dyslexia* Baltimore University Park Press, 1980

MASSARO, D W , JONES, R D , LIPSCOMB, C , & SCHOLZ, R Role of Prior Knowledge on Naming and Lexical Decisions with Good and Poor Stimulus Information *Journal of Experimental Psychology: Human Learning and Memory*, 1978, Vol 4, No 5, 498-512

MASSARO, D W , & SCHMULLER J Visual features, perceptual storage, and processing time in reading In D W Massaro (Ed) *Understanding Language. an information processing analysis of speech, perception, reading and psycholinguistics* New York Academic Press, 1975

MASSARO, D W , TAYLOR, G A , VENEZKY, R L , JASTRZEMBSKI, J E , & LUCAS, P A *Letter and Word Perception* Amsterdam North-Holland, 1980

MASSON, M.E.J. & SALA, L.S. Interactive Processes in Sentence Comprehension and Recognition. *Cognitive Psychology*, 1978, 10, 224-270.

McCLELLAND, J.L. Letter and Configuration Information in Word Identification. *Journal of Verbal Learning and Verbal Behavior*, 1977, 16, 137-150.

MEWHORT, D.J K. Sequential redundancy and letter spacing as determinants of tachistoscopic recognition. *Canadian Journal of Psychology*, 1966, 20 (4), 435-444.

MEYER, D E., SCHVANEVELDT, R W., & RUDDY, M.G. Loci of contextual effects on visual word recognition In: P.M.A. Rabbit & S. Dornic (Eds.), *Attention and Performance V*. London: Academic Press, 1975.

MORTON, J. Interaction of information in word recognition. *Psychological Review*, 1969, 76, 165-178.

MORTON, J. Facilitation in Word Recognition: Experiments Causing Change in the Logogen Model. In P.A. Kolers, M.E. Wrolstad & H. Bouma (Eds) *Processing of Visible Language, Volume 1*, New York: Plenum Press, 1979.

NAVON, D. Forest Before Trees: The Precedence of Global Features in Visual Perception. *Cognitive Psychology*, 1977, 9, 353-383.

NEISSER. U. *Cognitive Psychology*. New York, 1967.

NEISSER, U., & WEENE, P. A Note on Human Recognition of Hand-Printed Characters. *Information and Control 3*, 1960, 191-196.

NORMAN, D.A., & BOBROW, D G On data-limited and resource-limited processes *Cognitive Psychology*, 1975, *7*, 44-64

O'CONNOR, R.E , & FORSTER, K.I Criterion Bias and Search Sequence Bias in Word Recognition. *Memory & Cognition*, 1981, 9, 78-92.

PALMER, S E. Visual Perception and world knowledge. In D A Norman, D.E Rumelhart & LNR Research Group (Eds) *Explorations in cognition* San Fransisco W.H Freeman, 1975

PALMER, S E. Hierarchical Structure in Perceptual Representation. *Cognitive Psychology*, 1977, *9*, 441-474

POMERANTZ, J R Are complex visual features derived from simple ones? In· E. Leeuwenberg & H Buffart (Eds) *Formal theories of visual perception*. Sussex, England Wiley Ltd, 1978.

PRESSEY, L C., & PRESSEY, S C Analyses of three thousand illegibilities in the handwriting of children and adults *Educational Research Bulletin*, 1927

QUANT, L. Factors Affecting The Legibility of Handwriting. *Journal of Experimental Education*, 1946, 14 (4), 297-316.

REICHER, G.M. Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 1969, *81*, 275-280.

ROMEIN, J., & ROMEIN-VERSCHOOR, A. *Erflaters van onze beschaving*. Amsterdam: Querido, 1938-1940.

RUBENSTEIN, H., GARFIELD, L., & MILLIKAN, J.A. Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 487-492.

RUBENSTEIN, H. LEWIS, S.S., & RUBENSTEIN, M.A. Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 645-657.

RUMELHART, D.E. Toward an interactive model of reading. In S. Dornic (Ed.) *Attention and Performance VI*. Hillsdale, N.J.: Erlbaum, 1977.

RUMELHART, D.E., & SIPLE, P. Process of recognizing tachistoscopically presented words. *Psychological Review*, 1974, Vol. 81, No. 2, 99-113.

SCHINDLER, R.M., WELL, A.D., & POLLATSEK, A. Effects of segmentation and expectancy on matching time for words and nonwords. *Journal of Experimental Psychology*, 1974, Vol. 103, No. 1, 107-111.

SHIFFRIN, R.M., & GEISLER, W.S. Visual recognition in a theory of information processing. In R. L. Solso (Ed.) *Contemporary issues in cognitive psychology: The Loyola Symposium*. Washington, D.C.: V.H. Winston & Sons, 1973.

SHIFFRIN, R.M., & SCHNEIDER, W. Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory. *Psychological Review*, 1977, Vol. 84 (2), 127-190.

SIEGEL, S. *Nonparametric Statistics*. New York: McGraw-Hill, 1956.

SMITH, F. *Understanding Reading*. New York: Holt, Rinehart and Winston, Inc. 1971.

SPIRO, R.J., BRUCE, B.C., & BREWER, W.F. *Theoretical Issues in Reading Comprehension*. N.J.: Lawrence Erlbaum, 1980.

STERNBERG, S. High-speed scanning in human memory. *Science*, 1966, 153, 45-52.

STERNBERG, S. The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 1969, *30*, 276-315.

STERNBERG, S. Memory scanning: New findings and controversies. *Quarterly Journal of Experimental Psychology*, 1975, 27, 1-32.

SOMMMER, R. A note on 'graphonyms'. *American Journal of Psychology*, 1977, Vol. 90, No. 3, 527-528.

SPERLING, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, 74, (Whole No. 4).

SUEN, C.Y., BERTHOD, M., & MORI, S. Advances in recognition of handprinted characters. *Proceedings of the 4th International Joint Conference on Pattern Recognition*, Kyoto, nov. 7-10, 1978.

SUTHERLAND, N.S. Object Recognition. In: E.C. Carterette & M.P. Friedman (Eds.) *Handbook of Perception Vol III Biology of Perceptual Systems*. New York: Academic Press, 1973.

TAFT, M. Lexical Access via an Orthographic Code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behavior*, 1979, Vol. 18, No. 1, 21-41.

TEMPLIN, E.M. The Legibility of Adult Manuscript, Cursive, or Manuscript-Cursive Handwriting Styles. In: V.E. Herrick (Ed.) *New Horizons for Research in Handwriting*. Madison: The University of Wisconsin Press, 1963.

TERRY, P., SAMUELS, S.J., & LaBERGE, D. The Effects of Letter Degradation and Letter Spacing on Word Recognition. *Journal of Verbal Learning and Verbal Behavior*, 1976, *15*,577-585.

THOMASSEN, A.J.W.M. & HUDSON, P.T.W. Is de rechter hemisfeer het meest geschikt voor het lezen van handschrift? *Nederlands Tijdschrift voor de Psychologie*, 1982, in press.

THORNDIKE, E.L. Handwriting. *Teachers College Records*, 1910, 11, 83-175.

TIMM, N.H. *Multivariate analysis with applications in Education and Psychology*. Monterey: Brooks/Cole, 1975.

TINKER, M.A. *Bases for Effective Reading*. Minneapolis: University of Minnesota Press, 1965.

UIT DEN BOOGAART, P.C. *Woordfreqenties*. Utrecht: Oosthoek, Scheltema & Holkema, 1975.

WEISSTEIN, N., & HARRIS, C.S. Visual detection of line segments: An object superiority effect. *Science*, 1974, *186*, 752-755.

WERTHEIMER, M. Untersuchung zur Lehre von der Gestalt. II. *Psychologische Forschung*, 1923, 4, 301-350.

WEST, R.F., & STANOVICH, K.E. Automatic contextual facilitation in readers of three ages. *Child Development*, 1978, 49, 717-727.

WOLFORD, G. Perturbation model for letter identification. *Psychological Review*, 1975, 82, 184-199.

SUMMARY

In this study two aspects of reading handwriting are investigated One aspect deals with the segmentation of letters in cursive handwriting, the other with adaptation to the characteristics of individual handwritings Both aspects may potentially reveal differences between reading print and handwriting

In the introductory Chapter 1, a review of experimental research with respect to the recognition of handwriting is presented It suggests that reading of handwriting is more difficult and may require different recognition routines than reading print

Attention is called to certain methodological aspects of handwriting research It is pointed out that handwriting research may adopt two approaches or strategies that differ in attempted generalization across handwritings and therefore in sampling procedures A 'breadth' approach tries to establish whether certain processes are common to all handwritings or particular classes of handwritings In this approach, representative samples of handwriting will be required For reading research, representative samples of handwritings will be based on dimensions like physical characteristics or legibility Problems connected with sampling procedures, based on these dimensions, are discussed.

A depth approach is primarily concerned with the study of some particular phenomenon which may be considered characteristic for reading handwriting, but which does not have to occur in all handwritings or in some defined class(es) of handwritings Attempted generalizations will involve aspects of the phenomenon under study and not handwritings

Experiments 1-4, that are reported in Chapters 2 and 3, are concerned with segmentation procedures in cursive handwriting In Experiment 1, it is attempted to provide evidence for the greater complexity of segmenting cursive script The recognition of handwritten words in which letters are connected is compared with words in which letters are separate Contrary to expectations, no differences are found between the two conditions although significantly longer naming latencies are observed for a control condition in which line-segments are removed within letters These results indicate contours as an alternative for inter-letter spaces in the segmentation of handwriting

This is taken up in Experiment 2 where the formation of letter-contours is studied It is suggested that these contours are due to the operation

of Gestalt principles for grouping features (closed boundaries, spatial contiguity and similarity) within the configuration as a whole. The operation of the spatial contiguity principle is studied by means of varying the regularity of letter-widths and inter-letter distances while keeping constant overall stimulus length. In this way configurations of features are created that violate the spatial contiguity principle for grouping, because features belonging to adjoining letters are sometimes much closer together than features belonging to the same letter. The similarity of adjoining features is manipulated by placing next to each other, letters that share similar features (straight, vertical line-segments). Experiment 2 uses a *n*-detection task. Stimulus materials consist of illegal pseudo-words. Results show that both spatial contiguity and similarity of features have an effect on segmentation. Experimental conditions in which spatial relations between features are not consistent with the principle of spatial contiguity for assigning letter-boundaries result in generally longer latencies. The effect of similarity of adjoining features is dependent on spatial contiguity conditions In combination, the similarity and spatial contiguity of features create configurations that erroneously activate letter representations. Aspects of the results of Experiment 2 indicate that the speed of letter recognition is affected by the degree in which the letter approximates a standard form. In Experiments 3 and 4, it is studied whether this distinctiveness of letters also has an effect on segmentation.

The importance of clear letter forms for segmentation is (negatively) demonstrated by the existence of graphonyms'· handwritten words or segments that are ambiguous for segmentation. Graphonyms can be described as contours that are compatible with different sets of letters, which are themselves rather indistinctive Such graphonymic stimuli will be more difficult to segment, because of an increased probability that features will be combined in the wrong way.

In Experiment 3, artificially created graphonymic segments are introduced into the stimulus materials by placing undotted *i*'s next to *m*'s, *n* s, or *u*'s The effect of dotting the *i* for these graphonymic stimuli is compared with the effect for nongraphonymic stimuli in which the *i* is next to letters with quite different contours or features (like ascenders or descenders or curved line-segments). To assess whether linguistic context affects segmentation, Experiment 3 usesa lexical decision task which allows presentation of words and nonwords. When meaningful context facilitates segmentation, graphonymic nonwords will be especially difficult to identify

Although graphonymic stimuli result in significantly longer latencies than nongraphonymic stimuli and undotted stimuli are recognized slower than dotted ones, the expected larger effect of dotting the *i* for graphonymic stimuli is not observed In fact, dotting the *i* has a larger effect for the nongraphonymic stimuli A significant interaction between lexical status of the letter-sequence and graphonymy is obtained graphonymic nonwords result in significantly longer latencies than nongraphonymic nonwords

Experiment 4 is a replication of Experiment 3 in which the segmentability of configurations is manipulated in an additional way Stimuli are either segmented or connected forms (cf Experiment 1) In the former, connecting line-segments between letters are removed while they are present in the latter Because connected forms will be more difficult to segment than explicitly segmented forms, the effect of dotting the *i* is expected to be larger for connected forms

Segmented forms are recognized significantly faster than connected forms Moreover, the effect of dotting the *i* is significantly larger for the connected forms, making it likely that the dot on the *i* (i e , the distinctiveness of the letter) also facilitates segmentation of these stimuli

In the general discussion of the results of Experiments 1-4, three factors, that contribute to segmentation, are identified One factor has to do with the configuration of features in the stimulus as a whole Configurational aspects involving closed boundaries, spatial contiguity and similarity of adjoining features are especially relevant for determining letter-boundaries A second factor relates to the configuration of features at the letter level The distinctiveness of letters not only facilitates letter recognition but also segmentation It is argued that the interaction between the distinctiveness of the letter and configurational context in Experiment 4 indicates that segmentation and letter recognition belong to a common stage in processing Segmentation can, therefore, be immediate, i e based on specific configurations of features that activate letter representations directly A third factor identified is meaningful linguistic context The contribution of this factor consists in inducing guessing letters for configurations that are difficult to segment

In Chapters 4 and 5 experiments are reported that deal with perceptual learning of handwriting The experiments relate to the common observation that one gets used to the characteristics of particular handwritings. Two interpretations of initial practice as evidenced by increased reading speed are considered. Practice may reflect perceptual learning of the handwriting. This perceptual learning might consist in the discrimination of certain distinctive features or in the formation of higher-order perceptual codes. Increased reading speed may, however, also be an indication of more efficient top-down processing, made possible by the specific contents of the reading material or more optimal application of general top-down reading strategies.

For Experiments 5-7, that are reported in Chapter 4, handwritings are selected that differ in legibility. It is expected that practice effects will only be apparent for less legible handwritings. Handwritings with good legibility are assumed to be rather similar to print and for these handwritings large amounts of transfer are supposed.

In Experiment 5, seven (unrelated) prose-passages are read consecutively as practice materials, preceded and followed by a pretest/posttest involving single word recognition (word naming). Gain-scores on this latter task are used as a (context-free) measure of perceptual learning. Nine different handwritings are used, evenly distributed over a high, medium, and low legibility condition, together with a control condition with printed stimulus materials.

For handwritings in the high legibility condition, no practice effects of any kind are observed. Moreover, reading times in this condition are not significantly different from the printed control condition.

Handwritings in low and medium legibility conditions are less legible than print but still no consistent practice effects are observed. For handwritings in the low legibility condition, decreasing reading times for consecutive prose passages are not reflected in gain-scores for the pretest/posttest word recognition task. For handwritings in the medium legibility condition results are opposite: despite the fact that no significant decrease in consecutive reading times for the prose-passages is obtained, the pretest/posttest shows gain-scores significantly above zero.

An explanation of these results involves the conceptual difficulty of the text materials. As the prose-passages are rather easy to comprehend, they cause the adoption of elaborate top-down processing in reading handwritings with low legibility. Such processing means that the subject pays relatively little attention to the physical aspects of the handwritings. This strategy is, however, disadvantageous for performance in the single word recognition task in which no comparable top-down processing is possible. Handwritings with medium legibility will still be read in a predominantly bottom-up fashion, which results in perceptual learning as evidenced by the significant gain-scores.

In Experiment 6, this interpretation is tested by using conceptually difficult texts as practice materials It is reasoned that the use of such materials will limit the amount of top-down processing and will consequently force the reader to pay more attention to aspects of the handwriting itself Using the same design as in Experiment 5 (but excluding handwritings with high legibility), no significant decrease in reading times for consecutive prose-passages is found in the medium and low legibility conditions Significant gain-scores for the pretest/posttest word recognition are obtained for the low legibility condition but not for handwritings with medium legibility

In Experiment 7, conceptually easy and difficult texts are presented in handwritings with high, medium, and low legibility. A significant interaction between the conceptual difficulty of the text and the legibility of the handwriting is obtained Reading times for conceptually difficult texts increase markedly for handwritings with medium and low legibility The interaction provides direct evidence that in reading handwritings with reduced legibility, top-down information is used to overcome limitations on reading speed set by slow bottom-up processes.

Experiment 8, reported in Chapter 5, investigates whether long-term familiarity results in perceptual learning of handwriting. Six subjects are presented with six handwritings, one of which is their own Selected handwritings differ in legibility A significant overall faciliatation is observed for own handwriting and this facilitation tends to increase with reduced legibility Providers of less legible handwritings prove to be generally better readers of these handwritings than providers of handwritings with good legibility

Experiment 8 also studies general perceptual learning of handwriting. Librarians, who frequently read many different handwritings, are included as a separate group of subjects Their extensive experience might make them better readers of handwritings than university students Results, however, show the students to be the faster readers of handwritings They even appear to be significantly better than librarians in their ability to read less legible handwritings

In the general discussion, it is argued that improvements in reading speed, as observed in everyday reading of handwriting, are likely due to more efficient top-down processing made possible by the contents of the reading material Practice effects in reading less legible handwritings are understood as changes in the relative speed of top-down and bottom-up processes As the reader proceeds through the text, increasing amounts of semantic/conceptual knowledge will provide valuable assistance in the decipherment of the handwriting Such optimal use of top-down information is in line with the general goal of the reader to comprehend what has been written and may explain why perceptual learning in the initial reading of a handwriting is very limited. Because the reader is concerned with comprehension, the handwriting itself receives little attention It is pointed out that Experiments 5-8 demonstrate important differences in processing between handwritings differing in legibility. Such differences suggest that the distinction between handwriting and print may only be one manifestation of more fundamental distinctions like legibility.

SAMENVATTING

In deze studie worden twee aspecten van het lezen van handschrift experimenteel onderzocht Een aspect heeft betrekking op de segmentatie van lopend schrift in letters, het tweede op de aanpassing van de lezer aan de kenmerken van individuele handschriften Onderzoek van deze beide aspecten kan licht werpen op kenmerkende verschillen in het lezen van handschrift en het lezen van druk
In het inleidende Hoofdstuk 1 wordt een overzicht gepresenteerd van experimenteel onderzoek naar het lezen van handschrift Uit dit onderzoek blijkt dat het lezen van handschrift in het algemeen moeilijker is dan het lezen van druk en mogelijk een beroep doet op andere herkenningsprocedures.
De bestudering van de twee hierboven genoemde aspecten maakt gebruik van verschillende methodologische benaderingswijzen, die zich onderscheiden in beoogde generalisatie van experimentele bevindingen over handschriften en daarom ook in steekproef-procedures
Een 'diepte' benadering, zoals gebruikt voor de bestudering van segmentatie, is primair gericht op de bestudering van een bepaald verschijnsel dat als kenmerkend voor handschrift mag worden beschouwd, maar dat niet in elk handschrift of in een duidelijk omschreven klasse van handschriften behoeft voor te komen In deze benadering hebben generalisaties dan ook niet zozeer betrekking op handschriften zelf als wel op aspecten van het bestudeerde verschijnsel In een breedte' benadering, die gebruikt wordt voor de bestudering van gewenning, tracht men vast te stellen of bepaalde processen gemeenschappelijk zijn voor alle handschriften of voor bepaalde klassen van handschriften In deze benadering zijn representatieve steekproeven van handschriften gewenst In leesonderzoek zal men vaak dergelijke steekproeven baseren op dimensies zoals uiterlijke kenmerken of leesbaarheid Hoofdstuk 1 eindigt met een bespreking van problemen die zich voordoen bij het samenstellen van op deze dimensies gebaseerde representatieve steekproeven

De segmentatie van lopend schrift wordt onderzocht in vier experimenten, die beschreven worden in de Hoofdstukken 2 en 3 Experiment 1 beoogt aan te tonen dat segmentatie van handschrift moeilijker is dan de segmentatie van druk De herkenning van handgeschreven woorden met onderling verbonden letters wordt vergeleken met de herkenning van handgeschreven woorden met los van elkaar staande letters. Tussen beide condities wordt echter geen verschil gevonden. Een controle-conditie, waarin lijnsegmenten binnen letters zijn verwijderd, resulteert wel in significant langere herkenningstijden. Deze resultaten wijzen naar contouren als alternatief voor inter-letter spaties in de segmentatie van schrift.

Experiment 2 onderzoekt dan ook de vorming van lettercontouren binnen de hele configuratie van kenmerken (verschillende soorten lijnsegmenten). Gesteld wordt dat deze contouren het gevolg zijn van de werking van Gestalt principes zoals gesloten grenzen, spatiële nabijheid en gelijkenis van lijnsegmenten. De werking van het spatiële nabijheid-principe wordt bestudeerd door het onregelmatig maken van de onderlinge relaties tussen letterbreedten en inter-letter afstanden onder gelijkhouding van de totale lengte van het geschreven woord. Hierdoor ontstaan configuraties van kenmerken waarin het principe van spatiële nabijheid geschonden wordt omdat de afstand tussen kenmerken van aangrenzende letters soms veel korter is dan de afstand tussen kenmerken die tot dezelfde letter behoren. De gelijkenis van kenmerken wordt gemanipuleerd door het al dan niet naast elkaar plaatsen van letters met dezelfde kenmerken (rechte, verticale lijnsegmenten). Experiment 2 maakt gebruik van een *n*-detectie taak. Stimulus materiaal bestaat uit illegale pseudo-woorden. De resultaten geven aan dat zowel spatiële nabijheid als gelijkenis van aangrenzende kenmerken een effect hebben op segmentatie. Experimentele condities waarin spatiële relaties tussen kenmerken niet overeenkomen met het principe van spatiële nabijheid voor het toekennen van lettergrenzen resulteren in langere herkenningstijden. Het effect van gelijkenis van kenmerken hangt af van spatiële nabijheid condities. In combinatie kunnen spatiële nabijheid en gelijkenis van kenmerken leiden tot configuraties die ten onrechte letterrepresentaties activeren. Bepaalde aspecten van de resultaten van Experiment 2 geven aan dat de snelheid van de letterherkenning beinvloed wordt door de mate waarin de letter een standaardvorm benadert.

In Experiment 3 en 4 wordt onderzocht of deze distinctiviteit van letters ook bijdraagt tot de segmentatie. Het belang van duidelijke letters voor segmentatie wordt (negatief) gedemonstreerd door 'grafoniemen': handgeschreven woorden of segmenten die qua segmentatie ambigu zijn. Grafoniemen kunnen worden omschreven als contouren die interpreteerbaar zijn als verschillende letters. Deze letters zijn op zichzelf onduidelijk. Grafoniemen zullen in het algemeen moeilijker te segmenteren zijn vanwege een grotere kans dat kenmerken verkeerd gecombineerd worden.

In Experiment 3 worden grafonieme segmenten gecreëerd door een *i* zonder punt naast combinaties van *m*, *n* en *u* te plaatsen. Het effect van de punt op de *i* voor deze grafonieme stimuli wordt vergeleken met het effect voor niet-grafonieme stimuli, waarin de *i* naast letters staat, die een duidelijk andere contour hebben dan de *i* (zoals uitstekende lijnsegmenten of ronde contouren). Om te bepalen of een betekenisvolle context het segmentatie-proces beïnvloedt, wordt in Experiment 3 gebruik gemaakt van een lexicale decisie-taak waarbij zowel woorden als niet-woorden worden aangeboden. Als betekenisvolle context de segmentatie vergemakkelijkt, zullen grafonieme niet-woorden het moeilijkst te identificeren zijn.

Hoewel grafonieme stimuli resulteren in langere reactie-tijden dan niet-grafonieme stimuli en ook stimuli zonder punt langzamer worden herkend dan stimuli met een punt, laat Experiment 3 niet het verwachte grotere effect van het 'punten' van de *i* voor de grafonieme stimuli zien. De punt op de *i* blijkt zelfs een groter effect te hebben voor de niet-grafonieme stimuli. Wel wordt een interactie gevonden tussen *grafonymie en de lexicale status van een letter-sekwentie: grafonieme* niet-woorden resulteren in significant langere reactie-tijden dan niet-grafonieme niet-woorden.

Experiment 4 is een replicatie van Experiment 3, waarin de segmenteerbaarheid van configuraties op nog een andere manier gemanipuleerd wordt. Stimuli zijn ofwel gesegmenteerd ofwel verbonden (cf. Experiment 1). In gesegmenteerde vormen zijn de verbindende lijnsegmenten tussen de letters verwijderd, terwijl deze lijnsegmenten wel aanwezig zijn in de verbonden stimuli. Omdat verbonden stimuli moeilijker te segmenteren zullen zijn dan gesegmenteerde stimuli, wordt verwacht dat het effect van de punt op de *i* groter zal zijn voor de verbonden stimuli.

Uit Experiment 4 blijkt inderdaad dat verbonden stimuli langzamer worden herkend dan gesegmenteerde stimuli. Ook is het effect van de punt op de *i* groter voor de verbonden stimuli, wat erop wijst dat de punt in deze stimuli ook een bijdrage levert aan de segmentatie.

In de algemene discussie van de resultaten van Experimenten 1-4 worden drie factoren aangewezen die van invloed zijn op de segmentatie van lopend schrift. Een eerste factor heeft betrekking op de configuratie van kenmerken in de stimulus als geheel. Configurationele aspecten, die te maken hebben met gesloten grenzen, spatiële nabijheid en gelijkenis van kenmerken zijn in het bijzonder van belang voor het aanbrengen van lettergrenzen Een tweede factor is de configuratie van kenmerken op letterniveau Duidelijke letters vergemakkelijken niet alleen de letterherkenning zelf maar ook de segmentatie. Gesteld wordt dat de interactie tussen de distinctiviteit van de letter en zijn configurationele context, die gevonden wordt in Experiment 4, erop wijst dat letterherkenning en segmentatie tot eenzelfde stadium in de verwerking behoren Segmentatie kan daarom ook onmiddellijk zijn, d.w.z gebaseerd zijn op specifieke configuraties van kenmerken die direct letterrepresentaties activeren Een derde factor is de betekenisvolle linguistische context De bijdrage van deze factor zal vooral bestaan in het induceren van het gissen van bepaalde letters voor configuraties die moeilijk te segmenteren zijn

In de hoofdstukken 4 en 5 wordt verslag gedaan van vier experimenten, die de alledaagse ervaring van gewenning aan handschrift nader bestuderen Deze gewenning resulteert vaak in een verhoging van de leessnelheid Twee interpretaties van dergelijke oefeneffecten worden in overweging genomen Oefeneffecten kunnen het gevolg zijn van perceptueel leren, wat kan bestaan in het discrimineren van bepaalde distinctieve kenmerken of in de vorming van hogere-orde perceptuele eenheden. Een verbetering van de leessnelheid kan echter ook een indicatie zijn van meer efficiente top-down verwerking, die mogelijk wordt gemaakt door de inhoud van het leesmateriaal of die bestaat in meer optimale toepassing van algemene top-down leesstrategieen Voor de Experimenten 5-7, die beschreven worden in Hoofdstuk 4, worden handschriften geselecteerd die verschillen in leesbaarheid Verwacht wordt dat oefeneffecten alleen zullen optreden voor de minder leesbare handschriften Handschriften met goede leesbaarheid zullen veel overeenkomst vertonen met druk waarvoor perceptueel leren al heeft plaatsgevonden.

In Experiment 5 worden zeven niet met elkaar samenhangende prozapassages aangeboden als oefenmateriaal, voorafgegaan en gevolgd door een pretest/posttest die individuele woordherkenning inhoudt (woordbenoeming) Winst-scores voor deze pretest/posttest worden gebruikt als een (context-vrije) maat voor perceptueel leren In het experiment worden negen verschillende handschriften gebruikt die gelijkelijk verdeeld zijn over een hoge, middelmatige en een lage leesbaarheidsconditie Tevens wordt een controle-conditie met getypt materiaal gebruikt

Voor handschriften met een goede leesbaarheid wordt geen enkel oefeneffect gevonden Bovendien blijken leestijden in deze conditie niet significant te verschillen van die in de controle-conditie met getypt materiaal Hoewel de handschriften met slechte of middelmatige leesbaarheid duidelijk minder goed leesbaar zijn dan druk, worden voor deze handschriften geen consistente oefeneffecten gevonden Voor de handschriften met een slechte leesbaarheid worden afnemende leestijden voor het prozamateriaal niet weerspiegeld in significante winst-scores voor de pretest/posttest woordherkenning Voor de handschriften met een middelmatige leesbaarheid zijn de resultaten tegengesteld ondanks het feit dat in de leestijden voor het prozamateriaal geen afname optreedt, worden voor de pretest/posttest winst-scores gevonden, die wel significant van nul verschillen

Een verklaring van deze resultaten doet een beroep op de conceptuele moeilijkheidsgraad van het leesmateriaal Omdat de gebruikte prozapassages gemakkelijk te begrijpen zijn, leiden zij tot een uitgebreide top-down verwerking in het lezen van handschriften met slechte leesbaarheid Deze strategie is echter onvoordelig voor de pretest/posttest waarin geen vergelijkbare top-down verwerking mogelijk is.

In Experiment 6 wordt deze interpretatie getoetst door gebruikmaking van conceptueel moeilijke teksten als oefenmateriaal Gesteld wordt dat het gebruik van dergelijk materiaal het gebruik van top-down verwerking zal beperken en de lezer zal dwingen meer aandacht te besteden aan het handschrift zelf Met gebruikmaking van dezelfde experimentele opzet als in Experiment 5 (maar met uitsluiting van handschriften met goede leesbaarheid), wordt in Experiment 6 geen significante afname gevonden in de leestijden voor opeenvolgende prozapassages in de middelmatige en slechte leesbaarheidscondities Significante winst-scores voor de pretest/posttest woordherkenning worden wel verkregen voor de slechte leesbaarheidsconditie, maar niet voor handschriften met middelmatige leesbaarheid

In Experiment 7 worden conceptueel makkelijke en moeilijke teksten aangeboden, geschreven in handschriften met een goede, middelmatige en slechte leesbaarheid Tussen de conceptuele moeilijkheidsgraad van de tekst en de leesbaarheid van het handschrift wordt een interactie gevonden De leestijden voor conceptueel moeilijke teksten zijn aanzienlijk hoger voor handschriften met slechte leesbaarheid De interactie laat zien dat in het lezen van moeilijke handschriften top-down informatie gebruikt wordt om de lage leessnelheid, die veroorzaakt wordt door trage bottom-up processen, te vergroten

Experiment 8, dat in Hoofdstuk 5 beschreven wordt, onderzoekt of langdurige bekendheid met een handschrift leidt tot perceptueel leren van dat handschrift. Zes proefpersonen krijgen zes handschriften te lezen, waarvan een hun eigen handschrift is. De handschriften verschillen in leesbaarheid. Een significante facilitatie wordt gevonden voor het eigen handschrift en deze facilitatie neemt toe met afnemende leesbaarheid. Schrijvers van moeilijke handschriften blijken in het algemeen betere lezers van slechte handschriften te zijn dan schrijvers van handschriften met goede leesbaarheid.
Experiment 8 bestudeert ook algemeen perceptueel leren voor handschrift. Daartoe zijn bibliothecaressen als aparte groep proefpersonen in het experiment opgenomen. Hun uitgebreide ervaring in het lezen van uiteenlopende handschriften zou de bibliothecaressen tot betere handschriftlezers kunnen maken dan studenten. De resultaten laten echter zien dat de studenten de betere handschriftlezers zijn. Zij blijken zelfs aanzienlijk beter te zijn in het lezen van moeilijke handschriften.
In de algemene discussie van Experimenten 5-8 wordt gesteld dat verbetering van de leessnelheid, zoals die optreedt in het alledaags lezen van handschrift, te maken heeft met een meer efficiënte top-down verwerking die mogelijk wordt gemaakt door de inhoud van leesmateriaal. Oefeneffecten in het lezen van slechte handschriften zijn veranderingen in de relatieve snelheid van top-down en bottom-up processen. Naarmate de lezer in de tekst vordert zal toenemende semantisch/conceptuele kennis van de tekst een steeds grotere bijdrage gaan leveren aan de moeizame ontcijfering van het schrift. Een dergelijk optimaal gebruik van top-down informatie is in overeenstemming met de algemene doelstelling van de lezer te begrijpen wat geschreven is en kan ook verklaren waarom perceptueel leren bij het aanvankelijk lezen van een handschrift beperkt is. Omdat de lezer gericht is op het begrijpen van de tekst krijgt het handschrift weinig aandacht. De experimenten 5-8 laten belangrijke verschillen zien in de verwerking van handschriften die verschillen in leesbaarheid. Dergelijke verschillen wijzen erop dat het onderscheid tussen handschrift en druk slechts één manifestatie is van een meer fundamentele dimensie leesbaarheid.

CURRICULUM VITAE

Hendricus Johannes van Jaarsveld werd op 28 oktober 1947 geboren te Boskoop. Hij doorliep zijn middelbare schoolopleiding aan het St. Antonius College te Gouda en behaalde in 1965 het diploma HBS-A. Hij studeerde psychologie aan de Rijksuniversiteit te Leiden Na het kandidaatsexamen in 1968 legde hij het doctoraalexamen af in 1973 met als specialisatie funktieleer. Van september 1974 tot december 1975 was hij als onderzoekassistent van Prof Dr. C. E Osgood werkzaam aan de University of Illinois in Urbana-Champaign, Illinois Sinds april 1976 is hij als wetenschappelijk medewerker verbonden aan de vakgroep Algemene Taalwetenschap van de Katholieke Universiteit te Nijmegen. Hij verzorgt daar onderwijs in de psycholinguistiek en verricht onderzoek naar leesprocessen.

STELLINGEN

1. Variaties in het leesproces hangen niet zozeer samen met de tegenstelling handschrift/druk als wel met gradaties in leesbaarheid. (dit proefschrift)

2. Segmentatie van lopend schrift in letters kan zowel worden bepaald door globale als door locale configuraties van kenmerken. (dit proefschrift)

3. Verbeteringen in de leessnelheid die optreden in het aanvankelijk lezen van een handschrift, komen voort uit veranderingen in de relatieve snelheid van stimulus-gestuurde en kennis-gestuurde processen. (dit proefschrift)

4. De bepaling van de uitspreekbaarheid van letterreeksen op basis van fonotactische regels gaat voorbij aan het onderscheid tussen fonologische en articulatorische representatie.

   (Henderson, L. *Orthography and Word Recognition in Reading.* London: Academic Press, 1982)

5. Onderzoek naar categorische grafeemperceptie in woordcontext kan licht werpen op de temporele interacties tussen letter- en woordperceptie.

   (Yasuhara, M. & Kuklinski, T.T. Category boundary effect for grapheme perception. *Perception & Psychophysics*, 1978, Vol. 23 (2), 97-104)

6. Het schrijfonderwijs zou gediend zijn met onderzoek naar factoren die de leesbaarheid van handschriften bepalen.

7. De meeste inleidingen in de psycholinguistiek houden onvoldoende rekening met het feit dat linguistiekstudenten vaak weinig vertrouwd zijn met experimentele methoden.

8. Bij de aanstelling van tijdelijk wetenschappelijk medewerkers dient rekening te worden gehouden met de begeleidingscapaciteiten van wetenschappelijke staf in vaste dienst.

9. Het principe van verdelende rechtvaardigheid gebiedt de kwetsbaarheid van militaire commandoposten te laten toenemen met hun belangrijkheid.

10. Men besteedt als kind meer tijd aan het leren schrijven tussen regels dan aan het leren lezen tussen regels.

11. Het geruststellende 'Hij doet niets' van hondenbezitters veronderstelt vaak ten onrechte dat de baas meer vertrouwen wekt dan de hond.

Stellingen behorende bij

H.J. van Jaarsveld, On Reading Handwriting.

Proefschrift, Nijmegen, 1983.