WHY JOHNNY *CAN* READ CHINESE:

WORKING MEMORY, COGNITIVE PROCESSES,

AND READING COMPREHENSION

A DISSERTATION SUBMITTED TO THE

GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

EAST ASIAN LANGUAGES AND LITERATURES (CHINESE)

DECEMBER 2004

By
Jeffrey J. Hayden

Dissertation Committee:

Tao-chung (Ted) Yao, Chairperson
Hsin-I Hsieh
Ying-che Li
David Ashworth
Thom Hudson

© Copyright 2004
by
Jeffrey J. Hayden

DEDICATION

To my wife, Oanh, who has been by my side especially through the most trying times of this dissertation. To my parents, Mary Louise and Donald J. Hayden, for encouraging me to pursue my studies and supporting my decisions, both emotionally and financially.

Also, to the late Stephen Wang, my first professor of Chinese when I was an undergraduate at the University of Minnesota (1987-1990), who apparently revealed the hint of a smile many years later when he heard that I was pursuing a Ph.D. in Chinese.

ACKNOWLEDGMENTS

I owe a great debt of gratitude to the chair of my dissertation committee, Professor Tao-chung (Ted) Yao. I consider myself extremely fortunate to have had the opportunity to work with him as his student, teaching assistant, and advisee. Professor Hsin-I Hsieh is foremost of the people who contributed to my academic growth. Without his guidance, support, and encouragement, I would not have accomplished half as much as I have. Prof. Ying-che Li's contributions throughout my graduate career are simply too numerous and varied to recount, though I am grateful for having received them all. Prof. Ashworth has been a font of ideas and sound advice, and his just being there as a sounding board has been invaluable. Prof. Hudson's expertise in reading research and statistics has ensured that my interpretations of the various forms of data provided herein are as meaningful as possible. Finally, special thanks are in order for Prof. Terry Klafehn for coming through in a pinch and for those thinking-outside-of-the-box experiences.

From the time I arrived to the islands, Evelyn Nakanishi has always been there with the right answer. To the Department of East Asian Languages and Literatures office staff—Liz, Shirley, and Mira, who all retired just before seeing this dissertation completed—I must say that I appreciated your patience in dealing with one ornery *haole* for so many years. I feel I have personally grown in so many different ways by my interactions with you.

Finally, a word of thanks to the students. To my own, who provided me with a glimpse at and reminder of what I must have looked like when I was in their chairs, and also to those who volunteered to participate in all phases of my research, I extend my most heartfelt gratitude for contributions of time, effort, patience, and perseverance. This dissertation would be lacking significant substance without you.

ABSTRACT

For many native English speakers, learning to read Chinese as a second or foreign language is a long and oftentimes laborious process, fraught with seemingly endless and demoralizing frustrations. Unlike the relationships of English to French, German, or Spanish, which share varying but useful amounts of cognates, English and Chinese share no cognates, which means that from the start learners have nothing to rely on to help them navigate their way through seas of Chinese text. Yet, with time, non-native learners of Chinese do come to comprehend novel texts they are faced with on their own. In order to understand the paths non-native readers of Chinese take to comprehend the texts they read, this study examines and attempts to identify and describe effects of difficulty level on cognitive processing based on eye-tracking data. The data from three groups of participants (non-native Intermediate level [n=5], non-native Superior level [n=7], and native Chinese [n=13]) were gathered during a computer-adaptive test. Quantitative (descriptive) and qualititative analyses were conducted on the data.

The study shows that non-native readers of Chinese at an Intermediate level demonstrate different reading strategies than non-native readers of Chinese at a Superior level, who in turn use a similar strategies as native Chinese readers, but at different times during the reading process. There were also indications that there are two distinct types of readers within the non-native Superior level participants, i.e., those who were still more non-native in their approach to texts and those who displayed native-like reading strategies. The study shows that learners of Chinese go through phases of cognitive abilities as their knowledge of vocabulary and structures increases and they expand their reading experiences with a variety of text types. Theoretical, pedagogical, and methodological implications are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AOI	Area Of Interest
ASL	American Science Laboratories
CAT	Computer-Adaptive Test
CATRC	*Computer Adaptive (Proficiency) Test for Reading Chinese*
CFL	Chinese as a Foreign Language
CHS	Committee on Human Studies
CL1	Chinese as a 1st/native Language
CL2	Chinese as a 2nd Language
CLTA	Chinese Language Teachers Association
CSL	Chinese as a Second Language
EFL	English as a Foreign Language
ESL	English as a Second Language
EVF	Effective Visual Field
FL	Foreign Language
GRE	Graduate Record Examination
IPA	International Phonetic Alphabet
IRB	Institutional Review Board
JCLTA	*Journal of the Chinese Language Teachers Association*
L1	1st Language
L2	2nd Language
LTM	Long-Term Memory
LT-WM	Long-Term Working Memory
ms	milliseconds
MSC	Modern Standard Chinese
NNI	Non-Native Intermediate (participant group 1)
NNS	Non-Native Superior (participant group 2)
NS	Native Superior (participant group 3)
RFL	*Reading in a Foreign Language*
SOA	Stimulus Onset Asynchrony
STM	Short-Term Memory
TOEFL	Test Of English as a Foreign Language
UHM	University of Hawai'i at Mānoa
WM	Working Memory
wpm	words per minute

PREFACE

As a non-native reader of Chinese, I have wondered at myself and other students like me who learn to read such a complex script. For languages and orthographies that I have never studied and know nothing about—Arabic, Thai, and Tibetan, for example—I am still blind to their messages and can only admire their surface beauty. So, in a sense, when I look at text in those scripts I still know what it felt like before I learned my first Chinese character. However, to see how I acquired a skill over so many years of study, which I now sometimes take for granted, I decided to look at how students at differing skill levels come to comprehend short reading passages.

This study, then, is concerned with the processing of extended novel text by learners of Chinese as a second or foreign language (CL2/CFL), specifically as stimulated by items of a computer-adaptive test (CAT). The *Computer-Adaptive Test for Reading Chinese* (*CATRC*) was used as a measure of participants' reading levels, defined by the ACTFL (American Council on the Teaching of Foreign Languages) scale, and provided the target reading material used to gather participants' eye tracking data.

The research conducted in this dissertation is interdisciplinary and incorporates theories and practices from a several fields that contribute to Chinese language teaching and Second Language Acquisition (SLA) in general and to teaching and learning to read Chinese in particular. Specifically, I attempt to interweave general psycholinguistic models of reading and memory models from the cognitive sciences to address the issue of cognitive load theory as it applies to learners of Chinese as a second or foreign language.

Results from the eye-tracking research conducted in this study primarily illustrate how non-native learners at the Intermediate and Superior levels visually process different types and genres of texts as compared with native language readers, thus leading to suggestions for more effective teaching methods especially tailored to those language levels. Secondarily, in terms of the cognitive sciences, the research undertaken for this study sheds light onto the issue of mental processing during second language reading, especially as it involves both long-term working memory and sentence processing.

The study of a foreign language, especially one such as Chinese that is written so differently than English, requires commitment, determination, and discipline. It is demanding in terms of time and effort, but the joy of being able to read a Chinese text, whether it be a collection of essays written by a contemporary Chinese author or a Chinese web site, is definitely worth the late nights burning the midnight oil. By identifying the steps that need to be taken to get to a desired level of reading proficiency, a student can come to know that s/he does not need to "know" every single Chinese character before s/he can begin to enjoy the experience of reading Chinese.

CHAPTER 1

INTRODUCTION

In the very first issue of the *Journal of the Chinese Language Teachers Association* (*JCLTA*) in 1966, Professor John DeFrancis published the article "Why Johnny can't read Chinese". Perhaps its prominent placement as the vanguard article in that first issue of the Association almost 40 years ago was a foretelling of how it was to become a seminal and often-cited paper in the field of Chinese language teaching, with particular regard to the teaching of reading. The title of this dissertation pays homage to that article, and while the content herein continues in a direction which may not be the most logical or expected, I believe it still follows one of the paths made available, especially considering the increases in interest and research in the area of cognitive processing of language in the past twenty years.

Objectives

Not until relatively recently has reading comprehension been looked at in terms of cognitive processing of extended text, and still very little research has been done with reading second/foreign languages in this regard. The main purpose of this study is to try to examine the mental/cognitive processes involved during the reading of novel text as displayed on a computer screen. This dissertation then will attempt to look at the role working memory appears to play among readers of Modern Standard Chinese (MSC, or just simply Chinese, hereafter). The methodology for this research involves the use of eye-tracking equipment.

Statement of the Problem

One of the major concerns to theories of reading Chinese and Japanese has been whether Chinese characters are primarily phonologically or semantically encoded in the mental lexicon. This is especially of interest in terms of learning to read Chinese or Japanese as foreign languages, as there are direct pedagogical implications. Generally speaking, researchers agree that Chinese characters are first and foremost phonologically encoded. (Horodeck, 1987; Matsunaga, 1994) Then, through increasing familiarization with Chinese characters (both individually and in lexical compounds), particularly expert readers are able to bypass the phonological step and access semantic information directly. There remains, however, the question of whether these readers are actually directly accessing semantic information of all tokens or whether they are simply able to bypass (either by chunking or simply by skipping over) certain items—particularly function words—without adversely affecting comprehension efficiency.

While the question of the primacy of phonological encoding is important in both Chinese L1 and CL2/CFL research, this study is not concerned with when Chinese characters should be introduced in the curriculum to western students of the Chinese language. Also, while many eye-tracking studies have already been done in relation to the recognition and processing of individual Chinese characters (漢字, Chinese *Hànzì*, Japanese *Kanji*), or characters in compounds, there is still a need for more studies that use extended (or, natural, multi-sentential and paragraph-level) text samples as stimuli. While the Everson (1986) study—probably the first eye-tracking study to look at non-native readers of Chinese almost twenty years ago—used a paragraph as a stimulus and utilized a recall protocol to assess comprehension, his research dealt mainly with word-unit spacing and thus was still arguably more concerned with Hanzi at a sub-sentential level.

Therefore, this study is more concerned with the processing of extended novel text by learners of Chinese as a second or foreign language (CL2/CFL), specifically as stimulated by items of a computer-adaptive test (CAT). During a pilot study in Beijing during the academic year 1997-1998 (Hayden, 1998), I used the *Computer-Adaptive (Proficiency) Test for Reading Chinese* (CATRC) repeatedly as a measure of participants' reading levels over time, defined by the ACTFL (American Council on the Teaching of Foreign Languages) scale. Further experience (Hayden, 2003) with the CATRC, as well as data from two validity studies, suggest that the CATRC is a measure reliable enough to be used in a study such as the one currently under consideration.

The research presented here incorporates theories and practices from a variety of disciplines that could contribute to the fields of Chinese language teaching and Second Language Acquisition (SLA) in general and to teaching and learning to read Chinese in particular.

Results from the eye-tracking research conducted in this study help illustrate how students at the Intermediate to Superior levels visually process different types and genres of texts, which can thus be used to suggest more effective teaching methods especially tailored to those language levels.

Secondarily, in terms of the cognitive sciences, the research undertaken for this study sheds light onto the issue of mental processing during second language reading, especially as it involves both long-term working memory (LT-WM) (Ericsson & Delaney, 1999; Kintsch, 1994; and Kintsch, Patel, & Ericsson, 1999) and sentence processing (Su, 2001).

Assumptions

In the process of this investigation I have held to two assumptions: the first is that the ACTFL reading guidelines are sufficiently valid definitions of foreign language proficiency levels for curricular or assessment purposes. In light of Park's (1999) analyses, where the construct validity of a reading proficiency test based on the ACTFL reading guidelines was generally affirmed, the CATRC was determined to be a sound enough instrument of measurement for the purposes of this study.

The second assumption I hold through this research, intimately related to the first assumption, is that the CATRC is a reliable and valid test of Chinese as a foreign language (CFL) reading proficiency. Rather than assuming that a student has an intermediate level of reading proficiency simply because she or he may be in an intermediate-level language class, I felt it prudent to get a clearer idea by using a test instrument that has been or is currently being used as an entrance/exit assessment tool at several universities (including the East Asian Summer Language Institute at Indiana University, the Chinese School at Middlebury College, Stanford University, and the Department of East Asian Languages and Literatures at the University of Hawai'i).

Hypotheses

> "... observation of learner processing strategies via eye movement ... [provides] learner generated data [that] illustrate microlevel strategies used by individuals during reading." (Bernhardt, 1991, p. 17)

In this research there are two main hypotheses. The first is that L1 reading strategies do not transfer to L2 reading strategies in intermediate level non-native readers of Chinese as a foreign language. (Walter, 2000, pp. 113ff.) That is, even though the non-native participants of this study at this level are college-level students who have highly developed reading skills in their L1 (i.e., English), they are not necessarily able to use these same skill sets when they read in the foreign language. The second hypothesis is that L1 reading strategies do transfer to L2 reading strategies in superior level non-native readers of Chinese as a foreign language, while good native readers of Chinese use an interactive (a combination of bottom-up and top-down) processing strategy during reading.

Research Questions

The following three questions are guiding my research and will be addressed in my dissertation:

1. What does eye movement data tell us about how test takers take computer adaptive tests (CATs)?
2. Why do superior-level takers of CATs take longer to complete a CAT than low-level test takers?
3. Are there identifiable differences between the three groups (Non-Native Intermediate, Non-Native Superior, and Native) in terms of reading proficiency as evidenced during the administration of a CAT?

As mentioned briefly above, the main purpose of this study is to try to examine the mental/cognitive processes invoked during the reading of novel text displayed on a computer screen as evidenced by eye movement. What I will be looking for are indications of lexical and syntactic connections in the comprehension process during reading of extended text by analyzing eye-tracking data of participants taking a CAT of Chinese reading proficiency. Key indicators of mental processing during reading are indicated by so-called eye *fixations* and *saccades*, moments when and locations where the eyes have paused or retrieved visual stimulus, respectively. There seems to be a general consensus among eye movement experts that these moments and locations are the instances when the brain is processing information.

Limitations of the Study

There is one significant limitation to the present line of inquiry in that the population of non-native learners of Chinese is much lower than that of more commonly taught languages such as French, German, and Spanish. In particular, Superior-level non-native readers of Chinese are extremely rare. A cursory examination of students currently enrolled in First- through Fourth-year Chinese language classes shows that by the fourth year of instruction, most students still do not achieve an Advanced level of reading proficiency. This affected the number of possible participants available for the Superior level, seriously limiting the quantity of data available for comparison and analysis, as well as the generalizability of the analyses.

Overview of the Following Chapters

The remainder of this thesis is organized as follows: Chapter 2 explores the reading process, including looking at established models and guiding the reader through the process of a non-native learner of Chinese as a foreign language (CFL) learning to read Chinese. Chapter 3 discusses the role of working memory in reading and previous eye-tracking research with a concentration on reading Chinese. Chapter 4 presents the methodology for the present research, including descriptions of the participants, the materials used, the procedures conducted during the research, and descriptive analyses of the data collected. Chapter 5 shares results of the research and analyses. In Chapter 6, I discuss the results with a view towards clarifying the analyses and putting them in a foreign language learning context. Finally, in Chapter 7, I provide my conclusions based on the results and discussion, and look at some implications of the research in terms of reading research and Chinese language pedagogy.

CHAPTER 2

READING IN A FOREIGN LANGUAGE: THE READING PROCESS

Introduction

The process of reading, at one time thought to be a relatively simple act, has been shown to be quite complex. Early first language reading research focussed on the processes involved in reading alphabetic languages. Large-scale, in-depth research on the processing of a morphosyllabic[1] orthography such as Chinese has, to date, mainly been concerned with character recognition and phonological-syntactic sequencing, either individually or in character combinations (usually pairs) by native speakers/readers. Similarly, research in Japanese *Kanji* has, for the most part, also been concerned with investigating whether access is initially phonological or semantic (or both simultaneously) in nature, as it appears to be for alphabetic languages.

Reading Models

In this section I will very briefly discuss models of reading. There are three generally agreed upon reading models that all readers use in varying degrees and at different times—both in overall development of reading skills and during reading sessions (Kolers, 1970;

---

[1] The term "logograph(ic)" in reference to 漢字 (*Hànzì*, Chinese characters) is actually a point of contention with DeFrancis (1984), who argues that they should be described as "morphosyllabic" (p. 125-126, 187). DeFrancis (1989) explains that there are two reasons for this: 1) individual Chinese characters are at the same time a single syllable and a single morpheme, and 2) the majority of Chinese characters are composed of two elements, i.e., a semantic component and a phonetic component (p. 223).

Samuels & Kamil, 1984). These models are bottom-up, top-down, and interactive-compensatory. A reading process that begins with identification of sublexical features such as phonological cues as represented by letters is indicative of the so-called bottom-up process. Another separate process is the top-down process. In this process a reader approaches a text with expectations of using certain background knowledge to gain an understanding of the material. Both of these processes are said to be serial and linear in nature in that they tend to proceed in one direction only without interaction with the other. In addition, there is still some debate as to which process is called on first when faced with novel text: do readers start from trying to figure out the orthography (i.e., bottom-up), or do they go into a reading task with an idea already in mind (i.e., top-down)?

Further research into the reading process has persuasively argued that bottom-up and top-down processes are actually sub-components of a larger interactive process. An interactive reading process essentially calls on bottom-up and top-down strategies interchangeably and in parallel as needed during the reading process in order to achieve comprehension. This interactivity between sub-components has led to the suggestion that reading processes of native readers of different languages are actually so similar as to be universal.

Bottom-Up Process

Bottom-up processing is a linear way of piecing together the most basic building blocks of a written language in order to produce meaning. One major bottom-up model is Gough's (1972) information-processing model. He explained that in one second of reading the eyes begin from fixations on the symbols that make up the text and proceed through a decoding process in order to arrive at comprehension. In this model, "the reader is not a guesser" (Gough, 1972, p. 354), but rather pieces together letters to make words, words to make phrases, phrases to make meaning. Gough took us on a very detailed tour of one second of reading (English) through the eyes of a "moderately skilled adult reader" (p. 331) that included the initial fixation, recognition of a "precategorical" (p. 332) but patterned visual image, letter identification, phonological assignment, lexical search and mapping, comprehension, and vocalization.[2] This all happens, according to Gough, to every word that is read in a serial fashion.

Bottom-up processing is generally synonymous with lower-level processing, also referred to as data-driven and text-based. According to research done on reading processes, students beginning to learn to read in a foreign or second language such as Chinese or Japanese and many children learning to read in a native language such as English go through a similar initial process of focussing on individual letter (character) and word identification in order to compose meaning (Samuels & Kamil, 1984). For the native speaker of English learning to read, this means first identifying and decoding letters into their correspondent sounds In terms of reading Chinese characters, since Chinese is a morphosyllabic language and not an alphabetic language, this means that the basic pieces in need of decoding are strokes, radical components and phonological cues.

In the learning of languages such as Chinese and Japanese, instruction in the two orthographies (morphosyllabic and syllabary, respectively) often introduces cognitive hurdles not characteristic of alphabetic writing systems that keep learners dependent on this level of processing longer. To take a recent example using Japanese, Koda's (1992) study looked at

---

[2] In the same article, even before the sketch began, Gough admitted that his description was going to be "speculative" and "almost certainly flawed" (Gough, 1972, p. 331); but, as he hoped, those concessions turned out to be more virtues in light of the extended discussion it raised.

effects of lower-level processing skills of JFL (Japanese as a foreign language) students' reading performance. She gave 58 students a battery of tests that included a cloze test (120 words/20 deletions), a paragraph comprehension test (four paragraphs), and a multiple-choice sentence comprehension test (10 sets of 2 sentences). In her results, she found that faced with the task of deciphering the mixed orthography of Japanese (consisting of both two related but different syllabaries and Kanji, most of which have polysyllabic readings) her students tended to rely on the Kanji to provide meaning. Koda suggests that while the syllabary may provide direct symbol to sound mapping, stringing them together for meaning is more difficult. Kanji, on the other hand, seemed to provide the JFL students with a more direct route to the meanings they needed to comprehend the text.

Top-Down Process

About five years before Gough's (1972) explication of bottom-up processing, Goodman (1967, 1968) posited that reading was more of a top-down approach, "a rapid series of guesses, tentative information processing" (Goodman, 1968, p. 19). The top-down process, like the bottom-up process, is linear in nature, but rather than primarily combining smaller pieces of information to get to larger meanings as the bottom-up process does, the top-down process starts with knowledge of the language and ideas and experiences already existing in the mind of the reader. Ideas and concepts are then added to, modified, and built upon with the introduction of the new information. This is not to say, of course, that lower-level information such as sound-symbol correspondences is not utilized in this approach, only that it is secondary and resorted to only when necessary.

Top-down processing is generally synonymous with higher-level processing, also referred to as "conceptually-driven" (Samuels & Kamil, 1984, p. 212) and "knowledge-based" (Bernhardt, 1991, p. 72). As reading vocabulary increases and reading habits (behaviors, strategies) become more developed and fluent, readers are able to get meaning from connected text (context), even if they do not know every word. That is, the more background knowledge one has, the less information needs to be decoded from the text. This is a somewhat tougher obstacle to get past in learning to read a language like Chinese where learners do not have the phonetic transparency of an alphabetic language.

Studies utilizing miscue analysis, pioneered by Goodman, typically look at reading as a top-down process. Sergent (1990) conducted a study with six advanced and six highly advanced CFL learners orally reading a natural (extended) text and found that Goodman's miscue analysis and model of reading was not "adequate for predicting the reading behaviors of the readers." (p. 772) More recently, Li (1998) conducted a study using miscue analysis data from six beginning CFL learners to determine the nature of their reading acquisition. Li's qualitative data primarily included classroom observation, questionnaires, participant interviews, and read-aloud/think-aloud protocols, and secondarily included instructor interviews, participant discussions, test results and homework assignments, and exit notes. The miscue analysis was done on videotaped oral reading activities in the classroom. Based on this data, Li concludes that "As a result of the orthographic and linguistic differences between English and Chinese, when English speakers learn Chinese, the contextual cues and the word order are relied upon more than other cueing systems" (Li, 1998, p. 172).

Interactive-Compensatory Model

A third type of reading model, the interactive model, has bottom-up and top-down processing alternating depending on the task and difficulty level. Both LaBerge and Samuels' (1974) Model of Human Information Processing and Stanovich's (1980) Interactive-Compensatory Model are examples of interactive models. The key to LaBerge and Samuels' model is attention (including attention capacity and attention switching) and automaticity (Samuels & Kamil, 1984), while Stanovich's essentially says that a reader will compensate for weak subskills by utilizing whichever approach is stronger and will succeed in accomplishing the ultimate goal, i.e., comprehension. Put simply, every so-called "emergent reader" possesses strengths and weaknesses in their reading skills. For example, a beginning reader with good decoding (text-based) abilities will tend to rely on word-level processing, while a beginning reader with a good grasp of the mechanics of a language (i.e., syntax and semantics) will tend to rely on context and background knowledge to achieve comprehension.

Here we see the difference between good readers and poor readers (or, as Clarke, 1979, suggests, the difference between good and poor reading behaviors): poor (inefficient) readers almost exclusively use either bottom-up or top-down processes when they read, while good (efficient) readers generally have bottom-up and top-down interacting proportionally as necessary in order to reach that "click of understanding" (Barnett, 1989, p. 53; Samuels & Kamil, 1984, p. 185).

Clarke (1979) is generally credited with developing what is known as the "short circuit" hypothesis.[3] This hypothesis posits that presented with a text that demands more processing resources than are available, either at the lower word recognition level or at the higher concept integration level, a critical bottleneck will be created and the system will fail or break down.

Hudson (1982) identifies three components of reading breakdown shared by both bottom-up, lower-level approaches and top-down, higher-level processes. These components are (a) basal elements (i.e., letter and word recognition, sound-symbol correspondence, etc.), (b) hypothesis production and testing (including guessing and meaning identification, categorization and confirmation/ rejection of new information based on prior knowledge, etc.), and (c) affective features (such as motivation, response, environment, etc.). Hudson explains that generally only factors involving the first two components seem to cause short circuits in comprehension. That is, the quality or quantity of one's knowledge of a language defines a "language/linguistic ceiling" (pp. 3, 20) and the level of one's language proficiency creates a "proficiency ceiling" (p. 4) that can effectively restrict comprehension and increase the likelihood of a breakdown in the reading process.

Thus, an interactive-compensatory model provides for the ability of a reader to be flexible in their making use of either lower order or higher order processing depending on the difficulty of the task in an effort to avoid comprehension failure. For most beginning English-

[3] Actually, Goodman (1968) had used the term "short circuit" more than a decade before Clarke (1979), albeit in a slightly different context. In his discussion on distinguishing decoding and recoding and teaching beginning readers, Goodman suggests that "Preoccupation with teaching children to recode may actually short circuit the reading process and divert children from comprehension." (Goodman, 1968, p. 20)

speaking learners of Chinese as a foreign language this means alternating piecing together strokes, radical components, and phonetic elements in order to decode less familiar and unfamiliar characters with getting larger meaning from extended and connected text. The key with language learners, especially of a foreign language like Chinese, seems to be that the larger one's vocabulary and the greater and more developed one's character/word recognition skills—i.e., the more automatic the decoding skill becomes—the more ably they can gain the flexibility to incorporate a more interactive (i.e., efficient) approach to reading.

Reading as a Universal Process

The above models of reading are generalizable and may be applicable to the reading of any language. Based on their research, for example, both Goodman (1968) and Smith (1988, 1994) have suggested that the processes involved in reading are universal across languages, regardless of orthography. Recently the Goodman Psycholinguistic Model was shown to be applicable to Arabic (Al-Fahid, 2000), an alphabetic language that predominantly spells with consonants only and which is written right to left (Abu-Rabia, 1997; Khaldieh, 2001).

Most studies looking at whether reading is a universal process for all languages mainly investigate the role of phonological encoding and whether it is pre- or post-lexical access. This is especially important when looking at two languages as seemingly different as English and Chinese or Japanese. In fact, Zhang and Perfetti (1993), based on their experiments eliciting phonological activation in a preloaded memory paradigm, suggest that "once the reader accesses the lexicon, the processes of reading English and Chinese have more in common than any differences that are based on the writing systems" (p. 1092).

For many years, Chinese characters have been assumed to map directly to semantic meaning or syntactic function. Work by Cheng (1992), Horodeck (1987), Hue (1992), Matsunaga (1994), and Tzeng, Hung, and Garro (1978) have provided significant contributions to the debate of whether Chinese characters (Kanji) necessarily undergo phonological activation, and if so, when in the reading process it happens. Indeed, Zhang and Perfetti (1993) conclude that "phonological processes may be universal, serving text comprehension in both alphabetic and nonalphabetic systems" (p. 1092). Perfetti (2003) has even gone so far as to suggest that there is a "Universal Grammar of Reading", proposing that "Across writing systems, orthographies distribute their divergence from phonology in a way that minimizes the pain to the reader" (p. 10).

Recently, Matsunaga (1999), using the same data she gathered for her dissertation (Matsunaga, 1994, covered in more detail below), offered evidence that contributed to a universal hypothesis of the reading process. In her eye-tracking study of native speakers of Japanese reading natural texts for comprehension, she looked at the difference in effect of homophonic[4] and nonhomophonic[5] Kanji errors in texts. She discovered that her participants had a greater number of fixations and longer fixations on errors in the nonhomophonic stimuli, indicating phonological activation. She concludes by saying that the addition of her data to that of previous research

> clearly indicates that reading processes of the two types of scripts, alphabetic and nonalphabetic, are not as different as previously thought (Tzeng & Hung, 1995); speech mediation appears to be necessary most of the time, except when

---

[4] dissimilar appearance but same sound, e.g., 交換 (*kōkan*) and 交喊 (*kōkan*).

[5] similar appearance but dissimilar sound, e.g., 連続 (*renzoku*) and 連読 (*rendoku*).

identifying frequent English words (Seidenberg, 1985) or frequent Chinese characters with concrete meanings (Tan, et al., 1995) in isolation. (Matsunaga, 1999, p. 305)

It appears, then, that speech mediation is just one of the shared universal reading processes, in native as well as in a second or foreign language. There are also differences in the reading processes among native readers of Chinese and native readers of English (Jackson, Lu, & Ju, 1994). Of course, there is more to determining if reading processes are universal than just by looking at when and under what conditions phonological activation takes place, for example. However, considering that such activation—if it is activated at all—takes place very early on, studies such as these are critical steps in deepening our understanding of a very complex activity. Obviously, frequency of and familiarity with information at the character and word levels are also crucial factors in determining how generalizable a proposed universal model of reading is. The model should include operations for both native reading and second/foreign language reading as well.

Foreign Language Reading

While research seems to have shown that the essential processes of reading in a foreign language are similar to if not the same as reading in one's native language, especially if the orthographies are similar (Krashen, 1993), it still remains the case that there is something different about the experience of someone skilled in reading an alphabetic language beginning to read in a logographic (morphosyllabic) language. If we look at what language these studies are mainly investigating, we realize that they have been studying English as an L1 and English as the L2. Bernhardt (2001) forcefully concludes that "the Anglophilial nature of second-language research underlines the 'sameness'. That is, L1 reading research is almost exclusively 'English'; so is L2; therefore, the processes must be 'the same.'"

As Bernhardt (1991, pp. 22-25) points out in detail, "real" research into second language reading has not received a great amount of actual research in the past 30 years, certainly not as much as has been done on ESL reading. Bernhardt's (1983) early study on foreign language reading looked at three groups of readers: non-native inexperienced, non-native experienced, and native. Using what has become known as the Bernhardt Constructivist Model of Second Language Reading, Bernhardt (1983) used think-aloud protocols to assess reading comprehension while at the same time providing deep insight into what goes on in the mind of a foreign language reader.

A question that has yet to be answered is when is a foreign language reader a bilingual reader? Models of L2/FL reading processes by definition should include that category of language user and that part of their development. Most recently, Marian (2000) utilized advanced technological methods to investigate language processing in bilinguals. Several researchers have argued that there is a direct relation to a reader's reading skills, language and reading proficiency, and task demands (Bernhardt, 1991; Bernhardt & Kamil, 1995; Carrell, 1991). Research specifically related to text structure in relation to comprehension has also supported this aspect of foreign language reading development (Chun & Plass, 1997; Mecartty, 1998; Yano, Long, & Ross, 1994).

The above discussion on bottom-up, top-down, and interactive processes, as well as the possible extent of the universality of reading processes across first and second/foreign languages was presented as a precursor to a more focused look at reading in Chinese as a foreign language.

The following pages will examine what difficulties a non-native reader of Chinese faces when attempting to construct meaning from a given text.

Reading Chinese as a Second/Foreign Language

The Burden of Orthography

One of the main reasons often cited for causing so much difficulty among CL2/CFL students learning to read Chinese are the characters. In fact, five of Moser's (1991) nine reasons why Chinese is so hard deal with the characters:

1. Because the writing system is ridiculous.
2. Because the language doesn't have the common sense to use an alphabet.
3. Because the writing system just ain't very phonetic.
4. Because you can't cheat by using cognates.
5. Because even looking up a word in the dictionary is complicated.

(Moser, 1991, pp. 60-65[6])

There are generally three main concerns when discussing the instruction of Chinese characters: retention, recognition, and production. Retention deals with the "tricks", often mnemonic, that students use to keep the characters in memory. Recognition implies that students have managed to successfully embed a character or word into more long-term storage to the extent that at the least they will be able to provide a pronunciation or a meaning associated with it. Finally, production means that not only have students retained the character or word and can recognize it, but that they can recall and more-or-less faithfully reproduce a character in the right situation.

[6] The other four are: classical Chinese, romanization methods, tonal language, and isolated culture (Moser, 1991).

Similarly, for the concern of recognition (and, hence, comprehension), there are generally three areas researchers focus on when studying the effect of the orthography on the reading process: individual characters (including complexity, density, and frequency), character compounds (i.e., words, including word length, complexity, and frequency), and sentences. Much of this hinges on the concept of orthographic depth, which says that more attention will be required to process text that is further removed from phonological cues.

Chinese has traditionally been viewed as more of a meaning-based system because of the relatively irregular and imprecise orthographic pronunciation clues only sometimes available. DeFrancis (1984, 1989) and others (e.g., Hannas, 1997; Taylor & Taylor, 1995; Unger, 2004; W. S-Y. Wang, 1973) have written at length on the nature of Chinese orthography, and a detailed recounting would be beyond the scope of this dissertation. More to the point, however, are studies that have looked at character recognition and word recognition as performed by native and non-native readers.

First, as more related to the beginning study of Chinese as a foreign language, I will briefly digress and look at three studies that compare the reading of characters to the reading of an alphabet used to romanize the pronunciations of the characters. These three studies (Everson, 1988; Light, 1976; Sun, 1993) are similar in that they compare reading times of passages presented in both Chinese characters and Pinyin romanization.

Pinyin romanization is a tool used in the teaching of the (Mandarin) Chinese spoken language. It uses roman letters to spell pronunciations of Chinese characters.[7] Aside from a few "irregularities"[8], most of the sounds are similar to those as represented in the western alphabet. A few studies have looked how the reading of characters compares with the reading of Pinyin romanization. These studies need to be considered cautiously, as no one really reads Pinyin as a native. It is used in the instruction of (mainland) Chinese children up through about the 6th grade, and CL2/CFL students are almost invariably trained to use it from the first day of instruction (at least in the majority of US programs this seems to be the case).

Light (1976) is perhaps one of the first studies that looked at non-natives reading this romanization scheme. Starting from change he noticed in his own reading behavior when he was a student that he reached a point where he felt he could read Chinese characters faster than romanization, he conducted an informal academic year-end timed reading test and gathered data from 22 of his first-year students. Using two passages, each prepared in a character and Pinyin format, Light divided his students into two groups, where Group I read

---

[7] For example, 王 **wáng** ('king') and 美國 **Měiguó** ('America'). In this system, tones are indicated by diacritics iconically representative of their tonal contours placed over the nuclear vowel, a different diacritic for each of the four tones. In western terms, these are a macron (ˉ; e.g., **tāng** 'soup') for the first tone, an acute (ˊ) for the second tone (as in **wáng** and **guó**, above), a hačej/caron (ˇ) for the third tone (as in **Měi**, above), and a grave (ˋ ; e.g., **dì** 'earth') for the fourth tone. Also in this system, words are generally spelled together, where 'America' is **Měiguó**, not **Měi guó**, which reflects the non-monosyllabic nature of the modern language.

[8] These typically are the three letters standing for the palatals, **j**, **q**, **x** [in IPA: tç, tç^h, and ç, respectively]. However, considering these sounds do not exist in English, it seems only logical that their representations be less transparent.

Passage 1A (Pinyin/79 syllables) then Passage 2D (104 characters), and Group II read passage Passage 1B (79 characters) and Passage 2C (Pinyin/104 syllables). The Pinyin text for Passage 1 is provided below to illustrate its use in extended text.

> Wǒ cóngqián zhù zài Měiguó. Nèige shíhou wó de fùmǔ yě shì zài Měiguó. Kěshì tāmen qùnián qù Zhōngguó zhù le. Wǒ méi gēn tāmen yíkuàir qù yīnwèi tāmen yào líkāi Měiguó de shíhou, wǒ rènshi-le yíwèi hěn hǎokànde Fǎguó nǚrén. Wǒ yí kàn-le tā, jiù xiǎng gēn tā jiéhūn.

Figure 1. Extended Pinyin passage. (from Light, 1976, p. 9, with modifications)

Light (1976) found that even though his students said they felt they were reading characters faster and more efficiently, in fact they were generally still making more errors in comprehension of the character texts than with the Pinyin texts. Light suggests that this psycholinguistic phenomenon appeared to be more prevalent among his more proficient students, possibly because they were becoming more familiar with the orthography.

Following Light (1976), Everson (1988) tested 60 of his non-native Air Force Academy students, who had learned approximately 200 characters, to see which they read faster: Chinese characters or Pinyin romanization. Everson (1988) noted that previous research showed that "the featural makeup of Chinese orthography coupled with its irregular sound-to-symbol correspondence combine to make reading development a slow and laborious process for L2 learners" (Everson, 1988, p. 4). Based on Light's interpretations, Everson divided his participants into proficiency levels to attempt to confirm Light's (1976) observations that more proficient first-year students already became more comfortable reading characters and less willing to read Pinyin romanization. He found that his first-year students did indeed read romanized text more quickly and with greater accuracy than they did the character texts.

Everson attributes this to the possibility that his first-year students may still be dependent on the familiarity of the orthography of their L1.

Finally, as a point of contrast with the previous two studies, Sun (1993) conducted an eye-movement study of native Chinese primary school, high school, and graduate students reading a text prepared in Chinese characters and Pinyin. His data showed that without exception, the native readers fixated longer on the Pinyin, took in smaller and more frequent amounts of information during each fixation of the Pinyin text, and read the Pinyin text much slower than the character text (Sun, 1993, p. 252). Sun notes that "most of the books in China are printed in Hanzi (Chinese characters), so the greater difficulties associated with reading Pinyin may be due to a lack of practice" (p. 252) and concludes that, for the Chinese, "more efforts are required to master Pinyin" (p. 254).

Based on the results of all three of these studies, it seems that Pinyin is serving its designated function in facilitating learning of the phonological aspects of the Chinese language, but that once characters become the dominant form of reading passages, Pinyin becomes more secondary. It could also be argued that non-native CFL students might perform better than native Chinese readers because the non-natives encounter Pinyin more frequently (and more recently as compared to native Chinese adults) and are thus more familiar with it.

Phonological Coding of Chinese Characters

There are two prevailing theories and one emerging theory in the debate regarding sound and meaning of Chinese characters. One theory, supported by experimental data, claims that the meaning of a Chinese character is obtained from the mental lexicon only after a sound has been attributed to it. A second theory, also supported by experimental data, claims that the semantic meaning of a Chinese character (or word) is directly accessed in the mental lexicon without the need for phonological activation. Finally, a third line of thinking—as corroborated by experiments in Matsunaga (1994) and most recently Chen and Shu (2001)—seems to lend strong support for a combination of the two theories. In the pages that follow, I will look at each of the three perspectives in turn, followed by the pedagogical implications of the most recent trend.

The Phonological Mediation Theory (PMT): Meaning is derived via sound

As mentioned above, at one end of a continuum in the debate regarding the primacy of speech and the phonological encoding of Chinese characters are researchers who maintain that phonological mediation is a necessary first step in the reading process and successful reading comprehension. Experiments by Packard (1990) and Perfetti and Tan (1998), for example, conclude that Chinese characters represent "phonology-plus-meaning" (Perfetti and Tan, 1998, p. 102).

In an effort to test a component of their connectionist Interactive Multi-lexicon Model of Chinese Reading (Tan & Perfetti, 1997) as shown in Figure 2, Perfetti and Tan (1998) conducted two experiments utilizing primed-naming tasks of Chinese characters. A primed-naming task simply means that research participants are shown sets of pairs of Chinese characters, which the participants must name and their response times are logged. The first Chinese character of a pair assumes a priming function in relation to the second character of a pair, called the target.

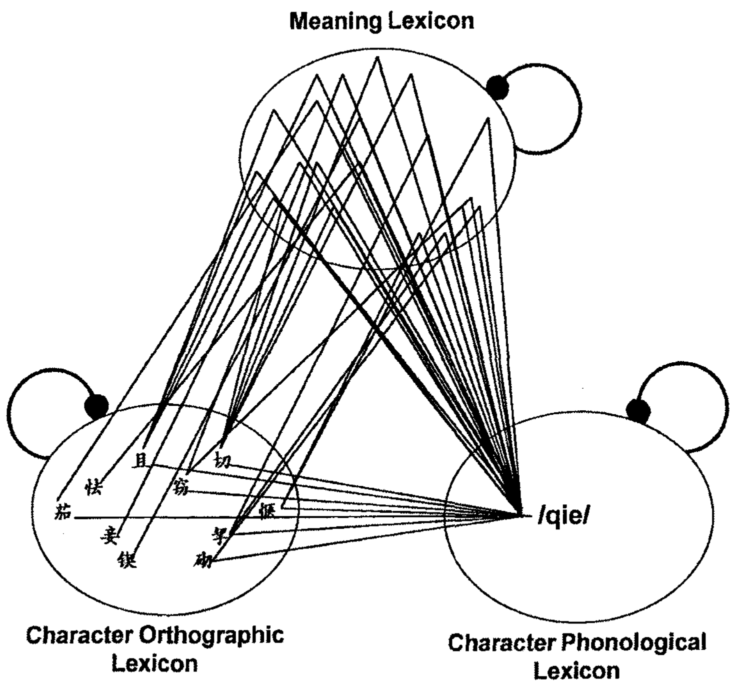


Figure 2. Interactive Multi-lexicon Model of Chinese Reading. A component from "a framework of the mental representations of orthographic, phonological, and semantic information for Chinese characters" (from Tan and Perfetti, 1997, p. 44).

In the Perfetti and Tan (1998) study, characters were paired according to one of four categories: 1) graphically-related prime–target, where the priming character is similar in form or shape to the target character; 2) phonologically-related prime–target, where the priming character is similar in sound (homophonic, in fact) to the target character; 3) semantically-related prime–target, where the priming character is similar in meaning to the target character; and 4) unrelated prime–target, where the priming character (or non-linguistic number symbol, '#') was not similar to the target character either graphically, phonologically, or semantically.

The first experiment in Perfetti and Tan (1998) report on, Experiment 1, involved 20 native Mandarin-speaking undergraduate students in Guangzhou. Experiment 2 involved 54 native Mandarin-speaking undergraduate students at the same college in Guangzhou as in Experiment 1, though none had participated in the previous experiment. Experiment 2 differed from Experiment 1 in that the time between the prime character disappearing and the target character appearing (called *stimulus onset asynchrony*, or SOA) was shortened. In both experiments, in addition to varying the relation of the prime to the target according to the above mentioned categories, Perfetti and Tan also introduced the variables of specificity and vagueness. Consequently, their results show that the more precise and certain the meaning of a character is, the faster the lexical information is accessed.

From their experiments, Perfetti and Tan (1998) conclude that there is a determined order to lexical information access. That order proceeds first from graphic (visual) recognition, then to phonological triggering, then semantic information is obtained. Perfetti and Tan do concede, however, that due to the nature of the task participants were asked to perform, there could be some residual effects.

The Direct-route Theory (DRT): Hanzi only represent meaning

Direct-route theorists such as Tzeng, Hung, and Wang (1977) and Henderson (1982) claim that, based on their (response-timed) research, phonemic coding seemed to play little if any significant role in accessing lexical information. That is, a Chinese character's lexical (semantic) information is directly accessed from the form of the character. This may be somewhat related to the "monosyllabic myth" attributed to Modern Standard Chinese (MSC), which says that every syllable has a meaning. To this, Chao (1968) states that the "so-called 'monosyllabic myth' is in fact one of the truest myths in Chinese mythology" (p. 139). He also says that "the predominance of monosyllabism is so great that a speaker or writer tends to read meaning into single syllables when there was none originally" (Chao, 1968, p. 139).

Recently, Chen and Shu (2001) reported convincing evidence that casts doubt on a phonemic-only theory of the coding Chinese characters. Chen and Shu replicated a primed-naming experiment done by Perfetti and Tan (1998). They say the results obtained by Perfetti and Tan were misinterpreted and decided to replicate the experiment. Chen and Shu conducted two replications on the same experiment (Perfetti and Tan, 1998, Experiment 2, as described above). Chen conducted one version of the experiment in Hong Kong with 60 native Cantonese speakers, while Shu conducted another version in Beijing with 60 native Mandarin speakers. Their results lead them to refute Perfetti and Tan's (1998) conclusions, followed by the following remonstration:

> "... even if the design and the stimuli of Perfetti and Tan (1998) are capable of generating reliable results, the interpretation of the obtained results should be made cautiously.... Thus, any difference in the pattern of results found in different conditions could have been due to different patterns of activation, different levels of prime–target similarity, or the interaction of the two factors.

The Dual-access Theory (DAT): Activation is relative

As Matsunaga (1994) discussed, most research done prior to the study she conducted concentrated on reading either at the individual character level (monosyllabic), the word level (typically polysyllabic), or the sentence level, leaving discourse (i.e., extended text) yet to be investigated. This is actually still the case ten years after her dissertation (and 18 years after Everson, 1986). Matsunaga's (1994) study, then, attempted to provide evidence for phonemic activation—or even the lack thereof—during reading texts of extended discourse (i.e., multi-sentential).

In her research, Matsunaga (1994) used eye tracking equipment to determine whether, and to what extent, participants relied on phonological encoding during reading of extended (newspaper) passages. Her dissertation experiment involved 65 native Japanese-speaking nationals living in Hawai'i. She concluded that the more familiar lexical items—typically function words—were processed without activating phonological codes, while deeper, more load-carrying lexical items—typically content words such as nouns—did activate their phonological coding. Thus, based on these results, she concluded that, for skilled readers, a dual-access theory approach to semantic information processing seemed most plausible.

Perfetti and Tan (1998), as detailed above, seem to begrudgingly corroborate Matsunaga's (1994) research when they conclude that "if our indeterminacy idea is correct, then, with contexts that highly constrain meanings, one might expect to observe semantic processes that temporally overlap the otherwise more rapid phonological processes" (p. 114). The fact that Chen and Shu (2001) also provide detailed and well-argued evidence for a "script-to-meaning" perspective also seems to lend support to Matsunaga's (1994) conclusion that, depending on the task, content, and context of the reading activity, semantic information is sometimes dependent upon phonological coding, but at other times lexical information can be accessed directly from the form of the character, essentially taking precedence over the phonological code.

The research into whether recognition of Chinese characters is phonologically activated prelexically, lexically, or postlexically is an important step in determining the essential nature of how the mind processes Chinese character information. As Everson (1998) notes:

> this emphasis on the critical contribution of phonology also is consistent with theories maintaining that (a) phonological encoding facilitates information processing in short-term memory during the reading of extended text (Kleiman, 1975), (b) provides a trace in short-term memory superior to what visual processing can offer (Baddeley, 1966), and (c) aids in the visual and semantic processing of orthographically unfamiliar words (Just & Carpenter, 1987). (p. 197)

While many researchers believe that phonological activation is prelexical for alphabetic languages and either lexical or postlexical for non-alphabetic (so-called logographic) orthographies (Horodeck, 1987; Huang & Hanley, 1995; Liu, 1988; Shen & Forster, 1999; Tan & Perfetti, 1998), the timing is so close that for simplicity's sake I have discussed phonological activation first and will now turn to character recognition and lexical activation.

The remainder of this section will look at two major areas that still dominate much of the research in reading Chinese, namely, orthographic effects in character recognition and word recognition. I will also briefly consider general notions of comprehension as they relate to these two research areas.

Character Identification

A large number of studies looking at the processing of the Chinese written language have focused on individual characters (Chen, 1986; Feldman & Siok, 1999; Hayes, 1987; Ke, 1996, 1998; Perfetti & Tan, 1998; Sergent & Everson, 1992; Sun, 1993; Tzeng, Hung, & Garro, 1978; Wang & Thomas, 1992; Zhou & Marslen-Wilson, 1999). Key among these studies has been the issue of to what degree the effects of complexity, density, and frequency play in processing the information presented by a character. Frequency can either be 'frequency of occurrence' (how often a character is used in the written language in general based on statistics gathered from corpus studies) or 'frequency of exposure' (how often a character is repeated in a language textbook) (Sergent & Everson, 1992). In either case, high-frequency characters are used more often and are thus more common than low-frequency, or rare, characters. Complexity differs from density in that complexity by definition is more generally limited to number of strokes while density refers to the proportion of space filled to construct a character.[9] As Yang and McConkie note, "additional complexity can arise in a number of different ways: by having nondecomposable characters with more strokes, or by having characters with additional components arranged in different ways" (1999, p. 218). For example, Figure 3 shows three pairs of characters where each pair has the same degree of complexity, yet differ in density:

6 strokes	8 strokes	10 strokes
回 好	京 事	笑 能

Figure 3. Visual difference between complexity and density in Chinese characters.

[9] In optical character recognition (OCR) terms this is technically called optical density, which refers to the percentage of ON-pixels in a defined (variable) unit area. In terms of a page of black text and white background, for example, all instances of black would be ON-pixels, the white background essentially being OFF, i.e., providing no stimulation. (See also Zheng, Liu, & Ding, 2002.)

There have been a few studies that have looked specifically at the three factors of frequency, complexity, and density in character identification.[10] Hayes (1987) looked at the nature of character complexity as it affected character identification. He used a character recall task that asked subjects (17 advanced non-natives and 17 natives) to match characters they believed were shown during a four-second display on an answer sheet provided. He divided 105 characters into four groups of varying complexity: Low (1-5 strokes), Medium (6-10 strokes), High (11-15 strokes), and Complex (15 or more strokes) (p. 49).[11] His results showed that accuracy of recall was not affected by character complexity.

Sergent and Everson (1992) used a character naming task to look at textual frequency and complexity in terms of both speed and accuracy in two levels of non-native learners of Chinese, beginning (n=17) and advanced (n=5). The three independent variables were instructional level (beginning, advanced), frequency (low frequency [<35.1] and high frequency [>35.1], based on occurrences in the language textbook), and density (low density [<7] and high density [>7], based on number of strokes). The two dependent variables were vocal latency and

---

[10] Chen (1986) also notes that display size seems to be "an important factor in character/word recognition" (p. 9).

[11] There is a potential problem with these groupings, however. If we look again at Figure 3, notice that two of the three pairs of characters represent the low and high ends of his Medium group. This would mean that both 回, and 能—characters of strikingly dissimilar densities—would be considered as being of equal complexity. I believe there is even a density distinction between 事 and 笑, where 事 could be identified as being more complicated than 笑 in a timed experiment even though 事 has fewer strokes. Thus, the same character in three different fonts would have the same complexity, obviously, but three different densities—e.g., 好 (standard text, 'Song' font), 好 (brush style, 'Kai' font), and **好** (gothic style, 'Hei' font), as well as a style change: **好** (standard text bold, 'Song' font). This sort of psycholinguistically perceived versus actual density would be an interesting study that has not been done yet.

accuracy. They observed that advanced readers were faster and more accurate in the character naming task than the beginners. Sergent and Everson do note, however, that

> "The longer reaction times for denser characters are congruous with the findings of Hayes (1987b) that non-natives were more likely than natives to attend to the visual aspects of characters when holding the character in short-term memory. The longer reaction times for the more visually complex characters indicate that non-natives do not process visually complex characters as rapidly as visually non-complex ones because they attend to the additional visual features of more visually complex characters. (1992, p. 41)

Recently, Xiao (2002) looked at the effect of complexity again. For her study, she defined her complexity levels as low (1-6 strokes), mid (7-11 strokes) and high (more than 12 strokes). Her data came from post-lesson vocabulary quizzes of 34 first semester non-native students. The vocabulary quizzes were comprised of three sections: character recognition, character production, and dictation. Like Sergent and Everson (1992), Xiao found that character complexity affected recognition and production but hypothesized that complexity effects diminish as students progress to higher language levels and incorporate chunking in the form of radicals and phonetic components as opposed to focusing on strokes at the beginning of Chinese character learning.

Because each of these studies used different subject populations and defined their parameters differently—Xiao's (2002) definition of density differed from that of Sergent and Everson's (1992), which differed from that of Hayes (1987). Sergent and Everson used frequency data, but the others did not—these studies came to contradicting conclusions which are difficult to usefully compare. Despite the fact that their conclusions seem to be at odds with one another, or perhaps because of that fact, more detailed study with more clearly defined parameters needs to be conducted to tease out what is exactly going on in the relationship between character complexity and recognition and production ability.

Finally, though it only deals with recognition, a recent study by Wang, Perfetti, and Liu (2003) corroborates what has been discovered in the previous studies above in that increased frequency of exposure is strongly correlated with faster response times and more accurate identification of characters in a naming task. Not surprisingly, since Chinese characters do not indicate tone, the 15 beginning-level CFL participants (composed of two dissimilar groups) performed poorly on a tone matching task. They conclude that their findings indicate a primacy of perceptual learning among students whose native written language is alphabetic (e.g., English) in learning a non-alphabetic written language such as Chinese.

Word Recognition

In order to identify whole words, it is argued that beginning readers must first decode their component parts. Experienced readers, it is argued, do this mainly with infrequent, unfamiliar characters and words. As discussed above, for less frequent and not-quite-yet-automatized lexical items, this typically involves the assignment of sound first, then accessing meaning. For English and other alphabetic languages, word identification requires the ability to construct meaning from varying strings of letters, which are representations of sounds. For Chinese, a morphosyllabic language (DeFrancis, 1989), prelexical identification of words seems to rely on the ability to compile meaning from jumbles of seemingly random strokes constrained within a limited square space.

Japanese Kanji share a similar process, but with an added orthographic tool in its syllabary (composed of two types, *hiragana* and *katakana*). Some early studies conducted at the University of Hawai'i determined that increased learning of and exposure to the orthography meant faster recognition times (Yamamoto, 1980), that there is no correlation between Kanji recognition and English whole-word ability in children (Ihara, 1983), and that there are significant but varying relationships between contextual, mnemonic, and rote learning on individual differences and learning style preferences (Okita, 1996). More recently, Mori (1999) looked at JFL learners' beliefs about English and Japanese word recognition, including degree of reliance on Kanji despite their perceived difficulty.

Everson (1998) defines word recognition as "deriving both the phonetic codes (or pronunciation) as well as lexical meaning from printed Chinese characters" (p. 194). He also notes the psycholinguistic difficulty in actually defining what a word is in Chinese in the first place (see also Packard, 2000, for an extensive investigation into possible ways of defining a word in Chinese).[12] By way of a simplified example, one could say that just as English has words of varying length (as measured by number of letters) so too does Chinese. The English words "a" (one letter) to "antidisestablishmentarianism" (28 letters) and everything in between can easily be identified as words in the written language by the use of spaces. Since the Chinese language does not use extra space boundaries[13], segmentation becomes a tricky issue. When talking about word recognition in Chinese, it must be remembered that while MSC (Modern Standard Chinese) is predominantly bisyllabic (approximately 60-65%), there are also many words that are represented by a single syllable, and hence a single character (Wong

[12] Packard (2000) discusses eight ways of defining a "word", including orthographic word, sociological word, lexical word, semantic word, phonological word, morphological word, syntactic word, and psycholinguistic word. See also Tsai (2001).

[13] Still, as Yang (1994) notes, there is still roughly two to four millimeters of evenly distributed space between characters in running Chinese text, depending on the font size.

& Weber, 1986). Even though Chinese words can range from one character (e.g., 人 'person') to three characters (e.g., 嚙齒類 'rodents')[14] in research on Chinese word recognition, as noted by Perfetti and Tan in 1996, because Chinese readers tend to adopt a "two character assembly strategy" (as cited in Tsai, 2001, p. 21),[15] the most common descriptor used to define a word is "two-character combinations" (Everson, 1998, p. 201*n*).[16]

This relates to the unitization effect discussed by both Chen (1986) and Tao and Healy (2002), which says that the more frequent and familiar a character or word is to a reader, the more likely it will be automatically (less deeply) processed based mainly on components (i.e., semantic radicals and phonetics) for individual characters or character combinations for compound words during the reading of extended (or "normal") text. In English this unitization effect is observed in particular with function words, such as "a" and "the" (and unique words like "alligator"), while in Chinese, these tend to be highly frequent semantic radicals such as

[14] According to Tsai's (2001, pp.40-41) data, the average length of a Chinese "word" is 1.33 to 2 characters, depending on tokenization. Average English word length ranges from 6 to 9 letters, depending on such factors as style, genre, etc. Although I understand Tsai's reasons for doing so (mainly as it fit his segmentation algorithm), I do not consider Chinese 'four-character phrases' (成語, **chéngyǔ**, 'set phrases') or groups of numbers such as 18 or 691,000,000 (Tsai, 2001, p. 36) as words here. Of course, this gets to the distinction between 辭 (cí, 'phrase') and 詞 (cí, 'word') in Chinese, where 辭典 and 詞典 —both pronounced **cídiǎn**—mean 'dictionary'.

[15] A similar observation was recently made by S. Biber (personal communication, August 19, 2004) who noticed that (mainland) Chinese computer users "are quite confused by the sentence-at-a-time entry mode and prefer to switch to a word-by-word system" of input.

[16] According to Huang and Liu's 1978 statistics based on a million-character running text corpus analysis, two-character compounds are said to account for 65% of the total (as cited in Yang, 1994, pp. 8, 28).

女 'woman' or 日 'sun', or grammatical particles such as 的 (de), or even words such as 中國 (zhōngguó, 'China').

Like the character identification studies just reviewed, there has been much research done in relation to word recognition. Koda (1996) has put together a valuable critical review of second language word recognition research, covering such topics as the significance of word recognition efficiency, connectionist accounts of performance improvement and cross-linguistic variations in word recognition processes in L1 word recognition, L2 word recognition frameworks and L2 processing experience effects, and L1-L2 orthographic distance effects. Some of the studies on word recognition focus on orthography (Wong, 2000; Wong & Chen, 1999; Yang, 2000; Zhou, Marslen-Wilson, Taft, & Shu, 1999), while others have looked at error detection (Xu, 1998), word length (Hayes, 1990), and comparisons of word recognition processes in Chinese as first and second languages (Sun, 1992).

With regard to Chinese word recognition, Tzeng, Hung, and Garro (1978) reported on an early study they had conducted that examined the possibility that Chinese word recognition in the context of a sentence was more likely to be phonemically related due to processing demands in working memory. They found that just as phonemic similarity of unrelated Chinese characters affected short-term retention, so too did it affect the reading of Chinese characters in meaningful sentences (p. 299).

Using a recall protocol rather than using eye-tracking equipment as Everson (1986) did, Hayes (1990) took the Everson study one step further in terms of looking at the effect of word length on memorability. The 13 non-native (Advanced) and 13 native (Taiwan) Chinese readers were presented with 10 sets of four short sentences which contained nouns varying in length from one to four characters.[17] Hayes observed that while there were quantitative differences between natives and non-natives in terms of how much text participants in each group recalled on average, both groups exhibited similar response trends in that four-character words were more likely recalled than one-, two-, or three-character words. Hayes hypothesized that "the memorability of one or more levels of the number of characters per word is strongly influenced by the level of proficiency of the reader" (p. 40), and also suggests the added context was a significant factor in this observation.

There has also been some relevant research done in the JFL context. For example, Koda (1992) pursues an investigation of 58 native English-speaking college students learning to read Japanese as a foreign language (JFL). She looks at the relationship between lower-level verbal processing skills and foreign language reading proficiency. She also notes that aspects of linguistic knowledge, background knowledge, cognitive and metacognitive skills, and lower-level verbal processing skills have also been suggested as major reader-related factors. Her research question examined letter identification and word recognition on foreign language (FL) reading performance. She uses three measures of reading comprehension: cloze, paragraph comprehension, and sentence comprehension. For verbal processing, Koda uses speeded word recognition and letter identification tasks. Her results show that there is a high correlation between the verbal processing scores and the reading comprehension scores that displayed an increase over time. Koda also concludes that "efficient lower-level verbal processing

[17] As with the Tsai (2001) data mentioned previously, Hayes (1990) considers combinations such as 少數民族 (shǎoshù mínzú, 'minority ethnic groups') as one word. Unfortunately, he only provides one example set, so it is difficult to determine the validity of the other four-character examples he considered "words".

operations are essential for successful performance in FL reading comprehension". Koda cites Ehri (1987) in arguing that the written language provides a mechanism of rehearsal that provides visual-spatial support for speech development.

One interesting aspect of eye-tracking research, discussed in more detail in the next chapter, is that even though we can detect fixations and say they are indicative of increased mental activity, we really do not know the reason the reader has fixated on that particular point. For example, it could be that s/he does not know the word in the first place, and so is trying to switch to a lower-level of processing in an effort to recompose it from its basic elements, or it could be that there is an error in the text that the mind has caught and is fixing in order to extract the information and integrate it into the overall whole. Specifically, eye-tracking research has been done in the area of word perception and identification. At this point it is only necessary to mention them as they relate to word recognition and phonological recoding.

In addition to phonological recoding, recent research in reading Chinese has looked more in-depth at word frequency, word length, launching and landing points of saccades (jumps), and contextual constraints. Yang (1994) determined that the optimal viewing position of a word in Chinese is at its center, which for a two-character word would be between the two characters in the combination (or on the right-most edge of the first character in the compound) and for a three-character word would be along a center line of the second character in the compound. Yang and McConkie (1999), however, conclude that "there appears to be no preferred landing position in words when reading Chinese text", and further that there is "no evidence for either a word- or character-based eye guidance process during the reading of Chinese text" (p. 215).

In looking at phonological effects versus orthographic effects, Wong and Chen (1999) conclude that "orthographic processing plays a prominent role in the early stage of reading Chinese, but phonological processing does not" (p. 475).

Most recently, Tsai (2001) developed the first theoretical computational model of Chinese reading that utilized segmentation algorithms in a word recognition study. Tsai's model took into consideration word frequency and structural complexity along with human eye movement behavior. While this model could be useful in replicating native reader habits when facing extended text, it does not cover the variables that present themselves when non-native readers attempt the same task. This is less a criticism of Tsai's work as it is of the field at large, in that very little research has been done to provide benchmarks based on eye movement data of non-native readers reading Chinese.

We will return to these studies in more detail in the review of eye movement research in the next chapter. Let me conclude this section by noting several factors related to word recognition that can influence reading comprehension. First, and probably foremost, is the factor of language competence or language proficiency (Horiba, 1996a; Mecartty, 1998). The longer one has studied a foreign language (assuming a certain degree of effort has also been invested), the more efficient one's L2 reading ability will become.

Another issue is the question of unknown vocabulary density (Hsueh-chao & Nation, 2000). As I have mentioned earlier in this chapter, the greater the extent of one's knowledge of vocabulary the more likely one's ability to read and comprehend a novel text. Of course, DeFrancis (1966) notes that even with a repertoire of 1200 individual characters, which represented 91.3% of characters in Chen's 1928 data (as cited in DeFrancis, 1966, p. 10), "knowing a character does not necessarily give a student access to the combinations of which it is a part" (p. 11). As Hayes (1990) argues, without a solid foundation of vocabulary (characters alone and in combination) as represented by a "memory network of Chinese" (p. 38), the ability to even exercise a top-down, conceptually-based, knowledge-driven approach to an extended piece of novel text is crippled with only limited potential for success at best.

Reading Strategies of Learners of Chinese

Equipped with a linguistic toolbelt packed with a developing store of vocabulary, a collection of grammar structures, and a spattering of cultural background knowledge, the CFL learner's efforts at processing information via print or screen can still be hindered by insufficient reading skills. One factor that is crucial to comprehending something that one is reading is the ability to overcome obstacles in the text, whether they be unknown or unfamiliar characters, words, or grammar (i.e. syntax). The strategies a reader uses to achieve some measure of comprehension are essentially learned behaviors as one gains experience through the frequent exposure to text. Reading strategies, as defined by Barnett (1989), are the problem-solving techniques readers employ to get meaning from a text (p. 36). In contrast, general learning strategies are used to acquire and improve skills with the second or foreign language tool, while reading strategies are used to obtain information by the most efficient use of available tools. The present study inferred reading strategies from eye movement data gathered during a single reading task session.

Several studies have looked into the relationships between foreign language proficiency level (Everson & Ke, 1997; Hayes, 1988), reading experience and skills (Bernhardt, 1983), and efficiency of reading processing and comprehension (Liu, 1999; Su, 2001). Everson (1994) reminds us that when readers are forced to grapple with a foreign language text, as opposed to reading it more or less fluently, they must "divide up their cognitive resources in a strategy termed 'code switching'" (p. 6). Further,

> When learners employ this strategy, they go through a text and circle characters and write in romanization and/or English in the margins so as to establish some baseline for establishing the meaning of the vocabulary. When this process is done, the learner has "freed up" a large amount of the attentional resources that were necessary to recognize and process much of the vocabulary, and that now may be used to establish grammatical relationships (syntactic feature recognition), as well as the relationships between different portions of the text (intratextual perception). (p. 6)

Other reading comprehension strategies include rereading, paraphrasing, monitoring understanding, making inferences, identifying important information, and integrating text (Kobrin, 2000, pp. 14-16). These strategies are arguably identifiable as top-down in nature. When non-native readers of Chinese encounter unfamiliar characters, words, or grammar structures, they almost invariably resort to more text-driven processes such as character component analysis, other attempts at character identification, and the guessing of possible word meaning from context, as introduced earlier in this chapter. Lee (1998) presents a reading strategy scheme based on her research involving think-aloud protocols which consists of two groups of strategies: Unit Identification Strategies (UIS) and Unit Assemblage Strategies (UAS). There are twelve strategies in the former category and eight strategies in the latter category. These strategies are divided roughly along the lines of bottom-up and top-down processing discusses earlier. Included among the UIS group are strategies such as scanning for unfamiliar words, annotating the text with pronunciation or meaning cues, and word-for-word translating. The UAS group include strategies such as using background knowledge, hypothesizing, and summarizing (Lee, 1998, p. 117).

Anxiety, of course, will have a negative effect on the learning, acquisition, and both short- and long-term retention of vocabulary, in addition to creating interference and being detrimental to cognitive processing and overall language proficiency and reading comprehension (Saito, Garza, & Horwitz, 1999; Watanabe, 1997). For beginning and inexperienced foreign language readers, by the very nature of the task, reading in a foreign language by oneself can be a nerve-wracking experience. Moving up another level of anxiety is what happens when students are asked to read aloud—or even silently—in class. Still another level of anxiety occurs when students are faced with a reading comprehension test. Each of these levels adds further degrees of difficulty that can potentially create hindrances to successful comprehension.

Finally, a note about test-taking strategies versus reading comprehension strategies. Test-taking strategies include whether to read the entire passage first or just some of it then moving on the the questions, or reading the questions first and then moving to the text to search for the answer. Even though the research conducted for this dissertation involved the use of a computer-adaptive "test" (CAT), the stakes involved were actually very low. Participants were paid a set fee regardless of their performance, and they were even told that they could discontinue their session for any reason without penalty. Also, unlike Kobrin's (2000) computer-based test passages and questions, the majority of items given by the CATRC (Computer-Adaptive Test for Reading Chinese), including text and questions, were presented on a single screen. In addition, participants were told there was no time limit for the test or for individual items and that they could take as long as they wanted and go at their own pace. For these reasons, and based on informal discussions with the participants after they had finished their sessions, I believe that test-taking strategies were actually of a secondary nature as compared with more primary reading comprehension strategies as the participants were more focused on the main task of reading for comprehension. Of course, the psychometric effects of knowing one is being tested or simply observed are difficult, if not impossible, to control for. Further details of the experiment sessions are provided in Chapter Four: Methodology.

Teaching Reading in CL2/CFL

In 1966, DeFrancis leveled a critical indictment against the situation of the teaching of reading in Chinese as a foreign language, citing the two main problems of the time as being a) "making too rigid a connection between reading and speaking" and b) "placing too much emphasis on learning characters" (p. 2). Around 20 years later observations made by Walker (1984) and Hayes (1988) indicate that the situation had not improved. DeFrancis (1966) talks about how pervasive this character-counting mentality among both teachers and students by repeating the question that students still ask today, "How many characters do you know?" (p. 6) In this type of environment, students are merely becoming highly skilled "decoders of texts" (Jorden, 1975, p. 118), ignoring the fact that reading is grammatical (Jorden, 1975, p. 111; Jorden, 1995, p. 97). Similarly, Jorden (1995) talks about how, trained to be obsessed by kanji, students become "kanji-hoppers" (p. 93), essentially jumping over and ignoring the clues and cues provided by the surrounding kana context.

Given what we currently know about the process of reading in a "foreign" language such as Chinese, teachers in the field have experimented with a variety of methods and have proposed a number of approaches to help make acquiring the skill more efficient. These methods and approaches have gone from research in general reading processes—including bottom-up models, top-down models, interactive models, and universals that apply across most all languages—to special considerations that need to be made for learning to read in a foreign or second language. The special considerations called for in learning to read and comprehend not just words in isolation but extended text in Chinese includes overcoming the burden of the orthography and learning the sound associations that go with mostly abstract symbols.

There have been occasional and varied debates within the field of Chinese pedagogy concerning the teaching of reading in Chinese. These include a range of suggestions from delaying character instruction (e.g., Packard, 1990, Dew, 1994), to *Pinyin* only instruction, to Pinyin versus *Zhuyin Fuhao* (a.k.a. *bopomofo*)[18] instruction, to tonal spelling versus diacritics (McGinnis, 1997), to simplified versus traditional character instruction, and to teaching cognitive and metacognitive reading strategies and schema theory at more advanced levels (Jiang, 2004).

Researchers and teachers that have suggested a delay in character instruction in order to first develop spoken language skills have varied in how long that delay should be: from as

---

[18] This system of character annotation is more similar to the Korean alphabet than it is to the Japanese syllabaries in the sense that it "spells" sounds. Recall the description of Pinyin provided earlier where the example of Měiguó was given for the characters 美國. In Zhuyin Fuhao that would be transcribed as ㄇㄟˇㄍㄨㄛˊ (where ㄇ would equal [in IPA] /m/, ㄟ = /ei/, ˇ = 3rd tone, ㄍ = /g/, ㄨ = /w/, ㄛ = /o/, and ˊ = 2nd tone. It is mostly used in educational materials in Taiwan and would appear either on top of horizontally running text or along the side of vertically running text, e.g.:

ㄇㄟˇ ㄍㄨㄛˊ
美 國

美 ㄇㄟˇ
國 ㄍㄨㄛˊ

little as a few weeks (Packard, 1990; Walker, 1984) into Chinese language instruction, to as much as a semester, to as much again as a year or more. No research has yet been published that has tested the feasibility of or the effects on student reading development that incorporates longer delays of character instruction, exclusive use of romanization in the early stages of instruction, or the use of either simplified or complex characters at the program level, so it remains to be seen whether these approaches would actually improve or enhance the CFL learner's overall reading skills development.

Essentially, it may be said that the more a student reads, the more a student learns (acquires more characters and words), the more a student learns, the sooner the student can just read for enjoyment (passing over some of the characters/words that they do not know and not letting them interrupt their reading flow—unless they have totally lost the meaning). This may require learners to spend quite a bit of time looking up characters and words in the dictionary in the beginning, but—especially if they are marking the words they have had to look up in the dictionary—they should notice that they do not have to look up the same words or as many words as time goes on.

Another issue that seems to be at the heart of student reading success or failure is their motivation—or lack thereof—to learn the characters in the first place. McGinnis (1995) asked (a) "how much do our students really want to learn to read and write Chinese", and (b) "how do our students actually learn to read and write, particularly at the earliest stages of study"? He found in answer to the first question that most beginning- and intermediate-level students are really only interested in learning to speak and listen. In answer to the second question, and based on the results of an ethnographic survey of 29 students at the beginning level, students adopt "highly idiosyncratic and thoroughly inconsistent" strategies for learning to read and write Chinese characters, despite the best efforts of their instructors to teach them. The results of this survey allowed McGinnis to identify two student-based myths related to the learning of Chinese characters, what he called the "non-necessity myth" and the "rugged individualist myth". Students who succumb to the first myth believe that Chinese characters are insignificant and useless to them and their personal goals for learning the language, while students who fall prey to the second myth feel that they alone know what the best way is for them to learn how to read and write and turn a deaf ear to the knowledge and experience of their teachers.

McGinnis (1999) later dropped the rugged individualist myth and switched it with the "radical-schmadical" myth (p. 163), but in my teaching experience, all three myths possess equal potential to rear their ugly heads, but appearing at slightly different times during the course of student learning. Speaking to the "radical-schmadical myth", McGinnis' data gave indications that, especially at the beginning of study, radical and phonetic components are simply too abstract to be of any practical use. It is not until later on in their studies that students are able to make connections with more frequent radicals and phonetics, thus seeing the value and usefulness in such an approach to learning characters.

If they have the drive and determination to continue learning Chinese characters, students are forced to come up with ways of retaining them in long term memory. In addition to creating flashcards and going through reams of paper writing and rewriting hundreds of characters in an effort to "burn them into the brain", teachers can help students to learn how to use other techniques such as creating mnemonic devices (Wang & Thomas, 1992) and tracing characters in the air (Yim-Ng, Varley, & Andrade, 2000).

Shu and Anderson (1999) note that in the PRC children are expected to know 2557 characters by 6th grade. Kennedy (1937) made a first guess that "mastery" of 1000-4000 represents a "'fair' knowledge of the language" (p. 5) while 4000-5000 would constitute an "unusually good education in Chinese" (p. 3). Later, Wang (1973) suggested that knowledge of 4000-7000 characters would be sufficient for reading a newspaper. Recently, Taylor and Taylor (1995) broke it down to 2000 characters for "limited literacy", 3500 characters for "functional literacy", and 6000 characters for "scholarly literacy" (p. 174). In light of these "guidelines", it is interesting to note that based on Chen's 1928 analysis of one million characters of running text there were only found 4719 unique characters (cited in DeFrancis, 1966, p. 10).[19]

Moving beyond character identification, and even beyond word recognition, training students to read extensively should remain one of the main goals of Chinese language education (DeFrancis, 1966, p. 20*n*). As Ho (1975) notes:

> Although a large vocabulary load in the early stages of learning to read is usually a hindrance, it is also an undesirable restriction on the student to require that he should be unable to understand any more than he is able to produce. This would put him at a considerable disadvantage in comparison with the native speaker who has a far larger 'passive' than 'active' vocabulary. It would be unnatural to limit the students to a pure 'active' vocabulary. It seems perfectly adequate that we should use texts in which some of the lexical content is intended for comprehension only. It is no argument against extensive reading to say that the learner may misunderstand, or even fail to place any meaning at all on some of the lexical items. In any case, this is an experience for which he needs to be prepared, since he will meet it as soon as he is exposed to the language outside the teaching situation. (1975, p. 9)

---

[19] Most recently, Da's (2004) corpus of 2.5 million characters of online Modern Chinese running text contained 9933 unique characters.

What is key here is that students need to be taught about this experience of failure in such a way so as not to cause even further demoralization. In order to maintain a sense of accomplishment, it is important to build both "passive" and "active" vocabulary. There seems to be a pedagogical priority shift, however, generally between traditionally defined intermediate and advanced levels of study (typically the third year of college study) where less attention is placed on the learning of character and word units and more on idea units (Jiang, 2004; Spring, 1999). These priorities include designing materials that use students' background knowledge (Everson, 1996), designing texts of varying difficulty levels (Yano, Long, & Ross, 1994), breaking down and analyzing text structures (Bai, 1997), and concentrating on narrative texts (Okurowski, 1987).

A few scholars have looked into purported benefits of certain types of composition exercises on the reading process (Light, 1975; Walker, 1984; Everson, 1994; Shen, 2000). Light (1975) lamented that "without too much exaggeration, it can be said that, in the main, we teach Chinese reading by exposure with little attention to active skills" (p. 70). He noted that there was, in the scholarship up to that time, the explicit assumption that "the four skills are not isolable" (p. 79) and that they offer mutual reinforcement. This was echoed almost a decade later by Walker (1984) who states that "reading is not an isolated, nor isolatable, language skill. Reading does not happen without a productive capacity in the language of the text" (p. 69).

Light (1975) suggests early on that curriculum designers explore the relationship between using writing Chinese to assist in CFL reading instruction, considering composition to be the active skill missing from the equation. The "controlled composition" (p. 71) exercises he recommends are to be of a paragraph in length or longer. Students are provided with text in

Chinese which they must read, comprehend, then manipulate in various ways and rewrite. Though his paper only offers his conviction but no research to back it up, it still shared a valuable idea on a component of an integrated program that is currently quite prevalent across the field of Chinese language pedagogy.

As Walker (1984) states, "in the long run, reading is probably the most important skill a learner can gain from formal instruction in Chinese" (p. 67). Ten years ago, Everson (1994) described an idea of a process view of teaching students how to read Chinese. This process first begins with students learning the sounds and some rudimentary vocabulary of the language through romanization (as discussed in the beginning of the "Burden of Orthography" section earlier in this chapter). After that foundation has been laid, students are then introduced to characters. From that point, students then continue to learn sounds and vocabulary simultaneously through the romanization system and the Chinese characters themselves. During this second step of the process, students are exposed and trained to recognize and recall characters and words quickly and accurately. Speed and accuracy are key indicators to successful readers, so training during CL2/CFL reading should include development of rapid decoding skills. The third step involves moving from rapid decoding of characters and words to assimilation and integration of larger chunks of text. Facilitation of this final step requires that Chinese language professionals collect and develop a wealth of reading materials that can be used at varying levels of language proficiency.

Summary

In this chapter I have attempted to bring together a variety of factors that I feel are significant to understanding the process of reading Chinese as a second or foreign language

(CL2/CFL). I began by sketching what we know about the essentially universal components that make up the reading process, including bottom-up, top-down, and interactive-compensatory processing models; then I focused on what are considered special issues related to reading foreign languages in general and reading Chinese in particular. I covered the orthographical topics of character and word recognition and the phonological coding/recoding processes inherent in silent and oral reading, as well as the still debated possibility of direct lexical mapping. Finally, I looked at the reading process from both the learner's perspective as well as the teacher's perspective. In Chapter 3, I will look at previous research on working memory and eye movement that will help to place the current study more firmly into its methodological framework.

CHAPTER 3

PREVIOUS RESEARCH AND EXPERIMENTATION

Introduction

In this chapter I will discuss previous theoretical and experimental research done in relation to working memory, eye-tracking, and reading. The literature discussed here will also include research that investigates theories on the relation between eye movement and mental processing. The first section will describe some current theories that attempt to explain the structure and process of working memory in general and during reading. The second section will look at eye-tracking research, providing both sketches of general concepts and descriptions of specific research studies involving Chinese and English.

Working Memory and Cognitive Processing

Models of Working Memory

Working memory (WM) is by all accounts and above all else understood as a workspace. This workspace is recognized as temporary in nature, as well as being cognitively based. The WM system is said to be limited in capacity, varying from individual to individual, and even subject to moods and emotions. Working memory, unlike long-term memory (LTM), is seen to hold onto the materials it is working with only for short periods of time.

Based on the ten theories of WM they discuss, Miyake and Shah (1999) propose the following extended definition:

> Working memory is those mechanisms or processes that are involved in the control, regulation, and active maintenance of task-relevant information in the

> service of complex cognition, including novel as well as familiar, skilled tasks. It consists of a set of processes and mechanisms and is not a fixed "place" or "box" in the cognitive architecture. It is not a completely unitary system in the sense that it involves multiple representational codes and/or different subsystems. Its capacity limits reflect multiple factors and may even be an emergent property of the multiple processes and mechanism involved. Working memory is closely linked to LTM, and its contents consist primarily of currently activated LTM representations, but can also extend to LTM memory representations that are closely linked to activated retrieval cues and, hence, can be quickly reactivated. (p. 450)

Working memory should not be confused with short-term memory [STM, called "extinct" now by Miyake and Shah (1999), despite continued research], which is only seen as a passive storage bin of extremely short duration. The difference between STM and LTM, then, is the duration of storage. However, while LTM does hold onto content for much longer spans of time, it is still subject to decay and loss. One key concept to remember, as alluded to in the definition above, is that working memory itself activates resources it needs to complete the processes the system is currently involved in. Neither STM nor LTM determine what gets stored nor what is done with those resources once they have been placed in storage.

When talking about capacity, working memory theorists are not talking about volume, i.e., how much working memory can hold and for how long it can hold on to it (although those are factors). Volume can be said to be more a characteristic of STM and LTM, and yet, as it is turning out, even the volume of these two components of memory in general are indeterminate and vary based on individual differences (Ericsson & Kintsch, 1995).

Rather, capacity refers to processing capability. This includes temporary storage (what Miyake and Shah, 1999, call "the old STM"), but only for as long as it needs to process. Working memory processing includes the encoding (e.g., phonological and semantic codes for language) of information for either immediate processing or storage for future processing, maintenance of information primarily through rehearsal mechanisms, and retrieval of information either currently in the working memory buffer or information that has been placed into LTM storage.

According to Miyake and Shah (1999), there are currently at least ten prevailing representations of working memory. These range from exclusively hierarchical to exclusively modular representations and several representations in between. Hierarchical representations posit that working memory processing is controlled and regulated by some central mechanism. On the other hand, modular representations hold that there are components within the working memory system that are interrelated, integrated, and work together to perform cognitive processing. Somewhere in the middle are theories that say that some parts of working memory are modular while the system overall operates under a hierarchy of control.

Just and Carpenter (1980) look at comprehension of discourse as stimulated by perception of visual cues on the page. A dozen years later, Just and Carpenter (1992) refined their earlier view somewhat to incorporate the construct of working memory as a multifunctional system of information maintenance and language processing. The capacity of this system of simultaneous maintenance and processing is, they propose, both finite for all and variable according to individual differences.

The key element to Just and Carpenter's (1992) Capacity Theory of Comprehension is activation, and "capacity" is defined as "the maximum amount of activation available in working memory" to support either storage or processing (1992, p. 123). In this theory, the storage and processing are running in parallel, but vie for control of limited resources. If more resources are needed for storage, processing must be cut back; if processing is more active, then storage is relinquished. Each of these conditions will vary depending on the degree of task difficulty. The more difficult or more demanding a task is, the more likely storage and processing will have to compete.

Memory Processing During Reading

Memory (and working memory in particular) is considered by many to be a significant, if not *the* most significant construct to language aptitude. The ability to take in, process, and recall language-related information all seem to hinge on the development and conditioning of the memory system. Reading comprehension, then, as a component of language aptitude, can be seen as depending quite extensively on the processing of information. Further, the efficiency and proficiency of that memory system varies greatly due to individual differences of learners/readers (Kintsch, Patel, & Ericsson, 1999; Kyllonen & Christal, 1990).

Reading is considered to be a highly demanding activity in terms of the requirements it places on memory capacity (Waters & Caplan, 1996). Theoretically, every individual's capacity (or, capability) to process, differs from every other individual's capacity (Cantor & Engle, 1993; Just & Carpenter, 1992). Conceptually, however, everyone has a potential 100% working memory capacity. How efficiently or proficiently the individual uses that capacity, however, is subject to other variables both internal and external (Chase & Ericsson, 1982; Ericsson & Delaney, 1999; Turner & Engle, 1989).

Reading in a foreign language is especially demanding on available resources. During the reading process, to pose a simplistic example, we can imagine that my reading of a particular Japanese text (a language which I have studied for less than two years) could experience a distribution of, say, 60% for lower-level, bottom-up processing, while only 40%

will be available for higher-level, top-down processing. Conversely, my reading of a particular English text (my L1) could experience a hypothetical distribution of 10% devoted to storage of vocabulary to be processed (so-called lower level) and 90% could be available for synthesizing, summarizing, and otherwise comprehending (so-called higher level processes).

What this basically says about students whose L1 is English (i.e., alphabet-based) learning to read an L2 orthography that relies on syllabaries or logographs (as Japanese and Chinese predominantly do, respectively) is that there is the introduction of an extra processing load. This extra processing load is virtually nonexistent, or at least significantly "lighter" for an L1 English student learning Spanish as an L2. Thus, more memory capacity can be devoted to higher-level processing much earlier in the course of L2 Spanish studies than can be afforded to L2 Chinese or Japanese reading.

Horiba, van den Broek, and Fletcher (1993) conducted a study comparing 47 L2 readers of English (L1 Japanese) with 72 L1 readers of English. Their study utilizes recall protocols to determine whether the structure of a text is preserved in memory (indicating high-level processing), or whether meaning chunks were preserved in memory (indicating low-level processing). Specifically, Horiba et al look at four structural factors and their effects on recall and comprehension. Those four factors were causal-chain status, number of causal connections, story-grammar category, and hierarchical position. From their results they discover that the L2 readers did make greater use of structure rather than meaning, implying that the readers may have created and remembered the structure *because of* gaps in vocabulary. In other words, the high-level processing in working memory compensated for low-level processing deficiencies, thus retaining the structural representation rather than the semantic representation in longer term memory.

Experiments on Memory and L2 Reading of Chinese/Japanese

Research into the role of memory during the first language (L1) reading process is quite extensive. There are also many studies that look into the question of memory as it relates to reading in a second language. Studies that look at the role of memory during the process of reading Chinese or Japanese, however, are not so numerous (e.g., Yu, Zhang, Jin, Peng, Zhang, & Simon, 1985; Zhang & Simon, 1985). As Everson (1986) noted earlier, and Matsunaga (1994) and Tsai (2001) both reiterated, this is especially the case for research for normal reading of Chinese and Japanese as a second language. By "normal reading", we mean reading more extended texts rather than simply digit, word, or sentence-length samples.

Most studies looking at reading Japanese and Chinese, however, still seem to be focusing on more discrete aspects of written language processing, i.e., strokes, radicals, character density, etc. These are believed to have more potential for providing insight into how WM and STM work. Hue and Erickson (1988) suggest that there are at least two types of short-term memory (STM), i.e., verbal and visual. They conduct four experiments that look at visual versus phonological coding of Chinese characters in STM. Experiment 1 of their investigation involves ten Chinese graduate students studying in the US and is thus a study of L1 reading skill. The participants were asked to recall lists of radicals (components of Chinese characters) that varied from high- to low-frequency. The results from Experiment 1 show that higher frequency radicals were recalled more correctly and at higher rates than lower frequency radicals. Hue and Erickson's (1988) Experiment 2 continues where Experiment 1 left off. Experiment 2 presented the lists of radicals in random order for both variables: frequency and density. After written recall, participants were asked to verbally recall the radicals. The results from Experiment 2 further corroborate those obtained in Experiment 1.

Hue and Erickson's (1988) Experiment 3 is a continuation of Experiment 2 where the assumption that, during processing, familiar, high-frequency radicals are stored in the verbal STM while unfamiliar, low frequency radicals are stored in the visual STM. Hue and Erickson come to similar conclusions for Experiment 3 as those arrived at for Experiments 1 and 2. Finally, Experiment 4 explores the "floor effect" that became apparent in Experiment 3. To test for this potentiality, Experiment 4 is the same as Experiment 3 with the only difference being that participants were given more time to complete the task. Hue and Erickson's results showed an increase in recall compared with Experiment 3 indicating that there is indeed dual code (visual and verbal) short-term storage.

While his study does not deal directly with nor mention the relation of the study to (working) memory issues, we can nevertheless analyze the results of Ke's (1996) study of perception and production of Chinese characters among a total of 60 students from four separate but geographically close programs in that light. In his data analysis and results section, the effect of instruction seems to play a significant role in answering his four research questions.

Ke's (1996) first research question looks for a correlation between Chinese character perception and production. His results show that instruction has a significant positive affect on both in increasing performance as well as the degree to which the two factors correlate over time. His second question asks whether students perform better at recognition or production. Here the data is inconclusive in that poor recognizers tend to also be poor producers, but that good recognizers are not necessarily good producers. Ke's third question asks if Chinese character density (i.e., the number of strokes in any given character) affects learning. Based on the data, Ke concludes (as Sergent and Everson [1992] did) that density more adversely affects production than recognition. Ke's final question asks whether there is any effect due to instruction on recognition and/or production. Based on an analysis of his data, his conclusion here is that there is a significant correlation between instruction and improvement in both modes of performance.

Looking at the data provided by Ke (1996), there seems to be some evidence suggesting that Chinese characters place high working memory processing demands on beginning learners of Chinese characters. Further, over time and with relatively constant exposure and use, lower-level processing does seem to give way potentially allowing for higher-level processing to take place more efficiently.

Tsai (2001) developed a computational modeling study based on a Chinese corpus of over 714,000 sentences that included over 4.6 million words, or, almost 7.5 million characters. In the first part of his study, Tsai's model "demonstrated how low level, primitive information, such as word lengths, word frequency counts, and the relative probabilities of boundary roles of individual characters, can effectively resolve most cases of ambiguity." (p. 47) This ambiguity was then restored for the second part of Tsai's research, the goal of which was to simulate human reading of Chinese. Towards this end, similar restrictions humans experience during reading—such as limited viewing area—were assigned to the model. The second part of Tsai's research thus focussed on uncertainty in eye movement control. Tsai admits that his approach "lacks empirical evidence", but maintains that "it is logical". (p. 53) Tsai's results show that, aside from the length of saccades being shorter than a human would make, he was able to reliably simulate human reading with his computational model.

Walter (2000): Working Memory and Reading in a Foreign Language

Walter (2000) conducted six experiments that looked at various aspects in the relationship between working memory (WM) and reading in a foreign language (in her research, French and English). While she mainly focuses on Baddeley and colleagues' (Baddeley, 1986; Gathercole and Baddeley, 1993; Baddeley, 1996) notions of the Central Executive, the Visuo-spatial Sketchpad, and the Phonological Loop, Walter also ties in important ideas concerning long-term working memory (LT-WM) as put forth by Ericsson and colleagues (Ericsson & Delaney, 1999; Kintsch, 1994; and Kintsch, Patel, & Ericsson, 1999).

The aim of the first of Walter's (2000) six experiments was "to construct and use an instrument for measuring WM in both languages of French-English bilingual subjects of differing levels of L2 attainment" (p. 15). Walter's participants for this study were two groups of French L1 students learning ESL. The first group were adolescents (aged 13-16 years old, mean age 14 years 7 months) and the second group were young adults (aged 16-19 years old, mean age 17 years 10 months). All participants were shown 100 declarative sentences six to eight words long with the Psyscope© software (Cohen, MacWhinney, Flatt, and Provost, 1993) and were asked to recall the final word of the sentence as well as provide a logicality judgment. Walter found significant correlations between WM and age and WM and school marks. She proposes, based on the results of her data, that "sentence-final words may have been stored in a more abstract form in L1 than in L2" (p. 50).

Walter's Experiments 2 and 3 further explored the storage of sentence-final words. These experiments looked more closely at the role of phonological encoding and word sequence (four words per sequence) recall, and she found that phonologically different sequences were recalled more correctly more often by both groups than phonologically similar sequences.

Walter claims that this further supports her proposal that abstract storage strategies are being used for L1 sequences while phonological storage is being used for the L2 sequences in the WM span tasks, most notably among the older students. Experiment 3 again looked at phonological storage, but it utilized wordlike and non-wordlike nonwords. Walter found that "Wordlike items were repeated much more accurately than Non-wordlike items by both groups" (p. 97).

The main purpose of Experiments 4-6 was to investigate the relation between the *L2 reading threshold phenomenon* and working memory. The L2 reading threshold hypothesis states that "there is a level of L2 proficiency below which readers do not transfer their L1 reading skills." Walter "aimed to provide a baseline measure of L1 and L2 reading comprehension" in Experiment 4 (p. 113). In this experiment, participants read short texts in both their L1 and L2 and then completed "gapped summaries" of the texts. Based on the results of Experiment 4, Walter concludes that below a certain level of proficiency, L1 reading skills are in fact not transferred to the L2. She also reports that there appears to be no relation of WM with reading comprehension in L1, but the relation appears to be clear for the L2.

Experiment 5 investigated the hypothesis that what was not being transferred was the skill that allows L2 learners to build reliable mental representations of text. Participants of both groups read one story in each language silently. The stories were embedded with sentences containing co-referential clauses. Walter's results were inconclusive in terms of the relation between WM and the construction of mental representations, but she says they did give indication that further analysis the thematic structure of a text could yield a clearer picture.

Experiment 6, Walter's final experiment had participants conduct self-monitoring of reading comprehension. Participants were asked to read texts ranging from between 100-300 words in length containing various anomalies, then write summaries and answer short questionnaires. From the results of these data Walter found no correlation between WM and anomaly detection despite high expectations.

L2 Reading Threshold

As a final note to the role of memory in foreign language reading, I should mention a little more about its restrictions. As I noted earlier in this chapter, there is a limit to what the human brain can take in and process at any given stage of the cognitive process (Laberge and Samuels, 1974). (As my students have often lamented to me, this limit is particularly vexing to the CFL learner.) In light of this, LaBerge and Samuels (1974) postulate that the more automatic the reading processes become, the less attentional resources must be devoted to the decoding of text, and the more efficient the process as a whole can be.

About a decade and a half later, Hue and Erickson (1988) observed that regardless of the complexity of a Chinese character's radical, if it was a highly familiar high-frequency component, recognition time of the character to which it belonged did not increase (i.e., it did not necessarily take longer to recognize a character that had a complex radical). In addition, they found that "low-frequency Chinese radical and characters, whose pronunciations and meanings are not apt to be known, seem to be stored in visual form in short-term memory" (p. 204), taking up precious attentional resources and negatively affecting the ability to efficiently process information. This is important to remember as the capacity[1] of short-term memory (STM) for Chinese, as noted by Hue and Erickson, is only about two or three characters, and indeed could be even as small as just one or two characters (1988, p. 204).

Thus, both familiarity and frequency of character components, characters, and words directly influence the load placed on short-term memory (STM), processing demands placed on working memory (WM), and storage and recall efficiency of long-term memory (LTM). Recall the study by Hayes (1990) mentioned briefly in the section on word recognition in the previous chapter. Hayes tested Carpenter and Just's 1987 claim that "the added context of multiple-character Chinese words facilitate comprehension, or at least memorability" (Hayes, 1990, p. 32). This facilitation of memorability possibly contributes to the freeing up of working memory processing resources that enables the formation of relationships and the strengthening of

---

[1] Note that speaking of STM capacity is different from WM capacity. Since STM is considered more of a storage device, and possibly an integral part of WM, STM (and LTM) capacity is *storage capacity*. WM, on the other hand, is process-oriented, thus WM capacity is *storage-and-processing capacity*. On some level I believe we can liken this to the situation of working on a jigsaw puzzle. Sometimes we hold on to a piece of the puzzle trying to figure out if there is some place to fit it or where it would best fit. Sometimes we hold on to the puzzle piece for a long time while we search, other times we either fit it into the overall picture almost immediately or quickly discard it for another piece knowing somehow that it is not relevant at that time. The longer we focus on fitting one piece, the more we delay processing the larger picture; the more automatic recognition of shapes and colors and their relations becomes, the quicker the puzzle can be completed. The only difference between this analogy and actual reading is that some pieces of the reading puzzle can be skipped and in the end we can still arrive at a fairly complete picture of the message. One interesting thing about this analogy, I think, is that, like STM, we are usually limited to holding only one or two pieces of the puzzle in our hand at a time as we search for a place to lay it. Also, like the completed portions of the puzzle that lay on the table as we continue to work, LTM is also there to contribute to the successful placement of pieces by maintaining context.

connections allowing for faster and more accurate recall. Hayes concluded that added context and increased proficiency level were significant contributors to processing efficiency.

An interesting connection possibly exists amongst STM, WM, and LTM. We can imagine STM as the first place where information is stored for the briefest of instances once it has been taken in to the processing system, then taken up by WM, piecing it together with additional new pieces of information and older established memory items taken from LTM. Hulme, Maughan, and Brown (1991), however, suggest that LTM has a line of communication, so to speak, with STM. This is further elaborated by Kintsch, Patel, and Ericsson (1999), who, in discussing Long-term Working Memory (LT-WM, a component of WM), hypothesize that "LT--WM is restricted to well practiced tasks and familiar knowledge domains" (p. 187). They do admit at the same time, however, that "it is not possible to explain *how* memory is used in many cognitive tasks, such as playing chess or text comprehension" within the standard WM framework (p. 187, italics added).

Typical measures of STM capacity research have been the digit span test and the word span test. Harrington and Sawyer (1992) used both of these, in addition to a reading span test to probe into WM capacity. In line with Just and Carpenter (1980, 1992), Harrington and Sawyer functionally define capacity in terms of "a trade-off between active processing and storage" (p. 25). While the Harrington and Sawyer study used English as a foreign language, they found that while the digit and word span tests were highly intercorrelated, they bore little relationship to L2 reading skills as displayed among 34 native Japanese college EFL student participants. They interpret the reading span test to be an index of working memory (p. 32), and suggest that future research into L2 reading comprehension focus their efforts on development of reading span test measures that may provide more accurate pictures of the L2 reading process.

Beyond memory capacity for characters and words, and more germane to the present study, there has been much research examining memory capacity in relation to complex sentences and extended text providing intriguing data for our consideration (Horiba, 1996b; Horiba et al., 1993; Kintsch et al., 1999; Swanson & Berninger, 1995; Walczyk, Kelly, Meche, & Braud, 1999; Zwaan & Radvansky, 1998).

Conducting two experiments designed to examine at the role of long term memory, Horiba looked at causally related sentences in extended text passages. The first experiment had 21 intermediate JFL students read sentence pairs during class sessions one week apart. The goal of one session was to study (memorize) the sentences and the goal of the other session was to elaborate the two sentences by providing context filler that connected them. Participants were given a cued recall test first immediately after and then again one day later. The session measures were counterbalanced among participants. Horiba found that differences in recall ability seemed to depend more on the nature of the encoding conditions rather than linguistic difficulty or comprehensibility of the sentences. To clarify her findings, Horiba conducted a second experiment similar to the first experiment but with 22 different intermediate JFL students. The degree of causal relatedness of the sentences used in the second experiment varied so she could investigate the relationship between text memory and encoding. The participants again performed better during the elaboration session than during the study-recall session. Horiba concludes that the richness of context provided compensation that allowed participants to more easily recall connections. This suggests that long-term memory seems to hold text structure primarily more firmly compared with lexical units.

Interestingly, although not a variable in the Horiba experiments, where participants were allowed to complete the sessions at their own pace, preliminary research suggests that certain time limitations of tasks actually contributes to more focused performance. In a study that compared the effects of no pressure, mild pressure, and severe pressure (as stimulated by no, low and high time constraints, respectively), Walczyk, Kelly, Meche, and Braud (1999) found that the slight stress as exerted by mild time constraints had a significant positive effect on the reading comprehension of 89 adult students during a task that involved the use of a sentence verification technique. They also noted, however, that previous research found that the application of severe time constraints caused participants to be "less likely to think critically about text information" (p. 157).

Cognitive Processing in Reading

In the previous chapter I described the essential components that make up the reading process, i.e., bottom-up, top-down, and interactive-compensatory processing. This was laid as a theoretical foundation for the present study since research in the cognitive processes involved in reading is mainly interested in "the manner in which humans acquire, interpret, organize, store, retrieve, and employ knowledge" (Smith, 1973, p. 2) and how different people come to comprehend and work with information and knowledge gained from text (Daneman & Carpenter, 1980; Kintsch, 1994).

In the past 16 years a few groundbreaking collections have been put together focusing on cognitive research of Chinese and other Asian languages (Chen, 1997; Chen & Tzeng, 1992; Liu, Chen, & Chen, 1988; Peng, Shu, & Chen, 1997; Wang, Inhoff, & Chen, 1999). As highlighted in the previous chapter, the specific areas of this cognitive research in terms of reading have mostly been character and character component identification and processing, word recognition, and to a lesser extent sentence processing and extended text comprehension. Miscue analysis is also a common way researchers try to probe into the mental processes of readers (Tien, 1983; Xu, 1998). Very often, however, this type of research takes the form of experiments in response latency as measured by reaction time (Chen, 1992; Shen & Forster, 1999), or more directly with eye movement data collected with eye-tracking equipment, and more recently in conjunction with the field of neuroscience using such apparatus as fMRI (functional Magnetic Resonance Imaging).[2]

Of direct relation to the present study is research utilizing eye-tracking. Perhaps especially since Gough's (1972) detailed educated guesswork about what goes on in the mind during one second of reading, based in part on eye-tracking data, contemporary researchers have increasingly tried to pick apart and describe the actual stages involved in the reading process (Ferreira, & Henderson, 1990; Henderson & Ferreira, 1993; Pollatsek, Raney, Lagasse, & Rayner, 1993; Rothkopf, 1978), even down to attempting to explain what the mind may be doing during 15 ms saccades. Even events such as pupil dilation have been determined to be another measure of cognitive processing (Just & Carpenter, 1993). More recently, Hyönä, Tommola, and Alaja, (1995) discovered that pupil dilation increased in relation to both hearing difficult words in simultaneous translation task, and repeating words back in the non-native language, suggesting that pupilary response is a good indicator of variations in processing load. This area of research will be discussed in further detail below.

[2] For an introduction to this fascinating area of research, see, e.g., Chee, Tan, and Thiel (1999) and Tan, Liu, Perfetti, Spinks, Fox, and Gao (2001).

Psycholinguistic Processes in Reading

In addition to being researched as a cognitive process, reading is also looked at as a psycholinguistic process (Goodman, 1968, 1994; Just & Carpenter, 1987; Smith, 1973, 1994). Smith (1973) identifies two kinds of information: visual information and nonvisual information. Visual information is external, that which is seen, while nonvisual is internal, that which is known. He explains that

> The first contribution of psycholinguistics is that there is a trade-off between visual and nonvisual information in reading—the more that is already known "behind the eyeball", the less visual information is required to identify a letter, a word, or a meaning from the text....
>
> The second major psycholinguistic contribution to reading is that there is a severe limit to the amount of information coming through the eye that the visual system can process. In other words, the trade-off between visual and nonvisual information is critical. The reader who relies primarily on visual information will simply overload his visual system. (p. 7)

Six years earlier, Goodman (1967) put forth the hypothesis that "reading is a psycholinguistic guessing game". One issue with this hypothesis is its limited applicability to lower-level foreign language learners, especially of a language like Chinese. Unlike learners of alphabetic foreign languages, generally with a much higher incidence of cognates, or at least a familiar alphabet, beginning learners of a morphosyllabic language like Chinese have little recourse to guess from surrounding context.

As discussed in Chapter 2, Kolers (1970) identified three stages in the reading process. Hoosain (1991, pp. 89-90), however, identifies four steps in the process of reading. The first step is the "analysis of the sequential visual structures". The second step is the "linking of visual and auditory-linguistic structures". The third step Hoosain identifies is the "establishment of regularities in grapheme-phoneme correspondences." The fourth step is "grouping words into larger units such as phrases and sentences". There is a final step in the process that follows even the fourth step and that is the step of comprehension (p. 89). These four steps to a certain extent parallel Goodman's (1976) "psycholinguistic guessing game", whereby a reader will take steps towards understanding a text based on trial and error and negotiation of visual and nonvisual information.

Hoosain's first step, analysis of sequential visual structures, is essentially the first instant of exposure to a text where the mind asks, "what are we looking at?" This identification of the text can confirm the expectation—for a native English speaker—that what will be seen are the familiar shapes of English letters, or it can inform the reader that what is being looked at is something that will require a different approach and different set of strategies. Once the mind has identified the type of text, it then moves on to Hoosain's second step, linking of visual and auditory-linguistic structures. For a native English speaker, the sound-symbol correspondence is transparent (albeit imperfectly). For a student of Chinese as a second/foreign language (CL2/CFL), for example, once they have identified the text as being in Chinese characters and not English, which they can do after very little training (Wang, Perfetti, & Liu, 2003), the next step they will take is try to assign sound values to the first characters that look familiar.

If we compare English and Chinese in terms of lower-level versus higher-level processing, as discussed briefly in the previous chapter, we might note that English could be characterized psycholinguistically as more of a top-down, knowledge-driven orthography and Chinese as a characteristically bottom-up, data-driven one. Based on the studies he reviewed, Hoosain (1991, p. 165) identifies Chinese language properties that, when compared with

English, seem to indicate that the Chinese, as a reflection of the way the language is read, "are more inclined to engage in tasks with well-defined goals and less in open-ended expression" (p. 164). Such properties, of perception in particular, that link reading to more typically bottom-up processing include visual acuity, directional scanning, perceptual unit, phonological access, phonological recoding, and lexical access (p. 165).

Eye-Tracking Research

Introduction

The two main ways humans take language in are by eye and ear. The key areas of the eye that actually acquire image signals are the fovea and the parafovea, as shown in Figure 4. The fovea has a 1- to 2-degree retinal area and the best acuity (sharpness). The parafovea surrounds the fovea occupying another 1 to 2 degrees of retinal area, but with much lower visual acuity. The cornea is the transparent, dome-shaped window covering the front of the eye, and is only about 0.5 millimeter thick. The cornea protects the pupil, the iris, and the lens. The pupil is at the center of the iris, with the lens just behind them.

The eye, it was discovered early on, does not acquire information smoothly, but rather "jumps" from place to place, stopping along the way to pick up new and additional information. The jumps are called saccades and vary in duration depending on length (the further the jump, the longer it takes to get there). The stop made as the eyes move along a length of text are called fixations. While saccades are generally very short in duration, usually less than 15 milliseconds (Just & Carpenter, 1987), fixations last much longer, with the average fixation lasting about 250 ms. Of course, depending on the task, fixations can vary widely, from quite short (100 ms) to much longer (over 600 ms). Fixations purportedly have three possible

RECEIVED

AS

FOLLOWS

functions: first, transmission of information can take place while the eyes are at rest; second, they allow the eyes to rest by essentially interrupting information intake; and third, by halting the intake of new or additional information, they allow the cognitive processes to piece together and interrelate the information that has already been taken in for comprehension (Kess & Miyamoto, 1999).

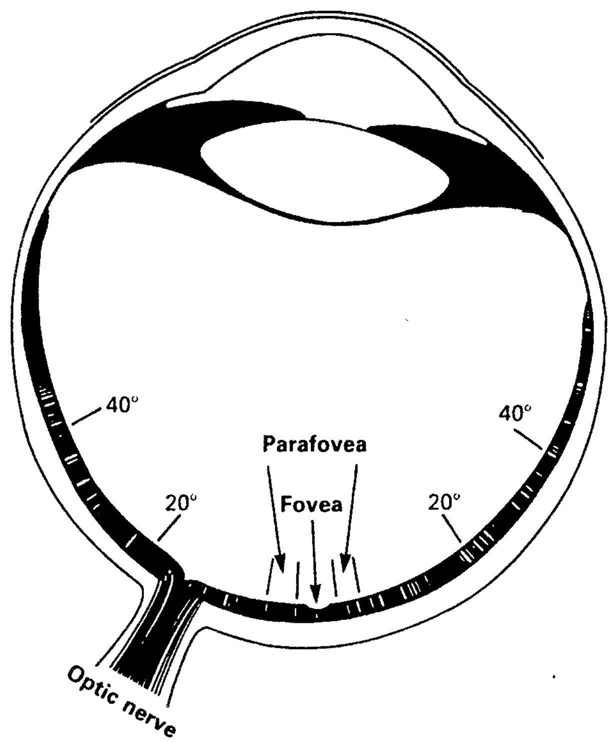


Figure 4. Cross-section of the eye. This view from the top shows the location of the fovea and the parafovea, as well as degrees of retinal area. (from Just & Carpenter, 1987)

Eye-Tracking Research of English L1 Reading

As noted in the previous chapter, most eye-tracking studies to date have focused on the English language (Fisher, 1975; Hyönä, 1995; Hyönä, Niemi, & Underwood, 1989; Just & Carpenter, 1978; Kliegl, 1981; Kowler, 1990; Yarbus, 1967). General eye-tracking studies show a direct relationship between mental processing and eye movement (Carpenter & Just, 1983; Daneman & Reingold, 1993; Just, & Carpenter, 1980) and the inferences that can be made (Rothkopf, 1978).

Points of fixation and individual fixation times and gazes are believed to indicate information gathering—longer for unfamiliar items or more complex concepts, shorter for more familiar items or less complex concepts. Areas of interest—compounds of time and occurrences/frequency of visitation—indicate something that has captured the reader's attention. Again, increased dedication of time and higher frequency of visitations indicate either unfamiliar items or complex concepts while less time and lower frequencies indicate familiar items or simpler concepts. Saccades are now believed to be when cognitive processing occurs. There is a special type of saccade that goes backwards and is called a regression. Regressions are said to occur when something happens during processing that stimulates a reconfirmation of information. Longer (in time or distance) saccades indicate integration of more, or possibly more dense, chunks of information, while shorter jumps indicate the piecing together of smaller bits of information.

Skilled, or experienced, readers are defined as those that accomplish the reading task quickly and efficiently (including success). The smaller the total time per test item and the more accurate the decision on an accompanying multiple-choice item, the more skilled the reader is judged to be. On the other hand, The larger the total time per test item, on average, and the less accurate the decision on an accompanying multiple-choice item, the less skilled the reader is judged to be.

Bottom-up processing is evidenced by the predominance of longer fixations and short and frequent saccades. Top-down processing appears as shorter fixations and longer and less frequent saccades overall. Skilled readers are defined as those that employ both top-down and bottom-up approaches (also called strategies), in apparently general balance to achieve efficient and successful comprehension, while unskilled readers are identified as those that employ exclusively only bottom-up or top-down strategies, slowing down their reading processing, and decreasing the likelihood of successful comprehension.

Of course, there are limitations to the so-called Eye-Mind Assumption, as pointed out by Underwood and Everatt (1992). They explain that there is the assumption that "the direction of our eyes indicates the contents of our mind" (p. 111). Based on eye-tracking data from word identification, syntactic processing, and semantic connections, Underwood and Everatt suggest that many eye movements occur independently from what the mind is doing. Also, what are known as parafoveal events—information such as end of line, etc.—may be picked up by the eye guidance mechanism in order to determine the next fixation location. Recently, Anderson, Bothell, and Douglas (2004) conducted a study that reiterated Underwood and Everatt's (1992) point and concluded that "eye movements do not necessarily reflect mental processes, but they do reflect ongoing processes to the extent that the processes depend on the encoding of information" (p. 230). That said, let us look at some examples of eye-tracking research in relation to Chinese.

Eye-tracking Research of Reading in Chinese

General eye-tracking research has a history of over 125 years, beginning with Emile Javal in 1879 (as cited in Bernhardt, 2001, Just & Carpenter, 1987, and Samuels & Kamil, 1984). The earliest eye-tracking studies that looked at reading Chinese limited their studies to native readers of Chinese and were typically conducted using tachistoscopic instruments, the apparatus of the day. Tachistoscopes generally present single lines of text scrolled through a limited field of vision or moving window. The setup allowed researchers to take pictures of subjects eyes at regular intervals, generally once every 200-250 ms (one-fifth to one-quarter of a second). In terms of Chinese, these types of studies mainly looked at differences between horizontal and vertical reading (Chen & Carr, 1926; Miles & Shen, 1925; Shen, 1927; Tu, 1930) and/or reading speed and comprehension during silent reading (Miles & Shen, 1925; Wang, 1935). Perhaps due to the moving away from the more traditional vertical layout of text to the now more common horizontal text layout, studies looking at vertical text processing of Chinese have also seen their time come and go.[3]

[3] This is definitely true for mainland Chinese (PRC) publishing practices. Taiwan recently passed a law saying that official printing should run horizontal from left to right (Ko, 2003). Popular literature in Taiwan, as well as in Hong Kong, however, may still be equally likely to be printed vertically as horizontally. Chen (2002) also notes that many Chinese media in the United States are also making a changeover to horizontal printing of text, due in large part to the increasing mainland Chinese readership. In Japan, however, there still remains a sense of tradition bound to the printing of vertical text, and so is still relatively accessible. Korean is still sometimes published vertically as well. Thus, recent studies of Japanese and Korean reading (and Sun, Morita, and Stark's 1985 study of Chinese reading) have included comparison between horizontal and vertical reading rates (e.g., Osaka, 1993; Seo & Lee, 2002), generally finding reading horizontally more efficient.

Despite Tzeng, Hung, and Garro's (1978) concluding remarks on the state of the art 25 years ago,

> Whether differences in writing systems would result in different information processing strategies during reading has been a major concern of reading specialists for a long time (Gibson and Levin 1975). However, most experimental investigations of this problem in the past have always focused on some gross measures such as reading rate (i.e., number of words or characters read per minute) and eye-voice span (the distance, usually measured in number of words ahead of the voice when reading aloud). The results of such cross-language and cross-writing system comparisons yield nothing but very crude guesses as to reading behaviors in general. Little can be said about whether there are indeed different problems faced by beginning readers in different language environments. (pp. 300-301)

the development of eye tracking equipment since then has allowed researchers to better pinpoint specific stages and has helped to identify strategies and behaviors in the reading process. While early studies used equipment that may have provided only limited results, they did reveal many characteristics that are at least peripherally useful to keep in mind as we prepare to look at more modern eye tracking studies of reading of Chinese, either as an L1 or CL2/CFL. For example, Miles and Shen (1925) looked at how 11 subjects read Chinese both horizontally and vertically and found that they performed better during vertical reading. They conclude that, "In general, we may say that the process of reading Chinese is essentially the same as for reading English" (p. 360), a conclusion echoed by Wang (1935, p. 539) ten years later. Miles and Shen also noted that

> As the horizontal method of reading is of comparatively recent origin as an alignment for Chinese texts, our subjects have undoubtedly had more experience in vertical reading. It is conceivable that long continued practice *may* have

> more than balanced any inefficiency for the vertical axis.... (1925, p. 362, italics as in the original)

Table 1 presents summaries of five early eye movement studies related to the reading of Chinese. We can see from these studies the effects of the increasing amount of horizontally-aligned printed matter beginning in the first quarter of the 20th century. While the statistics provided were not necessarily conclusive (e.g., Tu, 1930) and the arguments and conclusions seemed motivated more by some political agenda (in particular, Shen, 1927), many of these studies discussed similar factors that may have influenced the observed data on the reading processes, including individual differences, attitude, and age.

These early studies did not attempt to measure the effective visual field during reading of Chinese. The "effective visual field", alternatively called the "perceptual span", is defined as the visual field associated with the images of objects focused onto the fovea, and is usually measured in degrees of angle or number of characters or words.

An example of visual span during reading Chinese from Chen and Tang (1998) is shown in Figure 5. They also showed that revealing a character to the left of a fixated character did not increase reading efficiency. Perhaps even more significantly they found that during Chinese reading the effective visual field is "asymmetric to the right of fixation and extends from the fixated character to about two characters to its right" (Chen & Tang, 1998, pp. 251-252).

A previous study by Inhoff and Liu (1997) used an eye-movement-contingent display change technique in a window of varying size. Data from the first experiment of seven native Chinese participants suggested to them that "readers obtained useful visual information from at least three characters to the right of fixation" (p. 251). Their data also showed that "visibility of text to the left of fixation appeared to provide ... little, if any, benefit" (p. 251).

Table 1. Early eye movement studies for reading Chinese (1925-1935).

Study	Subjects	Materials	Task	Results
Miles & Shen (1925)	11 NS (college) (authors: Miles and Shen)	magazine articles: a) English (paragraph) (9pt, 14-15 words/line; 12pt., 13 words/line) b) Chinese (~12pt, ~14pt): vertical (35 chars./ line); horizontal (23 chars./ line)	read silently for comprehension comprehension questions asked after each passage	Vertical and horizontal eye movements are quantitatively different. Generally more regressions during reading of Chinese than English (due to saccadic errors: inadequate perception, faulty attention, misinterpretation). Reading processes for English and Chinese essentially the same.
Chen & Carr (1926)	64 NS	a) 2 Chinese passages, prose, vertical, horizontal b) mixture of Chinese, English, and Arabic numerals	a) read each passage 2x, verbal recall b) cancellation	Vertical reading faster and more accurate than horizontal. More training and practice in reading English (horizontal) weakened Chinese vertical and strengthened Chinese horizontal reading skills . Much of existing variance in data attributed to individual differences.

Note. NS= Native Speaker of Chinese, NNS = Non-Native Speaker of Chinese (i.e., native English); chars. = characters.

Table 1. (Continued) Early eye movement studies for reading Chinese (1925-1935).

Study	Subjects	Materials	Task	Results
Shen (1927)	13 NS	12 Chinese passages, varying styles (~500 chars./each) (vertical, horizontal; ~14pt, 32 chars./line)	read silently for comprehension comprehension questions asked after each passage	Vertical and horizontal eye movements are quantitatively different. Vertical reading faster and more accurate than horizontal. Generally more regressions during horizontal reading of Chinese.
Tu (1930)	8 NS 30 NNS (college) 74 NNS (gr. 5-7)	Exp. 1 "nonsense" and "sense" Chinese characters on flash cards: "nonsense" (3, 4, 5, 6 char. series, vertical, horizontal); "sense" (4, 5, 6, 7, 8 char. series; vertical, horizontal) Exp. 2 symbols (for NNSs) on flash cards	timed exposure, written recall, written report ("introspection")	Exp. 1 Horizontal reading faster than vertical Vertical and horizontal equally accurate Exp. 2 Speed and accuracy increase along with grade level and age Horizontal reading more efficient than vertical.
Wang (1934, 1935)	71 NS	10 120-word passages (5 literary, 5 vernacular, 20 characters/4 in. line), vertical only	normal, rapid, careful reading followed by a 10-item T/F test: Exp. 1 photographic Exp. 2 test	Vernacular read more quickly and comprehended more than literary. Quantitatively more and longer fixations per line of literary than vernacular text. Reading processes for English and Chinese essentially the same.

Note. NS= Native Speaker of Chinese, NNS = Non-Native Speaker of Chinese (i.e., native English); chars. = characters.

The second experiment of eleven native-speaking participants produced similar conclusions, where the effective visual field was found to be "no more than three characters to the right of fixation" and visibility of one character to the left provided marginal benefits, but that of two characters robust benefits (p. 257). Similar results were presented by Inhoff and Liu (1998).

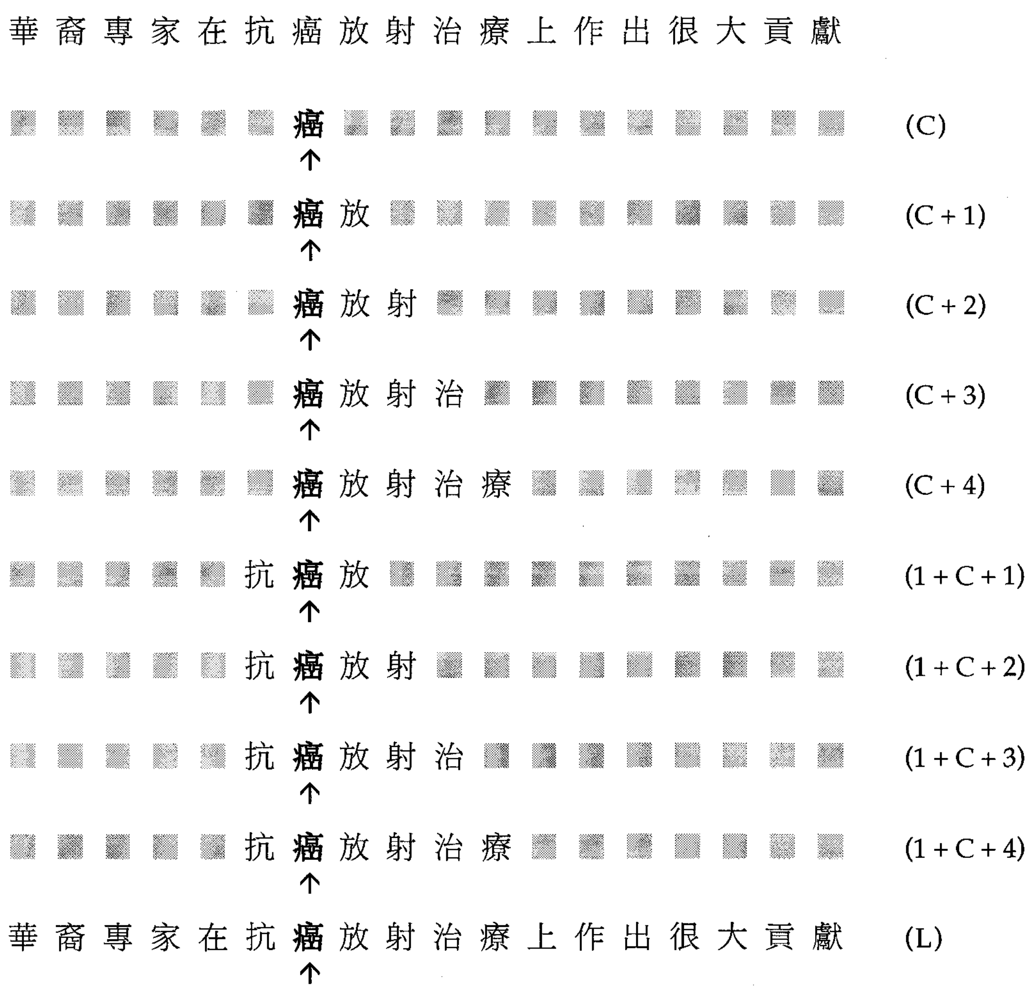


Figure 5. Visual field in reading Chinese. Fixated characters are marked with arrows, and each condition is labeled in parentheses. (from Chen & Tang, 1998)

A 1985 study by Sun, Morita, and Stark compared eye movement behavior details of native Chinese readers reading with that of native English readers reading English. In their study, Sun et al. (1985) gathered eye movement data from twelve native Chinese readers and eight native English readers; all readers had science backgrounds. The reading materials for both groups of subjects were semipopular scientific texts from *Scientific American* either in the original (English) or in translation (Chinese). Based on the eye tracking data, they determined that the reading span (forward jump) for skilled Chinese readers reading Chinese was approximately 1.7 Chinese "words" and 1.8 English words for native English readers reading English. For their calculations and analyses, Sun et al used an equivalency factor of 1.5 Chinese characters being equal to one English word.[4] This number was arrived at by measuring the number of Chinese characters in the translated text against the number of English words in the original text. In terms of words per minute (wpm), figured with the factor of 1.5, they found

[4] Sun et al (1985) figure that "on the average, the Chinese character group may be composed of approximately 1.5 characters, with the usual size being, perhaps, between one character (55%) and two characters (40%), with rather few three-character groups (5%)" (p. 502). Note how this differs with Wong and Weber's (1986) estimate mentioned in the previous chapter, who calculated the number of two-character words to be between 60% (p. 42). Unfortunately, neither Sun et al nor Wong and Weber tell us what they base their estimations on, so we can only assume these are subjective yet educated guesses. The breakdown of Tsai's (1996) collection of 134,450 unique entries, while admittedly rough, gives us a more reliable approximation. Based on his data, single-character "words" account for about 10%, two-character words about 54%, and three-character words about 19%. So-called "four-character phrases" (成語, chéngyǔ) and "idioms" (俗語, súyǔ), claim the majority of combinations of four or more characters that make up the remaining 17%. These are not easily dissected into smaller chunks, though it could probably be said that most four-character phrases are composed of four single-character words. Of course, these estimates only paint a picture of the language at large; none of this speaks directly to representative frequency and recurrence within a given text depending on genre or text type.

that the Chinese readers read (horizontally) at an average rate of 385-390 wpm and the English readers read at 380 wpm (p. 505). Similar results were reported in a later replication of this study by Sun and Feng (1999).

In terms of refixations and gaze durations, Yang and McConkie (1999) found striking similarities between Chinese and English readers. Refixations are defined as immediate second fixations, while gaze durations are defined as "the sum of fixation durations for the first consecutive sequence of fixations on the word" (p. 216). Yang and McConkie found that readers of Chinese and readers of English are equally likely to refixate on words of high complexity and low frequency. Similarly, gaze duration patterns were also affected by word complexity and word frequency.

By way of comparison, Stern (1978) noticed that "the informational packing density of Chinese per text line is considerably greater than that of English text" (p. 150). Because of this, Stern observed that his ten native Chinese readers reading Chinese made more saccades per line of text than the six native English readers reading English text. The Chinese readers averaged ten saccades per line of text while the English readers averaged four saccades per line (p. 151). Similar results were discovered by Peng, Orchard, & Stern (1983), who concluded that their 23 native Chinese generally "utilize significantly more fixation pauses, they read fewer lines without regressions, and they spend more time per fixation when reading English text"(p. 101).

Finally, Tang, Yeung, and Chen (1997), based on the results of a moving-window experiment with 18 native readers of Chinese concluded that "the effective visual field in the reading of Chinese may not be a fixed entity, but may vary according to, not only the nature of the parafoveal information, but also the contextual information that has been previously obtained" (p. 283).

Reading Chinese as L1

Compared with CL2/CFL eye-tracking reading research, which is still in its infancy (if not prenatal), much more research has looked at reading Chinese by native readers (Chen, 1992; Chen & Tzeng, 1992; Stern, 1978; Sun, 1993; Yang, 1994; Yang & McConkie, 1999). Most of that research has mainly been eye-tracking studies that deal with the recognition of individual characters rather than extended text although some research is beginning to be expanded into the area of how larger chunks of information are processed. The early studies conducted between 1925-1935 mentioned previously are the first examples of eye-tracking research on native readers reading Chinese.

Contemporary eye tracking research of the reading process in Chinese examines specific areas such as lexical and sublexical processing (Inhoff, Liu, & Tsang, 1999) and directional scanning preferences (Chen & Chen, 1988). Tsai (2001) has even gone from data collected from human and designed a computer model that can be used to predict eye movement behavior in different reading situations based on such factors as lexical frequency and a lexical parsing algorithm.

Reading Chinese as L2/FL

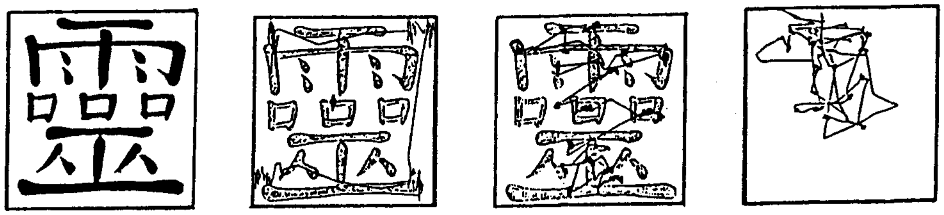
As we saw above, there has been a wealth of studies on native readers of Chinese. More recently, however, while eye tracking studies have still remained focused on native readers of Chinese—with some making comparisons with native readers of English (Chen &

Chen, 1988; Inhoff & Liu, 1998)—a few studies have made inroads to CL2/CFL reading research.

There have not been that many eye-tracking studies looking at the reading of Chinese as a second or foreign language (CL2/CFL). There are many possible reasons why this may be so, the most likely being the limited number of potential participants available to participate in such studies. Despite this dirth of resources, since the 1980s, the field has seen increased research activity in this area.

In a study that looked at reading times of text with and without spaces between word boundaries, Everson (1986) concluded that beginning non-native readers of Chinese showed little or no influence due to word spacing, while advanced non-native readers seemed to be negatively influenced by the artificial spacing. In other words, the advanced non-natives took longer to read the passage than the beginning non-native readers, and their eye movement data indicated more fixations than those exhibited by the beginning non-natives. Further, native Chinese readers' performance did not suffer because of the artificial spacing, and Hayes (1990) echoes Everson (1986) when he suggests this was possibly because of "the robustness of their reading strategies" (Everson, 1986, p. 85; Hayes, 1990, p. 32).

The main subjects in Sun's (1993) study were native speakers divided into three groups: primary school students, high school students, and graduate students. For comparison, the study also included two other groups: one consisted of presumably monolingual English speakers and the other of 10 Chinese-English bilinguals. The variety of reading materials included short paragraphs in both Chinese and English, short Chinese primary school texts written in both Chinese characters and Pinyin, individual characters, Chinese numerals. All five groups were shown individual Chinese characters.



(a) (b) (c) (d)

Figure 6. Three Eye Movement Patterns. (from Sun, 1993)

Sun's (1993) study clearly showed the difference between three different readers of the Chinese character 靈 (líng, 'soul'), a character made up of 24 strokes and combining seven elements (雨 ['rain'], 口 ['mouth', x3], 工 ['labor'], 人 ['person', x2]), and divisible into two components (top and bottom—the phonetic 霝 [líng] and the semantic 巫 [wú; 'wizard, witch']). In frame (b) above, Sun shows us how a native reader familiar with the character reads it, going from top-left to top-right, down to bottom-right, and finishing at bottom-left. In frame (c) we are shown how a (native?) reader unfamiliar with the character reads it by going through and identifying known elements, with fixations tracked generally proceeding diagonally from top-left to bottom-right. Finally, in frame (d) we see the eye-tracked data of a non-native reader with random fixations (as if the character were not even there). Sun determined that degree of familiarity with a character guided eye movements. More importantly, Sun concluded that:

> The reading eye movement data from bilingual subjects showed there are no real Chinese-English bilinguals. No one really possesses equal skill for reading Chinese and English quantitatively. For each bilingual subject, only one language is dominant, and the other has been suppressed. (Sun, 1993, p. 254)

Matsunaga (1994) included graphics of sample eye-tracking data. Recall that Matsunaga's study involved native and non-native readers of Japanese, with the native readers acting as her control group. She was also looking at phonological encoding, whereas the present study is looking at working memory and cognitive load. Despite the different focus of this study, the Matsunaga data revealed distinctions in the reading patterns of native and non-native readers. Figure 7 is an example of a native reader of Japanese reading a text containing 45% Kanji.

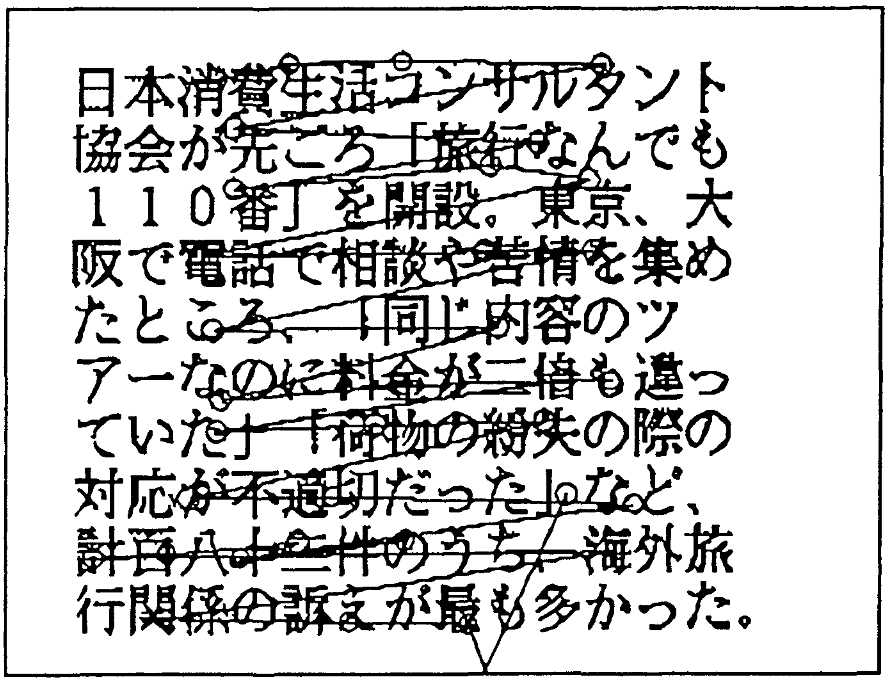


Figure 7. Native Japanese Reading Sample. (130 symbols, 45% Kanji). (from Matsunaga, 1994)

In this example we notice the relatively fluid progression of fixations from start to end. As discussed above, Matsunaga (1994) noted that despite the existence of homophones, native readers encountered little difficulty in read-through or disruption in reading speed, and virtually no regressions. In contrast, Figure 8 shows a non-native reader reading the same passage.

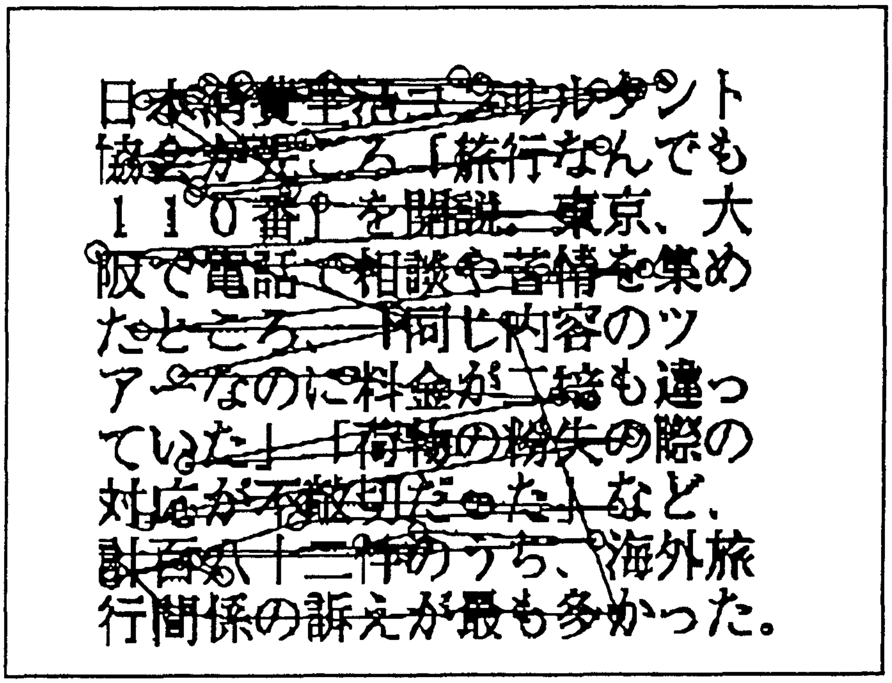


Figure 8. Non-native Japanese (JFL) Reading Sample. (130 symbols, 45% Kanji). (from Matsunaga, 1994)

In this example we notice a greater number of regressions due to, as explained by Matsunaga (1994), difficulty dealing with homophones. Taking both Figure 7 and Figure 8 into consideration for my purposes here, we can also infer increased cognitive load as discussed earlier in this chapter. Again, looking at the native reader's progress in Figure 7, we see from the eye-tracking data that working memory is not being taxed in the native Japanese reader as it is for the non-native reader in Figure 8.

Summary

The studies discussed above provide sketches of the available theoretical work on working memory and research on eye movement as they relate to second or foreign language reading in Chinese. As evidenced by the selection, the majority of studies mostly look at memory tasks in relation to reading skills at the component, single-character, and two-character word levels, or occasionally at the sentence level. As can be seen from the last section on eye-tracking research during reading of Chinese as a second or foreign language, very little attention has been paid to this area of investigation. Clearly, then, what is called for are more studies that look at the role memory, especially working memory and long-term working memory, plays in reading for comprehension tasks. These types of studies could also help us to understand which types of experiences various students have with different types of text. Further, Bai (1997) explains that "text structure refers to ways that authors organize their ideas. Examples of text structure include listing, comparison and contrast, cause-effect, and persuasion" (p. 31). If students are failing to comprehend particular types of text, this may give an indication of where methodologies currently employed in teaching students to read Chinese might need modification. Finally, following Just and Carpenter's (1980, 1992) research and suggestions, future research utilizing eye tracking equipment could shed light on the subject of learners of Chinese as a second or foreign language and the potentially problematic areas they encounter as they develop their L2 reading skills and strategies. The next chapter will describe the methodology of the present study which was conducted with this type of investigation in mind.

CHAPTER 4:

METHODOLOGY

Introduction

In this chapter I describe the research methodology and design of the study. This includes participant details as well as a description of the materials used (both equipment and test instrument). The main instrument and source of reading material for this research was the *Computer Adaptive (Proficiency) Test for Reading Chinese* (CATRC; Yao, 1994), so a brief look at that test specifically and computer adaptive testing in general is also provided. This will be followed by how the data were gathered, as well as what tools were used in the process of analysis.

Participants

Prior to any of the research conducted in this study, review of the study was conducted and approval for the project was granted by the University of Hawai'i Committee on Human Studies/Institutional Review Board (CHS/IRB) (cf. Appendix A).

All participants were college-level readers (including undergraduate, graduate, and faculty) who had either normal or corrected-to-normal vision. Their ages ranged from 18 to 55 years. They were all paid US$15 for their participation. The data were gathered between March and August, 2003.

Table 2. Participant background data summary.

ID#	ACTFL-R	Time-S	SA-Exp	SA-Loc	SA-Time	Use	Class
NNI01	IM	1	no	NA	NA	Reg	Jr
NNI02	IM	1	yes	Taiwan	3-4	Reg	Sr
NNI03	IH	2	yes	PR China	<1	Never	PhD
NNI04	IH	4	yes	PR China	<1	Seldom	Sr
NNI05	IH	5	yes	PR China	2-3	Reg	PhD
NNS01	S	4	yes	1,2	5	Reg	Jr
NNS02	S	3	yes	PR China	1-2	Reg	Sr
NNS03	S	5	yes	1	1-2	Reg	MA
NNS04	S	5	yes	PR China	<1	Often	Prof
NNS05	S	3	no	NA	NA	Never	PhD
NNS06	S	5	yes	PR China	5	Reg	PhD
NNS07	S	5	yes	1,2	5	Reg	Prof
NS01	S	6	yes	PR China	5	Often	PhD
NS02	S	8	yes	Hong Kong	8	Reg	Jr
NS03	S	11	yes	PR China	<1	Reg	MA
NS04	S	19	yes	PR China	<1	Often	MA
NS05	S	10	yes	PR China	<1	Reg	MA
NS06	S	14	yes	PR China	1-2	Reg	MA
NS07	S	11	yes	PR China	<1	Reg	MA
NS08	S	6	yes	PR China	1-2	Reg	MA
NS09	S	8	yes	Hong Kong	6	Reg	Sr
NS10	S	15	yes	PR China	2-3	Reg	PhD
NS11	S	9	yes	PR China	3-4	Often	Sr
NS12	S	27	yes	PR China	1-2	Reg	MA
NS13	S	<1	yes	Other	3-4	Never	Jr

Note. **ACTFL-R** = ACTFL Reading Level: IM = Intermediate Mid, IH = Intermediate High, S = Superior. **Time-S** = Time Spent studying Chinese (for native speakers of Chinese, Time Spent studying English): number of years. **SA-Exp** = Study Abroad EXPerience (for native speakers of Chinese, Time Spent living in the US). **SA-Loc** = Study Abroad (or native) LOCation: PR China (PRC) = People's Republic of China, NA = Not Applicable. **SA-Time** = Total Study Abroad Time: number of years, NA = Not Applicable. **Use** = present estimation of language Use: Reg = Regularly. **Class** = Year/Level in college: Jr = Junior, Sr = Senior, MA = Master's, PhD = Doctoral, Prof = Professor.

Data was gathered from a total of 25 participants[1] (12 native English speakers [i.e., non-native Chinese] and 13 native Chinese speakers) divided into three groups[2]: Group 1 consisted of 5 non-native students at the Intermediate level. Group 2 was made up of 7 non-native Chinese speakers at the Superior level. Group 3, the control group, consisted of 13 native readers. The Intermediate level participants had an average of two years of Chinese language study, while the non-native Superior level participants had an average of four years of formal language study. The native Chinese readers were of near-native proficiency or above in reading English.[3] Table 2 provides a summary of participant background data (as gathered mostly at the beginning of the CATRC), including assessed reading level (ACTFL-R) and level in college (Class).

Materials

Below I will briefly describe the equipment and test instrument used for this research. The equipment used was the ASL (American Science Laboratories) 501 series head-mounted eye tracker. The instruments included a set of two short reading passages (1 Chinese, 1 English) used during calibration and Yao's *Computer-Adaptive (Proficiency) Test for Reading Chinese* (CATRC), both presented on a computer screen.

---

[1] An additional 20 participants had to be rejected or otherwise removed from the participant population due to insufficient English language ability (for the Chinese NSs) or Chinese reading proficiency rating determined to be at either the Novice or Advanced levels (for the NNSs).

[2] The grouping was done based on assessment from the CATRC (see below).

[3] Based on a mean TOEFL score of 619 (n=10, min.=530, max.=663, *SD*=10.1) and a mean GRE score of 1982 (n=6, min.=1500, max.=2230, *SD*=6.2).

*Equipment: **Head-mounted Eye Tracker***

The ASL 501 head-mounted eye tracker (Appendix D) measures a person's eye line of gaze with respect to their head. The eye tracker is actually part of a larger, complex system of components. The system also includes the ASL 5000 control unit; a head-mounted sensor and halo, magnetic transmitter, and electronics unit (Flock of Birds); two pan/tilt optics cameras (one for eye, one for scene); two black and white video monitors (one to receive video from each camera); and a computer for interface between the eye tracker and data recording software (EYEPOS).

A ring of near infrared LEDs reflect off a monocle and track the movements of the eye (the eye tracker can be set up for tracking the movements of either the left or the right eye; in this study, the left eye was tracked). The eye tracker transmits the coordinates of the point of gaze (calculated from the center of the pupil and the corneal reflection) of the stationary scene plane (i.e., the computer screen) to the EYEPOS software on the computer (in this case, a PC running under Windows 2000). The equipment also measures pupil diameter. ASL's EYENAL and FIXPLOT software were used to process the data (see Data Analysis section below).

Instrument

Calibration Passages

Two reading passages were included in the calibration process at the beginning of each participant's session (Appendix E). The first passage was two paragraphs from an intermediate-level Chinese reader modified for length to fit two screens. The second passage was in English, also two paragraphs, and was a description of what was about to take place in the experiment and included a few words of encouragement for the participant's benefit. Each paragraph filled one computer screen (Apple 17-inch Studio Display), using a 34 point Chinese font ("Fang Song") with 1.25 line spacing, and a 24 point English font ("Palatino"), set in full text justification with 1.5 line spacing, thus providing no less than a 0.5-inch border around all four sides when displayed in Full Screen view.

Computer Adaptive Tests (CATs)

The use of a computer-adaptive test as the instrument for this research was chosen for two main reasons. The first was that it provided a more equal and standardized measure of participant reading ability and allowed participant assignment to groups based on these levels. After years of research (Bringjord, 2001; Brown, 1997; Dunkel, 1999; Hulstijn, 2000), concerns about possible performance differences between computerized and so-called pencil-and-paper tests have been addressed, and in general the conclusions have been that results are comparable. These studies have also compared the design of computer-adaptive tests with the designs of paper-and-pencil tests, including important global factors such as test validity and reliability, as well as local factors such as item type and task (Brown, 1996; Muller, 1972; Wolf, 1993), and matching content and proficiency level (Hughes, 1989; Mecartty, 1998).

Other researchers have looked at the cognitive and psycholinguistic experience of the test taker (Jennings, 2001; Kobrin, 2000). These studies have also shown that, especially with the increasingly ubiquitous nature of computers in our everyday lives—from desktops and laptops at home, school and in the workplace to information kiosks in public areas—the same fears that overcome test takers during paper-and-pencil tests are an issue with test takers during computerized tests. However, the research does not indicate additional demands on attention with computerized tests. If anything, the anxiety associated with taking the test itself seems to outweigh the apprehension associated with facing a computer screen.

A second reason for choosing to use the CATRC (described below) was the fact that it was readily available. Redesigning or creating a similar test with software such as PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993), while possible, would have created an unnecessary obstacle in the research schedule and would not have substantially added to the data collected.

Computer-Adaptive (Proficiency) Test for Reading Chinese (CATRC)

The *Computer-Adaptive (Proficiency) Test for Reading Chinese*, or CATRC, is, as its name implies, a computer-adaptive test (CAT) that tests examinees' proficiency in reading Chinese. The test used in the present study, developed by Yao, Ning, and Chi, is a Macintosh platform-only set of HyperCard stacks. Altogether there are currently over 600 items in the CATRC, roughly 50 items per level (former ACTFL scale of nine levels from Novice-Low to Superior) with more items available at the Superior level. These items cover the four main proficiency levels (Novice, Intermediate, Advanced, and Superior), including all sublevels (i.e., Novice-Low, Novice-Mid, Novice-High, etc.), as described in the ACTFL Chinese Proficiency Guidelines (1987).[4] Each level's set of items are stored in a HyperCard file called a "stack." The test randomly selects items from the different stacks to present to each

---

[4] The latest version of the Macintosh program, CATRC 1998, as used in this research, was designed using the pre-revised ACTFL Reading Guidelines, which defined only Advanced and Advanced High sub-levels, does not include the three sub-levels Advanced Low, Advanced Mid, and Advanced High. Since this research looked only at participants at the Intermediate and Superior levels, this change does not directly affect the data collected.

examinee. When an examinee answers a question correctly, s/he is given an item of similar or greater difficulty; if s/he answers a question incorrectly the program then randomly selects by way of an algorithm an item of similar or lesser difficulty. Through this process the application identifies an examinee's reading level.

The CATRC has been shown to be reliable and valid for all levels ranging from Novice-Low to Superior. In particular, see Hayden (1998, 2003), Watanabe and Yao (1998), and Yao (1994, 1995). Watanabe and Yao (1998) examined the validity of the CATRC by comparing it to the Preliminary Chinese Proficiency Test (Pre-CPT). They found that the CATRC correlated highly with the Pre-CPT (Watanabe & Yao, 1998). In the end, the CATRC fits into Bernhardt's (1999) idea of the essentials of comprehension assessment, in that it gives stable information about the abilities of learners so that their strengths and weaknesses can be identified; it ranks and orders comprehenders in their performance abilities; and, most important, it makes clear what they cannot *do* or *understand* (p. 1, italics as in the original).

The change of delivery of the TOEFL (Test of English as a Foreign Language) from a paper-and-pencil test to a computer-based test raised concerns regarding (mainly construct) validity. A recent technical report by Taylor, Jamieson, Eignor and Kirsch (1998), concluded that there was no significant difference between scores of examinees with low computer familiarity and examinees with high computer familiarity.[5] While the CATRC seems to be

[5] It should be noted that all of the examinees in the Taylor et al (1998) study underwent a forty minute tutorial session.

very simple and user-friendly in its design, it could, however, also be subjected to similar investigation. Finally, there is currently an online web version of the CATRC under de-velopment.[6]

Procedures

The participants were all tested individually. They were asked to try to come at some time during the day when they would not have to rush to make their appointment and to allow for a 90 minute session. They were told that the session would most probably not last longer than 60 minutes, but this "comfort buffer" was desired to ensure participants did not feel rushed during the experiment and could devote their full attention to the task. Stern (1978, p. 148) mentions concerns of possible effects due to time of day and time on task. While he limited his sessions to two 45 minute sessions once in the morning between the hours of 8-11 a.m. and once again in the evening between the hours of 7-10 p.m., the sessions for the present study were generally between the hours of 7:30 a.m. and 5:30 p.m., and generally for no more than 60 min.

Participants were invited into the Tracker Lab (a dedicated room for the eye tracker equipment) and asked to read and sign the CHS/IRB consent form (Appendix B) and fill out the background information sheet (Appendix C). This generally took about five minutes per participant.

After the participants had completed the CHS/IRB and background information forms, they were then asked to have a seat in front of the Macintosh (G4, MacOS9) computer as I explained the general idea of the nature of the research. Participants were then shown the ASL 501 head-mounted eye tracker unit as I explained the key components of the equipment

[6] This is being done by Alex Chapin at Middlebury College. I understand that the Middlebury Chinese School uses it as part of their placement test procedures.

(Appendix D). This included showing them the two cameras and what I would be seeing from each one: the scene camera showing me what the participant was looking at on the computer screen and the eye camera showing the participant's left eye (which is necessary to see in order to calibrate the eye tracker system). Initial setup and calibration for each participant took between 5 and 15 minutes.[7]

Instructions

Calibration required the participants to look at an image of nine numbered points specifically arranged in a 3 by 3 grid that corresponded to the area of the visual field on the screen (Appendix D). After calibration participants were asked to orally read two short passages split into two paragraphs, each of which occupied one computer screen (1024x768 pixels) of text. The Chinese text was read first, and participants were able to choose between traditional or simplified Chinese character forms (depending on whichever form they felt more comfortable with). The English text as described above was presented next (see Appendix E).

After reading these passages, participants were calibrated again using the calibration image described above before beginning the computer-adaptive test (CAT). Once the CAT program was run, the first screen they were presented with was a card with additional background questions mainly related to language study background. When they reached the last background question participants were told of the structure of the CAT and how to proceed

[7] Several factors created difficulties during calibration, the most common being hard contact lenses. Something about the material of the lenses seemed to interfere with the infrared light reflected off the monocle.

through it, essentially by clicking the mouse on the choice that best represented the answer they wished to select. Example item cards from the CATRC, including the background card, may be found in Appendix F. Because the CATRC, as its name suggests, is a computer-adaptive test, participants varied widely on the length of time required to complete the assessment.[8] Adding this variability in time to that of the variability in time needed for calibration makes it difficult to come up with a representative average time for participant sessions.

Data Collection

Background data—including age, gender, other language exposure, etc.—was gathered via a Participant Background Information Questionnaire adapted from the one used in language courses in the Department of East Asian Languages and Literatures (EALL) at the University of Hawai'i (see Appendix C). The CATRC also begins with a short background card that generally asks about the test-taker's language study history (Appendix F).

Data collected from the eye tracker includes points of fixations, fixation durations, inter-fixation times, angular distance between fixations, and fixation sequences. Using the ASL software tools I was able to define areas of interest (AOI, collections of fixations in a given area), calculated gaze times (total fixation times per AOI), identified saccades within the Chinese text as well as saccades between Chinese text, the English cue, and the English question and answers.

For each participant, the CATRC keeps a record of the number of items tested, item ID (including level and item number), the correct answer for each item and the answer provided by

---

[8] A summary of participant test duration times can be found in Chapter 5, Table 3.

the participant, a tally of correctly and incorrectly answered questions and each corresponding level, start and end times, a final reading proficiency level assessment based on the number of items answered correctly, and the total number of questions asked (see Appendixes G and H).

The CATRC only provides total time on task (i.e., how long it took to complete the whole test by subtracting end time from start time), but not time on individual items. The eye tracker recorded data that allowed me to identify the start and end time for each item (by placing a mark in the record at the point where a new item was presented while the participant was taking the test). Participant answers can be compared with correct answers for each item. Looking at the eye-tracking data in conjunction with the CATRC data, especially during items which were answered incorrectly, may be helpful in determining which topics or linguistic aspects caused a participant difficulty.

Data Analysis

Data gathered from the CATRC provided descriptive statistics about each participant's performance on the test, including total test time duration, correct and incorrect responses, item difficulty levels, and assessed reading proficiency level.

As discussed in Chapter 3, several previous studies in the literature have investigated the relationship between eye movements and higher and lower order processing. Recall that longer fixations, higher occurrences of regressions, and shorter saccades are generally agreed to be indicative of increased cognitive processes. Based on the research, identifying and measuring fixations, saccades, and regressions, allows us to infer mental activity. Increased mental activity for extended periods of time generally indicates higher demands on available but limited resources.

Thus, data gathered from the eye-tracking system was analyzed for duration of eye fixations within each AOI, and elapsed time on task, as well as saccades across AOIs. The data for each participant was recorded into their own separate file. As mentioned above, ASL's EYENAL and FIXPLOT software were used for the analysis. EYENAL produces data files in binary code that can be saved to ASCII code for use elsewhere, such as statistics, spreadsheet, or word processing applications.

Admittedly, the relatively small sample sizes of each group will limit the generalizabilty of the findings (see Chapter 5: Results); however, considering the population from which the sample came, there are highly suggestive trends that do make themselves apparent. As Brown (1988) notes:

> Questions often arise about how large a sample must be to be "big enough." There is no quick and easy answer to this question, but there is little doubt that the larger the sample, the better. I think it is reasonable that a sample that includes all but one member of a population is more representative than one that contains 50 percent, 10 percent, or 2 percent of it. But how many is enough? That depends on the situation and on the statistic or statistics that are involved. So rules of thumb that are proposed (e.g., *N* should equal at least 28 or 30 per group or per variable) are imprecise. Your best strategy is to check that ... the rationale/reasoning behind the sample size involved [is mentioned], that the population is clearly delineated, and that the sampling procedures make sense. Thus, the sample should seem large enough to be representative. All of which is to say that this issue is necessarily subjective but nevertheless important. (pp. 112-113)

Bearing this in mind, I must report that the reason for the small sample size is mainly due to the small initial population of potential subjects, and, for whatever reason, the lack of volunteers willing to participate in the study for the advertised compensation. Perhaps the fact that it was made clear on the flyers I had posted around campus that the session would require about an hour of time appeared too daunting even for people who might otherwise have participated. Another possibility is that perhaps the $15 compensation was perceived as too little and did not represent significant incentive. The question remains that if I could have designed the sessions to be of shorter duration or if I could have offered greater remuneration, would I have seen an increase in the number of willing participants?

In terms of clearly delineating the population that did answer my call for participants, as stated in the Research Questions section of the Introduction to this dissertation, I have explained that I was mainly looking for three groups of participants: non-native Intermediate-level readers of CFL (NNI), non-native Superior readers (NNS), and native (Superior) readers (NS). As shown in Table 2 at the beginning of this chapter, there were five NNI participants, seven NNS participants, and thirteen NS participants.

As has been explained elsewhere in this dissertation, the CATRC was the main tool in sampling the population that volunteered. For reasons given earlier in this chapter and the previous one, I decided not to simply follow a "class level" as is commonly done in much foreign language research. In using the CATRC to place the participants in a group of participants that shared a similar level of Chinese reading proficiency as defined by a single assessment instrument, I believe I was able to maintain a certain degree of consistency.

While the above explanation may appear to be a very elaborate excuse for less than ideal sample size, it nevertheless represents a salient reality when trying to conduct this type of research. I believe the sample size that I was finally able to work with is large enough to be representative of the respective populations, especially considering the amount of data gathered from each participant. As Everson (1986) noted almost twenty years ago, which is still very much the case with a less commonly taught language (LCTL) as Chinese, researchers doing these types of studies "will probably have to resign him/herself to the inevitability of small sample sizes" (p. 91).

CHAPTER 5:

RESULTS

Introduction

In this chapter I provide the results from the CATRC (*Computer-Adaptive [Proficiency] Test for Reading Chinese*) and from the eye-tracking data. Descriptive data from the CATRC include performance data of participants in terms of time on task and comparisons between groups of number of items. Descriptive statistics from the eye-tracking data includes total and mean fixation times and durations, total and mean fixation frequencies, and fixation frequencies per item.

CATRC Results

Figures and tables detailing each participant's progression through the CATRC can be found in Appendixes G and H, respectively. In this chapter, Table 3 presents a summary of the participants' CATRC performance data in terms of time on task. English proficiency data is also provided here for the native Chinese participants (NS). Table 4 provides mean time on task, average number of items, and mean time per item for each of the three groups under study.

Table 3. Participant CATRC performance data: Time on task.

ID#	Start	Stop	Total CATRC Time	# Items	Mean Time/Item	TOEFL	GRE
NNI01	8:44	9:19	0:35:00	33	01:03.6		
NNI02	11:13	11:41	0:28:00	29	00:57.9		
NNI03	11:53	12:40	0:47:00	31	01:31.0		
NNI04	11:12	12:24	1:12:00	45	01:36.0		
NNI05	13:18	14:01	0:43:00	26	01:39.2		
NNS01	14:05	14:43	0:38:00	36	01:03.3		
NNS02	15:25	16:48	1:23:00	38	02:11.1		
NNS03	8:13	9:31	1:18:00	28	02:47.1		
NNS04	15:17	15:46	0:29:00	26	01:06.9		
NNS05	14:18	15:45	1:27:00	30	02:54.0		
NNS06	14:33	15:19	0:46:00	26	01:46.2		
NNS07	16:01	17:18	1:17:00	26	02:57.7		
NS01	11:03	11:38	0:35:00	42	00:50.0	603	1500
NS02	14:29	15:11	0:42:00	38	01:06.3	NA	NA
NS03	17:39	17:54	0:15:00	29	00:31.0	640	2050
NS04	17:22	17:48	0:26:00	26	01:00.0	650	2100
NS05	11:58	12:22	0:24:00	35	00:41.1	625	2230
NS06	12:00	12:21	0:21:00	31	00:40.6	653	1960
NS07	9:24	9:59	0:35:00	26	01:20.8	NA	NA
NS08	9:15	10:02	0:47:00	*59*	00:47.8	623	NA
NS09	11:02	11:35	0:33:00	27	01:13.3	NA	NA
NS10	11:21	11:46	0:25:00	26	00:57.7	633	2050
NS11	13:21	14:00	0:39:00	27	01:26.7	570	NA
NS12	9:14	9:38	0:24:00	30	00:48.0	663	NA
NS13	14:20	14:51	0:31:00	35	00:53.1	530	NA

Note. NNI = Non-Native Intermediate, NNS = Non-Native Superior, NS = Native Superior; NA = Not Available.

Table 4. Group CATRC performance data: Mean times on task.

Group	*n*	**CATRC Time**	**# of Items**	**Time/Item**
NNI	5	00:43:00	33 (6.5)	01:17.7
NNS	7	01:02:34	30 (5)	02:06.6
NS	12*	00:29:10	31 (5.3)	00:57.4

Note. Values in parentheses are *SD*s.
* One outlier (NS08) given 59 items due to a technical problem was removed from this group's calculations for this data set. This participant's information is otherwise listed in descriptive statistics throughout this dissertation.

The mean calculations in Table 4 show that even though the average numbers of items for all groups were similar, the total mean times on task for each group to complete the test and the mean times per item varied greatly. However, caution must be exercised when making comparisons here. When comparing NNI with NS, for example, degree of item difficulty must be kept in mind. The majority of items presented to the NNI group were at the Intermediate (Low, Mid, High) level, while those presented to the NS group were predominantly at the Superior level. Thus, as I have done in my arguments throughout this dissertation, comparisons should be limited to between the NNI and the NNS groups and between the NNS and the NS groups.

For instance, it is interesting to note that the mean times (both whole test and per item) for the NNS group averaged more than twice as long as the mean time for the NS group. Similarly, as could be expected due to item difficulty, the mean times for the NNI group were shorter than those of the NNS group, despite the larger average number of items for the NNI group.

In looking at each NNI participant's performance during their own session, something else interesting occasionally happened. That is, the complexity of the Chinese test item often belied its actual difficulty. In other words, some test items that were not especially complex in terms of total number of characters caused noticeable increases in cognitive load as evidenced by eye movement data and extended time on task. For example, for participant NNI02, the Intermediate-Mid level item #04 only contains 10 characters, but the participant took over 34 seconds to decide on their answer. On the other hand, the Intermediate-Low level item #06 has 50 characters, but this same participant only needed 33 seconds to choose an answer. There may be at least two possible explanations for this observed phenomenon. First, for the Intermediate-Mid item, the participant could not get past certain characters to piece together a correct answer. Second, at that level, with that particular test item (im04), it is possible that not enough context was provided for the participant to compensate for something unknown. Though not done as part of this research, a recall protocol immediately after the end of the CATRC assessment would have provided clues to possible reasons behind variations in time on task and and other useful data from the test taker's position.

Table 5 presents summary data about each participants CATRC performance in terms of item count. The table includes the form of Chinese characters (either Traditional, T, Simplified, S, or both, Z) as participants indicated they were more familiar with ("Pref"), the number of items answered correctly and incorrectly (each further divided in terms of character form), and the total number of items answered.

Table 5. Participant CATRC performance data: Item count.

		Correct				Incorrect				
ID#	Pref	T	S	Z	Total	T	S	Z	Total	# Items
NNI01	T	1	2	11	**14**	2	6	11	**19**	**33**
NNI02	T	0	0	15	**15**	2	1	11	**14**	**29**
NNI03	S	3	5	7	**15**	3	8	5	**16**	**31**
NNI04	S	7	4	13	**24**	9	5	7	**21**	**45**
NNI05	Z	8	6	1	**15**	5	5	1	**11**	**26**
NNS01	T	9	11	1	**21**	11	4	0	**15**	**36**
NNS02	S	12	15	3	**30**	7	1	0	**8**	**38**
NNS03	T	15	7	0	**22**	1	5	0	**6**	**28**
NNS04	Z	12	8	0	**20**	2	4	0	**6**	**26**
NNS05	T	12	11	0	**23**	4	3	0	**7**	**30**
NNS06	Z	11	11	0	**22**	3	1	0	**4**	**26**
NNS07	Z	10	13	0	**23**	3	0	0	**3**	**26**
NS01	S	15	10	2	**27**	9	6	0	**15**	**42**
NS02	T	15	7	0	**22**	4	12	0	**16**	**38**
NS03	S	14	7	0	**21**	5	3	0	**8**	**29**
NS04	S	11	9	0	**20**	5	1	0	**6**	**26**
NS05	S	8	16	0	**24**	3	8	0	**11**	**35**
NS06	S	4	18	0	**22**	4	5	0	**9**	**31**
NS07	S	13	11	0	**24**	2	0	0	**2**	**26**
NS08	S	17	20	2	**39**	14	6	0	**20**	**59**
NS09	T	9	12	0	**21**	0	6	0	**6**	**27**
NS10	S	9	10	0	**19**	5	2	0	**7**	**26**
NS11	T	7	14	0	**21**	3	3	0	**6**	**27**
NS12	S	9	14	0	**23**	3	4	0	**7**	**30**
NS13	S	14	9	0	**23**	8	4	0	**12**	**35**

Note. NNI = Non-Native Intermediate, NNS = Non-Native Superior, NS = Native Superior; Pref = Chinese character form most familiar with, T = Traditional Chinese character form; S = Simplified Chinese character form; Z = Both Traditional and Simplified Chinese character forms.

As with Table 4, Table 6 provides mean item counts for each group, including correct, incorrect, and total items.

Table 6. Group CATRC performance data: Mean item count.

Group	*n*	Correct	Incorrect	Total
NNI	5	16.60	16.20	32.80
NNS	7	23.00	7.00	30.00
NS	12*	22.25	8.75	31.00

* One outlier (NS08) given 59 items due to a technical problem was removed from this group's calculations for this data set. This participant's information is otherwise listed in descriptive statistics throughout this dissertation.

We notice that the mean total for each of the groups is similar to the others, but that overall the NNI group had roughly equal means of correctly and incorrectly answered items. Looked at in terms of items answered, we see that NNI participants selected an almost equal number of correct answers as they did selecting incorrect answers. NNS and NS participants, on the other hand, chose correct answers about 75% of the time (77% and 72%, respectively).

Eye-Tracking Data Results

Quantitative Analyses

Table 7 presents data gathered from the eye-tracking system. Data gathered includes the total number of fixations and total fixation time, total number of items per participant, mean fixation time and mean fixations per item.

The total fixations listed in Table 7 are related to the total time on task provided in Table 3. Generally, longer time spent on completing the CATRC increased total fixations and time, as well as mean fixations per item and mean fixation time per item. Table 8 shows mean group fixation data.

Table 7. Participant fixation data.

ID#	Total Fixations	Total Fixation Time (*s*)	# Items	Mean Fixations/Item	Mean Fixation Duration/Item (*ms*)
NNI01	5127	1564	33	155	346
NNI02	3661	1391	29	129	371
NNI03	6392	2380	31	206	372
NNI04	12749	3542	45	240	277
NNI05	5730	1933	26	220	325
NNS01	8040	2194	36	223	276
NNS02	12562	3869	38	330	306
NNS03	9593	3489	28	343	361
NNS04	4846	1461	26	186	298
NNS05	8859	4753	30	295	532
NNS06	7509	2196	26	289	290
NNS07	9858	3060	26	379	377
NS01	6639	1338	42	158	200
NS03	3164	667	29	109	208
NS04	5085	1275	26	248	248
NS05	4519	1014	35	129	221
NS06	3436	668	31	111	192
NS07	6751	1801	26	259	264
NS08	8887	1724	*59*	150	186
NS09	4214	1189	27	120	278
NS10	4964	1167	26	190	278
NS11	6333	1537	27	234	248
NS12	4223	985	30	140	231
NS13	5210	985	35	173	187

Note: The data for native participant NS02 had to be eliminated due to excessive noise.

Table 8. Group mean fixation data.

Group	*n*	Total Fixations	Mean Fixations/Item	Mean Fixation Duration/Item (*ms*)
NNI	5	6738	190	338.00 (39.32)
NNS	7	8752	292	348.43 (89.12)
NS	11*	4913	168	227.91 (29.70)

Note. Values in parentheses are *SD*s.

* One outlier (NS08) given 59 items due to a technical problem was removed from this group's calculations for this data set. This participant's information is otherwise listed in descriptive statistics throughout this dissertation. The data for native participant NS02 had to be eliminated due to excessive noise.

Table 8 shows that non-native Superior participants (NNS) required higher total fixations and higher mean fixations per item compared with Intermediate participants (NNI). This is most likely due to the fact that the NNS participants faced longer and more complicated CATRC items (including both passage and question) than the NNI participants. On the other hand, NNS participants had higher mean fixations and higher mean fixation times per item compared with native Superior (NS) participants. This was expected, and would seem to indicate a point of language development of an emerging distinction between non-skilled and skilled readers (i.e., the NNS participants the NS participants, respectively).

Table 9. Average Fixation Frequency: Correctly vs. Incorrectly Answered Items. Means and standard deviations of average fixation frequency as a function of Chinese language ability and items answered correctly or incorrectly.

	Correct		Incorrect		Overall	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
NNI	2961.40	1907.89	2688.80	1305.63	2825.10	1547.92
NNS	5931.40	2758.33	2573.80	652.17	4252.60	2588.84
NS	4411.62	2891.64	1564.33	732.94	2987.98	2524.05
Overall	4427.43	2751.88	2049.32	1036.54	3238.37	2371.89

N=24

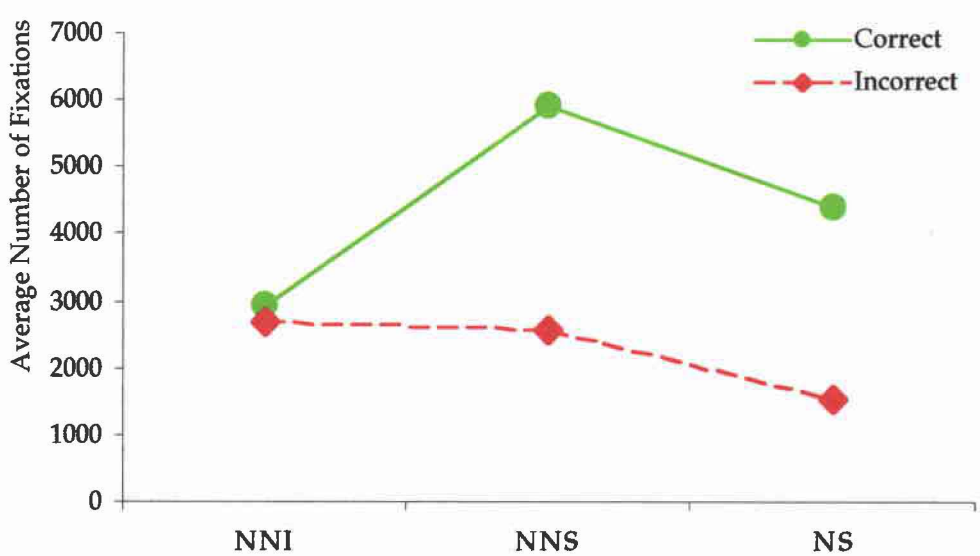


Figure 9. Fixation Frequency: Correct vs. Incorrect Answer. Interaction between Chinese language ability and correct/ incorrect answer as measured by fixation frequency.

Figure 9 reveals that there is distinct variation between the NNI and NNS groups in terms of fixation frequency for questions answered correctly versus questions answered incorrectly. In addition to a similar mean number of items for the NNI group as noted in Table 8, mean fixation frequencies were similar for the NNI group for correctly and incorrectly answered items. The NNS and NS groups, on the other hand, showed less disparity with each other in mean fixation frequencies. That is, both the NNS and NS groups fixated much more frequently for correctly answered items than they did for incorrectly answered items, possibly suggesting a "guess and move on" strategy.

Table 10. Average Fixation Frequency: Traditional vs. Simplified Characters. Means and standard deviations of average fixation frequency as a function of Chinese language ability and Chinese character form.

	Traditional		Simplified		Overall	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
NNI	1992.80	1967.04	2102.00	1461.55	2047.40	1634.74
NNS	4354.60	1270.31	4037.00	1753.88	4195.80	1453.40
NS	2571.08	1271.91	2651.92	925.08	2611.50	1088.44
Overall	2845.00	1627.93	2841.73	1627.93	2843.36	1495.94

N=24

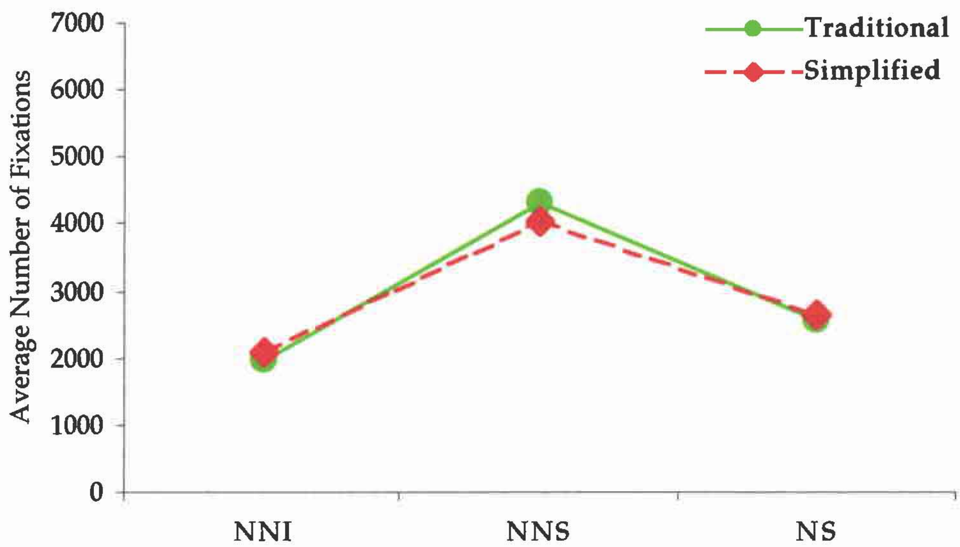


Figure 10. Fixation Frequency: Traditional vs. Simplified Characters. Interaction between Chinese language ability and traditional/ simplified characters as measured by fixation frequency.

While Figure 9 shows similar behavior for the NNS and NS groups, Figure 10 shows similar performance for both the non-native groups (NNI and NNS) and the superior groups (NNS and NS). Although the mean fixation frequencies for the NNI group were about half as much as those for the NS group, mean fixation frequencies for traditional and simplified characters within each group were about the same. In other words, the mean frequencies for the NNI group and the mean frequencies for the NNS group each reading traditional Chinese characters were roughly equal to their group mean frequencies for reading simplified Chinese characters—meaning, perhaps, that both groups of readers found both forms equally difficult.

It is interesting that the NNS group is also similar to the NS in mean fixation frequencies for traditional and simplified characters. On one hand this suggests that both the NNS and NS participants dealt with both forms equally in the sense that neither form was more or less difficult to process per se. On the other hand, it also seems to suggest that the NNS participants had more difficulty in general with both forms of characters, although with a slight indication of strength for reading traditional Chinese characters (as could be expected from looking at their selected preference in Table 5).

Qualitative Analyses

The very small participant populations for each of the three groups in this study precluded the application of more complicated statistical measures of significance, thus the tables and figures summarizing the results in this chapter must simply stand as they are. The purpose of this section, then, is to actually look at example items of a selection of participants to see if we can identify differences in reading or test-taking strategies based on the observed eye-tracking data. The key indicator for analyses will be fixation duration as represented by filled circles of varying sizes, i.e., smaller circles indicate shorter fixation durations while larger circles indicate longer fixation durations. Where true images (as presented in Appendixes I, J, and K) of eye-tracking data show leftmost fixations typically indented about two Chinese characters (about 2-3 Chinese characters, or 8 letters) and rightmost fixations generally falling off the text (Magloire, 2002), the alignment of the circles have been slightly readjusted to facilitate the observations.

Non-Native Intermediate Readers

We will first look at a few items from participants in the NNI group. The complete set for participant NNI03 is presented in Appendix I.

Character Complexity and Density

Recall from Chapter Two that there are at least two main ways to consider difficulties in single character recognition: the first is a character's complexity, which is generally (and imprecisely) measured in number of strokes; the second is a character's density (or, how dark or "full") a character is perceived to be. Thus, two characters can have the same complexity (i.e., number of strokes), but will vary in density, which may more likely be the source of difficulty in recognition, especially for non-native readers.

Figure 11 shows a large concentration of increased fixation frequency and of fixations of very high duration. Each dot represents one fixation and the size of the dot indicates the duration of the fixation (i.e., larger dots represent longer durations). The lines indicate saccades, or the jumps between fixations. What is difficult to show with static pictures is the progression through the Chinese text and the reader's alternation between the Chinese text and the English question and answers. In the case of Figure 11, the participant actually gathers a chunk of information, then jumps down to the English question, then jumps back up to where they left off and gathers another chunk of information, then jumps down to the English answers, and so on until s/he decides that s/he has comprehended the text and can make a selection from the four choices.

Figure 11 is typical of all participants in the NNI group (regardless of whether they were Intermediate-Low, -Mid, or -High), and shows a high degree of "traffic" between the two areas of interest (AOIs), i.e., the Chinese area and the English area. This behavior seems to be evidence of a type of reading strategy where the participant gathers as much information as they can hold and process in working memory before s/he needs to check her/his comprehension as hinted at by the English clues provided.

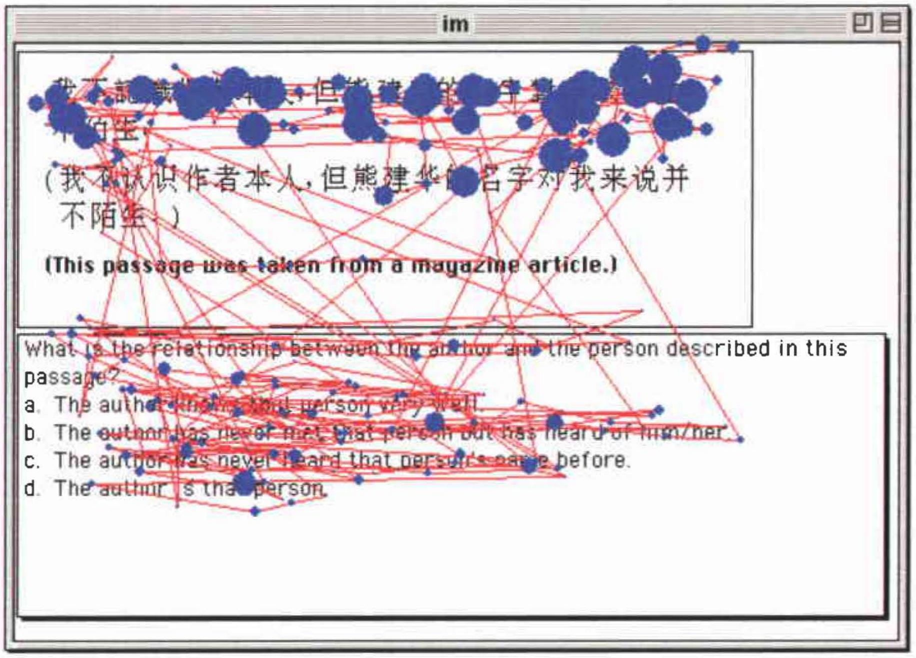


Figure 11. Eye-tracked Sample 1: NNI(T). An example item from a participant who declared a preference for traditional Chinese characters.

Traditional Characters vs. Simplified Characters

The best examples available for the analysis of this aspect of reading Chinese come from items at the Intermediate-Mid level and below. The majority of these items in the CATRC simultaneously present both traditional and simplified Chinese characters, with the former on top and the latter underneath in parentheses.

It is interesting to note that, depending on their preferred form of Chinese characters as selected at the beginning of the CATRC,[1] NNI participants either stayed with or changed to

[1] Remember that participants received a mixture of forms—traditional Chinese characters, simplified Chinese characters, or both (up to the Intermediate-Mid level)—regardless of their selection of preference; this was simply for background information only.

the form they were more comfortable with. In addition to showing the influence of character complexity on fixation frequency and duration, Figure 11 also shows that a participant who has declared a preference for traditional Chinese characters will almost exclusively read that form when presented with a choice, and only rarely check the other form for clues to aid comprehension. Figure 12 shows how a NNI participant who declared a preference for simplified Chinese characters stayed almost exclusively with that form.

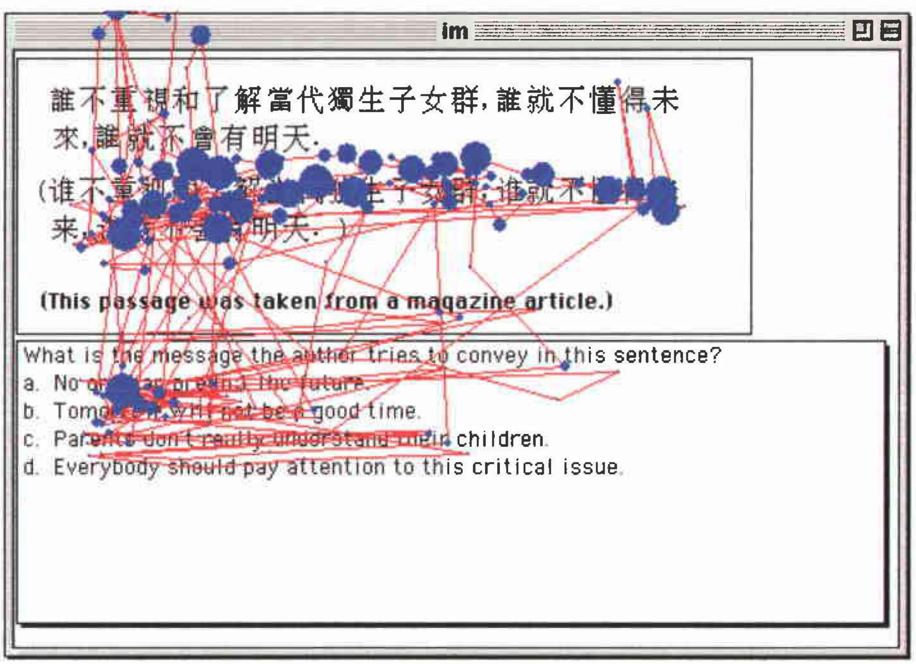


Figure 12. Eye-tracked Sample 2: NNI(S). An example item from a participant who declared a preference for simplified Chinese characters.

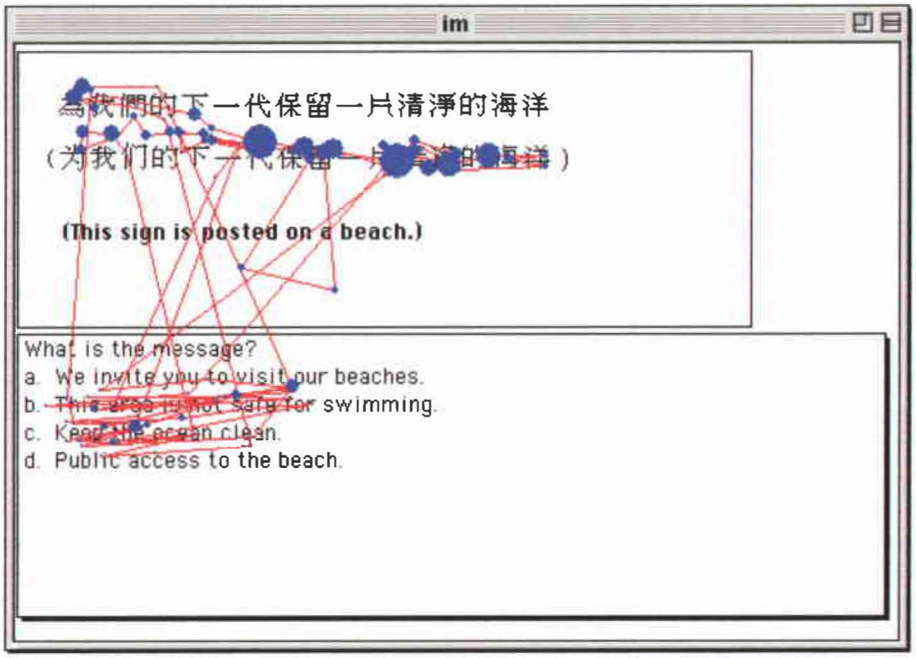


Figure 13. Eye-tracked Sample 3: NNI(S). An example item from a participant who declared a preference for simplified Chinese characters.

In contrast, as shown in Figure 13, this same participant (as used for Figure 12) changed mid-sentence from the text given in traditional Chinese characters on top to the simplified Chinese character text provided in parentheses below it. Curiously, this happened after two occurrences of traditional (or, in Chinese, 繁體 **fántǐ**, lit. 'complex form') characters.[2]

[2] The Chinese text reads as follows:

為我們的下一代保留一片清淨的海洋

(为我们的下一代保留一片清净的海洋)

In the first row, traditional characters #1 and #3—為 **wèi** and 們 **men**, respectively—have been simplified to 为 and 们. The only other character in this text that has been simplified is 淨 **jìng**, which becomes 净.

Instances such as this would have benefited the addition of a follow-up recall protocol to tease out from the participants more in-depth analyses about such observed reading behavior.

Non-Native Superior Readers

We will next look at a few items from participants in the NNS group. The complete set for participant NNS03 is presented in Appendix J.

Text Complexity and Density

As with the inherent properties of single characters that introduce difficulties in their recognition, the properties of complexity and density can also be applied to larger text passages. The complexity of a text passage can be measured mainly by number of characters, frequency ratings of vocabulary, and difficulty levels of grammar used within the passage. The density of a text passage will of course include the above properties and, in addition, will also include the overall occupation of white space by the text. This is to say that, generally speaking, a passage written in simplified Chinese characters will be less dense than its traditional character counterpart.

Figure 14 shows the performance of a participant from the NNS group. Despite a declared preference for traditional Chinese characters, there is a high frequency of fixations, and the durations of those fixations are generally quite long, as indicated by the large dots. The difference in fixation duration becomes especially evident if we compare the sizes of the dots in the Chinese AOI with those exhibited in the English AOI. Similarly, we also note the predominance of saccades, or jumps, in the English AOI (as represented by the lines) as compared with the predominance of fixations in the Chinese AOI. This is strongly indicative of the increased cognitive load on processing required by this NNS participant. In fact, this is typical of most of the participants in the NNS group.

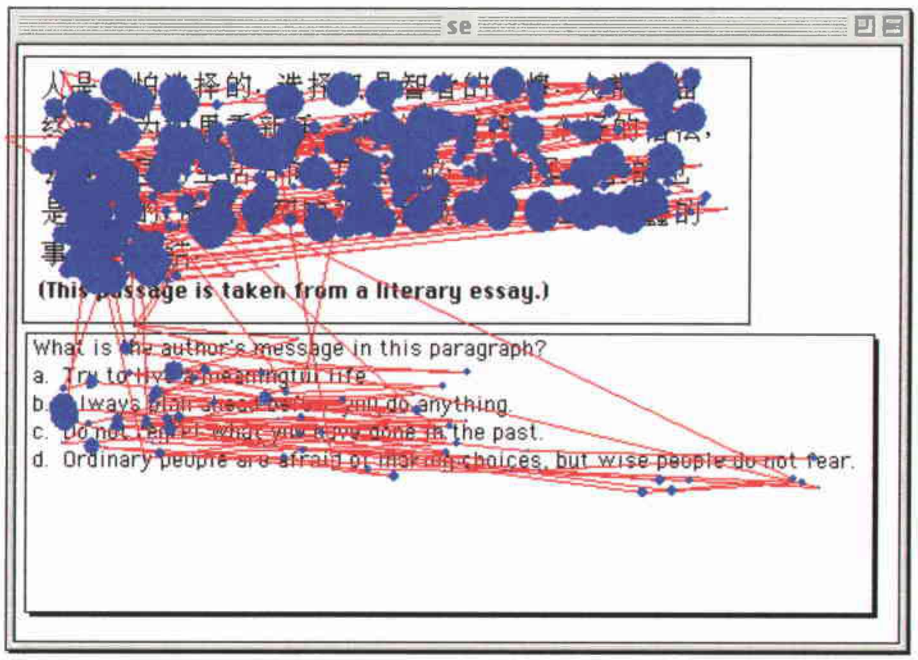


Figure 14. Eye-tracked Sample 4: NNS(T). An example item from a participant who declared a preference for traditional Chinese characters. (Chinese text is in traditional Chinese characters.)

Figure 15 shows how this same participant (as used for Figure 14) performed while reading a text in simplified Chinese characters. As we noted in Table 10 and Figure 10, participants at each level generally exhibited fixations of equal frequency on traditional characters and texts as they did on simplified characters and texts. Thus, we notice similarities in performance between Figure 14 and Figure 15 despite their texts being either traditional or simplified Chinese characters, respectively. Both figures reveal frequent fixations of long duration, possibly indicating that both forms present a similar degree of difficulty for the non-native reader, even at the superior level. Put another way, however, assuming non-native readers are ultimately unaffected by the form of the Chinese characters used in a text, text complexity could be causing the increases in fixation frequency and duration, which can be interpreted as the cost of additional processing requirements needed to comprehend the text.

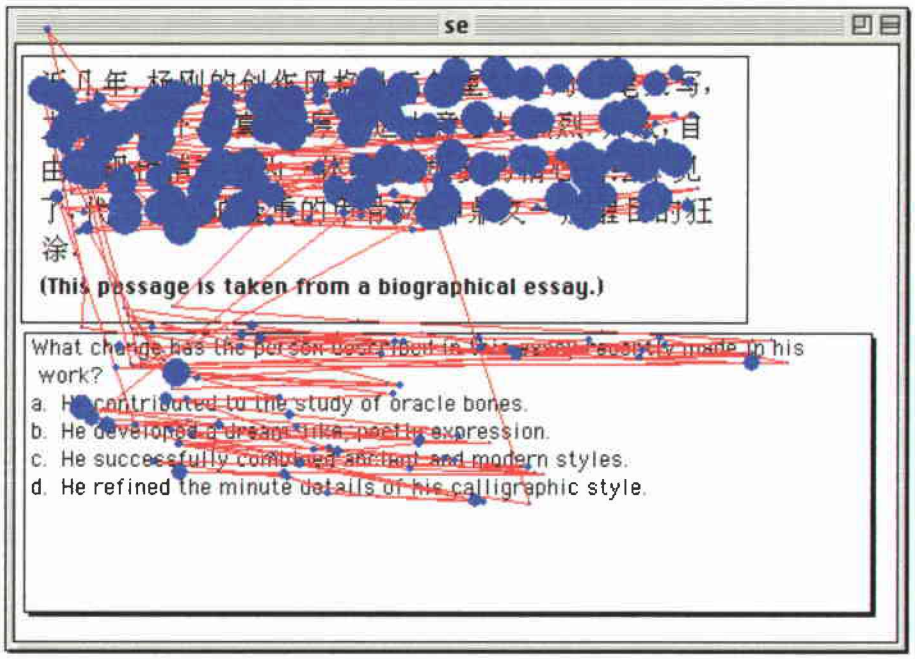


Figure 15. Eye-tracked Sample 5: NNS(T). An example item from a participant who declared a preference for traditional Chinese characters. (Chinese text is in simplified Chinese characters.)

Native Superior Readers

We will next look at a few items from participants in the NS group. The complete set for participant NNS07 is presented in Appendix K.

Working Memory (WM)

Recall that, as discussed in Chapter 3, working memory (WM) is, most simply put, an information processing mechanism of the brain. During reading, the eyes gather information, which is transmitted to the brain, a meaningful part of which is sent to WM for processing. WM includes mechanisms for holding on to chunks as well as processing the chunks either separately or by attempting to piece them together in an effort to consolidate them and comprehend them. When necessary, long-term memory (LTM) can also be tapped by WM for supplemental details or larger concepts previously encountered or learned and held in storage there. Expert knowledge is believed by some (e.g., Ericsson & Kintsch, 1995; Kintsch et al, 1999) to be the workspace called Long-Term Working Memory (LT-WM) where this interaction and processing occurs.

Figure 16 shows a participant from the NS group. The main difference we notice between the NNS example and the NS example is the low fixation frequency and the relatively short fixation durations. The text behind the eye-tracked example of the NNS participant in Figure 15 is virtually obscured by the superimposed eye movement data, while that of the NS participant in Figure 16 is essentially unobscured.

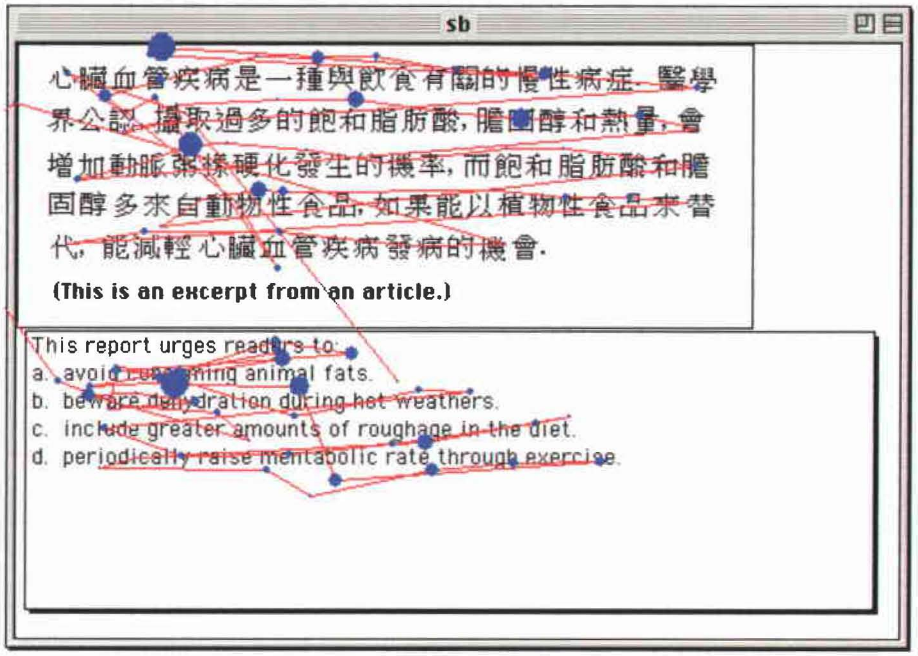


Figure 16. Eye-tracked Sample 6: NS(S). An example item from a participant who declared a preference for simplified Chinese characters. (Chinese text is in traditional Chinese characters.)

Figure 17 shows a view similar to Figure 16, but with increased activity in the English AOI—indicative of a native Chinese, non-native English reader. In contrast with the relative ease this NS participant handled the item in Figure 16, the English question and choices in Figure 17 could possibly be phrased unclearly or could be asking for more complex, higher order thinking. Also in contrast with Figure 16, which shows a general one-time read-through of the Chinese and only one saccade between the Chinese AOI and the English AOI, in Figure 17 we notice more read-throughs of the Chinese as well as more saccades between the Chinese and the English AOIs. This suggests increased item difficulty even for the NS participant.

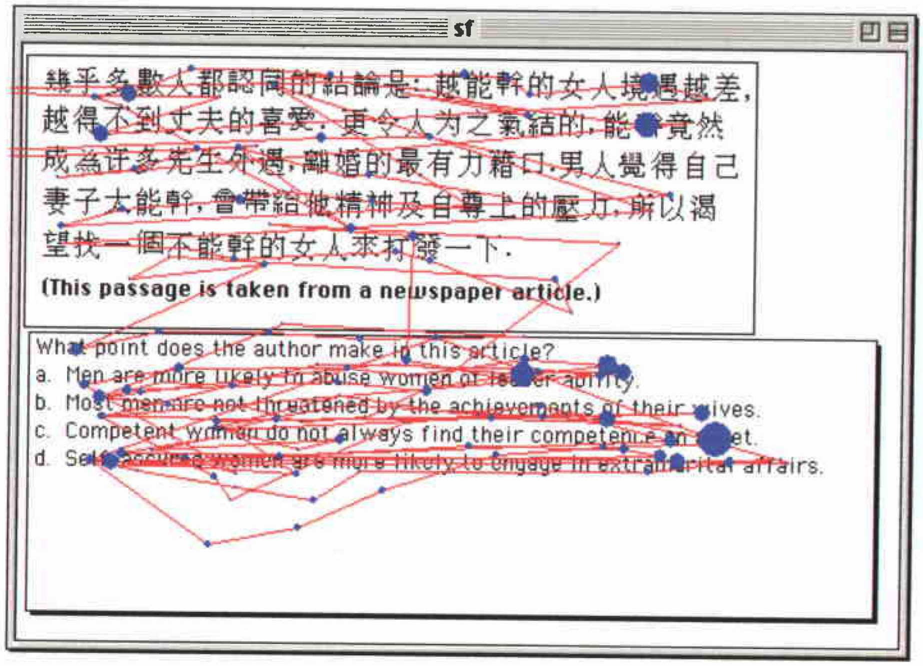


Figure 17. Eye-tracked Sample 7: NS(S). An example item from a participant who declared a preference for simplified Chinese characters. (Chinese text is in traditional Chinese characters.)

Cognitive Processing

Finally, in Figure 18, we notice two main differences. The first main difference is between the NNI example and the NNS example. If we look at the NNI example (topmost in Figure 18) and compare it with the NNS example (middle), we note that despite the increased complexity of the item, the frequency of saccades between the Chinese AOI and the English AOI for the NNS is actually lower than those exhibited by the NNI participant. Similarly, there is evidence that while relative fixation frequency and fixation duration decreases from the typical NNI participant to the typical NNS participant, the latter still requires highly frequent fixations and extended fixation durations in order to comprehend the given text.

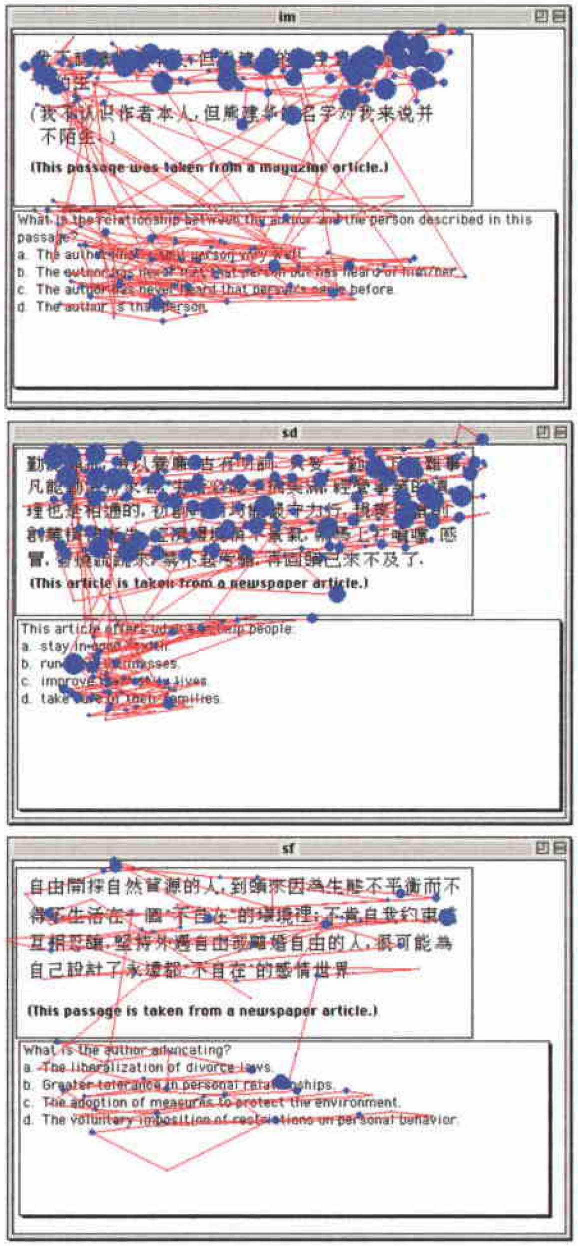


Figure 18. Eye-tracked Sample 8: NNI-NNS, NNS-NS Comparison.
NNI (top), NNS (middle), NS (bottom).

The second main difference in Figure 18 that seems quite obvious is between the NNS example (middle) and the NS example (bottom) in terms of fixation frequency and fixation duration during the reading of an item at a similar level of difficulty and complexity. The fact that the typical NNS reader spends a large duration of time in the English AOI as well as in the Chinese AOI when compared with the typical NS participant.

CHAPTER 6:

DISCUSSION

In this chapter I discuss the results and findings of the previous chapter as they relate to the research questions presented in Chapter 1. For the sake of convenience, let us recall the hypotheses and the research questions below and discuss each one in turn. Also, remember that the main purpose of this study was to try to examine the mental/cognitive processes invoked during the reading of novel text as presented during a CAT displayed on a computer screen and as recorded in eye movement and test data.

There were two main hypotheses prefacing the research conducted here. Due to the limited nature of the data, this hypothesis can neither be confirmed nor disconfirmed. In other words, while the non-native Intermediate-level CFL participants of this study are college-level students who have highly developed reading skills in their L1 (i.e., English), they were not necessarily able to use these same skill sets when they read the Chinese test item passages. As shown in Appendix H, different reading strategies can be observed

The second hypothesis states that L1 reading strategies do transfer to L2 reading strategies in Superior level non-native (NNS) CFL readers, while Superior level native (NS) Chinese readers use an interactive-compensatory (a combination of bottom-up and top-down) processing strategy during reading. Again, due to the limitations in the data, this hypothesis can not be addressed. However, an examination of the eye-tracking data in Appendix J does give the indication that for some NNS readers, native-like reading strategies during the comprehension process.

The first research question asks, "What does eye movement data tell us about how test takers take computer adaptive tests (CATs)?" Looking at the fixation data provided in the previous chapter, we can see that NNI participants completed the CATRC with fewer fixations and in less time than the NNS participants. Also, for NNS participants, generally speaking, longer passages resulted in more fixations of longer durations. This, however, was not the case with NS participants encountering the same level, length, and difficulty of material, who generally required fewer fixations and less time to complete their assessment than NNS participants.

The second research question asks, "Why do superior-level takers of CATs take longer to complete a CAT than low-level test takers?" The most obvious reason for this becomes apparent through examination of the item set of each group upon completion of a test. Test items given to NNI participants were for the most part less complex (as measured in length of passage as determined by number of characters) and less difficult (in terms of content and subject matter) than those test items that the NNS and NS participants faced. Given the higher frequency of eye fixations among NNS participants during their time on task, we can infer increased demand on cognitive processes both in comprehending the text of a given item as well as making a determining which answer would most likely be correct

Longer eye fixation times (average total fixation time or fixation duration) and higher fixation frequencies for incorrect answers could indicate increased processing load. This increased demand on processing resources could be due to either one of two factors. First, as argued for by Matsunaga (1994), participants could be attempting to associate a sound to characters as part of the lexical access process. The hypothesis emphasizing the primacy of speech, and a position of language instruction that begins first with speaking (and listening) followed by reading (and writing) would be validated in this case, especially if it could be shown that doing so improved students' reading perfomance and comprehension. Second, as exhibited by the higher fixation frequencies and longer fixation durations, inability to access lexical meaning alternatively by visual cues could also contribute to comprehension failure.

However, shorter eye fixation times and fewer frequencies for incorrect answers were also observed. This could indicate abandonment of motivation on task. That is, in attempting to answer a question but failing to achieve a click of comprehension in a given period of time (varying from participant to participant and from question to question) a participant reaches a point where the final strategy they choose is simply to give up. This particular strategy could become increasingly likely the longer a participant has been on task. Quite possibly, then, fatigue and impatience begin to present themselves as factors influencing motivation and success.

Finally, the third research question asks, "Are there identifiable differences between the three levels of reading proficiency as evidenced during the administration of a CAT?" As can be seen in the descriptive data provided in the previous chapter in Table 3, the eye tracking data shows us that, despite the relatively shorter passage content, non-native Intermediate readers (NNIs) of CFL spend long spans of time working through the test items. For example, NNI01, with a total test time of 35 min. for 33 items and an average of 1 min. 6 sec. per item. This suggests that NNIs are generally concerned with processing each and every character for meaning, if not sound, before they decide to select an answer. Failure to come up with either meaning or sound or both may cause a bottleneck in comprehension processes delaying the click of comprehension that signals to the reader that they have comprehended something (based on the question being asked) and can then select an answer among the choices given and then move on to the next test item.

CHAPTER 7:

CONCLUSION

This study took as its starting points places where Everson (1986) indicated future research was needed. Like the Everson study, the research conducted here involved three groups of readers. Where Everson examined the effects of word-spacing among non-native beginning, non-native advanced and native Chinese readers, this study examined differences in reading strategies among non-native intermediate, non-native superior, and native Chinese readers during the administration of a computer-adaptive test. Unlike the Everson study, however, which used a single paragraph read by all subjects alike, participants in the present study were presented disparate items of varying difficulty as determined by reading comprehension performance (i.e, the Intermediate-level readers in general did not receive Superior-level items and the Superior-level readers did not receive Intermediate-level items). One methodological improvement this study has, I believe, over the Everson study is in its incorporation of the ACTFL Reading Guidelines, which had not yet been codified at the time of his study. This attempted to address one limitation acknowledged by Everson in that it "establishes group membership through rigorous categorization" (p. 91).[1]

As with the Everson (1986) study, the purpose of this study was multifaceted: first, to add to the still paltry collection of research that looks at non-native readers of Chinese. While many studies have recently added much to the eye-tracking research on native readers

---

[1] Of course, it could be admitted that considering the depth and breadth of participant backgrounds, as well as the still formative state of the Guidelines, the rigorousness of such a measure might still be called into question.

of Chinese, there is still much work to be done. Second, in this study I have attempted to examine the cognitive processes involved during the reading of novel text as presented via a computer-adaptive test (CAT). By using a CAT, it was possible to expose readers to both orthographies (i.e., English and Chinese), as well as two forms of a single orthography (i.e, traditional and simplified Chinese characters), requiring them to vary their reading strategies during the administration of a single bilingual reading proficiency test.

In concord with Everson's (1986) conclusions, the data gathered and analyzed in this dissertation supports perceptual models of reading which takes efficiency in the perception of the printed word as a measure of reading development (p. 81). Based on the reading performances of non-native Intermediate (NNI) and non-native Superior (NNS) level readers in comparison with native readers of Chinese, it can be concluded that the NNI readers still do not have the familiarity or experience with the written Chinese language necessary to allow for automatic transitioning between perceptual subskills that would make for more efficient reading comprehension. Similarly, as with Everson's Advanced readers, the NNS readers in the present study also generally do not have a level of automaticity needed for efficient reading at a native level as exhibited by the NS readers in this study.

Similar to Everson's findings, both groups of non-native participants in the present study exhibited more frequent fixations and fixations of longer duration than their native counterparts. In addition, when comparing the non-native superior (NNS) readers with the native Chinese readers (NS), we find that for both groups both fixation frequency and duration increased for items that were ultimately answered correctly. This suggests that motivation and attention can be maintained over extended periods in order to devote more time and resources to comprehension processing by systematically switching between the various subskills involved in that process. Among the participants whose fixation frequencies were lower and fixation durations were longer for incorrectly answered items, we can infer that this may be indicative of a "goal abandonment" strategy, where the reader simply gives up trying to successfully resolve a certain problem in favor of a quicker route to the completion of the CAT. On the other hand, it could also simply imply that there was a mistaken assumption on the part of the participant that s/he knew the answer when in fact they did not.

Based on the data collected, both from the Computer-Adaptive (Proficiency) Test of Reading Chinese (CATRC) and the eye-tracker system, it seems that there are distinct differences in the reading strategies of various types of readers of Chinese. Specifically, the way Non-Native Intermediate readers (NNI) read test items that were at or slightly above their proficiency level differed from the way Non-Native Superior readers (NNS) read test items at or above their level of proficiency. Those differences, however, do not warrant dismissal of the "fixation duration plateau" concept suggested by Everson "which hampers their development towards native-like automaticity" (1986, p. 83). In fact, an examination of NNS participant fixation durations as compared with those of NNI participants seems to show a shortening of durations despite the increase in textual difficulty. Indeed, another trend that seemed to emerge from the data suggests that within the NNS participants there is a split in ability between highly skilled and less skilled readers. A more precise reading instrument than any curently available would be necessary to investigate the parameters of this, as explained in the Future Research section below.

Implications

The main question this dissertation may be able to answer, at least in part, is "How can learners of CL2/CFL be helped to develop efficient (i.e., 'good and appropriate') reading skills and strategies?" Based on the examination of the quantitative differences between how non-native and native readers process text, and comparing that with indicators of processing (i.e., fixations) exhibited by lower-level readers, we may be able to determine what we can do to help students of Chinese acquire more advanced reading skills. Everson (1986) posits that "classtime can be devoted to exercises designed to encourage as well as enhance the learner's *silent* reading skills" (p. 90, emphasis mine). In contrast, Dew (1994) notes that "reading and the expansion of vocabulary is for the most part a solitary activity, undertaken by the student in his own unsupervised time" (p. 40). It may be that both of these points of view coexist on one continuum of reading instruction whereby during initial stages of language instruction students receive explicit guidance and practice in how to effectively navigate the reading process, and moving towards autonomy and independence only occasionally needing to consult common resources such as dictionaries and online search engines.

From the data presented in the present study, the reading strategies of non-native Superior readers generally differ from that of non-native Intermediate readers, but is still not as efficient as that of native Chinese readers. So, we may well ask the following question: "How can learners at lower levels of CL2/CFL reading proficiency be best taught to approach novel (i.e., unfamiliar, or novel) reading tasks, based on the processing strategies exhibited by native readers?" Again, based on the data gathered and analysed in this dissertation, it would seem that a wide variety of graded readers needs to be collected and/or created in order to allow novice-, intermediate-, and even advanced-level CFL students more opportunities to gain more exposure to the written language. The greatest hindrance to incorporation of extensive reading in programs for languages such as Chinese and Japanese is the general dearth of materials available that will actually turn emerging readers into avid readers. Everson's (1986) advice given almost twenty years ago still holds true today:

> Most Chinese language learners begin the task of learning to read by attempting to automate their character recognition skills through the use of flashcard drills employing individual or combinations of characters on one side, with the romanization and English translation on the reverse. Although this "brute force" approach to literacy seems to be inevitable during the initial stages, instructors would do well to encourage students to read paragraph-level texts sooner, and to read these same texts often. This enables the learner to take advantage of the context and redundancy inherent in longer passages, and can have the potential of instilling in the learner that the passage itself can provide clues to the meanings of characters not in their vocabulary. (p. 90)

There are two further implications that can be noted in relation to those raised by Everson (1986). First, we need to recognize the daunting task facing students needing to familiarize themselves with and automate recognition of several thousand characters in order to make their Chinese reading experiences as efficient and effective as possible. Based on the observed eye-tracking data gathered for the research in this dissertation, specifically mean fixation durations and fixation frequency averages, we notice that even non-native learners achieving Superior ratings still have measures as much as two to three times that exhibited by Native readers of Chinese. Second, instructors must ensure that the passages are of sufficient length to provide for context and redundancy. As observed from the eye-tracking data in the present study, shorter passages of only a sentence or two typically do not establish enough context or make use of enough redundancy to create a meaningful reading experience.

Another pedagogical implication of the research conducted in this study may be the need to devote some time in lower level language classes for metacognitive instruction in reading and testing strategies. Especially as CATs become more prevalent as instruments of assessment, one potential benefit of training students to be more efficient during reading for comprehension activities could be higher assessed levels of proficiency. Obviously, students need also to be reminded that "haste makes waste", that they need to stop short of adopting a blanket "guess and move on" strategy towards reading and reading tests, and instead need to become conscious of their limits, as well as their abilities, and push them when possible and practical.

Future Research

Finally, recall that the ACTFL Guidelines for the Distinguised level say

> Able to read fluently and accurately most styles and forms of the language pertinent to academic and professional needs. Able to relate inferences in the text to real-world knowledge and understand almost all sociolinguistic and cultural references by processing language from within the cultural framework. Able to understand a writer's use of nuance and subtlety. Can readily follow unpredictable turns of thought and author intent in such materials as sophisticated editorials, specialized journal articles, and literary texts such as novels, plays, poems, as well as in any subject matter area directed to the general reader. (American Council on the Teaching of Foreign Languages, 1986)

Since the CATRC does not include items at the Distinguished level of the ACTFL reading proficiency guidelines, this suggests either a limitation in the precision of the CATRC or of the ACTFL reading guidelines themselves, as suggested by Park (1999). Future research in this area might need to develop a more precise assessment instrument that has a large item bank (perhaps at least 100 items) for each level of proficiency and that includes items at the Distinguished proficiency level. However, as Horodeck noted (1987) noted, tactics such as "skimming, skipping around, and guessing" are what real readers do. So, in that sense, the ideal assessment instrument would allow for participants to use the same skills as they would in "real reading", as opposed to a test that was designed differently (e.g., to see whether readers could spot errors where their reading strategies would mainly be those characteristic of proofreading). Advances in eye-tracking equipment and technology now allow researchers to gather data across multi-screen and multi-page scenes, which would mean that studies could be done that look at non-native and native readers of Chinese reading lengthy paragraphs, pages, chapters, and whole books.

RECEIVED
AS
FOLLOWS

APPENDIX A

COMMITTEE ON HUMAN STUDIES (CHS) APPROVAL

UNIVERSITY OF HAWAI'I

Committee on Human Studies

MEMORANDUM

February 14, 2003

TO: Jeffrey John Hayden
Principal Investigator
East Asian Languages and Literatures

FROM: William H. Dendle
Executive Secretary

SUBJECT: CHS #12199- "Reading in Chinese"

Your project identified above was reviewed by the Chair of the Committee on Human Studies through Expedited Review procedures. The project qualifies for expedited review by CFR 46.110 and 21 CFR 56.110, Category (4) of the DHHS list of expedited review categories.

This project was approved on February 14, 2003 for one year. If in the active development of your project you intend to change the involvement of humans from plans indicated in the materials presented for review, prior approval must be received from the CHS before proceeding. If unanticipated problems arise involving the risks to subjects or others, report must be made promptly to the CHS, either to its Chairperson or to this office. This is required in order that (1) updating of protective measures for humans involved may be accomplished, and (2) prompt report to DHHS and FDA may be made by the University if required.

In accordance with the University policy, you are expected to maintain, as an essential part of your project records, all records pertaining to the involvement of humans in this project, including any summaries of information conveyed, data, complaints, correspondence, and any executed forms. These records must be retained for at least three years from the expiration/termination date of this study.

<u>The CHS approval period for this project will expire on February 14, 2004</u>. If your project continues beyond this date, you must submit a continuation application to the CHS at least four weeks prior to the expiration of this study.

We wish you success in this endeavor and are ready to assist you and your project personnel at any time.

Enclosed is your certification for this project.

Enclosure

Figure 19. Committee on Human Studies (CHS) approval memo.

APPENDIX B

PARTICIPANT CONSENT FORM
(Research Agreement)

Agreement to participate in the project: Reading Comprehension in Chinese

The experiment that you are going to participate in is part of a one-time research project, designed and conducted by Jeffrey J. Hayden, graduate student in the Department of East Asian Languages and Literatures. The primary objective of the research is to investigate how students read and comprehend in Chinese a second or foreign language. The experiment lasts no longer than one hour. In the experiment, you will read selected items from an assessment measure for reading Chinese presented on a computer screen. You will be presented with short sentences or passages in Chinese along with a four-choice multiple-choice question in English about the content of the text. After reading the text, you will announce your selected answer aloud to the investigator. At the end of the test you will be asked to answer a brief questionnaire concerning your study of the Chinese language. You are one of approximately 30 participants from which the study is planning to collect data.

There is no known risk to you in participating in this experiment. The information collected from this experiment will remain confidential and will be used only for the purpose of the research. Your name will not be identified in any report of the results. Your participation is voluntary and there is no penalty for refusing to participate. You can withdraw at any time you wish without penalty. You will receive US$15 for completing the experiment. If you choose to withdraw before the end, you will receive partial compensation based on the percentage of the experiment you have completed.

Although it is unlikely that the results of this research will be of direct or immediate benefit to you, research in this area is of general benefit to society. The results should help us to understand the processes involved in comprehending text. This knowledge might aid in such tasks as further developing computerized reading comprehension tests, effective teaching of a second language, or treating individuals with language disorders.

If you have any questions concerning the research and participant's rights, please refer to the investigator, Jeffrey J. Hayden. If you cannot obtain satisfactory answers to your questions or have comments or complaints about your treatment in this study, please contact: Committee on Human Studies, University of Hawai'i, 2540 Maile Way, Honolulu, HI 96822. Phone: (808) 956-5007.

The language of this form is a formality required by the Committee on Human Studies, University of Hawai'i for all research involving human participants.

Jeffrey J. Hayden
Graduate Student, Department of East Asian Languages and Literatures
University of Hawai'i at Manoa
(808) 956-6881

RESEARCH AGREEMENT CONSENT SIGNATURE FORM

Agreement to participate in the project: Reading Comprehension in Chinese

I certify that I have read and that I understand the forgoing, that I have been given satisfactory answers to my inquiries concerning project procedures and other matters, and that I have been advised that I am free to withdraw my consent and to discontinue participation in the project at any time without prejudice.

I agree to participate in this study and permit the researcher, Jeffrey J. Hayden, to use the data for his study.

Name (please print): ______________________________

Signature: ______________________________ Date: ____________

APPENDIX C

BACKGROUND INFORMATION QUESTIONNAIRE

ID# (from **Consent Form**):____________________ Today's Date: ______________

Major/Field: ____

Class Standing (please circle): Freshman, Sophomore, Junior, Senior, MA, PhD, other

Total Number of Years of Formal Chinese Language Study: ____________________

Your native language/dialect: ____________________________________

Other language(s)/dialect(s) you can read: ____________________________

If either of your parents or anyone else with whom you are currently living or with whom you have lived for a substantial length of time is a ***native*** speaker of Chinese, indicate that person's native dialect:

mother: __________ father ____________ other (specify): ______________

If you have lived for substantial lengths of time in any country which includes Chinese as one of its major languages, fill in the following:

country: ________________________ Years of residence: _____, from age _____ to age _____

country: ________________________ Years of residence: _____, from age _____ to age _____

APPENDIX D

ASL 501 SERIES EYE-TRACKING EQUIPMENT

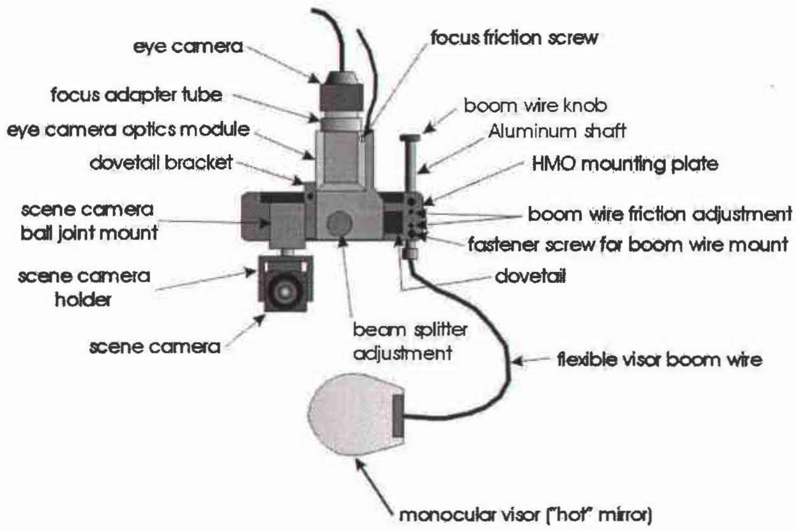


Figure 20. Front view of ASL 501 eye-tracking equipment.

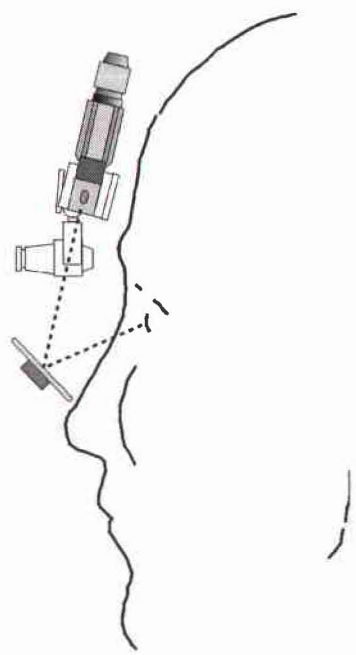


Figure 21. Side view of ASL 501 equipment placement on head.

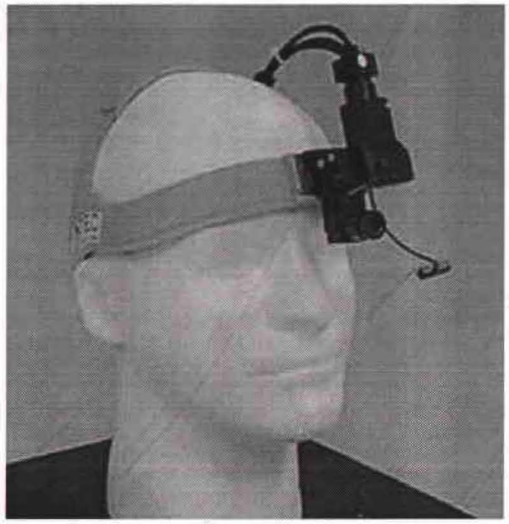


Figure 22. Eye-tracking equipment mounted on headband. The total setup weighs approximately 8 ounces and sits comfortably on the participant's head.

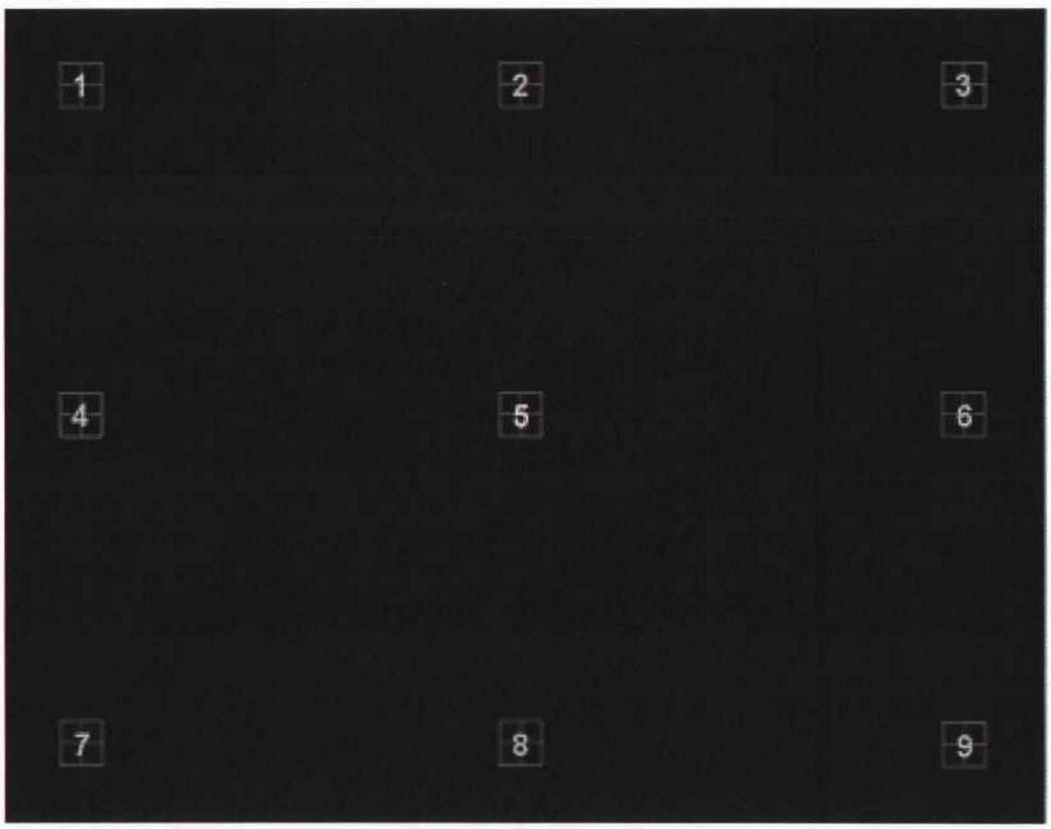


Figure 23. Calibration graphic. (50% original size.)

APPENDIX E

CALIBRATION READING: CHINESE TEXT

(Adapted from Fred Fang-yü Wang's Chinese language
Preface to *The Lady in the Painting* (pp. 1-2).
New Haven, CT: Far Eastern Publications.)
(Tradiational Character version.)

(Screen 1)

中國字又難寫，又容易忘。學過的字，幾天沒看見，就忘了。有時候，你看見一個字，跟看見一位朋友一樣。有的，是老朋友，認識很多年了。一看見就知道：他是誰，他在什麼地方住，他的父親母親是誰，父母多大歲數，有孩子沒有，孩子是男的是女的；不但知道這些，還知道他心裡想什麼，他是什麼意思。也有的，你就認識他的臉，你知道你見過他，可是他姓什麼，叫什麼名字，是在什麼地方見過的，怎麼想，也想不起來。

(Screen 2)

還有的字平常是一個樣子，有一天，他穿了一見大衣，你就不認識他了。也有的字，看著跟老朋友差不多，可是有一點兒不對，也說不出來什麼地方不對。就知道昨天看見他的時候，不是這個樣子。你也許想，這位朋友今天有一點兒病，穿了大衣了，你就過去叫他，可是一叫他，你才知道他不是你的朋友，你又錯了。

CALIBRATION READING: ENGLISH TEXT

(Screen 1)

The test you are about to take is what is known as a computer-adaptive test (CAT). This type of test adapts to each test taker's answer and so is different for each person. The items the program will give you will be random and their level will depend on whether you answered the previous item correctly or incorrectly. If you answer a question correctly, the test is programmed to give you an item at either a similar level of difficulty or a more difficult item. If you answer an item incorrectly, the test is programmed to give you an item either at the same level (in case you made a mistake when selecting your answer) or at a slightly easier level.

(Screen 2)

Since the items are random and each test is different for each person, the amount of time to complete the test will vary from person to person. Generally speaking, it will take about 30-40 minutes to finish this test. One thing to keep in mind when taking the test is not to worry about the time. The more you focus on just reading the items and answering the questions, the quicker the test will complete its assessment. If you are not sure of an answer, don't be afraid to guess and move on. The test works on averages, so guessing won't count against you. Also, you cannot go back to previous items. Again, it may seem like the test goes on forever, but on average the test gives only about 20-25 items per person. So, remember, just answer the questions to the best of your ability and it will be over before you know it.

APPENDIX F

CATRC SAMPLE CARDS AND TEST ITEMS

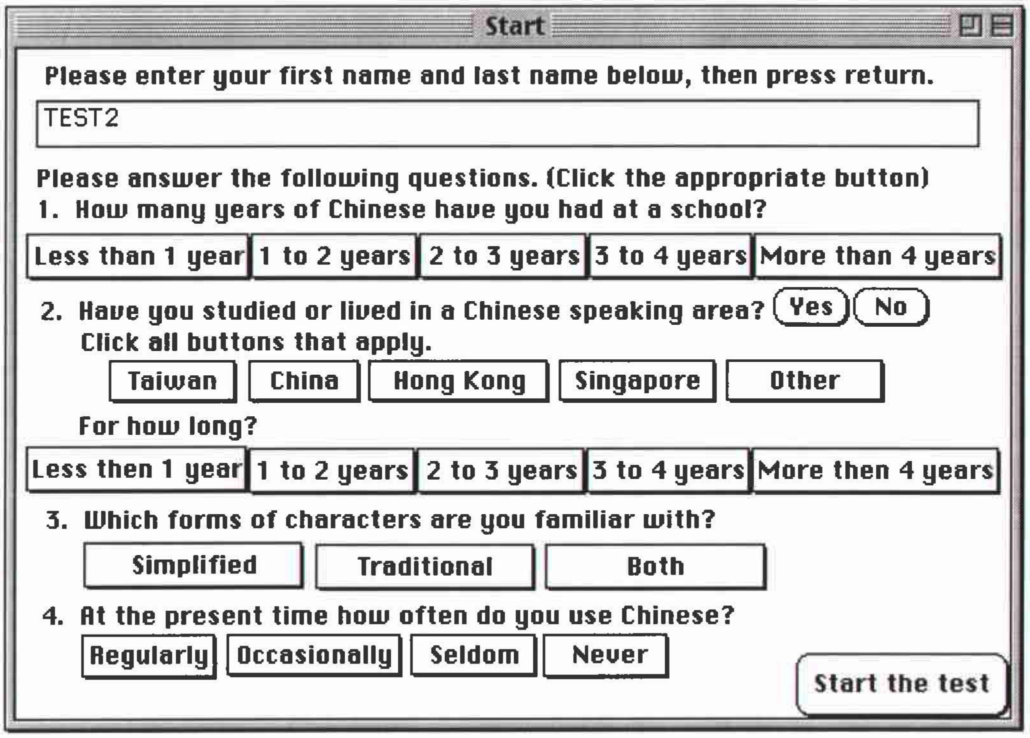


Figure 24. CATRC background card.
(Scaled down to 80% of original size)

Figure 25. Sample test item 1. Intermediate Low (il) level card.
(Scaled down to 80% of original size)
Note provision of both traditional and simplified Chinese character forms.

[Text reproduced for clarity. Romanization and translation provided for convenience.]

請走人行道謝謝您

(请走人行道谢谢您)

[Qǐng zǒu rénxíngdào, xièxie nín.]

[Please use sidewalk. Thank you.]

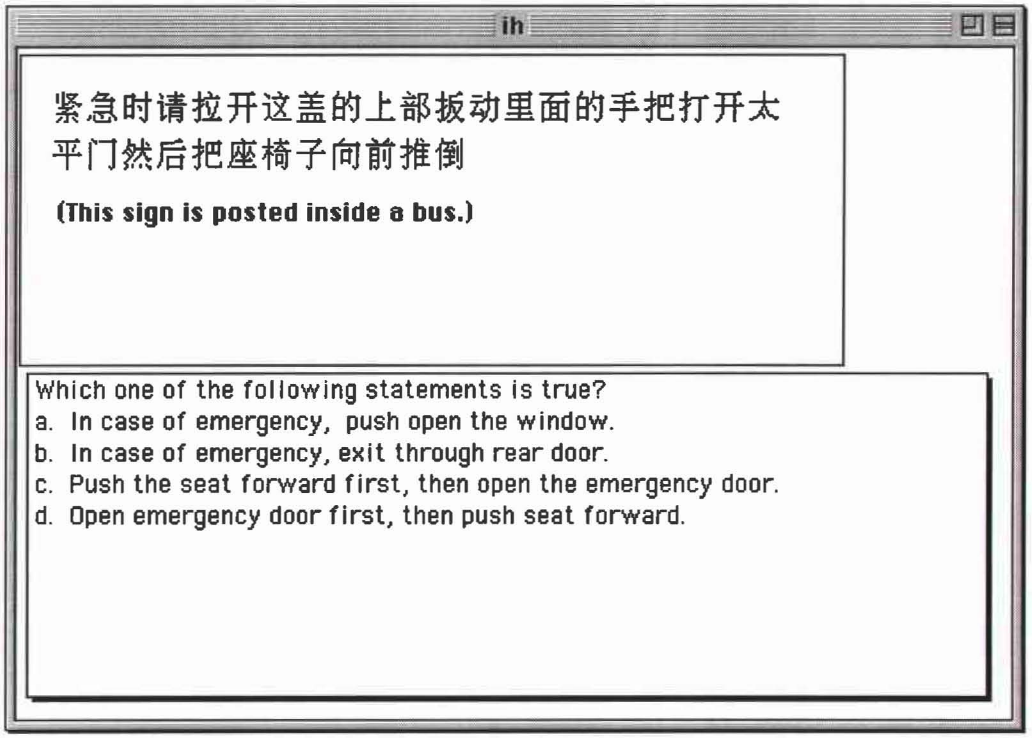


Figure 26. Sample test item 2. Intermediate High (ih) level card.
(Scaled down to 80% of original size)
Note provision of only simplified Chinese character forms.

[Text reproduced for clarity. Romanization and translation provided for convenience.]

紧急时请拉开这盖的上部扳动里面的手把打开太平们然后把座椅子向前推倒。

[Jĭnjí shí qĭng lākāi zhè gài de shàngbù bāndòng lĭmiàn de shŏubă dăkāi tàipíngmén ránhòu bă zuòyĭzi xiàng qián tuīdăo.]

[In an emergency, please pull open the top of this cover and pull the handle inside to open the emergency exit door, then push the seat forward.]

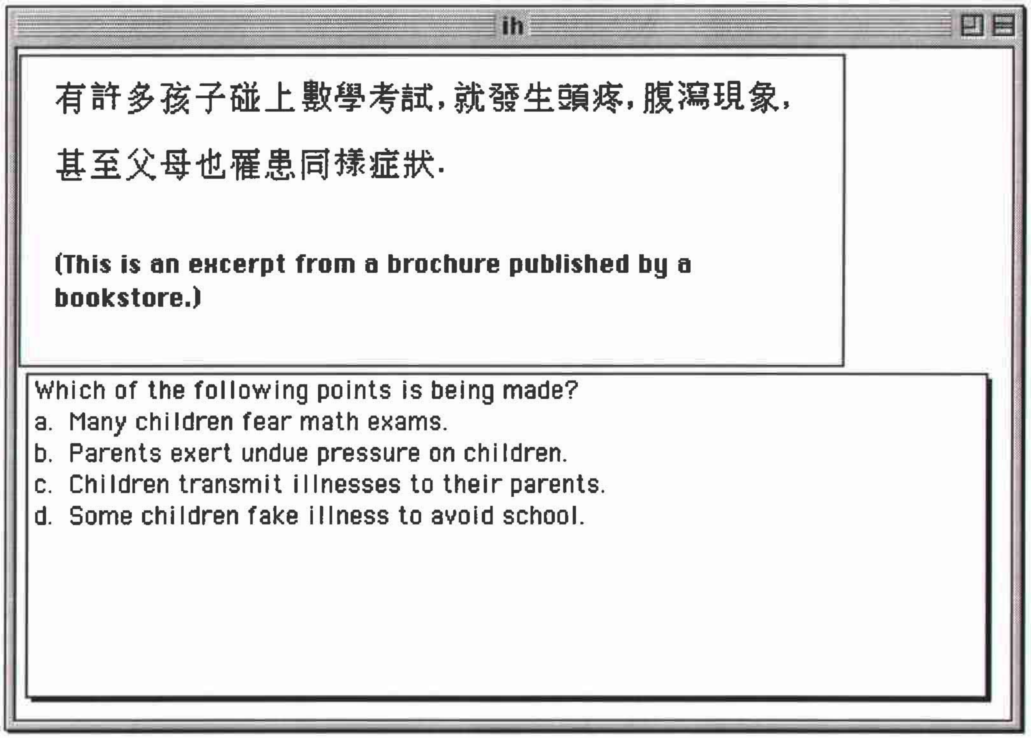


Figure 27. Sample test item 3. Intermediate High (ih) level card.
(Scaled down to 80% of original size)
Note provision of only traditional Chinese character forms.

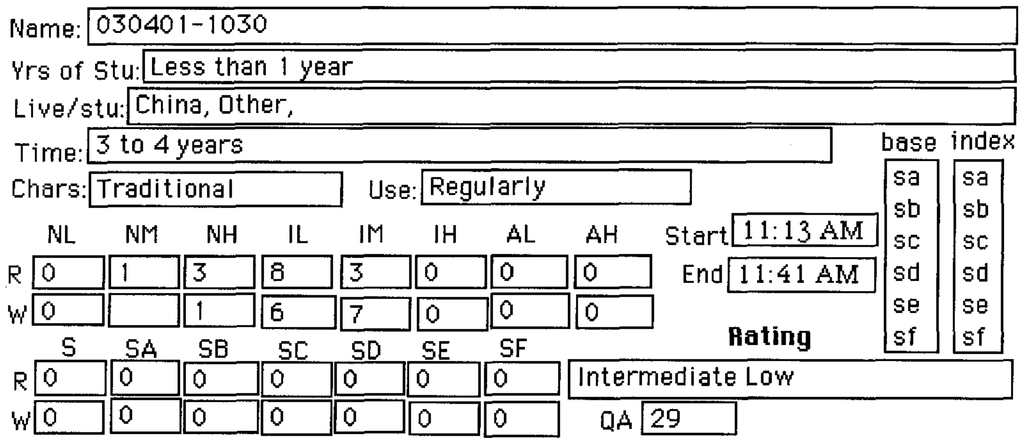
[Text reproduced for clarity. Romanization and translation provided for convenience.]
有許多孩子碰上數學考試，就發生頭疼，腹瀉現象，甚至父母也罹患同樣症狀。

[Yǒu xǔduō háizi pèngshang shùxué kǎoshì, jiù fāshēng tóuténg, fùxiè xiànxiàng, shènzhì fù-mǔ yě líhuàn tóngyàng zhèngzhuàng.]

[Many children develop headache and diarrhea phenomena when they encounter a math test, even parents suffer from the same symptoms.]

APPENDIX G

PARTICIPANT CATRC BACKGROUND AND PROGRESS REPORTS



Name: 030401-1030
Yrs of Stu: Less than 1 year
Live/stu: China, Other,
Time: 3 to 4 years
Chars: Traditional Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	1	3	8	3	0	0	0
W	0		1	6	7	0	0	0

Start 11:13 AM
End 11:41 AM

base	index
sa	sa
sb	sb
sc	sc
sd	sd
se	se
sf	sf

	S	SA	SB	SC	SD	SE	SF
R	0	0	0	0	0	0	0
W	0	0	0	0	0	0	0

Rating: Intermediate Low
QA 29

Figure 28. Example CATRC Report.

Abbreviations:	Name	=	Examinee Name (or User ID, as in this study)
	Yrs of Stu	=	Years of Study (Less than 1 year, 1-2 years, 2-3 years, 3-4 years, more than 4 years)
	Live/stu	=	Chinese-speaking countries Lived in or Study abroad experience (or, for native speakers, country of origin) (China, Taiwan, Hong Kong, Other)
	Time	=	Length of Study abroad experience (or, for native particpants, time in US) (Less than 1 year, 1-2 years, 2-3 years, 3-4 years, more than 4 years)
	Chars	=	Chinese character form most familiar with
	Use	=	Rough estimation of current foreign language use
	Start	=	Test administration Start Time
	End	=	Test administration End Time
	base	=	base set at Superior level
	index	=	indexed set at Superior level
	Levels	=	NL (Novice Low), NM (Novice Mid), NH (Novice High); IL (Intermediate Low), IM (Intermediate Mid), IH (Intermediate High); AL (Advanced Low), AH (Advanced High); S (Superior—General), SA-SF (Superior—Subject Area-Specific)
	R	=	Right (correct answer)
	W	=	Wrong (incorrect answer)
	Rating	=	Participant's assessed rating (ACTFL levels)
	QA	=	Number of Questions Asked

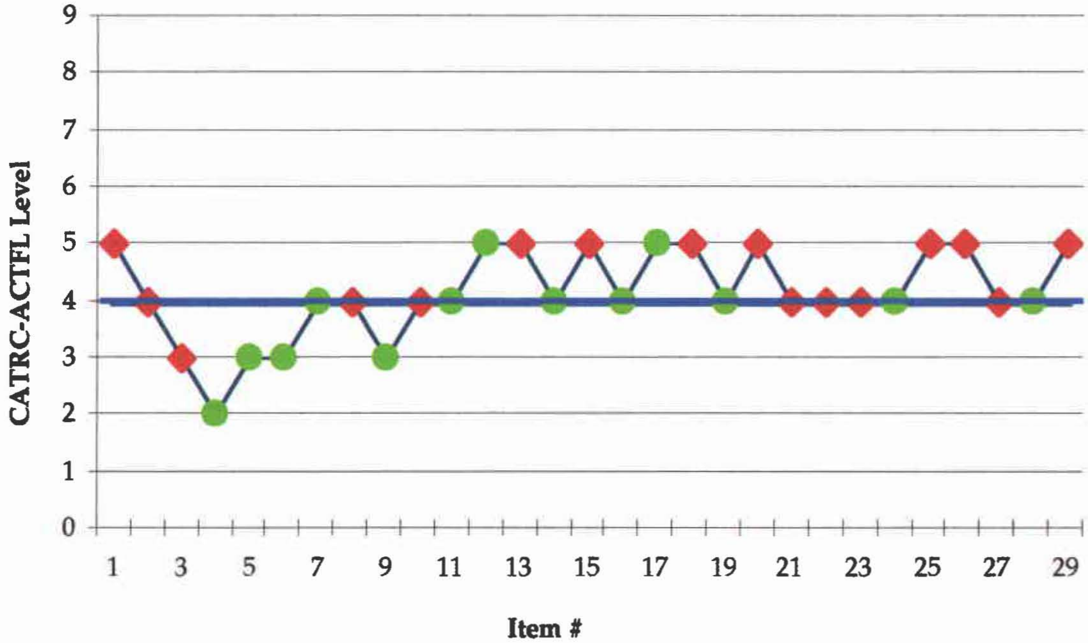
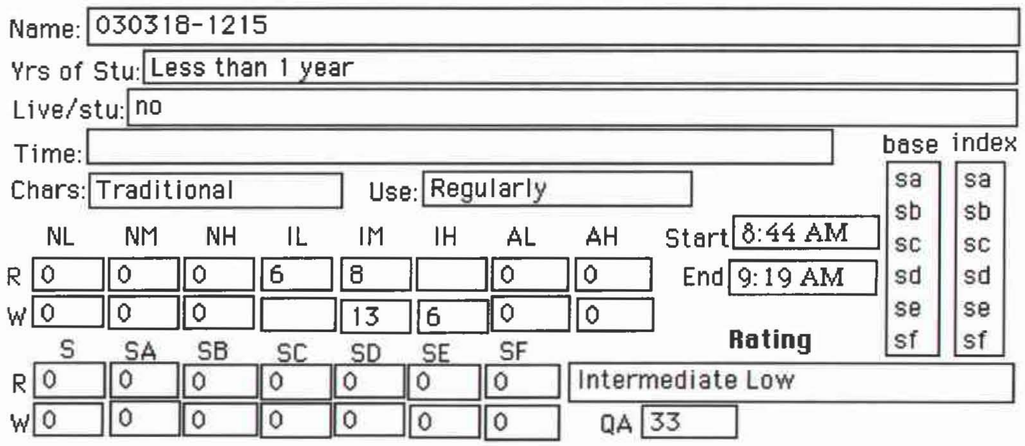


Figure 29. Example CATRC Progression.

Abbreviations:	CATRC-ACTFL:	1 = Novice Low, 2 = Novice Mid, 3 = Novice High; 4 = Intermediate Low, 5 = Intermediate Mid, 6 = Intermediate High; 7 = Advanced Low, 8 = Advanced High; 9 = Superior
	Item #:	Question Asked (sequential order)
	●:	Correct Answer
	◆:	Incorrect Answer
	▬ :	CATRC-Assessed ACTFL Rating (cf. above)

Non-native Intermediate (NNI) Participants



Name: 030318-1215

Yrs of Stu: Less than 1 year

Live/stu: no

Time:

Chars: Traditional Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	6	8		0	0
W	0	0	0		13	6	0	0

Start 8:44 AM

End 9:19 AM

	S	SA	SB	SC	SD	SE	SF
R	0	0	0	0	0	0	0
W	0	0	0	0	0	0	0

Rating: Intermediate Low

QA 33

base	index
sa	sa
sb	sb
sc	sc
sd	sd
se	se
sf	sf

Figure 30. Participant #NNI01 CATRC Report.

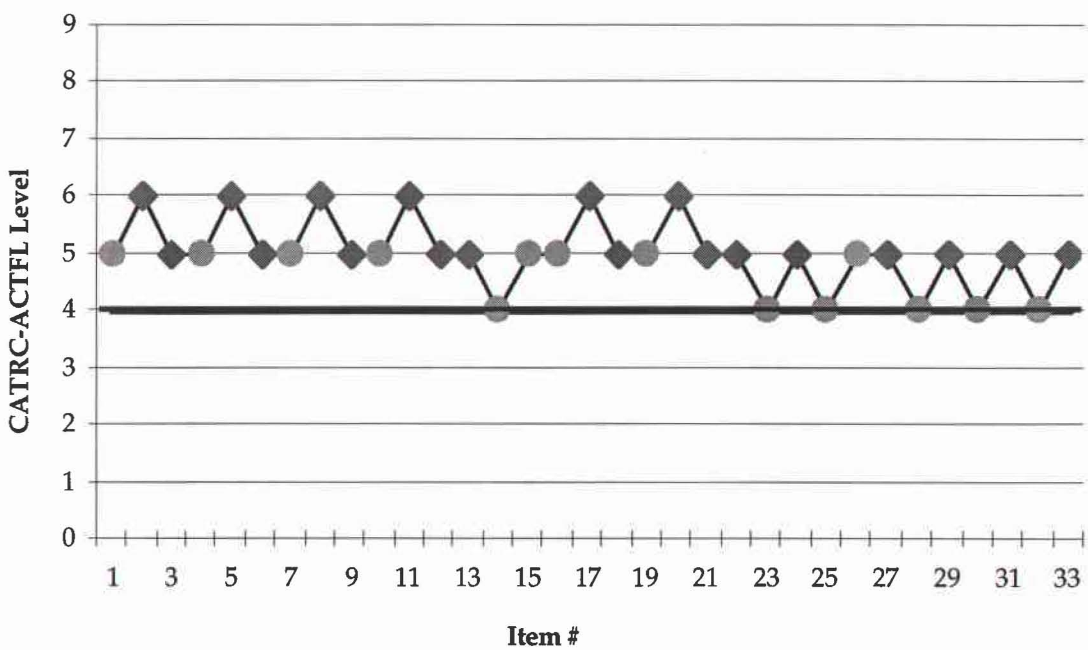


Figure 31. Participant #NNI01 CATRC Progression.

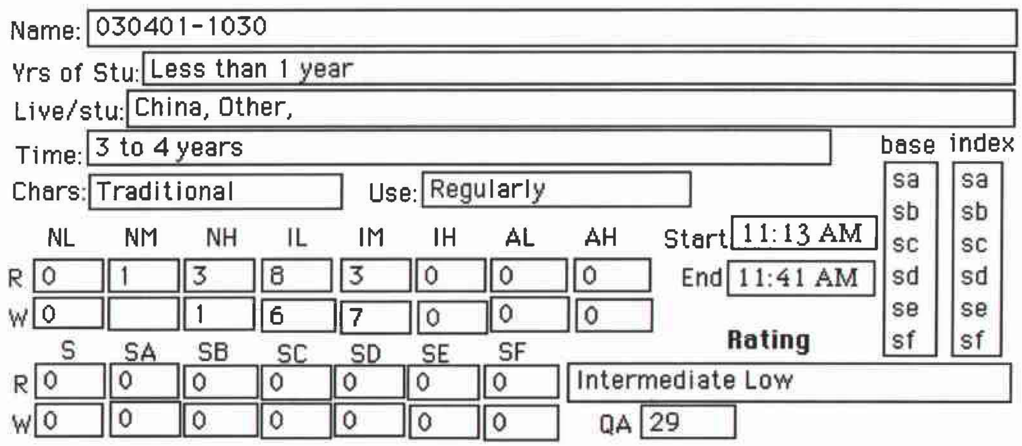


Figure 32. Participant #NNI02 CATRC Report.

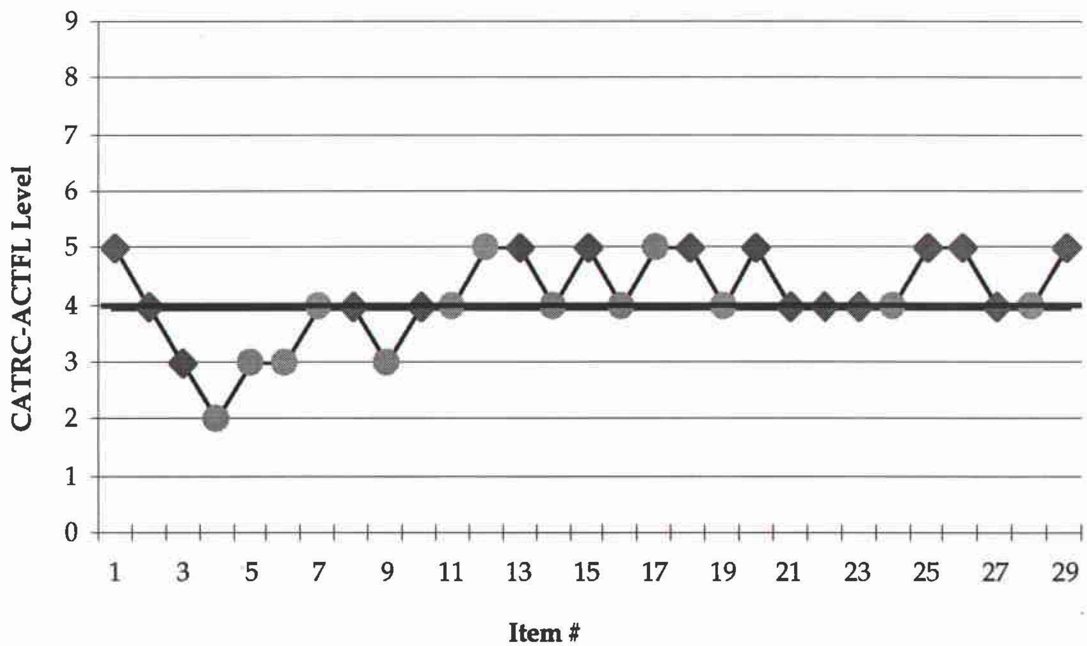


Figure 33. Participant #NNI02 CATRC Progression.

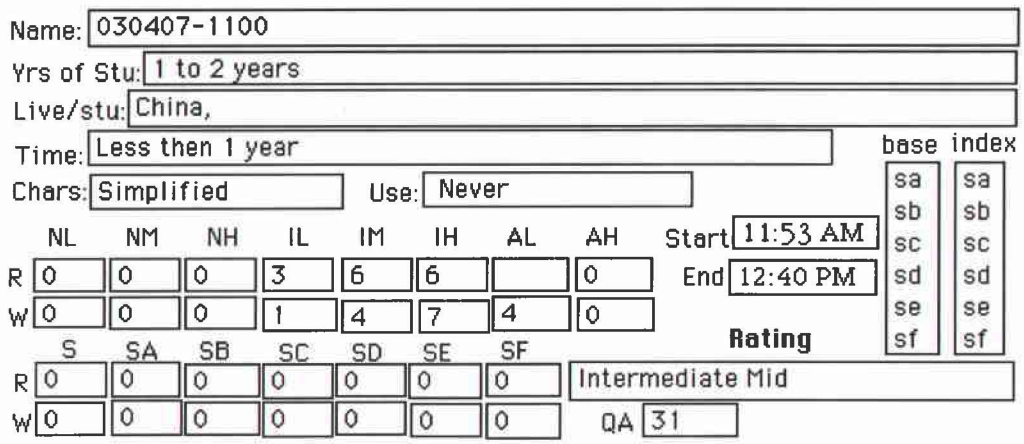


Figure 34. Participant #NNI03 CATRC Report.

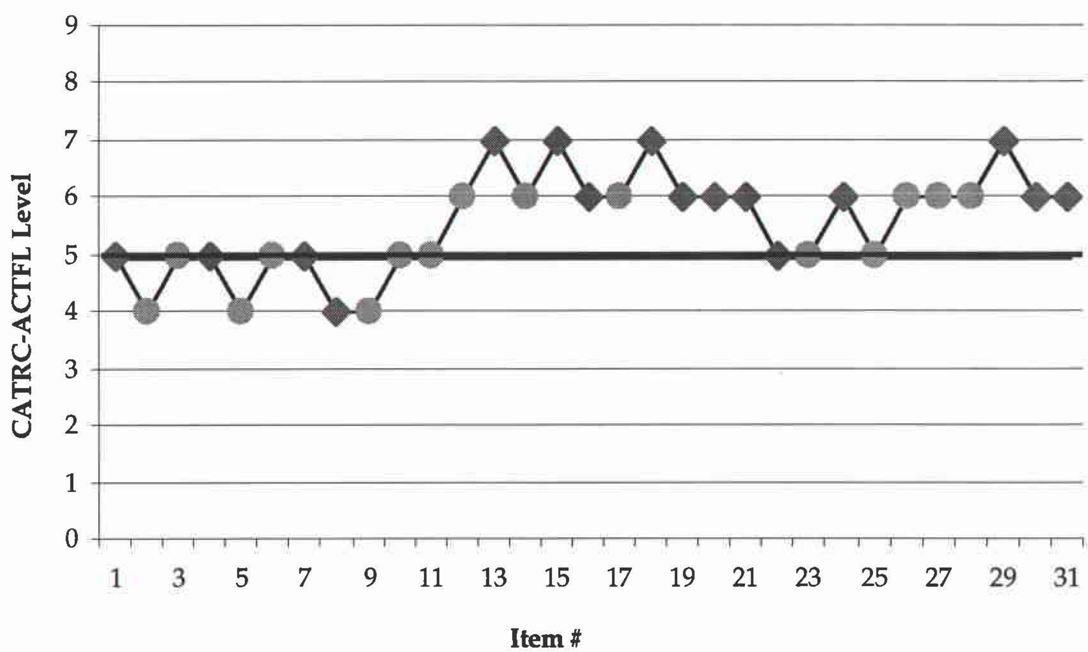


Figure 35. Participant #NNI03 CATRC Progression.

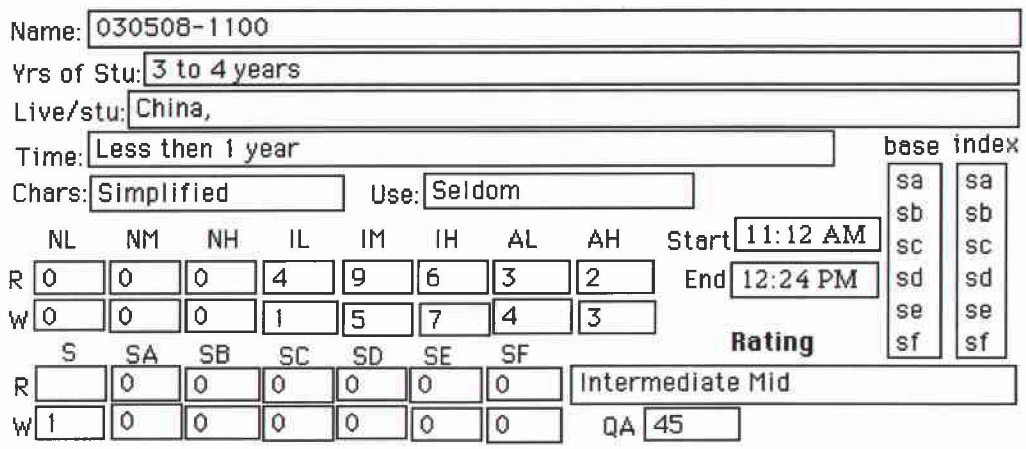


Figure 36. Participant #NNI04 CATRC Report.

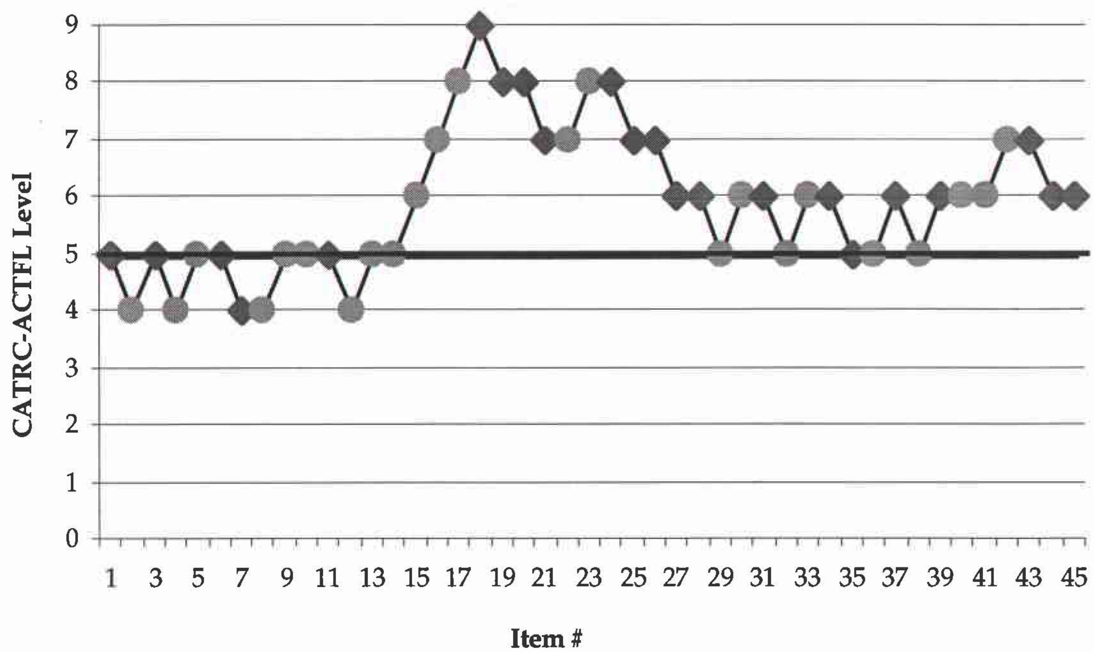
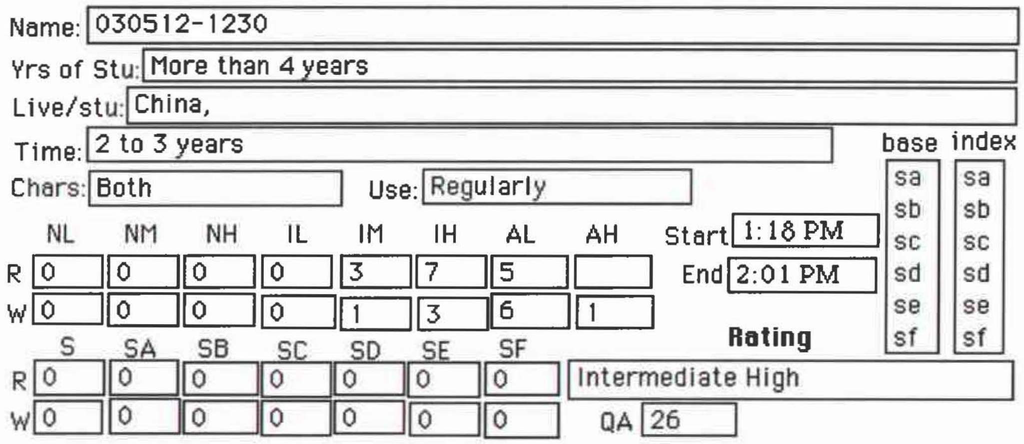


Figure 37. Participant #NNI04 CATRC Progression.



Name: 030512-1230
Yrs of Stu: More than 4 years
Live/stu: China,
Time: 2 to 3 years
Chars: Both
Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	3	7	5	
W	0	0	0	0	1	3	6	1

	S	SA	SB	SC	SD	SE	SF
R	0	0	0	0	0	0	0
W	0	0	0	0	0	0	0

Start 1:18 PM
End 2:01 PM

Rating
Intermediate High
QA 26

base	index
sa	sa
sb	sb
sc	sc
sd	sd
se	se
sf	sf

Figure 38. Participant #NNI05 CATRC Report.

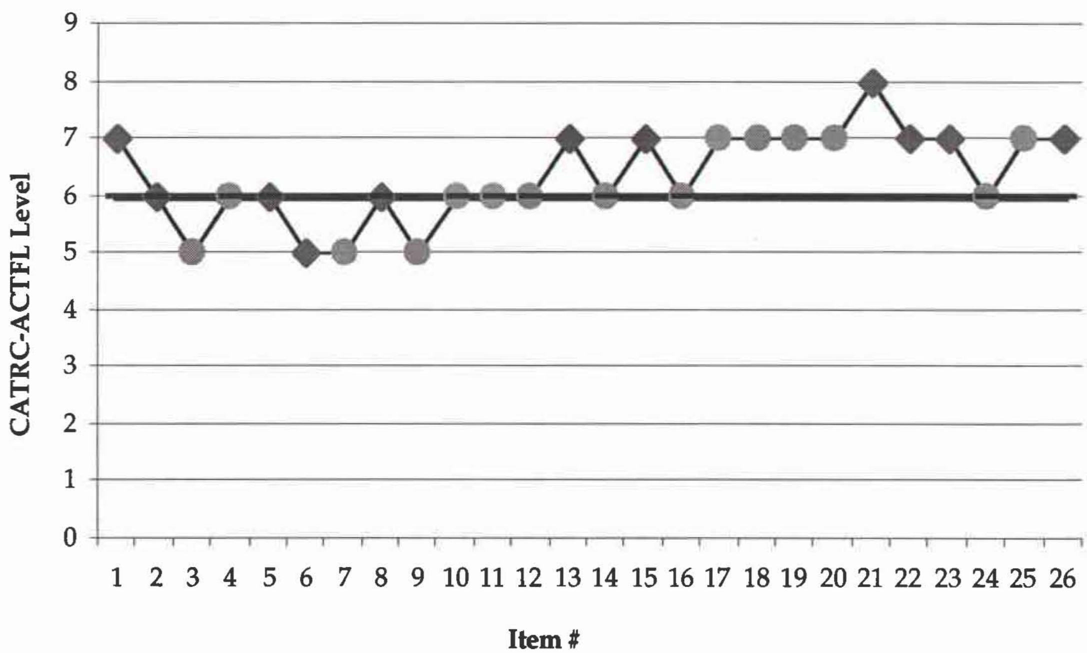
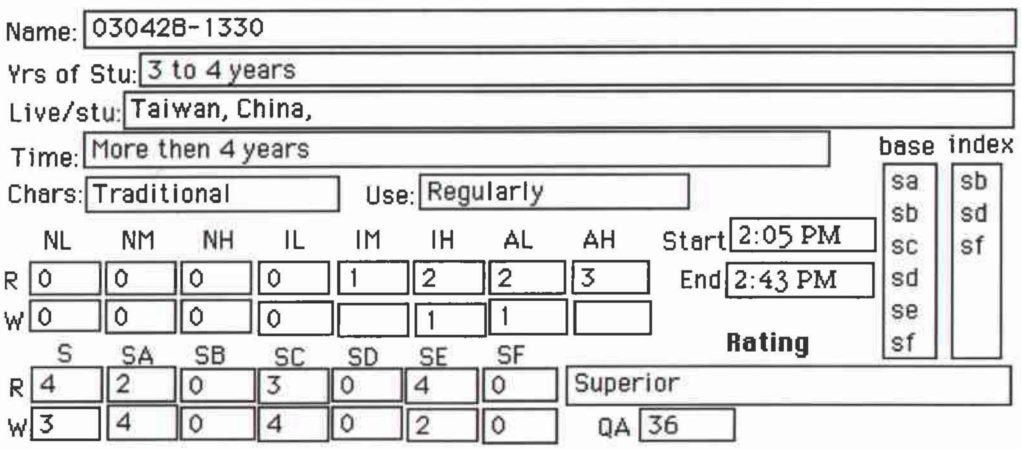


Figure 39. Participant #NNI05 CATRC Progression.

Non-Native Superior (NNS) Participants



Name: 030428-1330
Yrs of Stu: 3 to 4 years
Live/stu: Taiwan, China,
Time: More then 4 years
Chars: Traditional   Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	1	2	2	3
W	0	0	0	0		1	1	

Start 2:05 PM
End 2:43 PM

	S	SA	SB	SC	SD	SE	SF
R	4	2	0	3	0	4	0
W	3	4	0	4	0	2	0

Rating: Superior
QA 36

base	index
sa	sb
sb	sd
sc	sf
sd	
se	
sf	

Figure 40. Participant #NNS01 CATRC Report.

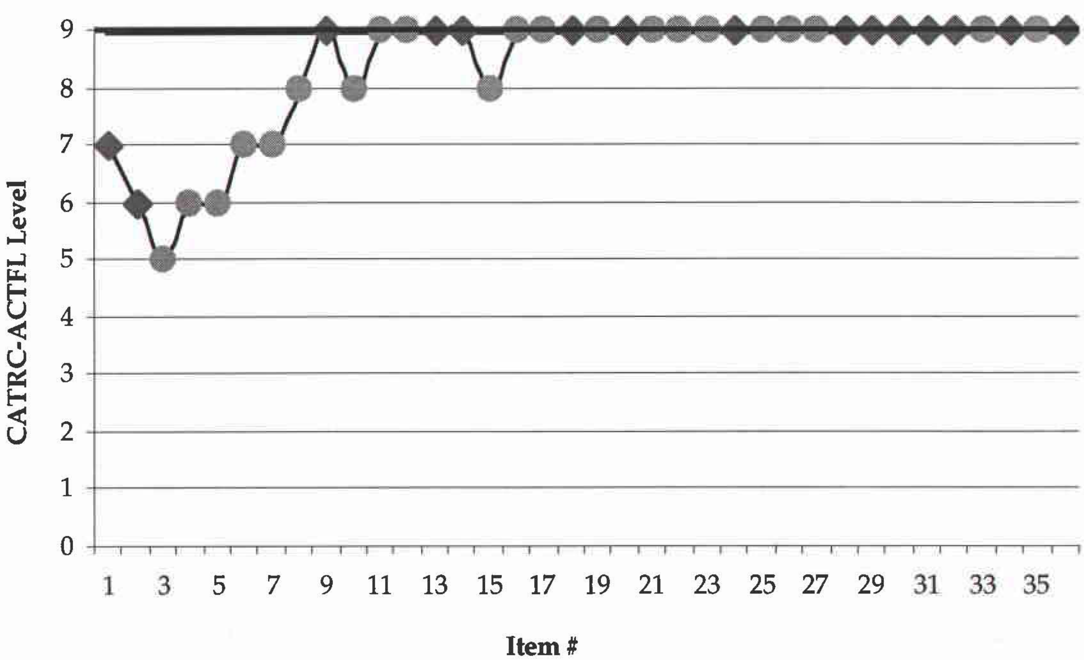
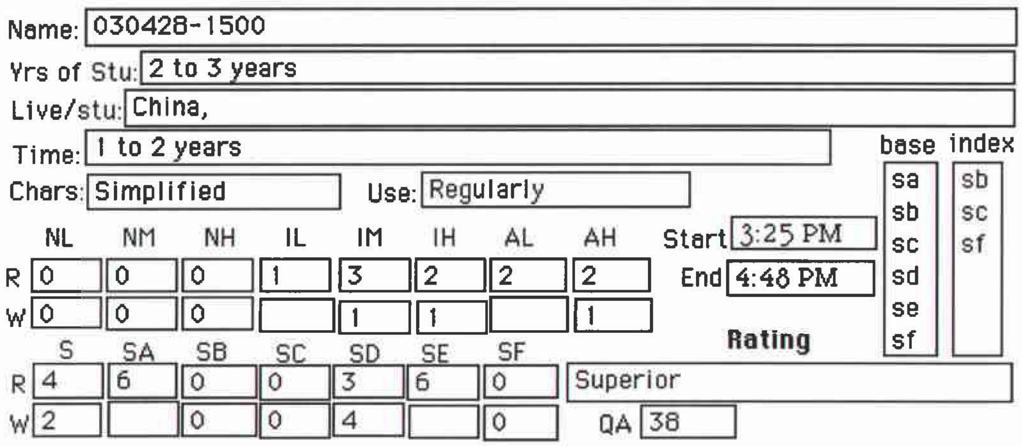


Figure 41. Participant #NNS01 CATRC Progression.


Name: 030428-1500

Yrs of Stu: 2 to 3 years

Live/stu: China,

Time: 1 to 2 years

Chars: Simplified   Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	1	3	2	2	2
W	0	0	0		1	1		1

Start 3:25 PM

End 4:48 PM

	S	SA	SB	SC	SD	SE	SF
R	4	6	0	0	3	6	0
W	2		0	0	4		0

Rating: Superior

QA 38

base: sa, sb, sc, sd, se, sf

index: sb, sc, sf

Figure 42. Participant #NNS02 CATRC Report.

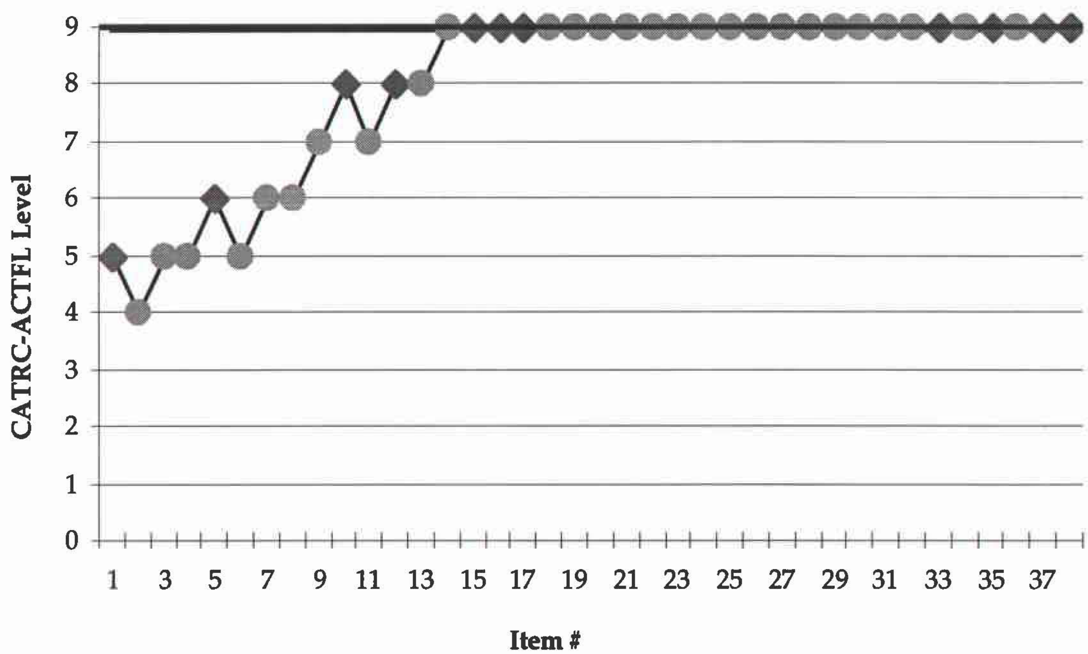


Figure 43. Participant #NNS02 CATRC Progression.

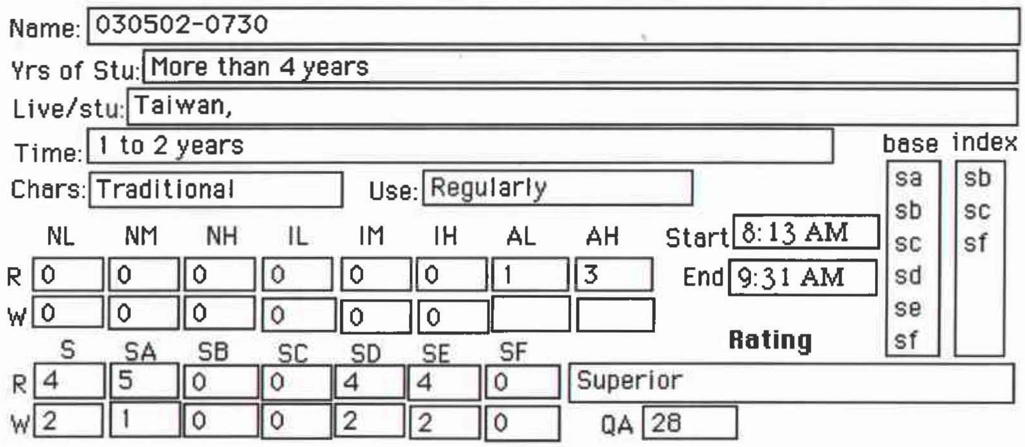


Figure 44. Participant #NNS03 CATRC Report.

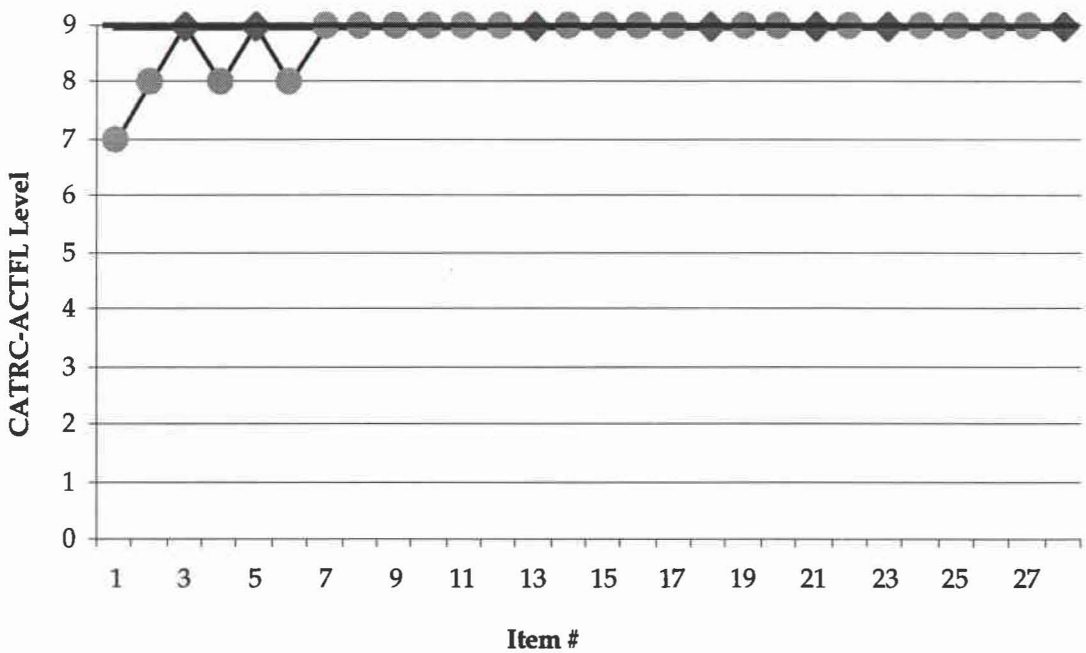
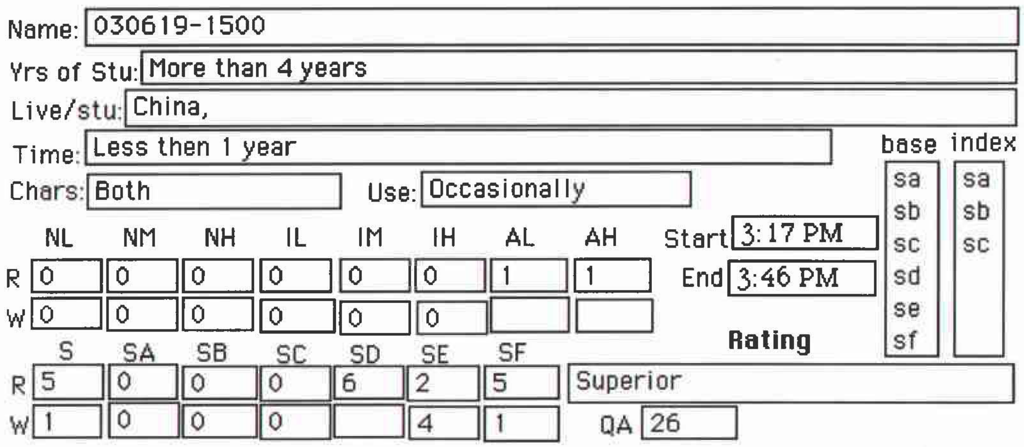


Figure 45. Participant #NNS03 CATRC Progression.



Name: 030619-1500
Yrs of Stu: More than 4 years
Live/stu: China,
Time: Less then 1 year
Chars: Both
Use: Occasionally

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	0	1	1
W	0	0	0	0	0	0		

Start: 3:17 PM
End: 3:46 PM

base	index
sa	sa
sb	sb
sc	sc
sd	
se	
sf	

	S	SA	SB	SC	SD	SE	SF
R	5	0	0	0	6	2	5
W	1	0	0	0		4	1

Rating: Superior
QA: 26

Figure 46. Participant #NNS04 CATRC Report.

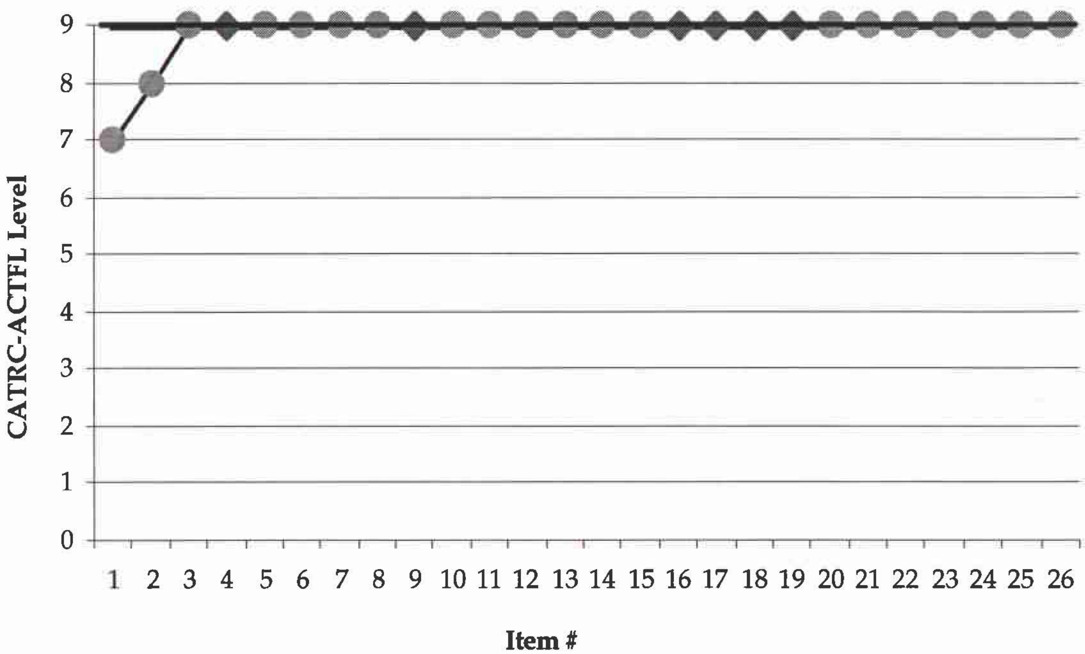


Figure 47. Participant #NNS04 CATRC Progression.

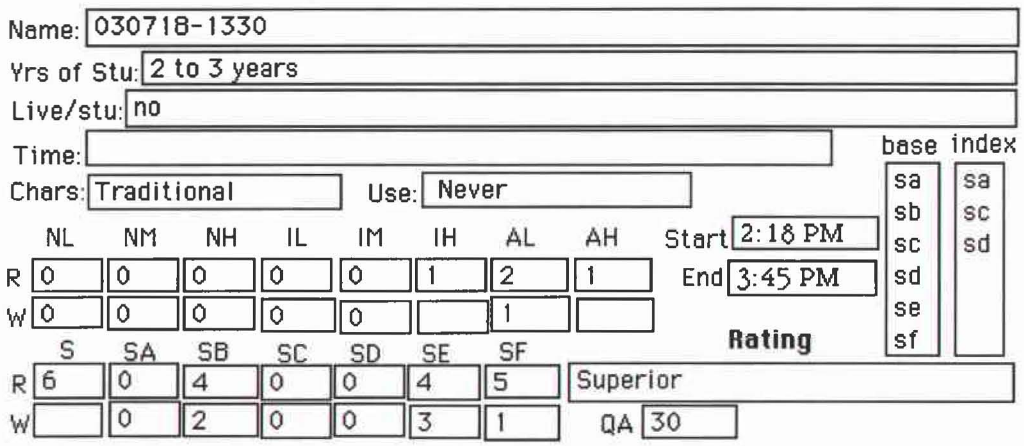


Figure 48. Participant #NNS05 CATRC Report.

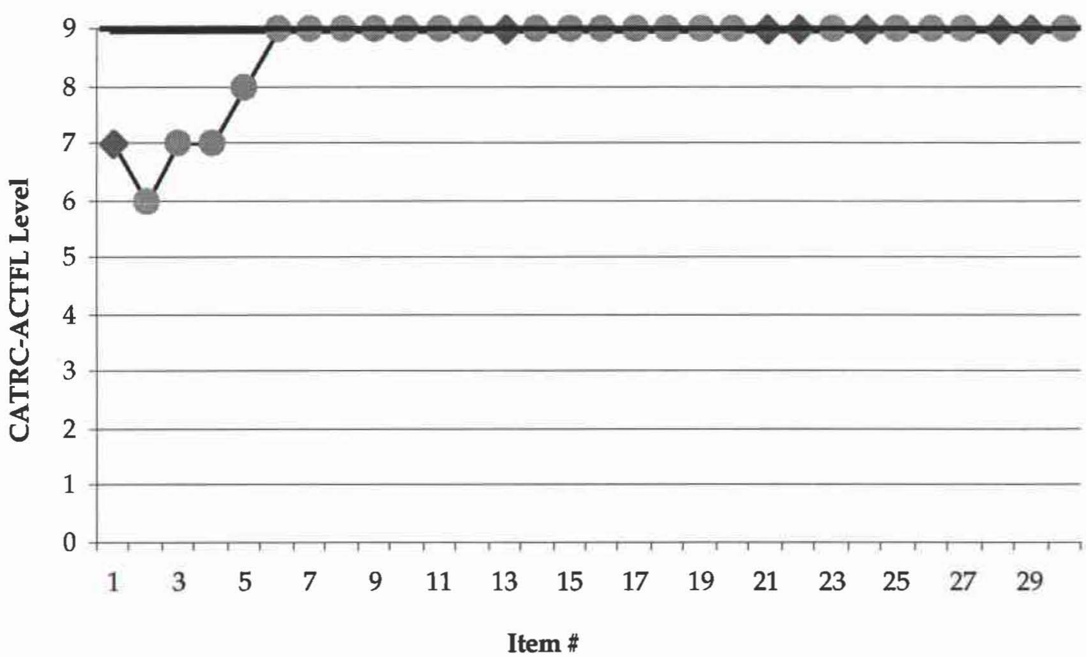
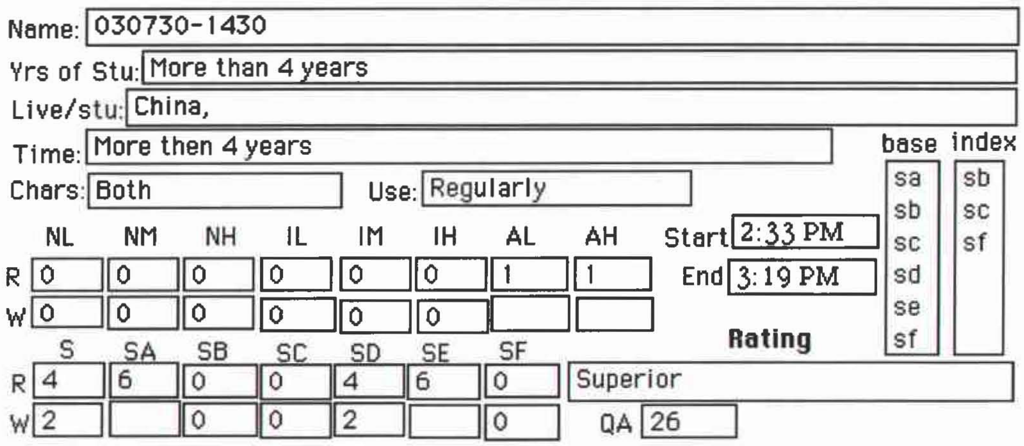


Figure 49. Participant #NNS05 CATRC Progression.



Name: 030730-1430
Yrs of Stu: More than 4 years
Live/stu: China,
Time: More then 4 years
Chars: Both
Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	0	1	1
W	0	0	0	0	0	0		

Start 2:33 PM
End 3:19 PM

base	index
sa	sb
sb	sc
sc	sf
sd	
se	
sf	

	S	SA	SB	SC	SD	SE	SF
R	4	6	0	0	4	6	0
W	2		0	0	2		0

Rating: Superior
QA 26

Figure 50. Participant #NNS06 CATRC Report.

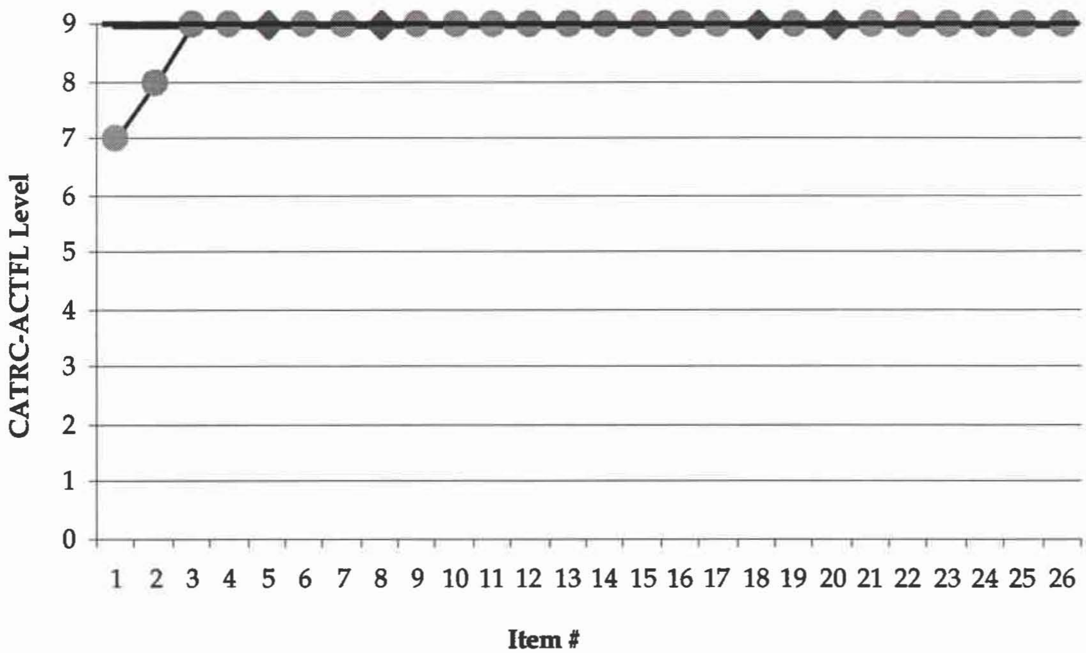
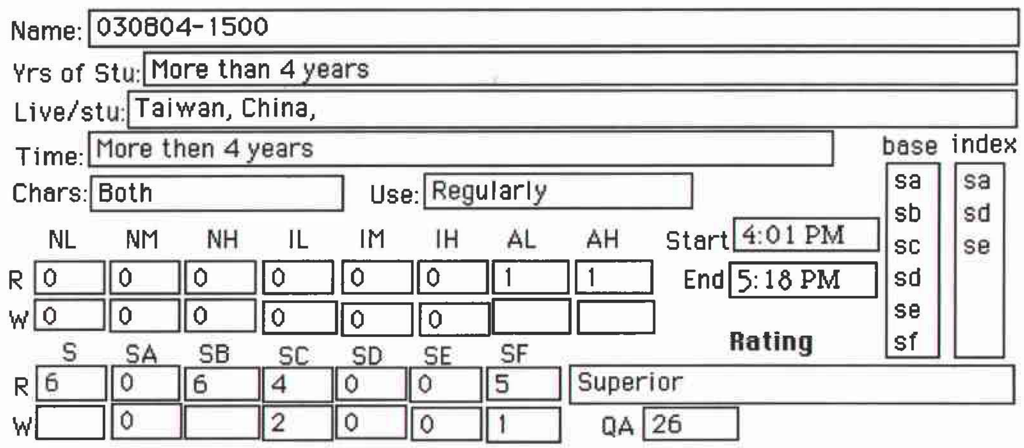


Figure 51. Participant #NNS06 CATRC Progression.



Name: 030804-1500

Yrs of Stu: More than 4 years

Live/stu: Taiwan, China,

Time: More then 4 years

Chars: Both

Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	0	1	1
W	0	0	0	0	0	0		

Start: 4:01 PM

End: 5:18 PM

	S	SA	SB	SC	SD	SE	SF
R	6	0	6	4	0	0	5
W		0		2	0	0	1

Rating: Superior

QA: 26

base	index
sa	sa
sb	sd
sc	se
sd	
se	
sf	

Figure 52. Participant #NNS07 CATRC Report.

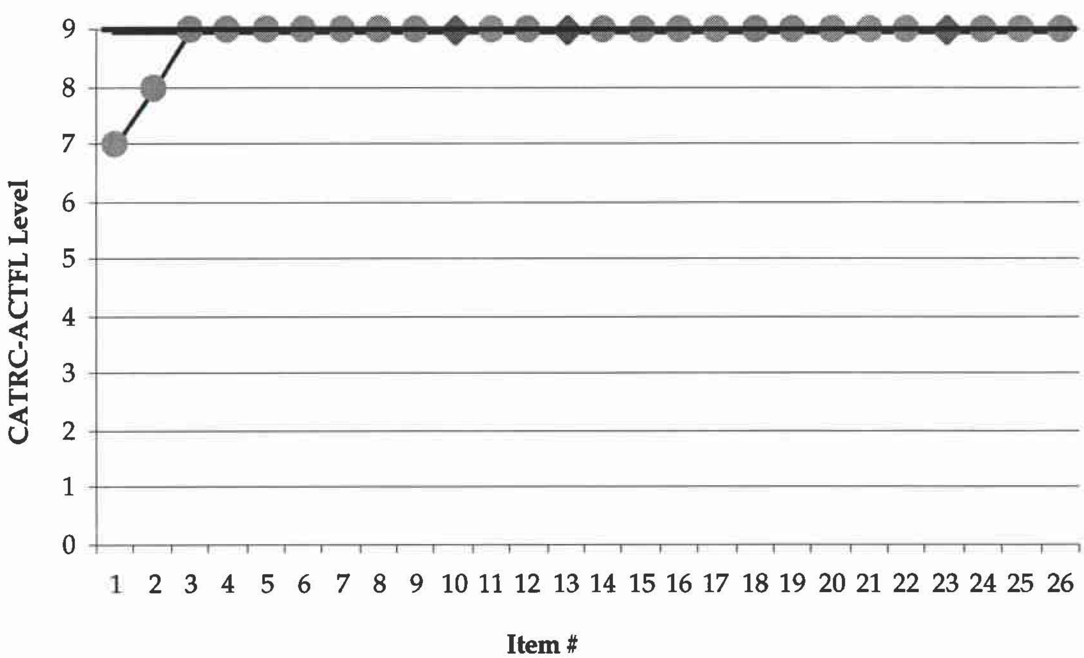


Figure 53. Participant #NNS07 CATRC Progression.

Native Superior (NS) Participants

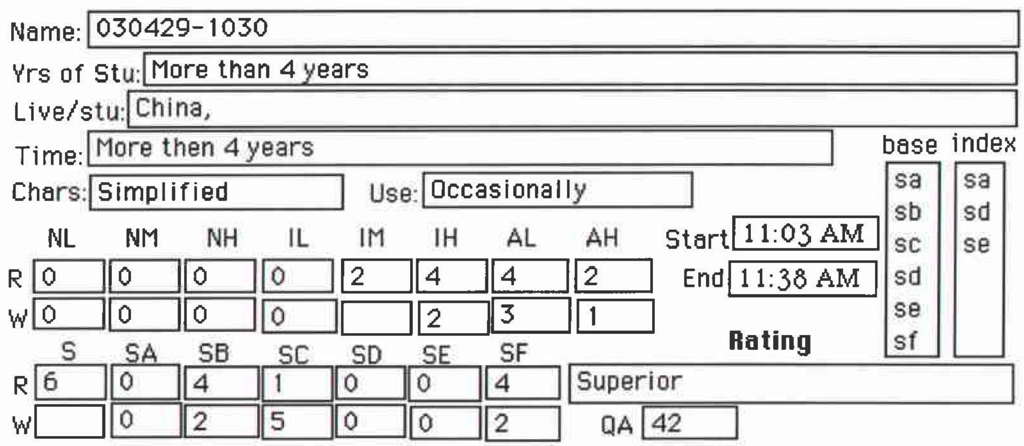


Figure 54. Participant #NS01 CATRC Report.

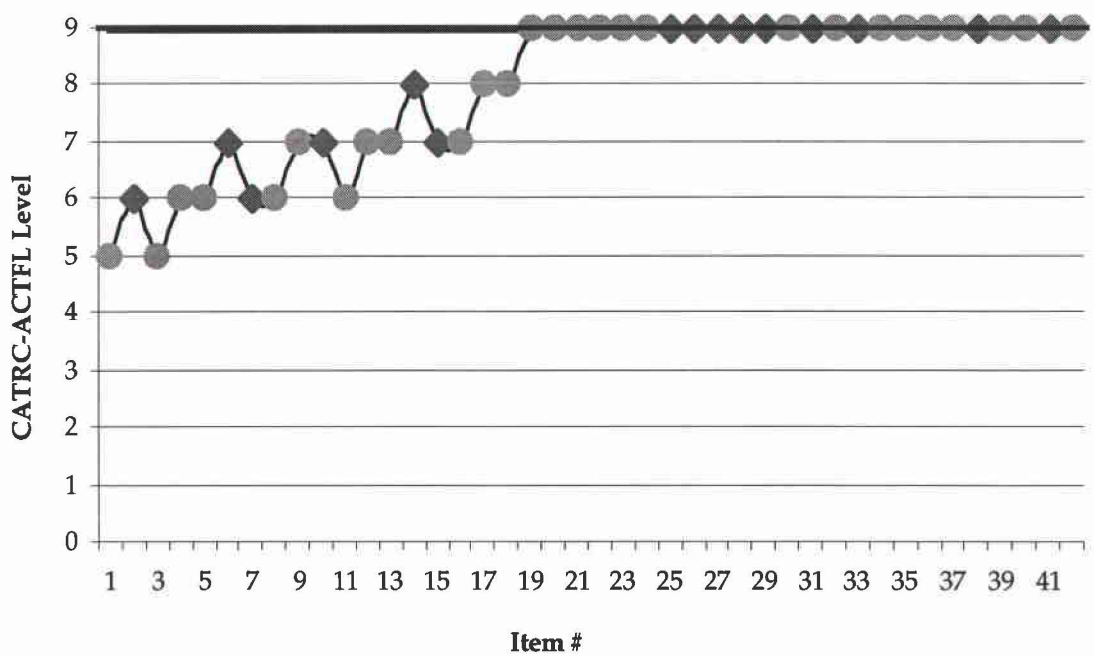
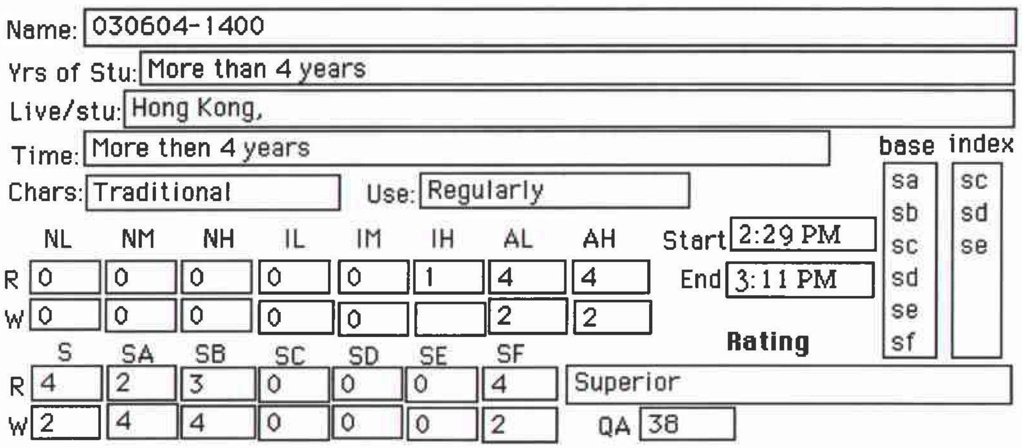


Figure 55. Participant #NS01 CATRC Progression.



Name: 030604-1400

Yrs of Stu: More than 4 years

Live/stu: Hong Kong,

Time: More then 4 years

Chars: Traditional Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	1	4	4
W	0	0	0	0	0		2	2

	S	SA	SB	SC	SD	SE	SF
R	4	2	3	0	0	0	4
W	2	4	4	0	0	0	2

Start 2:29 PM

End 3:11 PM

Rating: Superior

QA 38

base: sa, sb, sc, sd, se, sf

index: sc, sd, se

Figure 56. Participant #NS02 CATRC Report.

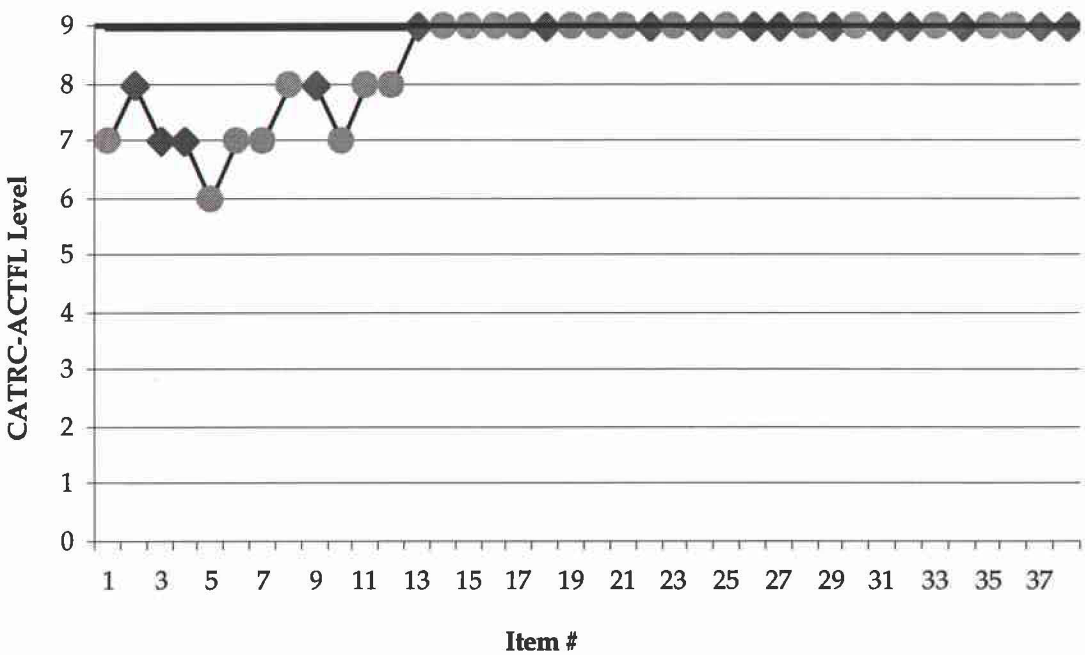
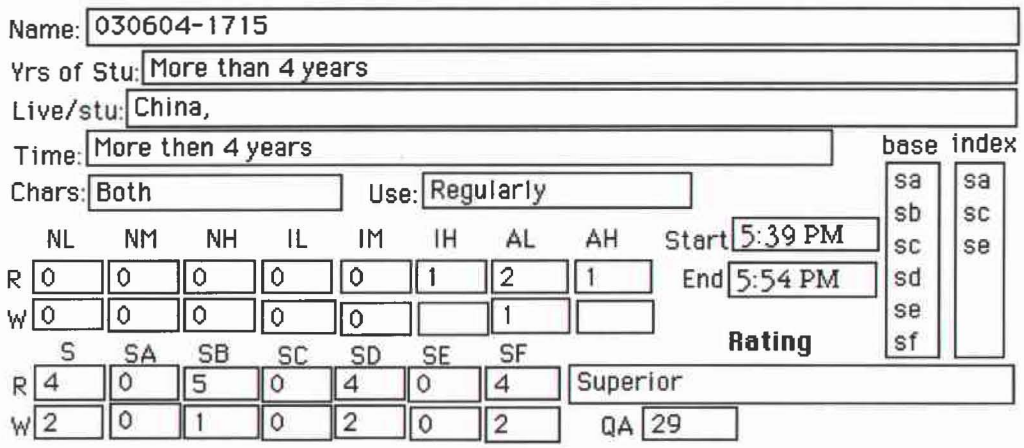


Figure 57. Participant #NS02 CATRC Progression.



Name: 030604-1715
Yrs of Stu: More than 4 years
Live/stu: China,
Time: More then 4 years
Chars: Both
Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	1	2	1
W	0	0	0	0	0		1	

Start 5:39 PM
End 5:54 PM

	S	SA	SB	SC	SD	SE	SF
R	4	0	5	0	4	0	4
W	2	0	1	0	2	0	2

Rating: Superior
QA 29

base	index
sa	sa
sb	sc
sc	se
sd	
se	
sf	

Figure 58. Participant #NS03 CATRC Report.

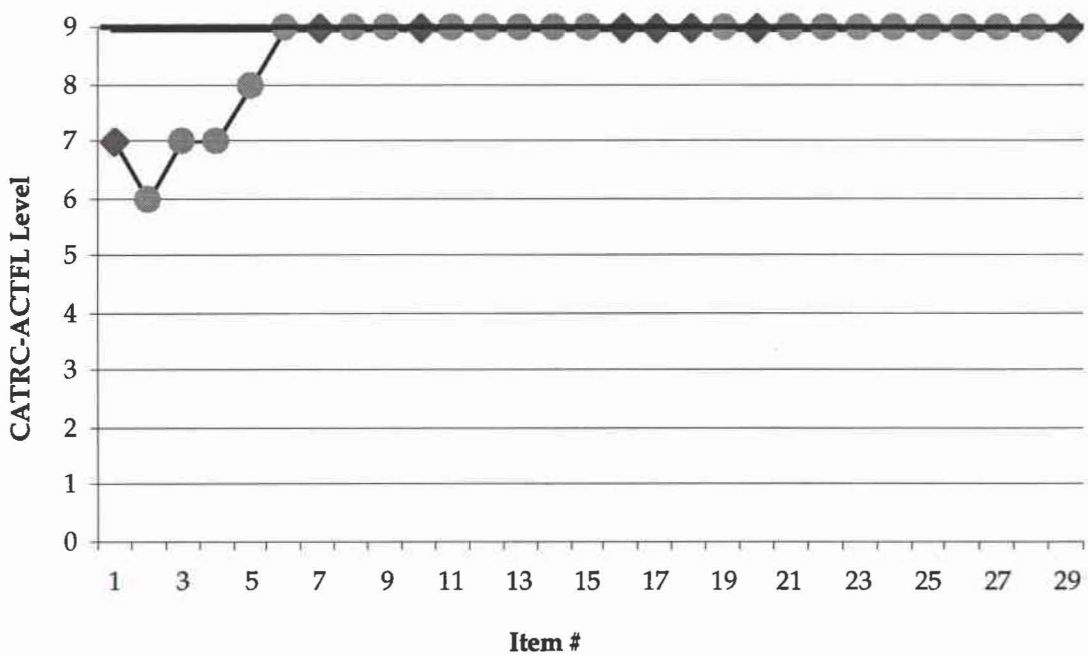
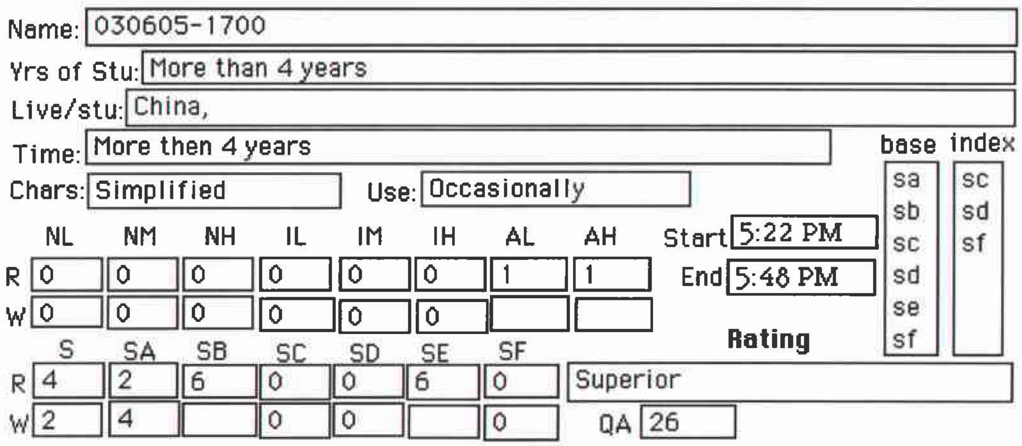


Figure 59. Participant #NS03 CATRC Progression.



Name: 030605-1700
Yrs of Stu: More than 4 years
Live/stu: China,
Time: More then 4 years
Chars: Simplified
Use: Occasionally
Start 5:22 PM
End 5:48 PM

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	0	1	1
W	0	0	0	0	0	0		

	S	SA	SB	SC	SD	SE	SF
R	4	2	6	0	0	6	0
W	2	4		0	0		0

Rating: Superior
QA 26

base: sa, sb, sc, sd, se, sf
index: sc, sd, sf

Figure 60. Participant #NS04 CATRC Report.

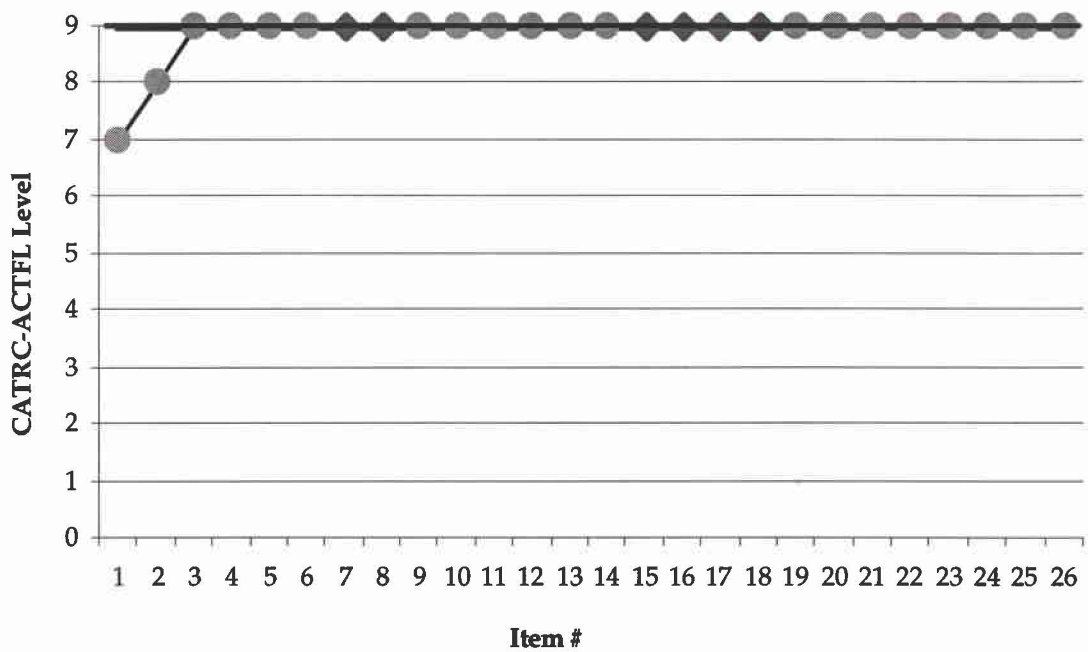
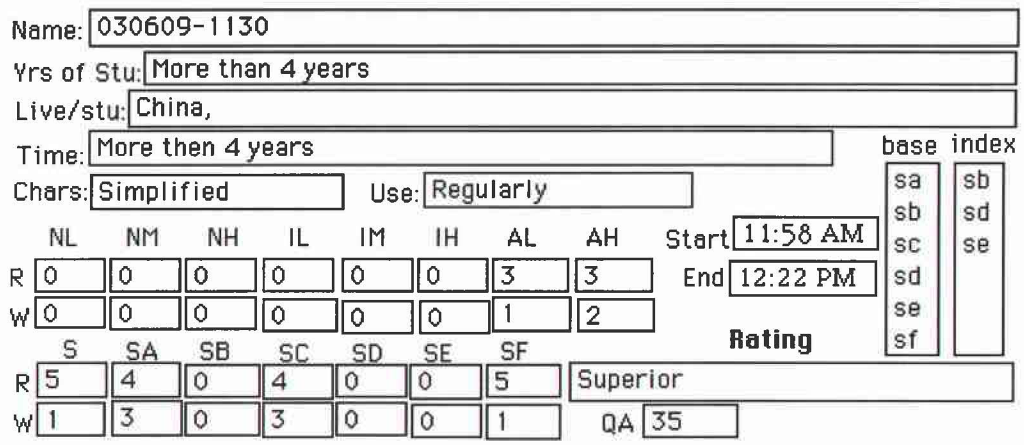


Figure 61. Participant #NS04 CATRC Progression.



Name: 030609-1130
Yrs of Stu: More than 4 years
Live/stu: China,
Time: More then 4 years
Chars: Simplified
Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	0	3	3
W	0	0	0	0	0	0	1	2

	S	SA	SB	SC	SD	SE	SF
R	5	4	0	4	0	0	5
W	1	3	0	3	0	0	1

Start 11:58 AM
End 12:22 PM
Rating
Superior
QA 35

base	index
sa	sb
sb	sd
sc	se
sd	
se	
sf	

Figure 62. Participant #NS05 CATRC Report.

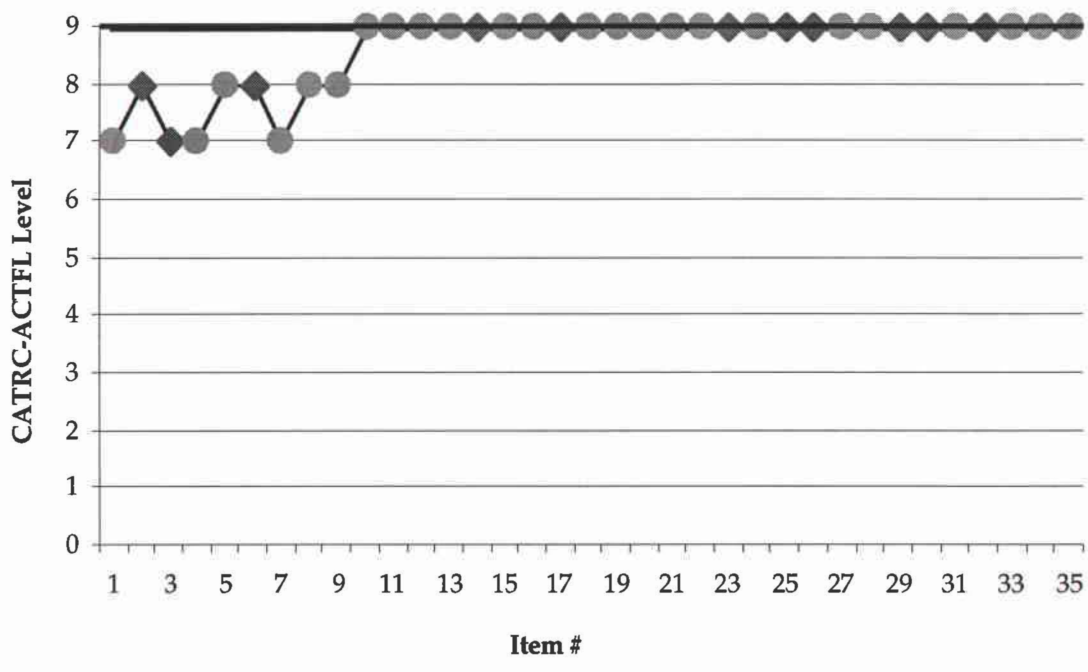
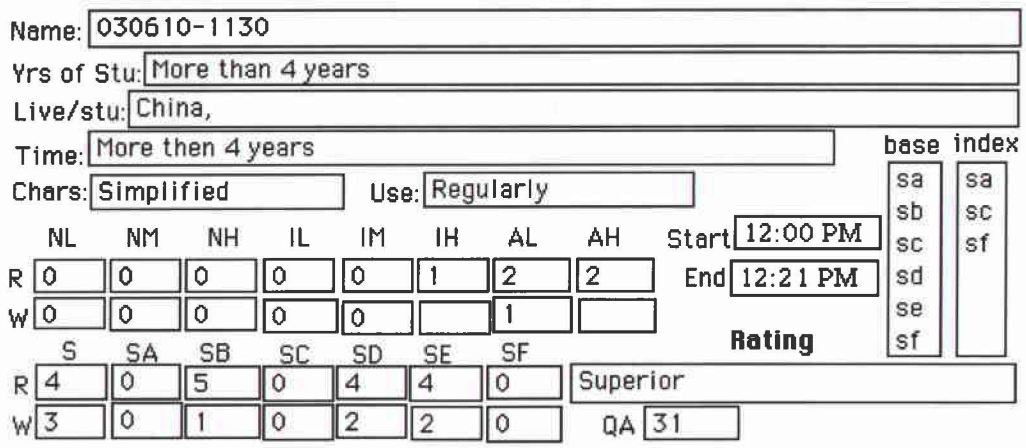


Figure 63. Participant #NS05 CATRC Progression.


Name: 030610-1130
Yrs of Stu: More than 4 years
Live/stu: China,
Time: More then 4 years
Chars: Simplified   Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	1	2	2
W	0	0	0	0	0		1	

Start 12:00 PM
End 12:21 PM

	S	SA	SB	SC	SD	SE	SF
R	4	0	5	0	4	4	0
W	3	0	1	0	2	2	0

Rating: Superior
QA 31

base	index
sa	sa
sb	sc
sc	sf
sd	
se	
sf	

Figure 64. Participant #NS06 CATRC Report.

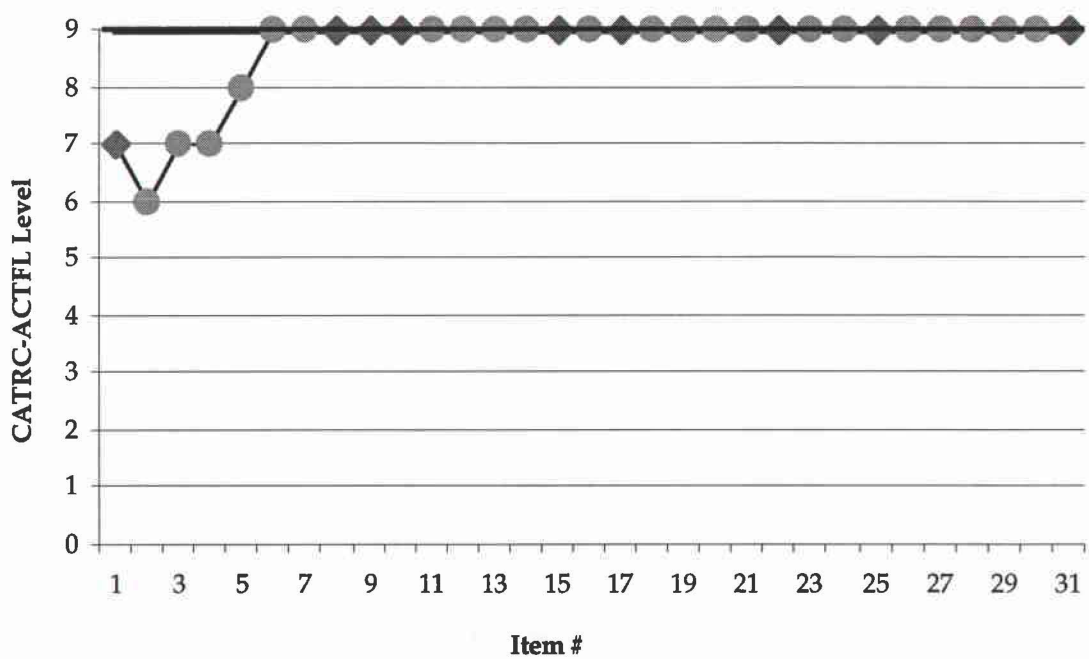


Figure 65. Participant #NS06 CATRC Progression.

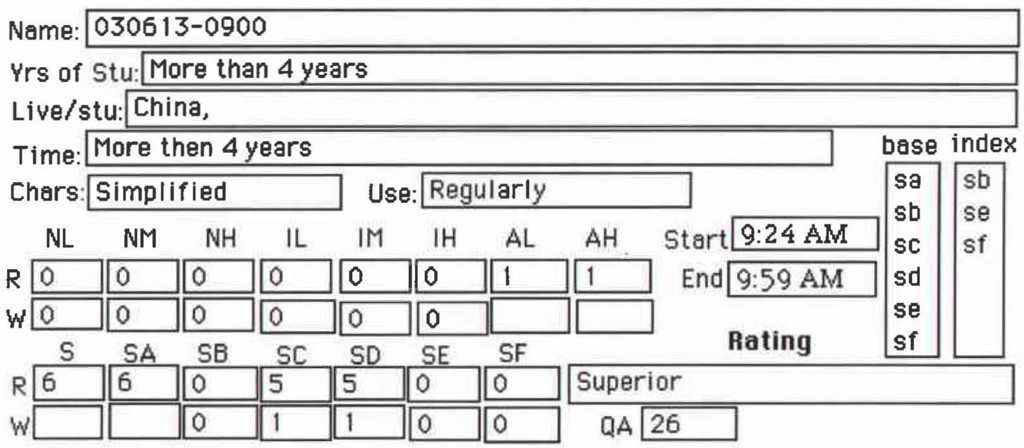


Figure 66. Participant #NS07 CATRC Report.

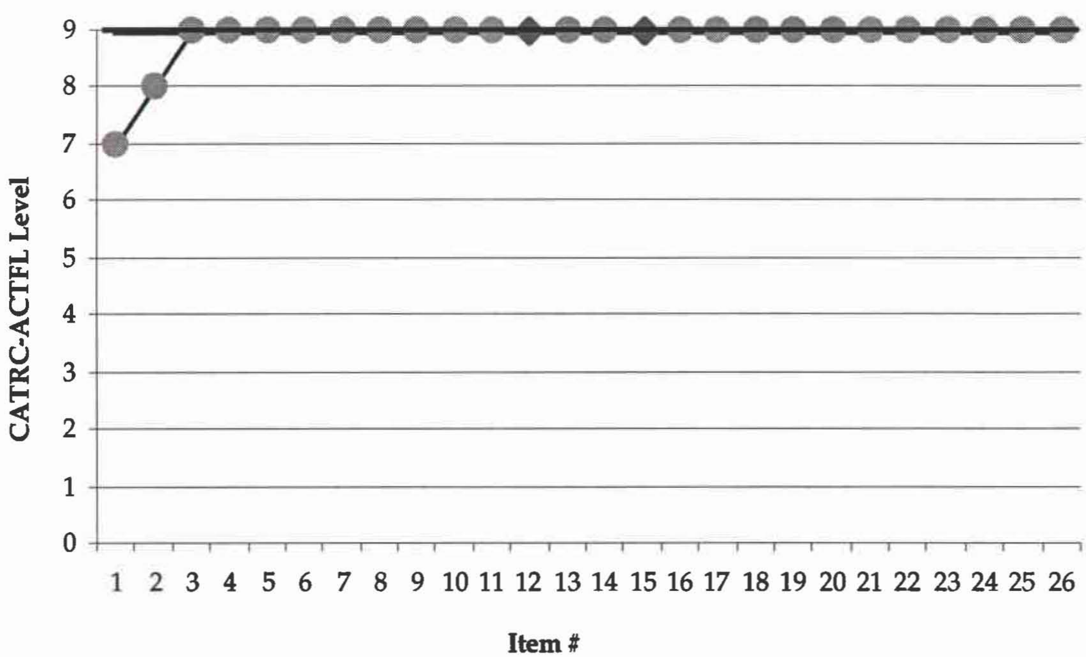
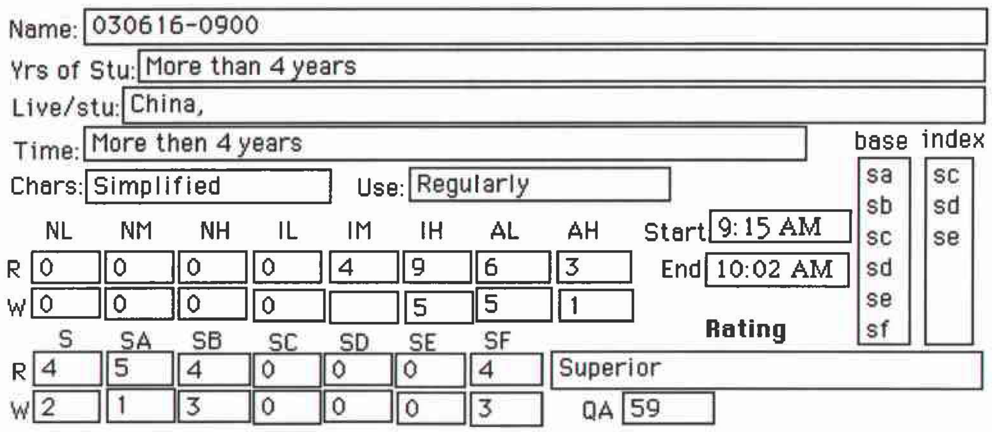


Figure 67. Participant #NS07 CATRC Progression.



	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	4	9	6	3
W	0	0	0	0		5	5	1

	S	SA	SB	SC	SD	SE	SF
R	4	5	4	0	0	0	4
W	2	1	3	0	0	0	3

Figure 68. Participant #NS08 CATRC Report.

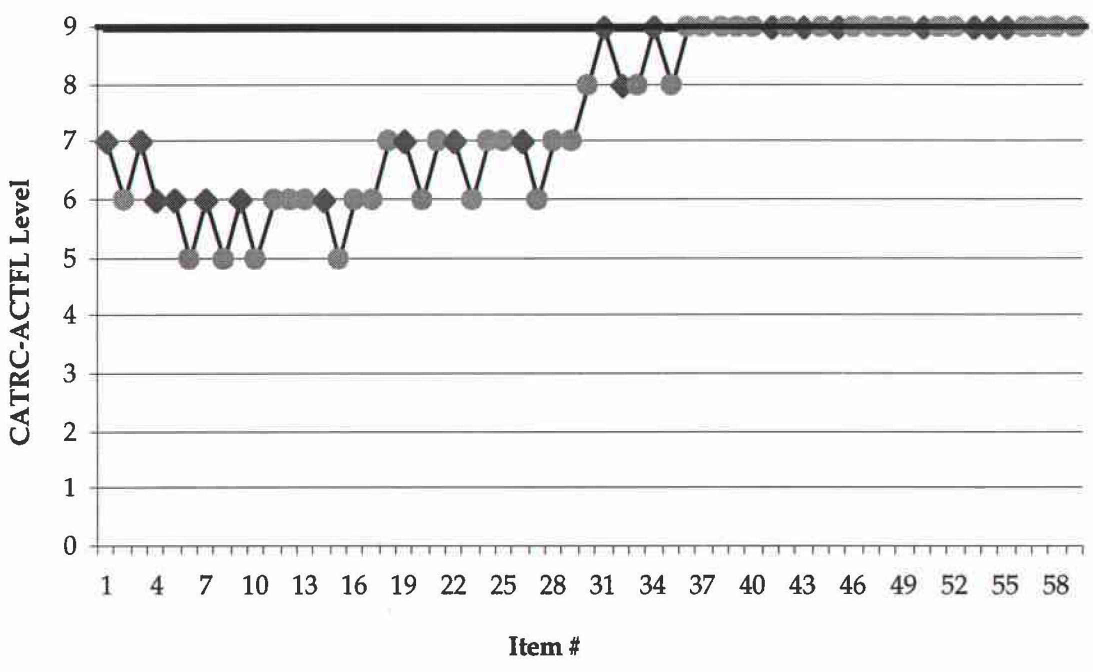


Figure 69. Participant #NS08 CATRC Progression.

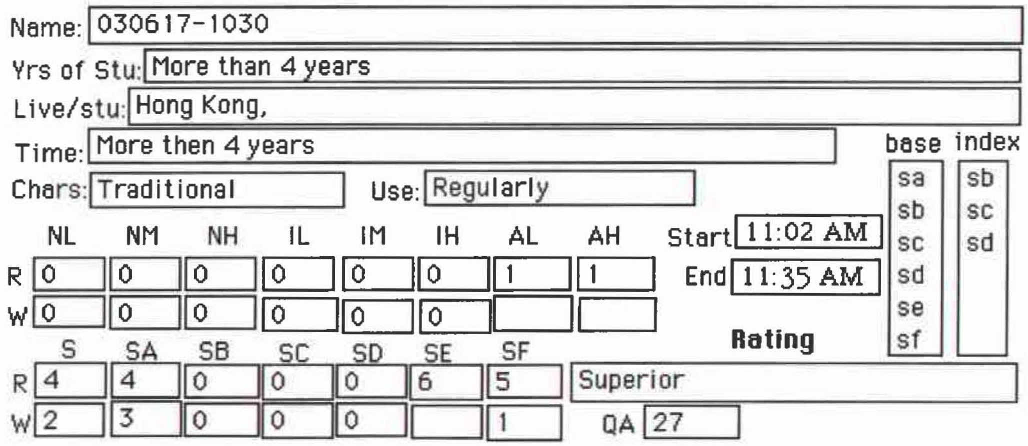


Figure 70. Participant #NS09 CATRC Report.

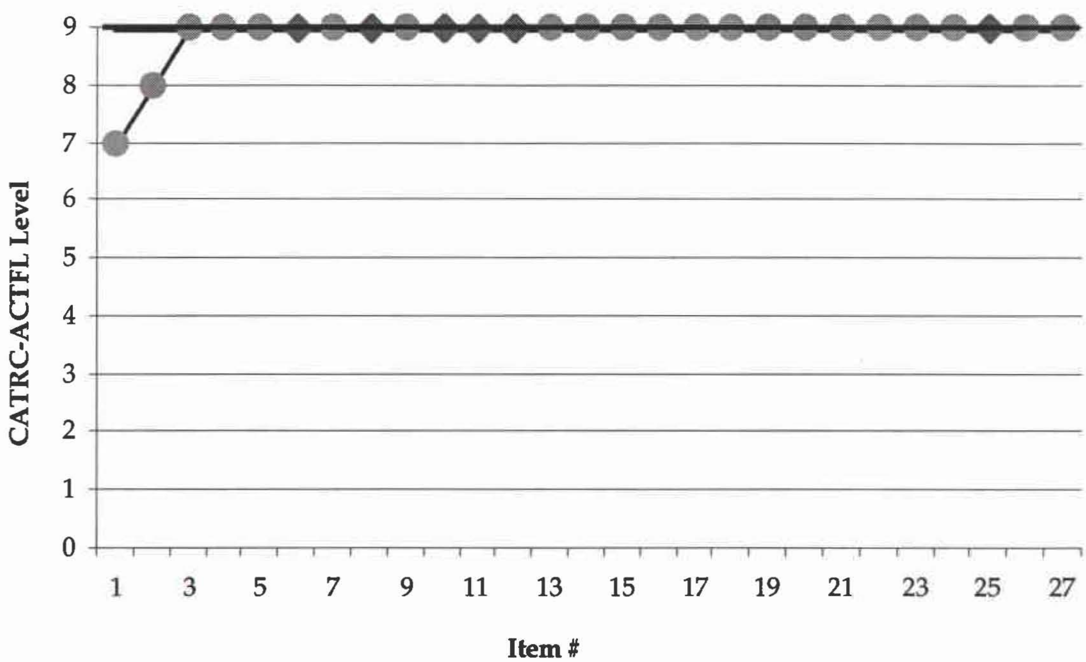


Figure 71. Participant #NS09 CATRC Progression.

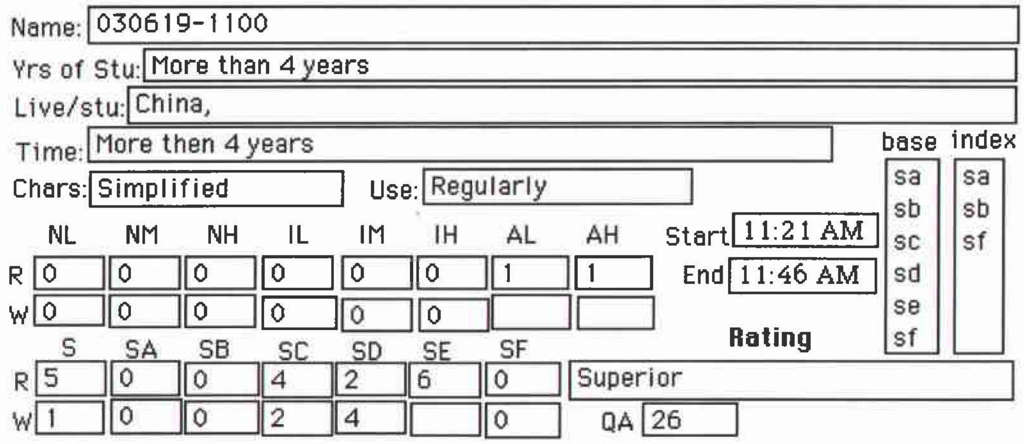


Figure 72. Participant #NS10 CATRC Report.

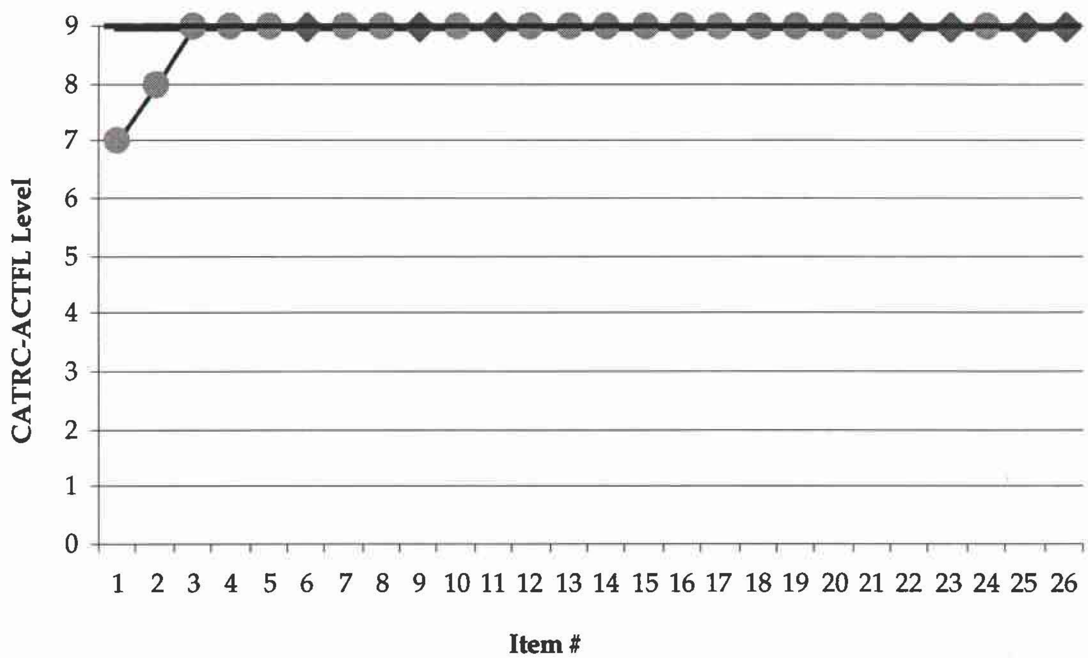


Figure 73. Participant #NS10 CATRC Progression.

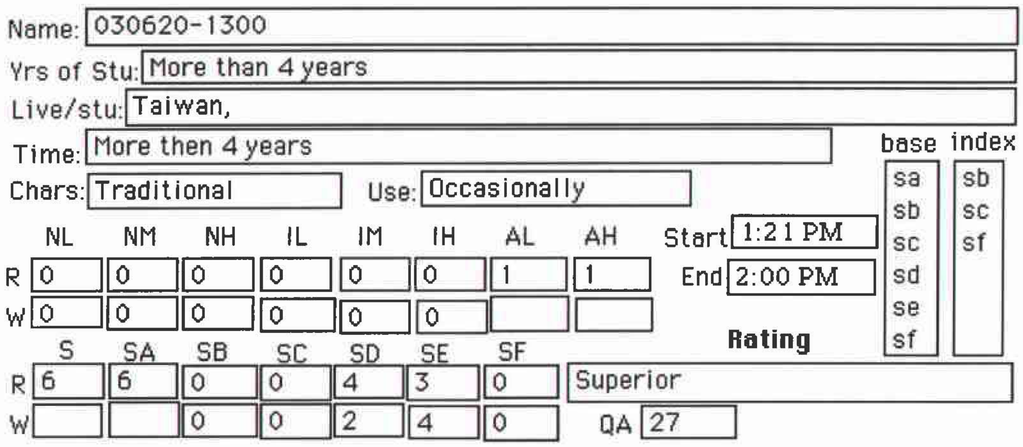


Figure 74. Participant #NS11 CATRC Report.

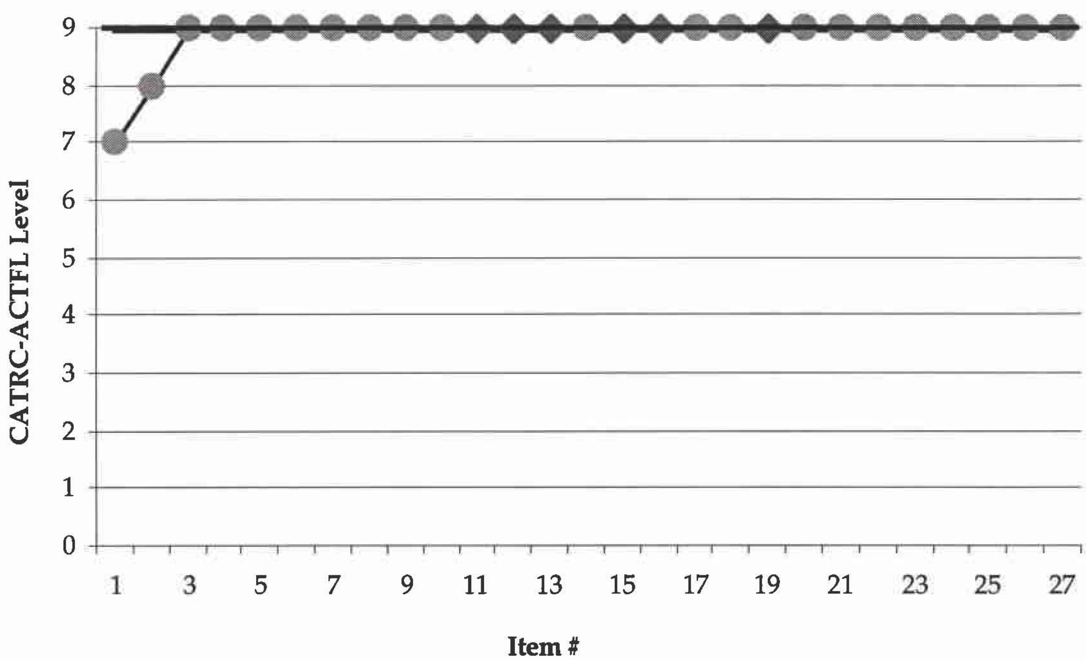
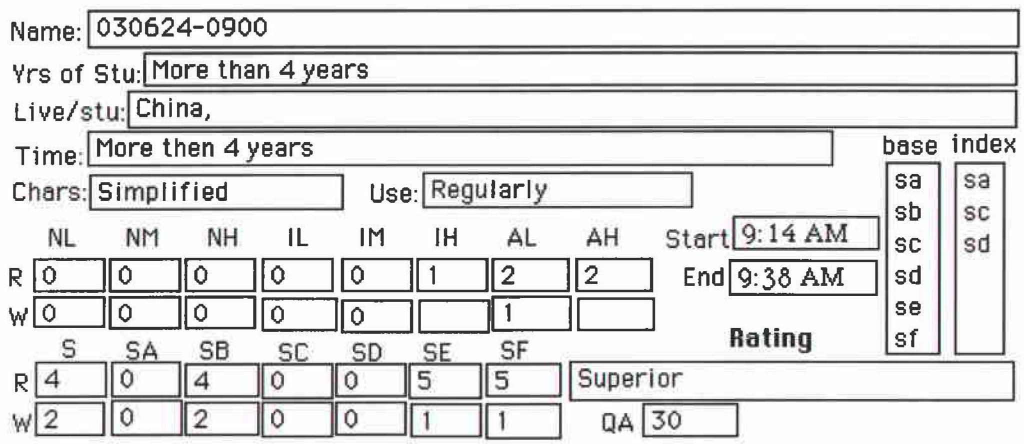


Figure 75. Participant #NS11 CATRC Progression.


Name: 030624-0900
Yrs of Stu: More than 4 years
Live/stu: China,
Time: More then 4 years
Chars: Simplified
Use: Regularly

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	1	2	2
W	0	0	0	0	0		1	

	S	SA	SB	SC	SD	SE	SF
R	4	0	4	0	0	5	5
W	2	0	2	0	0	1	1

Start 9:14 AM
End 9:38 AM
Rating
Superior
QA 30
base: sa sb sc sd se sf
index: sa sc sd

Figure 76. Participant #NS12 CATRC Report.

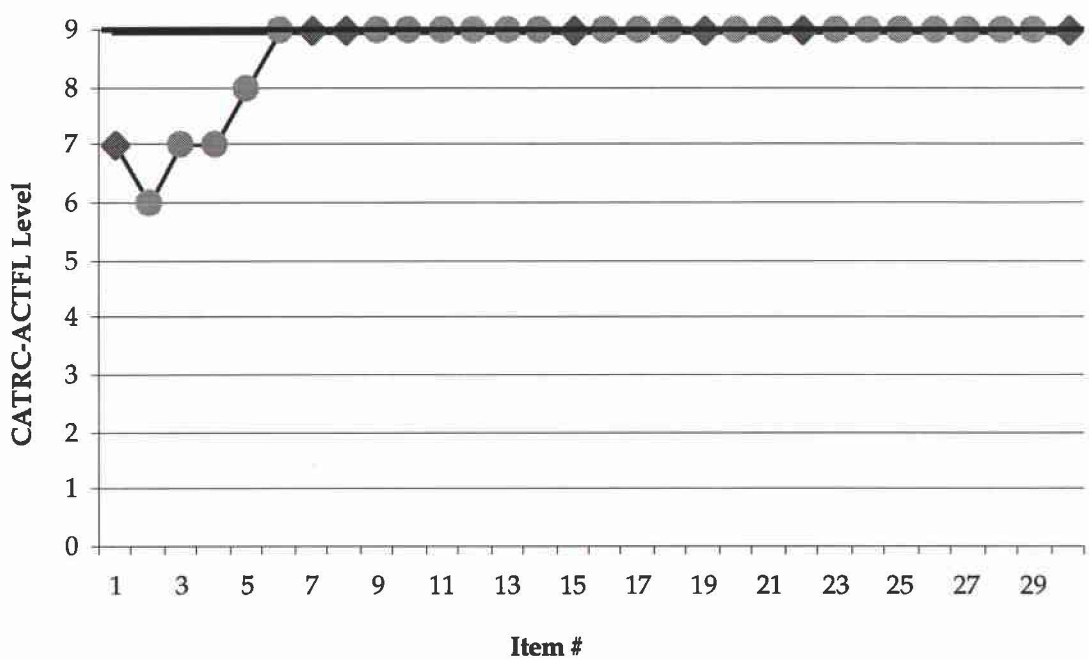
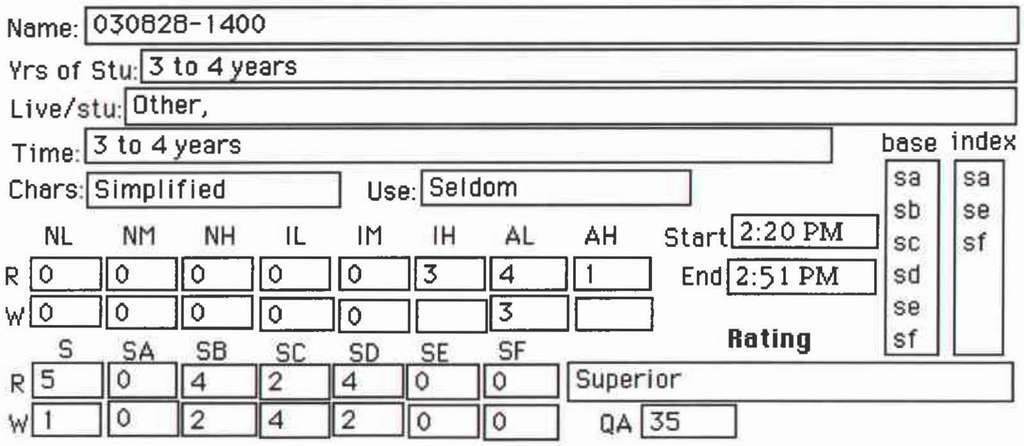


Figure 77. Participant #NS12 CATRC Progression.



Name: 030828-1400
Yrs of Stu: 3 to 4 years
Live/stu: Other,
Time: 3 to 4 years
Chars: Simplified Use: Seldom
Start 2:20 PM
End 2:51 PM

	NL	NM	NH	IL	IM	IH	AL	AH
R	0	0	0	0	0	3	4	1
W	0	0	0	0	0		3	

	S	SA	SB	SC	SD	SE	SF
R	5	0	4	2	4	0	0
W	1	0	2	4	2	0	0

Rating: Superior
QA 35

base	index
sa	sa
sb	se
sc	sf
sd	
se	
sf	

Figure 78. Participant #NS13 CATRC Report.

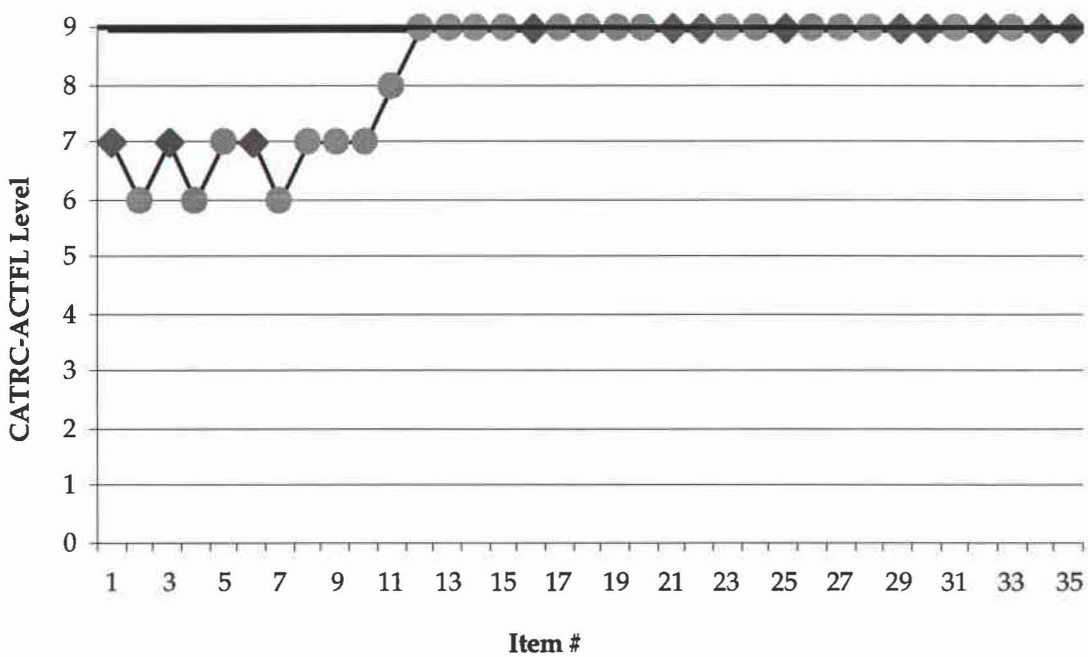


Figure 79. Participant #NS13 CATRC Progression.

APPENDIX H

PARTICIPANT CATRC ITEM SETS

Abbreviations:

QA = Question Asked (i.e., sequential order)
ST = Simplified or Traditional Chinese characters (s, t, or z)
- s = Simplified
- t = Traditional
- z = Both Simplified and Traditional

CA = Correct Answer (a, b, c, or d)
SA = Student Answer (a, b, c, or d)
CN = Card Number (i.e., Item ID)
- nl = Novice Low
- nm = Novice Mid
- nh = Novice High
- il = Intermediate Low
- im = Imtermediate Mid
- ih = Intermediate High
- al = Advanced Low (i.e., ACTFL Advanced)
- ah = Advanced High (i.e., ACTFL Advanced Plus)
- sX = Superior (X stands for a-f, depending on category)

Non-native Intermediate (NNI) Participants

Table 11. Participant #NNI01 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	z	d	d	im49	0018	z	d	c	im16
0002	t	b	a	ih81	0019	z	c	c	im42
0003	z	a	b	im23	0020	t	d	c	ih73
0004	z	b	b	im32	0021	z	d	a	im22
0005	s	c	b	ih39	0022	z	b	d	im15
0006	z	a	c	im33	0023	z	c	c	il23
0007	s	b	b	im39	0024	z	b	a	im19
0008	s	a	b	ih44	0025	z	d	d	il08
0009	z	d	a	im53	0026	s	d	d	im08
0010	z	c	c	im29	0027	z	c	b	im20
0011	s	d	b	ih08	0028	z	d	d	il41
0012	z	d	a	im10	0029	z	a	c	im48
0013	s	d	c	im37	0030	t	a	a	il01
0014	z	a	a	il39	0031	s	b	a	im35
0015	z	c	c	im50	0032	z	a	a	il35
0016	z	c	c	im59	0033	z	a	b	im07
0017	s	a	b	ih71					

Table 12. Participant #NNI02 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	z	a	b	im43	0016	z	b	b	il36
0002	z	a	b	il18	0017	z	c	c	im58
0003	z	b	c	nh47	0018			b	im03
0004	z	c	c	nm15	0019	z	b	b	il21
0005	z	b	b	nh08	0020	z	b	c	im40
0006	z	c	c	nh49	0021	z	b	a	il25
0007	z	c	c	il16	0022	t	d	b	il06
0008	z	a	d	il31	0023	z	d	b	il12
0009	z	c	c	nh61	0024	z	b	b	il19
0010	z	c	c	il13	0025	z	b	b	im11
0011	z	b	b	il44	0026	z	a	c	im51
0012	z	c	c	im52	0027	z	b	d	il17
0013	z	b	c	im57	0028	z	a	a	il45
0014	z	b	b	il37	0029	t	d	b	im04
0015	s	a	d	im02					

Table 13. Participant #NNI03 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	z	b	d	im57	0017	t	c	c	ih05
0002	z	a	a	il45	0018	t	b	d	al02
0003	z	a	a	im33	0019	s	b	a	ih29
0004	z	a	c	im51	0020	s	b	d	ih25
0005	z	a	a	il39	0021	s	d	a	ih17
0006	z	b	b	im32	0022	z	d	c	im46
0007	z	a	c	im07	0023	s	b	b	im54
0008	z	b	c	il27	0024	s	c	d	ih32
0009	z	c	c	il34	0025	z	a	a	im41
0010	z	b	b	im11	0026	s	d	d	ih10
0011	s	d	d	im08	0027	t	a	a	ih78
0012	s	d	d	ih26	0028	t	c	c	ih54
0013	s	b	a	al36	0029	t	c	b	al39
0014	s	d	d	ih50	0030	t	b	d	ih52
0015	s	c	b	al07	0031	s	d	c	ih67
0016	s	d	b	ih08					

Table 14. Participant #NNI04 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	z	d	c	im16	0024	t	d	b	ah02
0002	z	b	b	il36	0025	t	d	a	al12
0003	z	c	d	im42	0026	s	b	c	al31
0004	z	b	b	il25	0027	s	c	a	ih48
0005	z	c	c	im59	0028	t	b	a	ih62
0006	z	d	a	im22	0029	z	d	d	im49
0007	z	b	a	il30	0030	t	d	d	ih02
0008	z	c	c	il04	0031	z	c	b	ih11
0009	z	d	d	im10	0032	z	c	c	im30
0010	s	d	d	im36	0033	t	c	c	ih77
0011	z	a	c	im48	0034	t	b	d	ih01
0012	z	b	b	il21	0035	z	c	a	im58
0013	z	d	d	im53	0036	z	a	a	im43
0014	z	c	c	im52	0037	t	c	b	ih03
0015	z	a	a	ih15	0038	z	b	b	im15
0016	t	d	d	al44	0039	s	b	b	ih09
0017	s	d	d	ah15	0040	t	b	b	ih64
0018	t	b	d	sq19	0041	t	c	c	ih07
0019	s	b	c	ah42	0042	s	d	d	al11
0020	t	d	c	ah14	0043	t	d	b	al58
0021	t	d	b	al01	0044	s	c	b	ih30
0022	t	b	b	al47	0045	s	a	b	ih44
0023	t	b	b	ah30					

Table 15. Participant #NNI05 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	d	b	al01	0014	t	b	b	ih60
0002	s	b	a	ih35	0015	s	c	d	al10
0003	s	b	b	im35	0016	s	a	a	ih12
0004	t	b	b	ih74	0017	t	d	d	al19
0005	s	a	d	ih37	0018	t	c	c	al45
0006	z	d	b	im01	0019	t	c	c	al04
0007	z	c	c	im56	0020	s	a	a	al28
0008	t	a	d	ih51	0021	s	a	b	ah35
0009	s	a	a	im02	0022	t	b	c	al03
0010	t	a	a	ih75	0023			c	al54
0011	t	a	a	ih69	0024	s	c	c	ih80
0012	s	c	c	ih23	0025	t	c	c	al05
0013	t	b	d	al60	0026	t	b	d	al37

Non-native Superior (NNS) Participants

Table 16. Participant #NNS01 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	b	a	al53	0019	s	b	b	se19
0002	t	c	d	ih19	0020	t	b	c	se21
0003	z	a	a	im23	0021	t	c	c	se09
0004	t	a	a	ih68	0022	s	c	c	se12
0005	s	a	a	ih71	0023	s	b	b	se14
0006	t	c	c	al62	0024	s	a	c	sc04
0007	s	d	d	al34	0025	t	c	c	sc16
0008	t	c	c	ah04	0026	t	b	b	sc18
0009	t	d	b	sq18	0027	s	b	b	sc07
0010	s	a	a	ah16	0028	s	c	b	sc06
0011	t	a	a	sq08	0029	t	c	b	sc15
0012	t	c	c	sq12	0030	t	c	b	sc17
0013	t	b	d	sq09	0031	t	a	c	sa21
0014	s	b	c	sq21	0032	s	c	b	sa22
0015	s	a	a	ah41	0033	s	b	b	sa16
0016	s	b	b	sq23	0034	t	d	c	sa08
0017	t	d	d	sq04	0035	s	c	c	sa11
0018	t	d	b	se03	0036	t	d	c	sa15

Table 17. Participant #NNS02 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001			d	im03	0020	s	a	a	sa19
0002	z	c	c	il23	0021	s	c	c	sa13
0003	t	d	d	im04	0022	s	b	b	sa06
0004	s	b	b	im39	0023	t	c	c	sa09
0005	t	c	b	ih63	0024	s	a	a	sa12
0006	z	b	b	im19	0025	t	d	d	sa20
0007	t	d	d	ih73	0026	s	c	c	se01
0008	t	d	d	ih66	0027	t	c	c	se05
0009	s	c	c	al24	0028	t	b	b	se07
0010	t	b	c	ah05	0029	s	a	a	se15
0011	t	b	b	al08	0030	t	c	c	se10
0012	t	a	a	ah56	0031	s	c	c	se13
0013	t	c	c	ah24	0032	s	d	d	sd11
0014	s	b	b	sq11	0033	t	a	c	sd09
0015	s	c	c	sq22	0034	s	c	c	sd16
0016	t	a	b	sq26	0035	t	b	c	sd04
0017	t	b	a	sq28	0036	t	d	d	sd23
0018	s	b	b	sq01	0037	t	a	b	sd02
0019	s	a	a	sq02	0038	s	d	a	sd17

Table 18. Participant #NNS03 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	d	d	al40	0015	t	d	d	sa04
0002	t	d	d	ah33	0016	s	a	a	sa05
0003	s	c	b	sq03	0017	s	a	a	se17
0004	t	b	b	ah08	0018	t	a	d	se04
0005	s	b	d	sq16	0019	s	d	d	se20
0006	t	c	c	ah48	0020	t	c	c	se02
0007	t	a	a	sq07	0021	s	b	d	se16
0008	t	b	b	sq10	0022	t	b	b	se06
0009	t	b	b	sq20	0023	t	b	a	sd18
0010	t	a	a	sq14	0024	t	b	b	sd15
0011	s	b	b	sa02	0025	s	c	c	sd22
0012	t	d	d	sa14	0026	s	d	d	sd12
0013	s	b	d	sa01	0027	t	b	b	sd01
0014	s	c	c	sa03	0028	s	d	c	sd07

Table 19. Participant #NNS04 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	s	c	c	al46	0014	t	b	b	sf20
0002	t	d	d	ah02	0015	s	c	c	se13
0003	s	b	b	sq25	0016	t	b	a	se07
0004	s	b	d	sq16	0017	t	b	c	se21
0005	t	a	a	sq07	0018	s	a	d	se18
0006	t	a	a	sq14	0019	s	c	b	se01
0007	s	c	c	sq22	0020	s	c	c	se12
0008	t	b	b	sq10	0021	t	b	b	sd03
0009	s	c	b	sf05	0022	t	c	c	sd19
0010	s	a	a	sf03	0023	t	b	b	sd01
0011	t	d	d	sf12	0024	t	b	b	sd04
0012	t	d	d	sf13	0025	s	d	d	sd12
0013	s	a	a	sf02	0026	t	d	d	sd14

Table 20. Participant #NNS05 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	s	b	c	al26	0016	s	a	a	sf01
0002	s	c	c	ih30	0017	s	d	d	sf04
0003	t	a	a	al55	0018	t	d	d	sb06
0004	s	c	c	al32	0019	s	c	c	sb16
0005	t	a	a	ah32	0020	s	b	b	sb02
0006	t	d	d	sq04	0021	t	b	a	sb22
0007	s	c	c	sq22	0022	t	b	a	sb10
0008	t	b	b	sq10	0023	t	b	b	sb09
0009	t	a	a	sq15	0024	s	c	b	se13
0010	t	a	a	sq08	0025	s	a	a	se17
0011	t	d	d	sq29	0026	s	c	c	se12
0012	s	d	d	sf08	0027	t	b	b	se21
0013	t	a	d	sf19	0028	s	d	c	se11
0014	s	d	d	sf17	0029	t	b	a	se07
0015	t	b	b	sf15	0030	t	d	d	se03

Table 21. Participant #NNS06 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	b	b	al56	0014	s	a	a	sa17
0002	t	d	d	ah02	0015	s	c	c	sd22
0003	s	b	b	sq01	0016	t	d	d	sd06
0004	s	b	b	sq01	0017	t	c	c	sd19
0005	s	a	d	sq02	0018	t	b	d	sd05
0006	t	d	d	sq29	0019	t	b	b	sd03
0007	s	c	c	sq03	0020	t	a	c	sd09
0008	t	b	a	sq28	0021	t	c	c	se02
0009	s	d	d	sa18	0022	s	b	b	se16
0010	t	a	a	sa10	0023	s	c	c	se01
0011	t	d	d	sa04	0024	s	c	c	se01
0012	s	c	c	sa03	0025	t	d	d	se03
0013	s	d	d	sa23	0026	t	b	b	se21

Table 22. Participant #NNS07 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	s	c	c	al24	0014	s	a	a	sc12
0002	s	d	d	ah04	0015	t	d	d	sb06
0003	t	a	a	sq14	0016	t	b	b	sb19
0004	s	b	b	sq21	0017	s	c	c	sb16
0005	s	b	b	sq25	0018	s	d	d	sb21
0006	s	d	d	sq24	0019	s	b	b	sb13
0007	t	d	d	sq29	0020	s	b	b	sb02
0008	t	b	b	sq10	0021	t	c	c	sf09
0009	t	a	a	sc08	0022	t	c	c	sf16
0010	t	c	d	sc15	0023	t	a	b	sf07
0011	s	a	a	sc05	0024	s	a	a	sf02
0012	s	b	b	sc13	0025	t	c	c	sf14
0013	t	c	d	sc17	0026	t	a	a	sf11

Native Superior (NS) Participants

Table 23. Participant #NS01 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	z	b	b	im40	0022	t	c	c	sq05
0002	t	a	b	ih61	0023	s	d	d	sq13
0003	z	c	c	im50	0024	s	b	b	sq25
0004	s	c	c	ih39	0025	s	b	a	sc13
0005	t	c	c	ih24	0026	s	d	c	sc11
0006	s	b	c	al26	0027	t	d	a	sc20
0007	t	c	a	ih55	0028	t	d	b	sc02
0008	s	b	b	ih70	0029	t	c	a	sc19
0009	t	d	d	al01	0030	s	a	a	sc12
0010	s	d	c	al09	0031	s	d	c	sf04
0011	s	b	b	ih47	0032	t	a	a	sf19
0012	t	a	a	al55	0033	t	c	b	sf14
0013	s	d	d	al22	0034	t	b	b	sf18
0014	t	d	b	ah17	0035	s	a	a	sf01
0015	s	b	c	al38	0036	t	d	d	sf10
0016	t	d	d	al16	0037	s	b	b	sb13
0017	t	b	b	ah06	0038	t	c	a	sb08
0018	t	c	c	ah23	0039	t	b	b	sb23
0019	t	d	d	sq27	0040	t	a	a	sb07
0020	t	d	d	sq06	0041	t	b	a	sb10
0021	t	a	a	sq15	0042	s	c	c	sb18

Table 24. Participant #NS02 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	b	b	al47	0020	s	a	a	sf03
0002	t	c	a	ah55	0021	s	a	a	sf02
0003	t	c	d	al43	0022	t	d	b	sf13
0004	s	d	c	al27	0023	t	d	d	sf12
0005	t	a	a	ih68	0024	s	a	b	sf01
0006	s	b	b	al36	0025	t	b	b	sf18
0007	t	c	c	al62	0026	s	c	a	sa22
0008	t	d	d	ah45	0027	s	c	a	sa11
0009	s	c	b	ah38	0028	t	d	d	sa08
0010	s	c	c	al33	0029	s	d	c	sa18
0011	t	b	b	ah37	0030	t	d	d	sa14
0012	t	c	c	ah36	0031	s	d	a	sa07
0013	s	b	d	sq16	0032	t	c	b	sb08
0014	t	b	b	ah44	0033	t	d	d	sb06
0015	s	c	c	sq22	0034	s	b	a	sb02
0016	t	b	b	sq10	0035	t	a	a	sb07
0017	t	d	d	sq18	0036	s	d	d	sb17
0018	s	d	a	sq13	0037	s	d	b	sb21
0019	s	b	b	sq11	0038	s	c	b	sb16

Table 25. Participant #NS03 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	a	c	al57	0016	t	a	b	sf07
0002	t	c	c	ih77	0017	t	b	a	sf15
0003	t	b	b	al08	0018	s	c	b	sd13
0004	t	d	d	al01	0019	t	a	a	sd08
0005	s	a	a	ah35	0020	s	d	c	sd07
0006	s	a	a	sq02	0021	s	c	c	sd16
0007	s	c	d	sq03	0022	t	d	d	sd10
0008	t	a	a	sq14	0023	t	d	d	sd23
0009	t	d	d	sq29	0024	s	c	c	sb01
0010	t	a	b	sq07	0025	s	c	c	sb15
0011	s	b	b	sq21	0026	t	a	a	sb20
0012	t	c	c	sf14	0027	t	b	b	sb10
0013	t	a	a	sf11	0028	s	a	a	sb03
0014	t	a	a	sf19	0029	t	b	c	sb19
0015	t	c	c	sf16					

Table 26. Participant #NS04 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	d	d	al06	0014	t	d	d	se03
0002	t	c	c	ah04	0015	s	a	b	sa19
0003	t	a	a	sq08	0016	t	a	d	sa10
0004	s	b	b	sq01	0017	t	d	c	sa04
0005	t	d	d	sq04	0018	t	c	b	sa09
0006	s	d	d	sq24	0019	t	a	a	sa21
0007	t	d	c	sq06	0020	t	d	d	sa20
0008	t	c	a	sq17	0021	t	b	b	sb09
0009	t	b	b	se21	0022	s	b	b	sb13
0010	s	d	d	se11	0023	s	c	c	sb18
0011	t	c	c	se02	0024	t	b	b	sb22
0012	s	a	a	se17	0025	s	c	c	sb01
0013	s	d	d	se20	0026	s	c	c	sb01

Table 27. Participant #NS05 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	d	d	al40	0019	t	c	c	sf09
0002	s	a	c	ah54	0020	t	b	b	sf20
0003	s	c	b	al07	0021	s	d	d	sf17
0004	s	a	a	al25	0022	t	a	a	sc08
0005	t	a	a	ah49	0023	t	c	b	sc17
0006	t	d	b	ah57	0024	s	b	b	sc13
0007	s	d	d	al11	0025	s	a	d	sc12
0008	s	a	a	ah21	0026	t	c	a	sc15
0009	s	a	a	ah41	0027	s	d	d	sc11
0010	t	a	a	sq15	0028	t	b	b	sc18
0011	s	b	b	sq25	0029	s	b	a	sa02
0012	s	b	b	sq01	0030	s	d	b	sa23
0013	s	b	b	sq01	0031	s	b	b	sa06
0014	s	a	c	sq02	0032	s	a	b	sa17
0015	s	b	b	sq01	0033	s	b	b	sa16
0016	s	a	a	sf06	0034	s	a	a	sa12
0017	s	d	b	sf04	0035	s	a	a	sa05
0018	t	d	d	sf10					

Table 28. Participant #NS06 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	d	b	al58	0017	t	b	a	se07
0002	s	b	b	ih29	0018	s	b	b	se16
0003	t	a	a	al20	0019	s	a	a	se18
0004	t	a	a	al14	0020	s	d	d	sd17
0005	s	b	b	ah18	0021	t	d	d	sd14
0006	s	b	b	sq01	0022	t	c	a	sd20
0007	s	b	b	sq01	0023	s	d	d	sd12
0008	s	a	d	sq02	0024	s	c	c	sd22
0009	s	b	a	sq01	0025	t	c	a	sd19
0010	s	b	a	sq01	0026	s	c	c	sb01
0011	t	c	c	ah11	0027	s	c	c	sb01
0012	s	b	b	sq01	0028	s	c	c	sb01
0013	s	b	b	sq01	0029	s	c	c	sb01
0014	s	b	b	se14	0030	s	c	c	sb01
0015	s	a	c	se15	0031	s	b	c	sb02
0016	s	c	c	se13					

Table 29. Participant #NS07 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	c	c	al13	0014	t	a	a	sd08
0002	t	a	a	ah51	0015	t	c	b	sc17
0003	t	a	a	sq08	0016	s	b	b	sc13
0004	t	d	d	sq18	0017	s	a	a	sc12
0005	t	d	d	sq29	0018	s	a	a	sc04
0006	s	b	b	sq16	0019	s	a	a	sc05
0007	s	c	c	sq22	0020	s	c	c	sc06
0008	t	c	c	sq17	0021	t	d	d	sa14
0009	t	d	d	sd23	0022	s	b	b	sa06
0010	s	d	d	sd07	0023	s	a	a	sa05
0011	t	b	b	sd15	0024	t	c	c	sa09
0012	t	b	c	sd04	0025	t	d	d	sa04
0013	t	b	b	sd01	0026	s	c	c	sa13

Table 30. Participant #NS08 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	s	a	b	al28	0031	t	a	c	sq07
0002	s	d	d	ih50	0032	t	d	a	ah45
0003	t	d	b	al12	0033	s	b	b	ah22
0004	s	d	c	ih28	0034	s	a	c	sq02
0005			b	ih63	0035	t	b	b	ah06
0006	t	d	d	im04	0036	t	b	b	sq10
0007	t	a	c	ih49	0037	s	b	b	sq01
0008	s	d	d	im37	0038	t	d	d	sq06
0009	t	b	d	ih52	0039	s	c	c	sq03
0010	z	b	b	im19	0040	t	b	b	sb10
0011	t	a	a	ih78	0041	t	b	a	sb09
0012	s	c	c	ih48	0042	s	b	b	sb02
0013	s	b	b	ih41	0043	s	c	d	sb15
0014	t	d	b	ih18	0044	t	a	a	sb07
0015	z	b	b	im11	0045	t	b	c	sb19
0016	s	c	c	ih80	0046	s	d	d	sb21
0017	s	c	c	ih23	0047	s	b	b	sa01
0018	t	c	c	al04	0048	t	d	d	sa20
0019	t	b	a	al37	0049	s	c	c	sa22
0020	s	d	d	ih40	0050	t	a	b	sa10
0021	s	b	b	al23	0051	s	b	b	sa16
0022	s	b	a	al26	0052	t	a	a	sa21
0023	s	d	d	ih13	0053	s	d	a	sf17
0024	s	d	d	al34	0054	t	b	a	sf15
0025	t	c	c	al62	0055	t	a	d	sf07
0026	t	d	b	al01	0056	t	b	b	sf18
0027	s	a	a	ih12	0057	s	c	c	sf05
0028	t	d	d	al58	0058	t	a	a	sf19
0029	t	a	a	al57	0059	t	d	d	sf12
0030	t	b	b	ah40					

Table 31. Participant #NS09 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	a	a	al55	0015	t	d	d	sa08
0002	t	d	d	ah27	0016	s	b	b	se14
0003	t	d	d	sq04	0017	s	a	a	se17
0004	s	b	b	sq25	0018	s	d	d	se20
0005	t	a	a	sq14	0019	s	b	b	se16
0006	s	b	a	sq11	0020	s	a	a	se18
0007	t	a	a	sq15	0021	t	b	b	se21
0008	s	d	b	sq24	0022	s	a	a	sf06
0009	s	c	c	sa03	0023	t	b	b	sf20
0010	s	d	b	sa07	0024	t	d	d	sf13
0011	s	d	b	sa18	0025	s	d	c	sf08
0012	s	b	c	sa02	0026	s	a	a	sf02
0013	s	a	a	sa19	0027	s	a	a	sf03
0014	s	a	a	sa17					

Table 32. Participant #NS10 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	b	b	al03	0014	t	c	c	sc16
0002	t	d	d	ah17	0015	s	b	b	se14
0003	t	a	a	sq15	0016	t	d	d	se03
0004	s	b	b	sq21	0017	s	d	d	se11
0005	t	a	a	sq08	0018	s	b	b	se19
0006	t	d	b	sq04	0019	s	a	a	se15
0007	s	b	b	sq11	0020	t	c	c	se02
0008	t	d	d	sq29	0021	s	d	d	sd11
0009	t	c	b	sc17	0022	s	d	b	sd07
0010	s	b	b	sc13	0023	t	b	c	sd05
0011	t	d	a	sc20	0024	t	d	d	sd10
0012	s	d	d	sc11	0025	s	c	d	sd13
0013	s	a	a	sc04	0026	t	b	c	sd18

Table 33. Participant #NS11 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	s	d	d	al30	0015	t	b	a	se21
0002	s	a	a	ah41	0016	t	d	b	sd23
0003	s	b	b	sq01	0017	s	d	d	sd17
0004	s	b	b	sq01	0018	s	c	c	sd16
0005	s	a	a	sq02	0019	t	a	c	sd09
0006	t	d	d	sq29	0020	t	d	d	sd21
0007	s	c	c	sq03	0021	t	d	d	sd06
0008	t	d	d	sq27	0022	s	d	d	sa23
0009	s	d	d	se20	0023	t	c	c	sa09
0010	s	a	a	se17	0024	s	b	b	sa02
0011	s	b	a	se16	0025	t	a	a	sa10
0012	s	c	b	se01	0026	s	c	c	sa11
0013	s	c	b	se01	0027	s	d	d	sa07
0014	t	c	c	se02					

Table 34. Participant #NS12 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	a	d	al57	0016	s	a	a	se18
0002	s	d	d	ih13	0017	s	c	c	se01
0003	t	c	c	al43	0018	s	b	b	se19
0004	s	d	d	al22	0019	s	c	b	sb15
0005	t	b	b	ah5	0020	s	c	c	sb01
0006	s	b	b	sq21	0021	t	b	b	sb19
0007	s	b	c	sq11	0022	s	b	c	sb13
0008	t	d	c	sq06	0023	s	d	d	sb17
0009	t	d	d	ah1	0024	t	a	a	sb20
0010	t	a	a	sq14	0025	s	a	a	sf03
0011	s	b	b	sq25	0026	t	c	c	sf16
0012	s	d	d	sq13	0027,	t	c	c	sf14
0013	s	a	a	se15	0028	t	b	b	sf20
0014	s	d	d	se20	0029	s	c	c	sf05
0015	s	b	d	se14	0030	t	d	b	sf13

Table 35. Participant #NS13 Item Set

QA	ST	CA	SA	CN	QA	ST	CA	SA	CN
0001	t	a	d	al57	0019	t	a	a	sd02
0002	s	c	c	ih39	0020	t	b	b	sd03
0003	t	c	d	al05	0021	t	d	c	sd23
0004	t	a	a	ih49	0022	t	d	b	sd14
0005	s	c	c	al33	0023	t	d	d	sd21
0006	s	b	c	al38	0024	s	c	c	sb15
0007	s	d	d	ih50	0025	s	b	c	sb13
0008	s	c	c	al49	0026	t	b	b	sb22
0009	s	c	c	al07	0027	t	a	a	sb20
0010	t	c	c	al45	0028	s	a	a	sb03
0011	t	d	d	ah12	0029	t	b	c	sb09
0012	s	c	c	sq03	0030	s	b	d	sc13
0013	t	b	b	sq10	0031	t	b	b	sc18
0014	s	d	d	sq13	0032	t	a	d	sc08
0015	t	d	d	sq06	0033	t	c	c	sc15
0016	t	d	b	sq18	0034	s	a	c	sc04
0017	t	a	a	sq15	0035	t	d	a	sc20
0018	t	b	b	sd01					

RECEIVED

AS

FOLLOWS

APPENDIX I

EYE-TRACKING SAMPLE SET: NON-NATIVE INTERMEDIATE (NNI03)

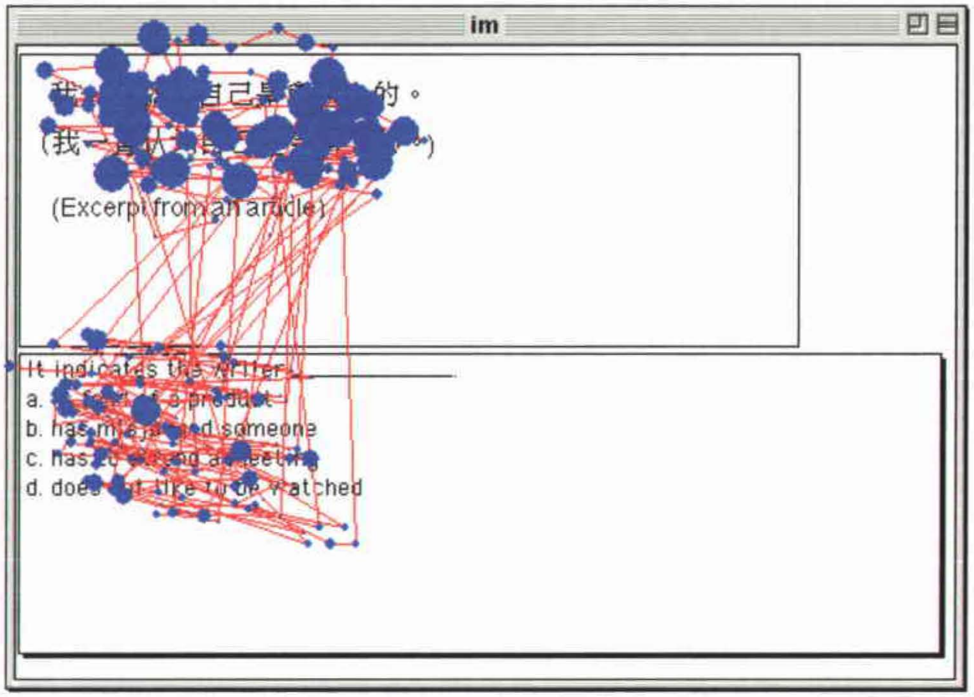


Figure 80. NNI03-Item 1 (im57).

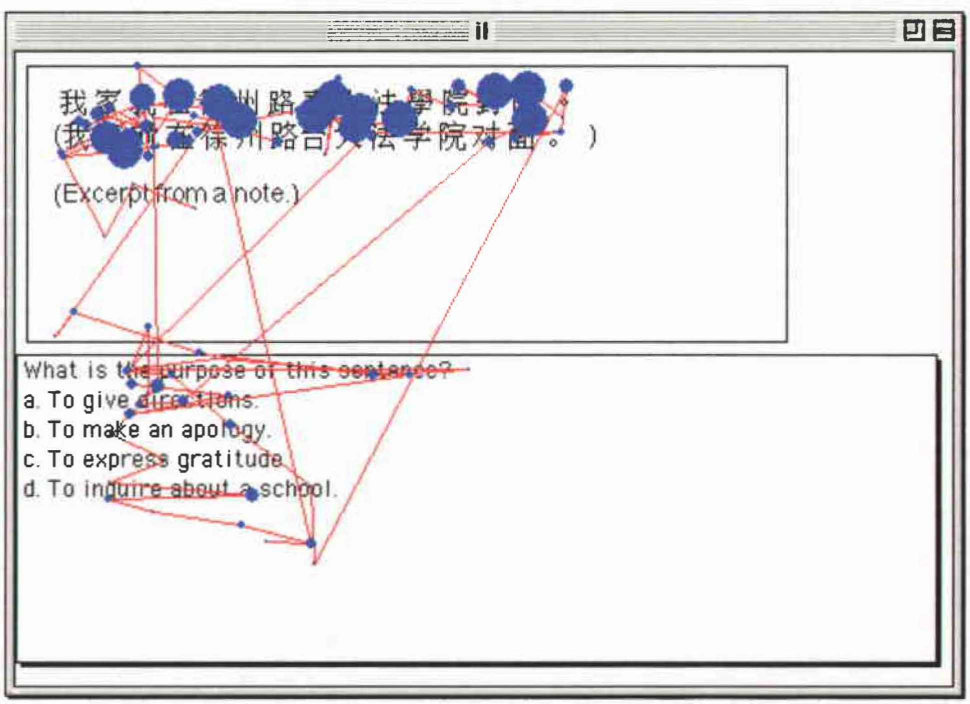


Figure 81. NNI03-Item 2 (il45).

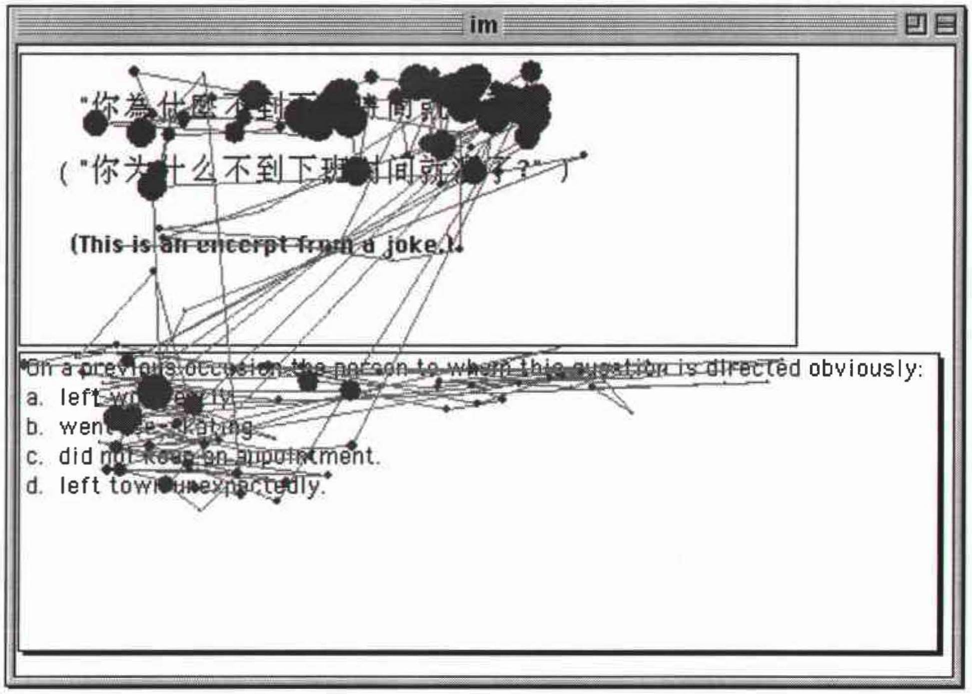


Figure 82. NNI03-Item 3 (im33).

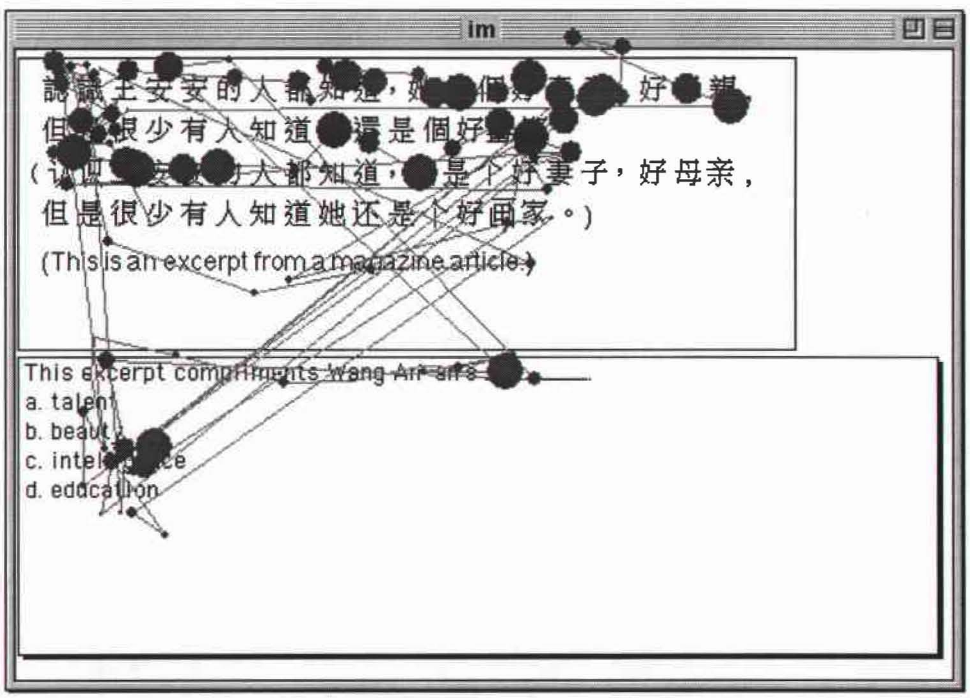


Figure 83. NNI03-Item 4 (im51).

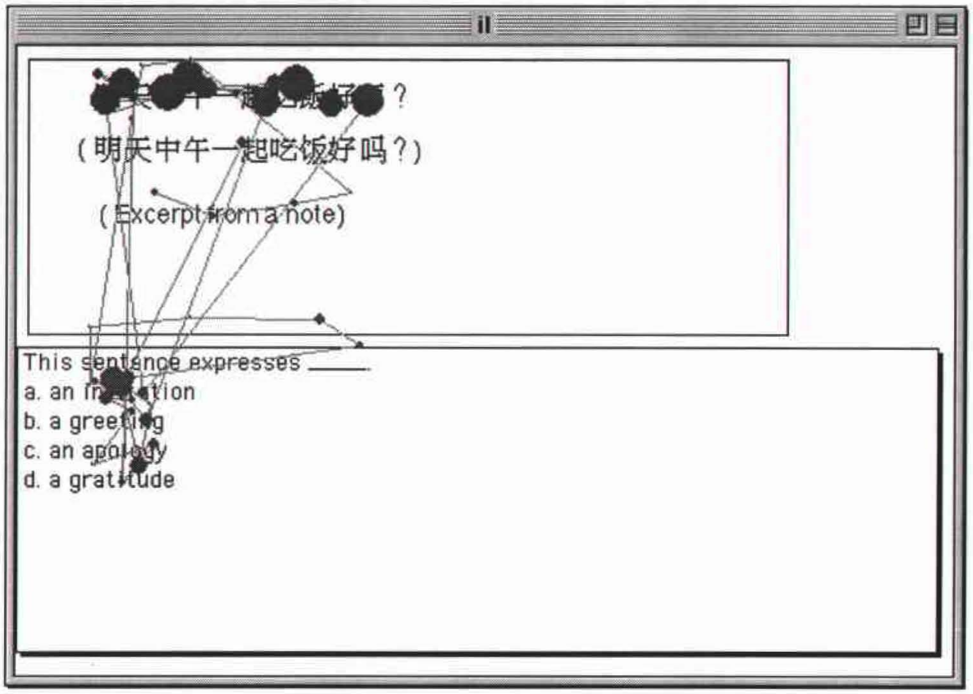


Figure 84. NNI03-Item 5 (il39).

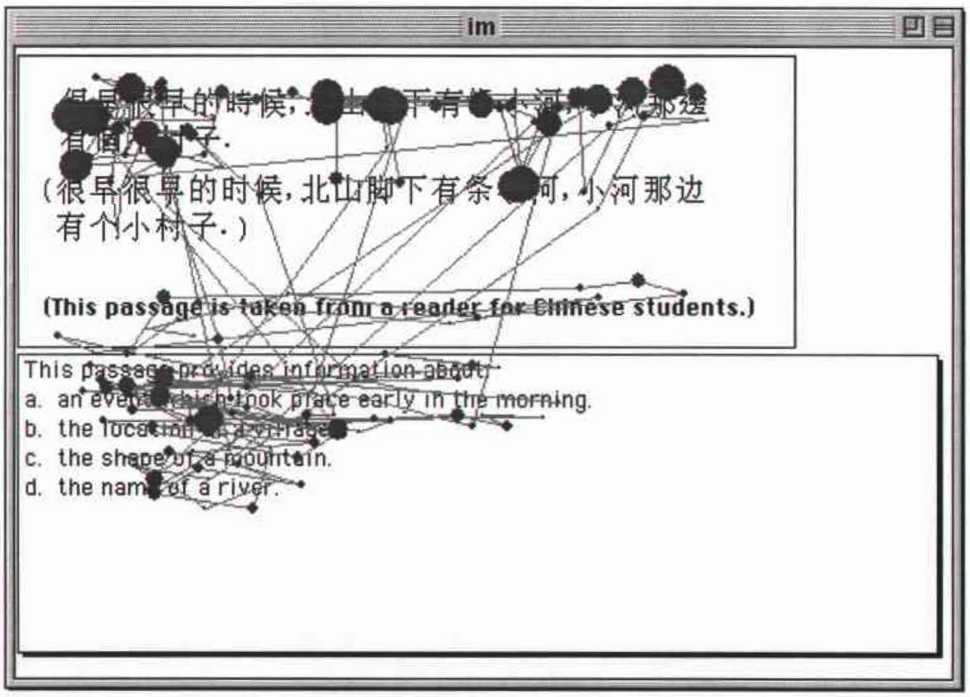


Figure 85. NNI03-Item 6 (im32).

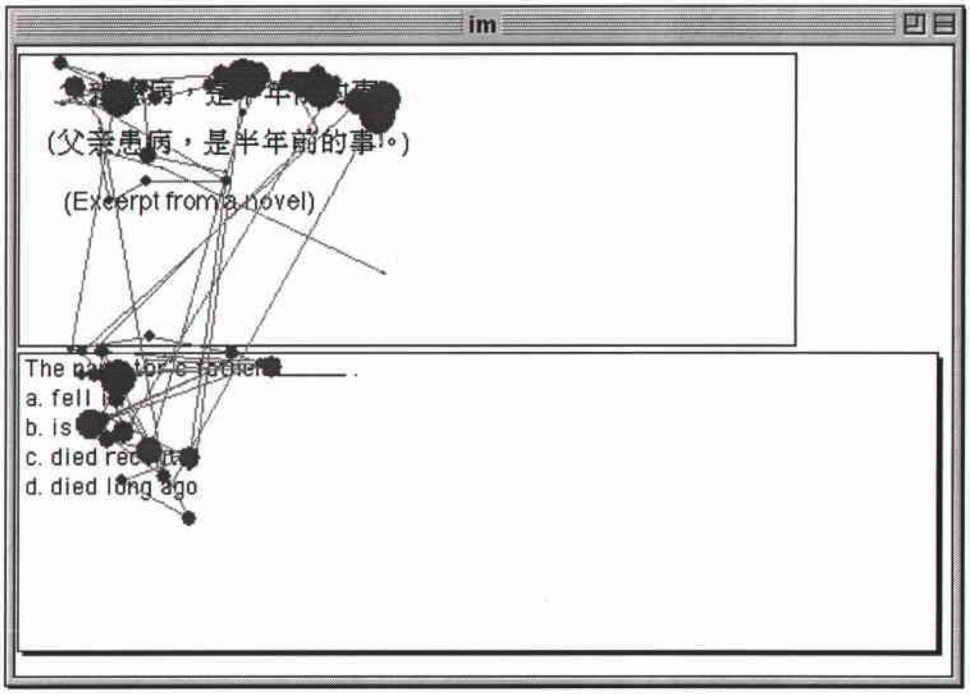


Figure 86. NNI03-Item 7 (im07).

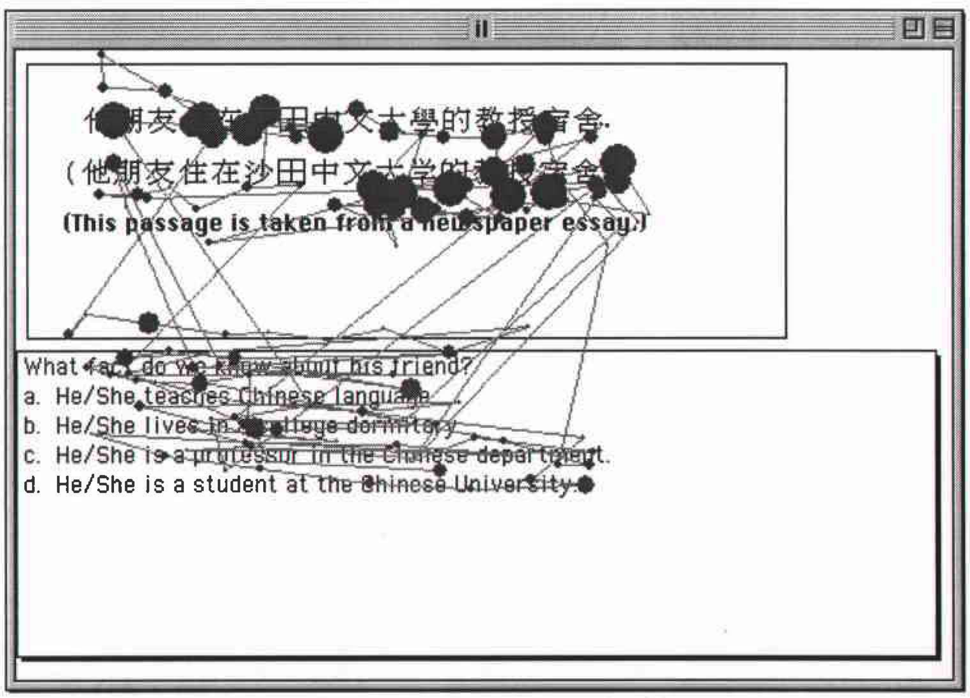


Figure 87. NNI03-Item 8 (il27).

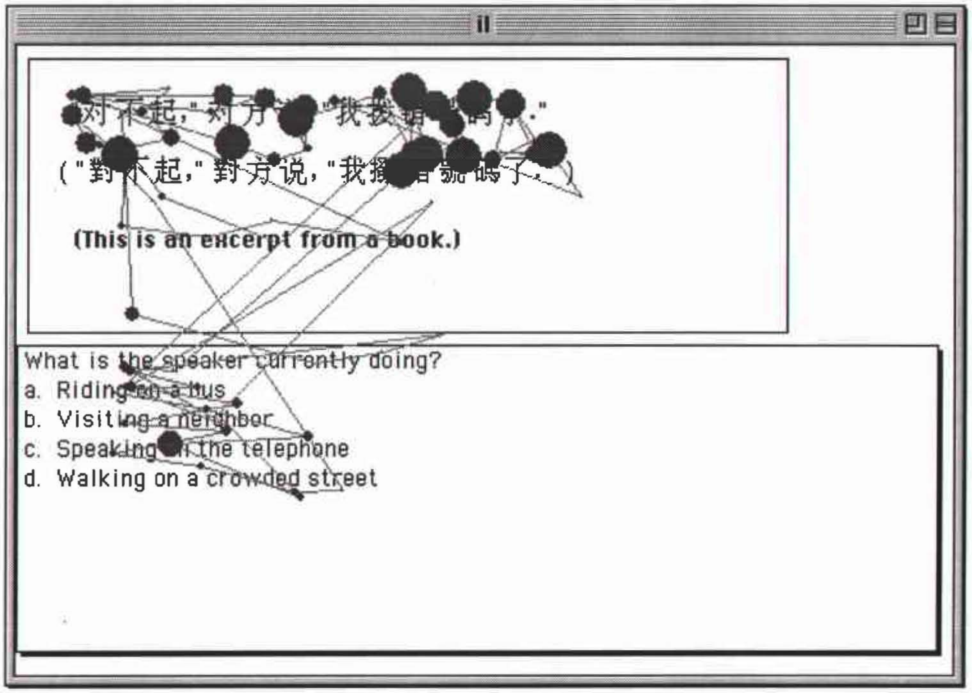


Figure 88. NNI03-Item 9 (il34).

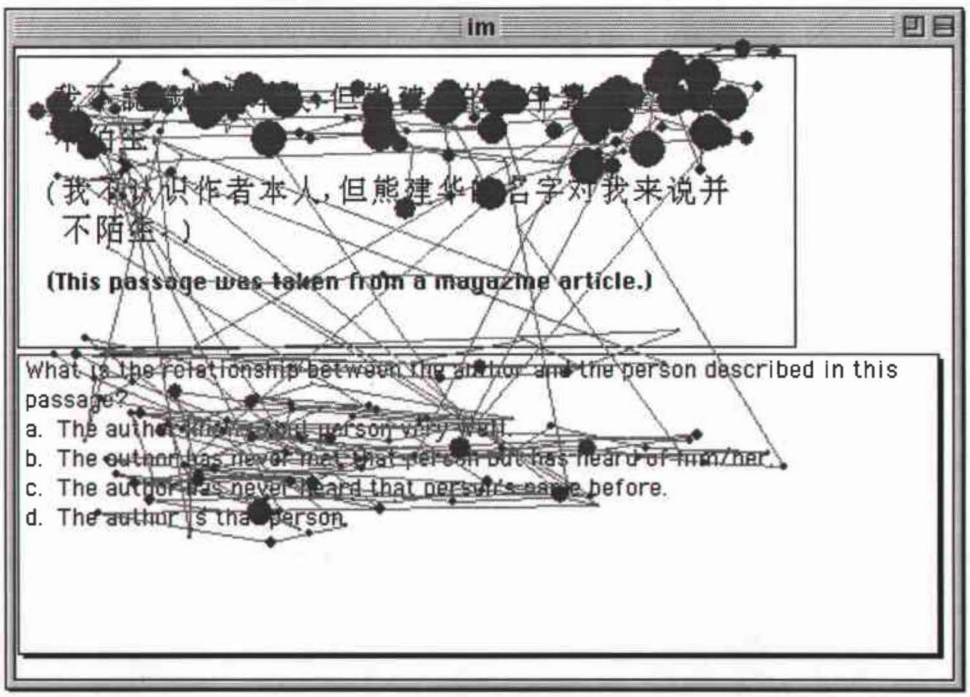


Figure 89. NNI03-Item 10 (im11).

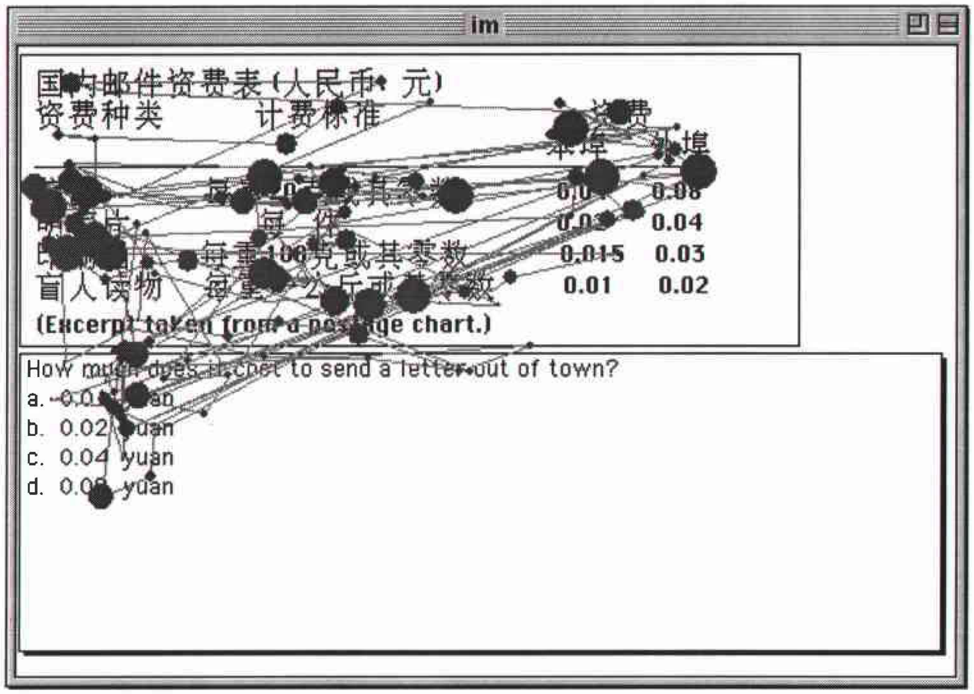


Figure 90. NNI03-Item 11 (im08).

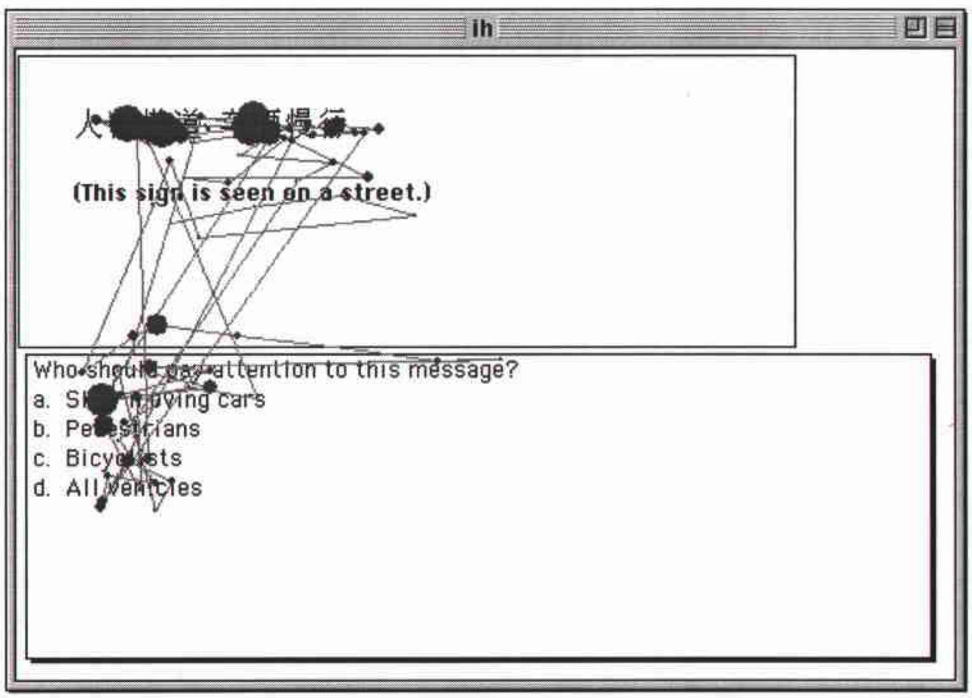


Figure 91. NNI03-Item 12 (ih26).

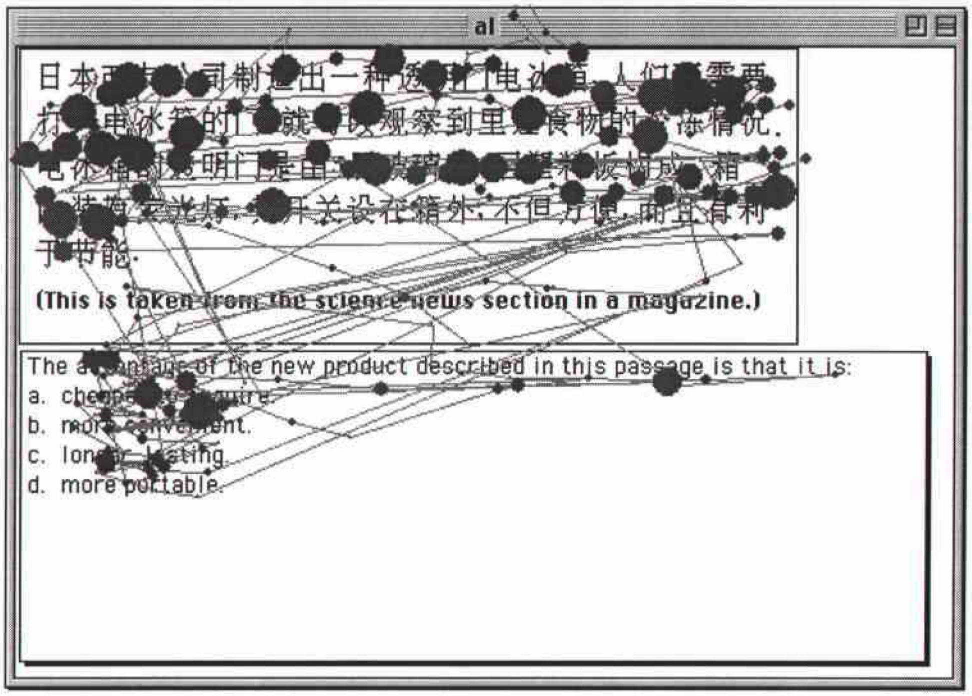


Figure 92. NNI03-Item 13 (al36).

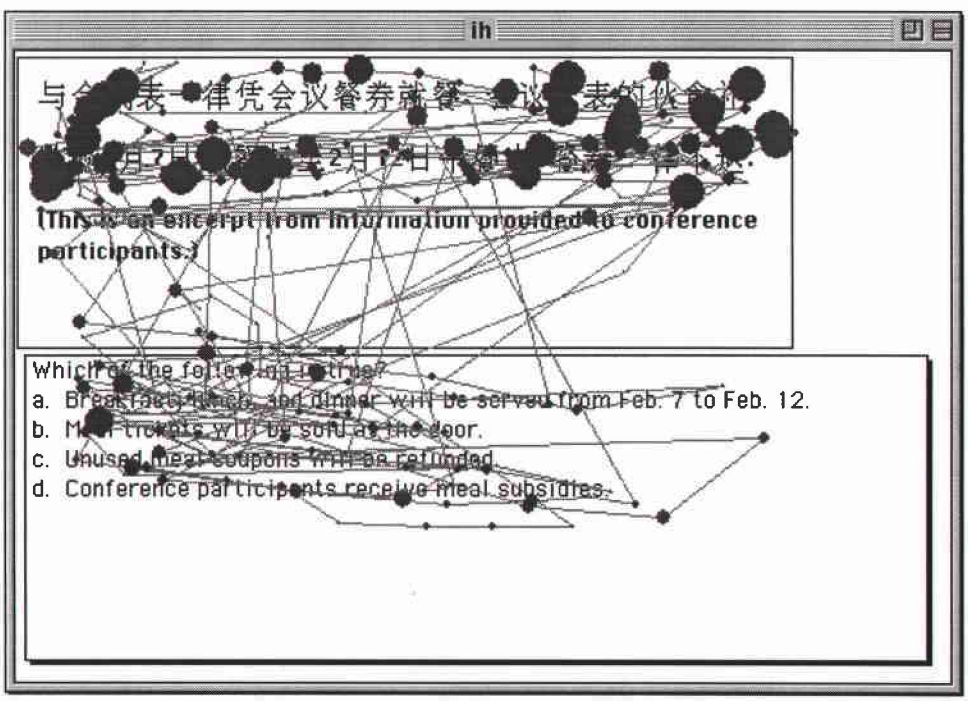


Figure 93. NNI03-Item 14 (ih50).

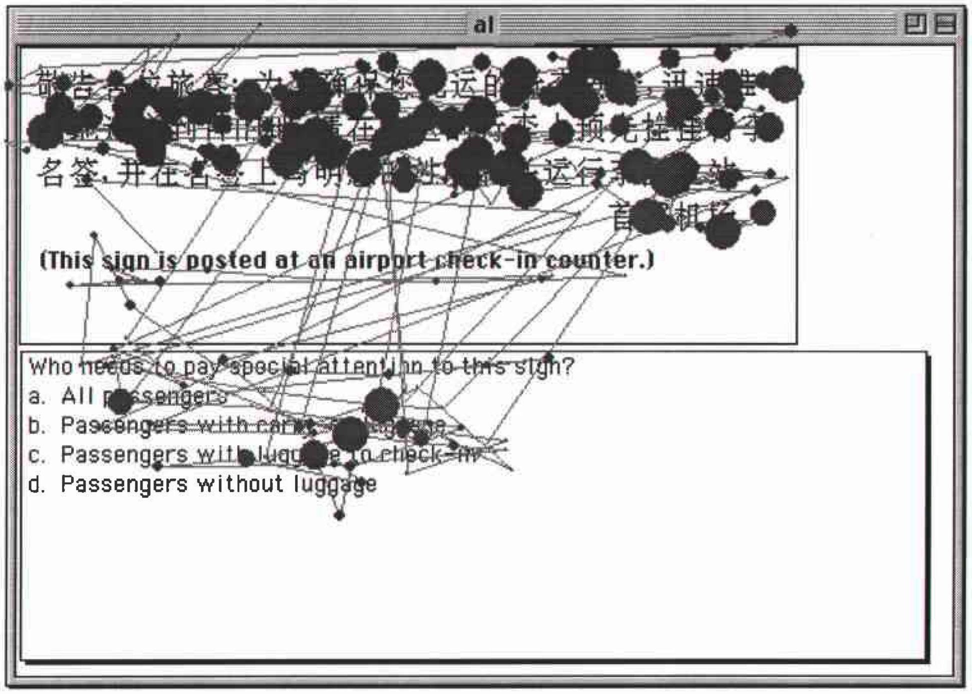


Figure 94. NNI03-Item 15 (al07).

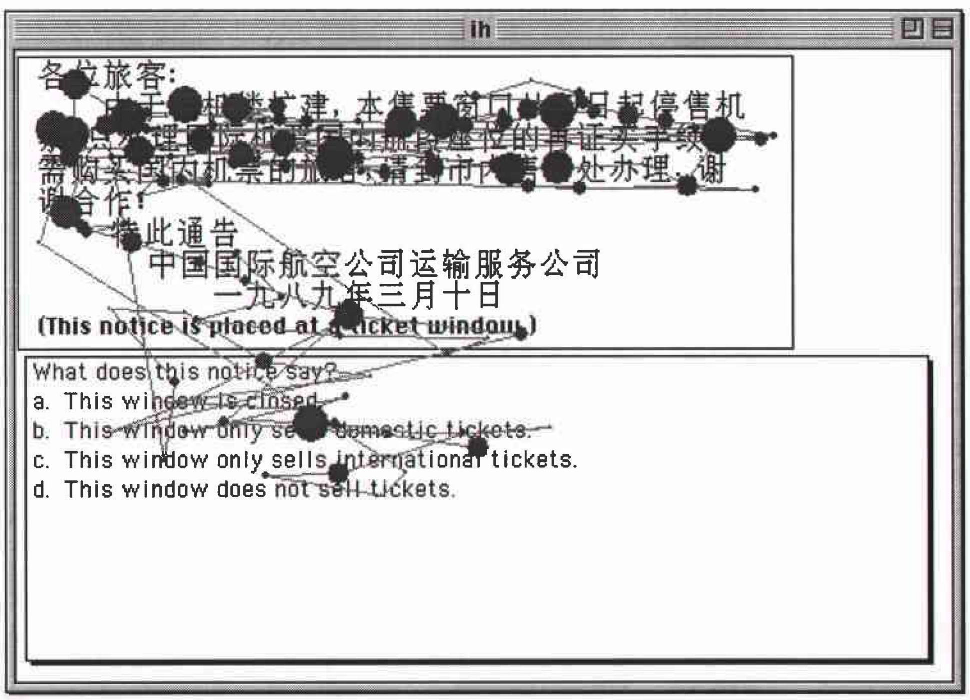


Figure 95. NNI03-Item 16 (ih08).

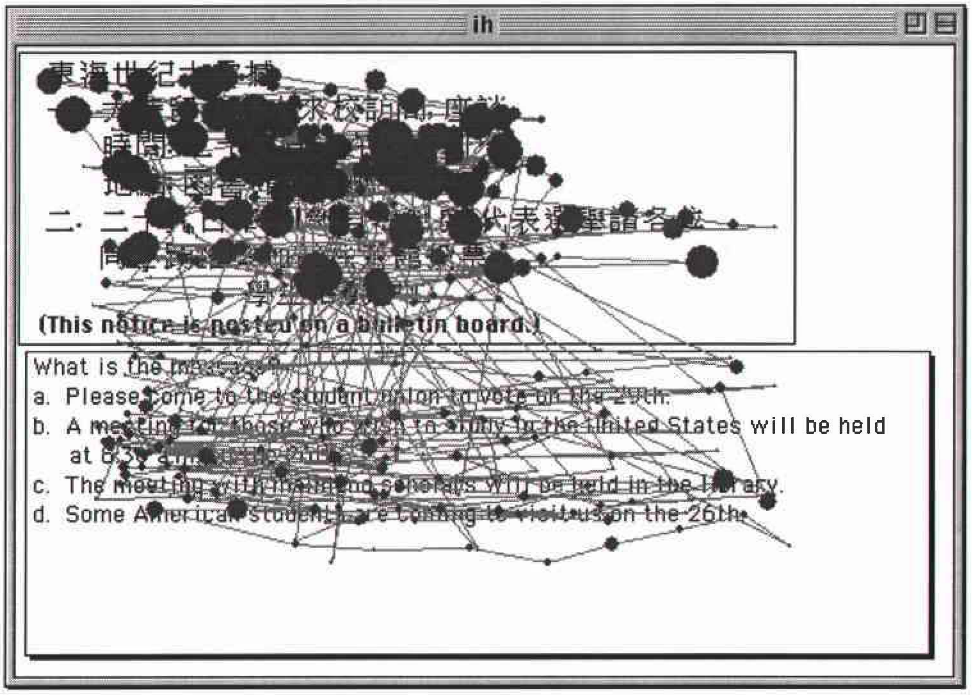


Figure 96. NNI03-Item 17 (ih05).

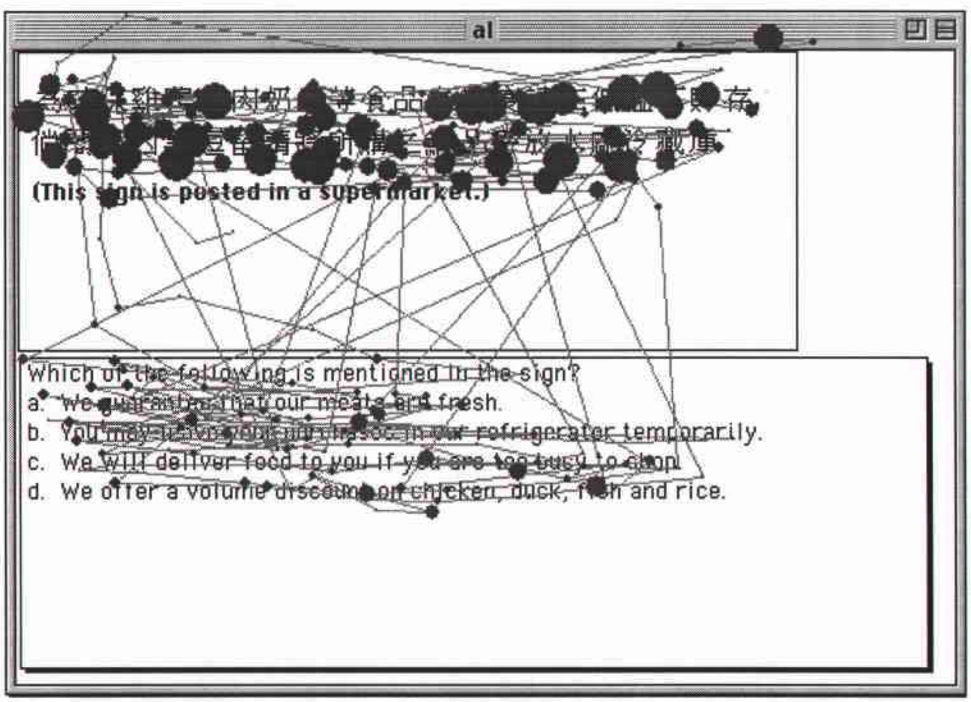


Figure 97. NNI03-Item 18 (al02).

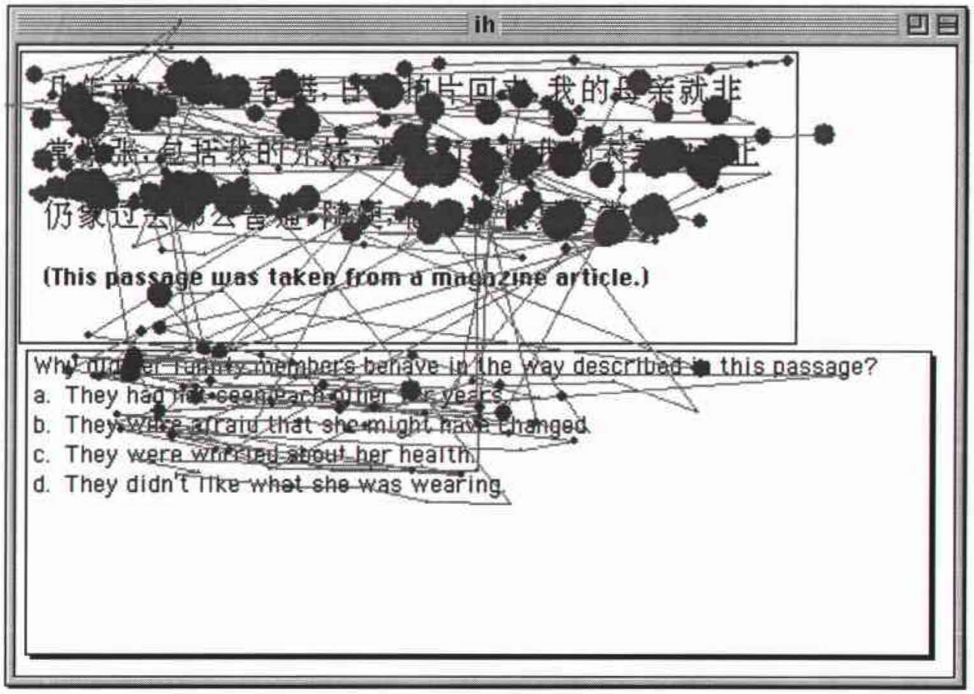


Figure 98. NNI03-Item 19 (ih29).

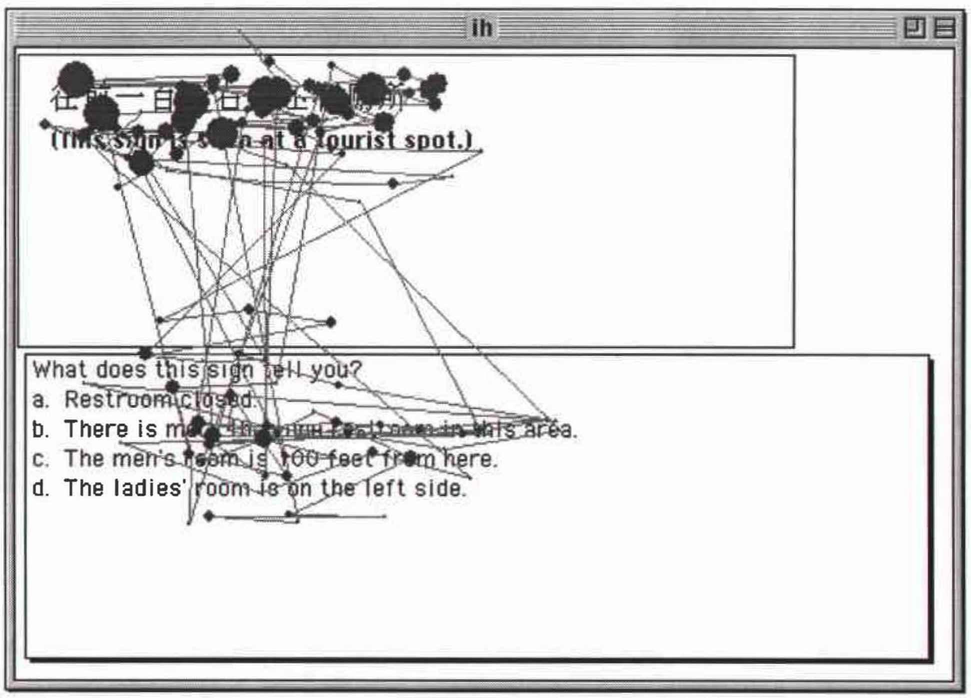


Figure 99. NNI03-Item 20 (ih25).

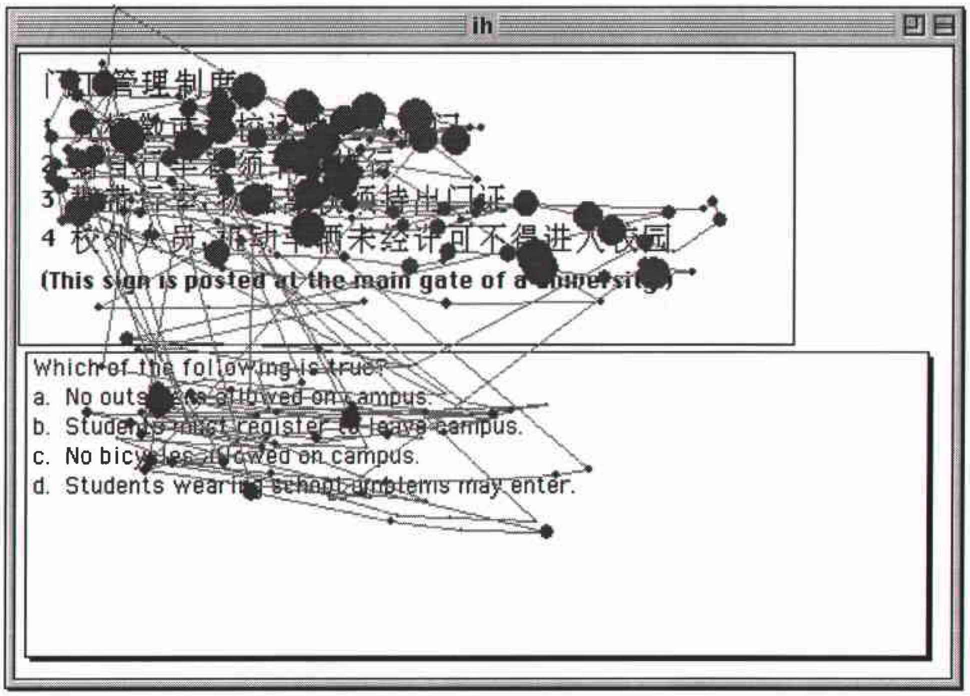


Figure 100. NNI03-Item 21 (ih17).

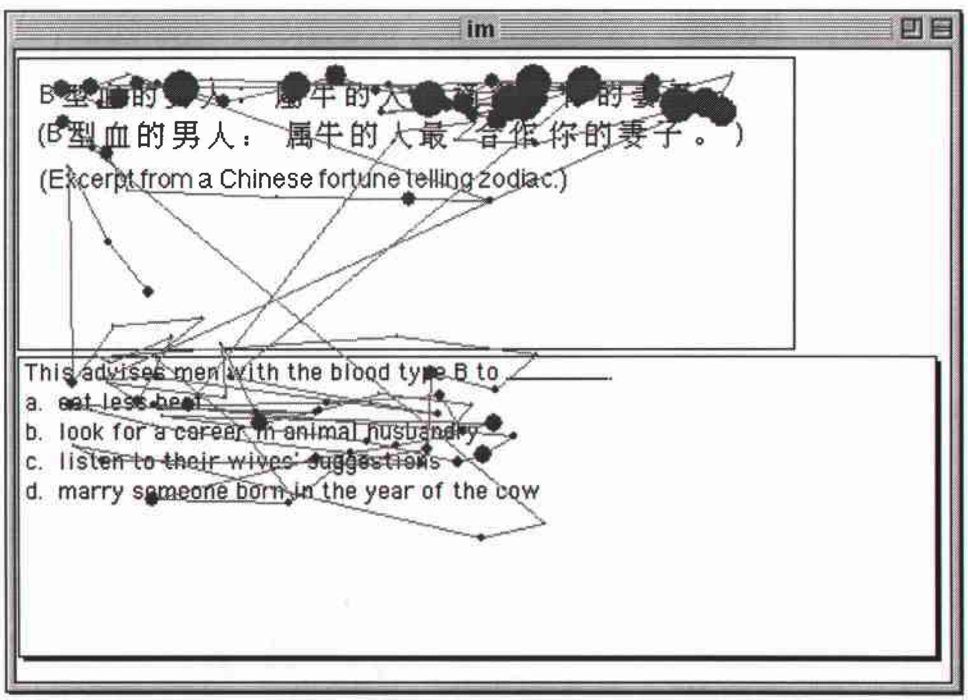


Figure 101. NNI03-Item 22 (im46).

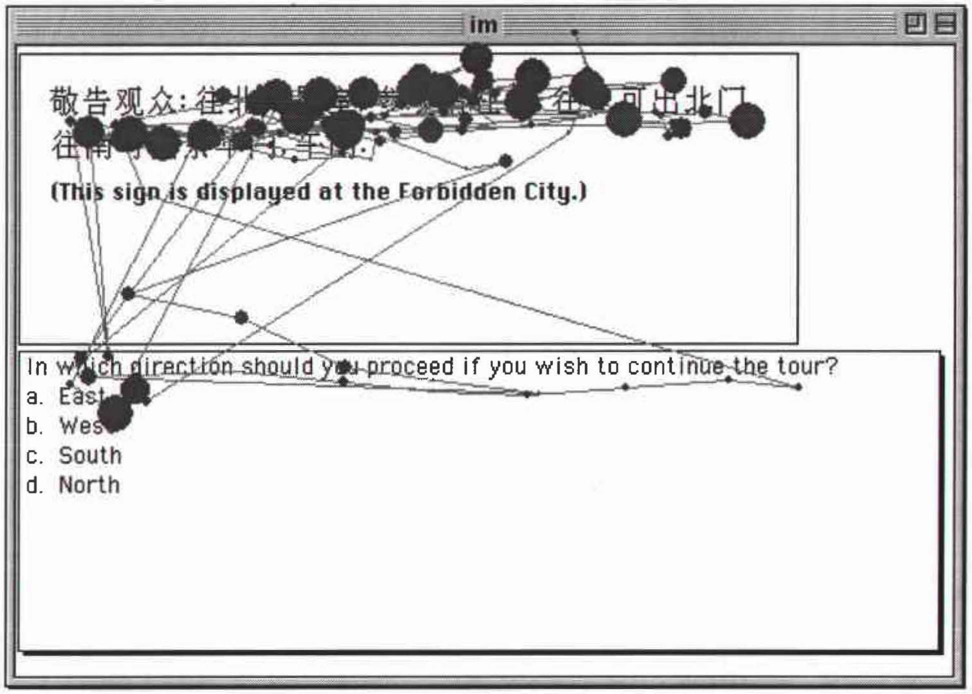


Figure 102. NNI03-Item 23 (im54).

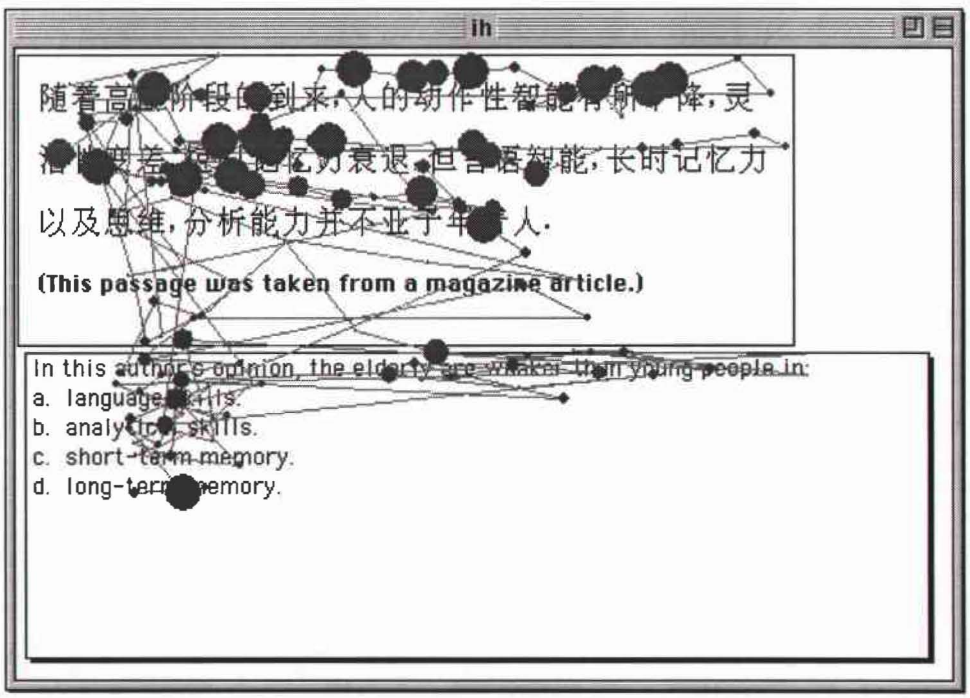


Figure 103. NNI03-Item 24 (ih32).

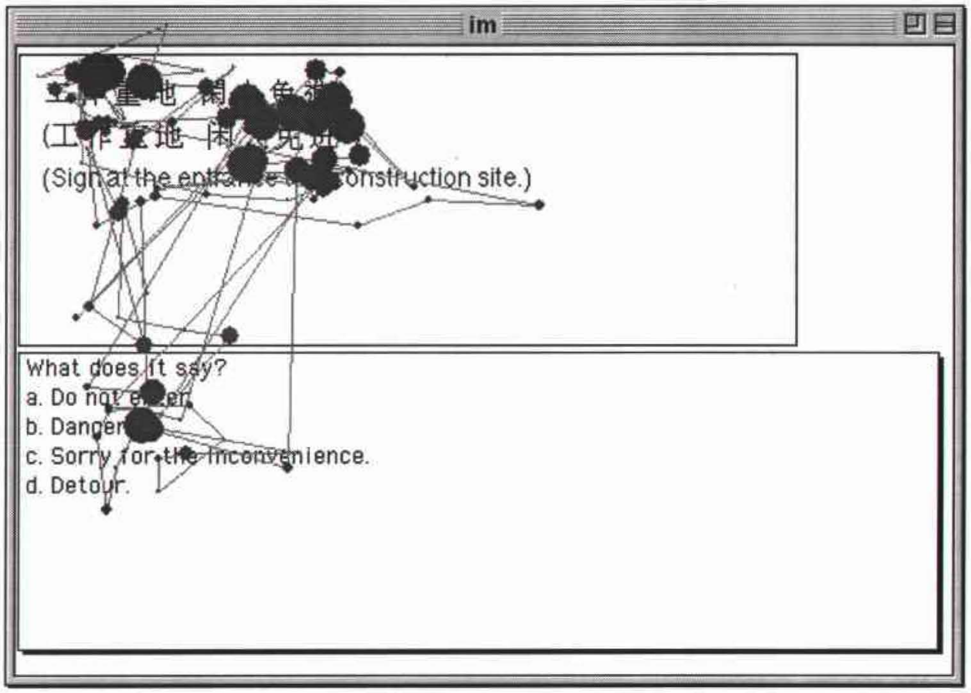


Figure 104. NNI03-Item 25 (im41).

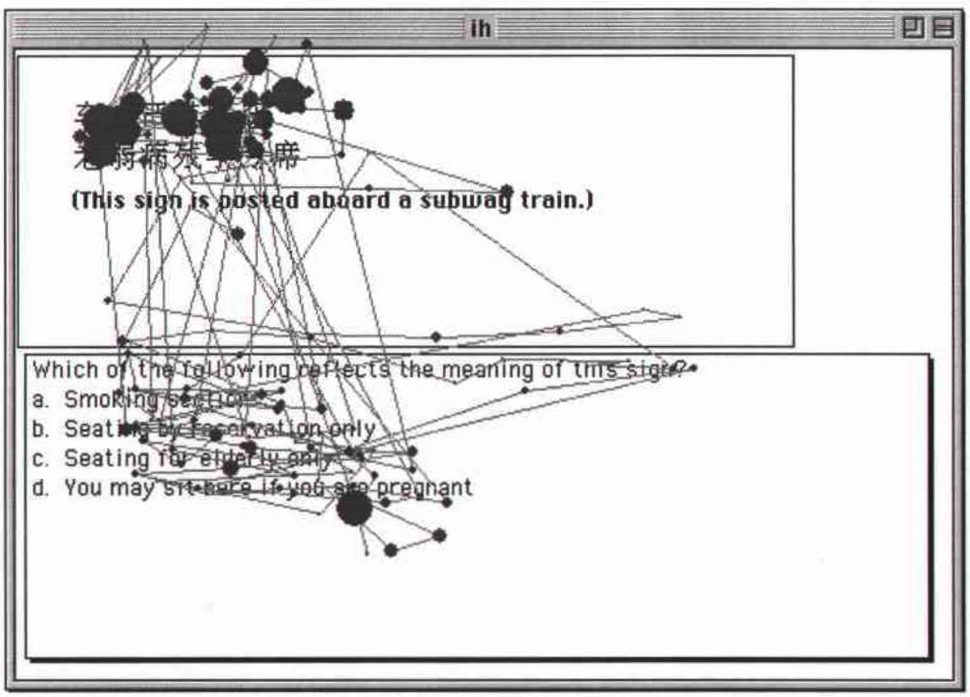


Figure 105. NNI03-Item 26 (ih10).

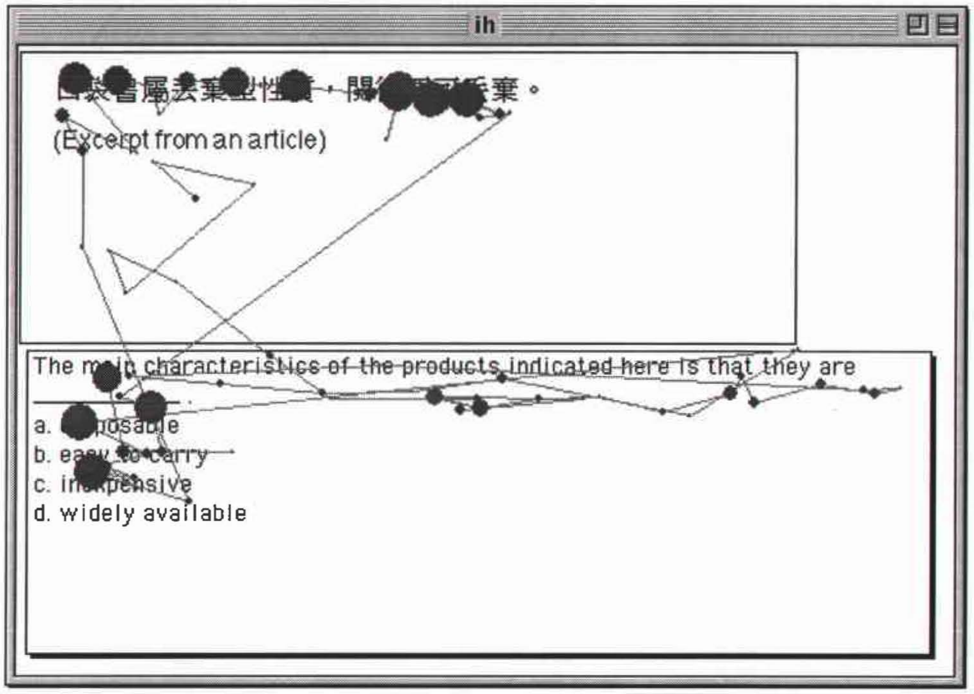


Figure 106. NNI03-Item 27 (ih78).

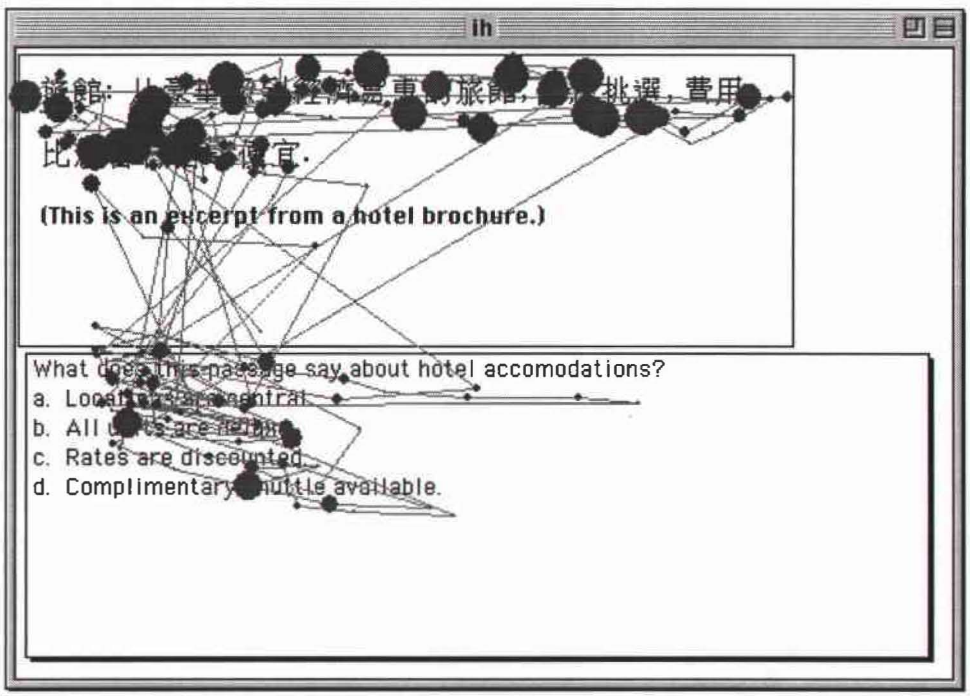


Figure 107. NNI03-Item 28 (ih54).

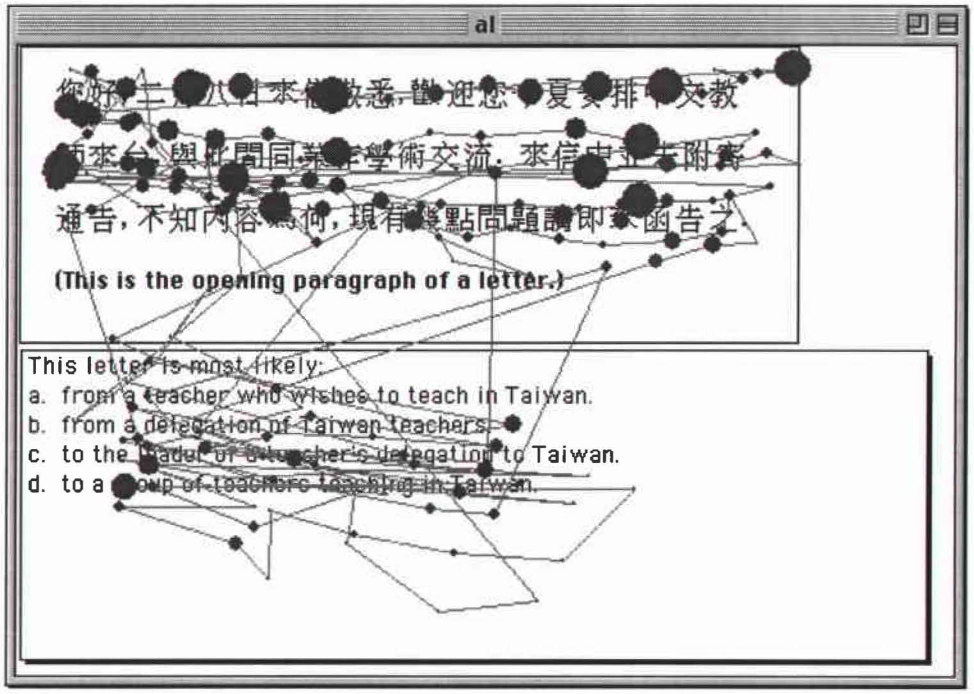


Figure 108. NNI03-Item 29 (al39).

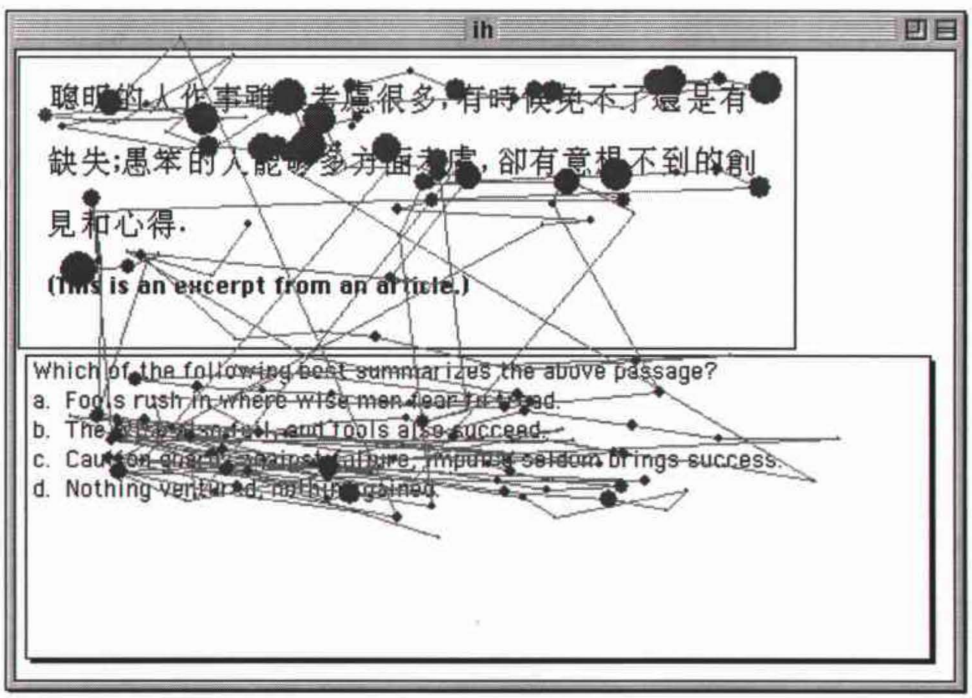


Figure 109. NNI03-Item 30 (ih52).

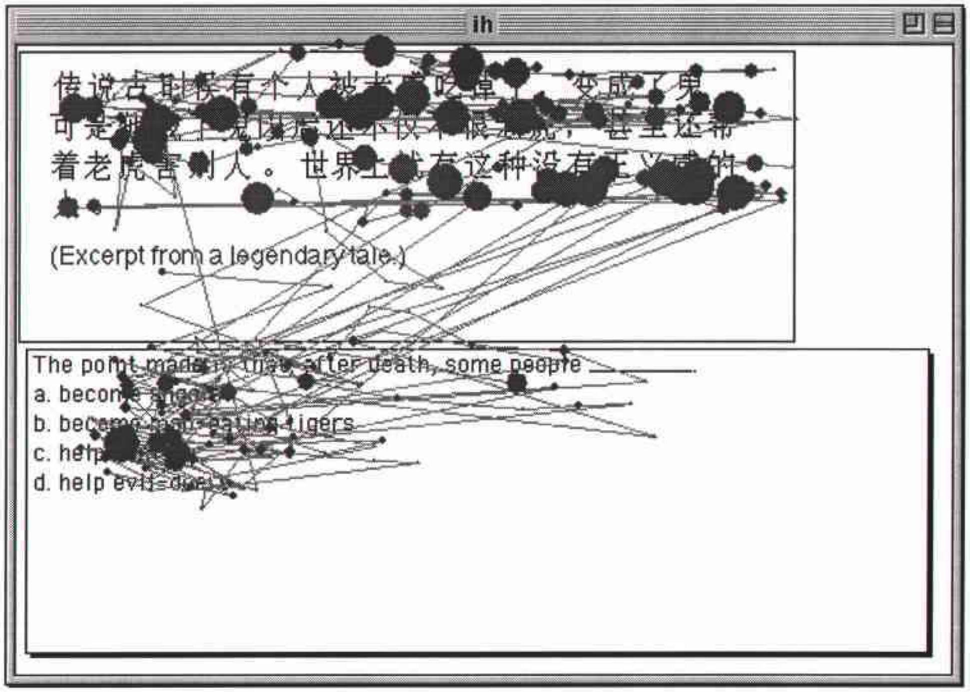


Figure 110. NNI03-Item 31 (ih67).

APPENDIX J

EYE-TRACKING SAMPLE SET: NON-NATIVE SUPERIOR (NNS03)

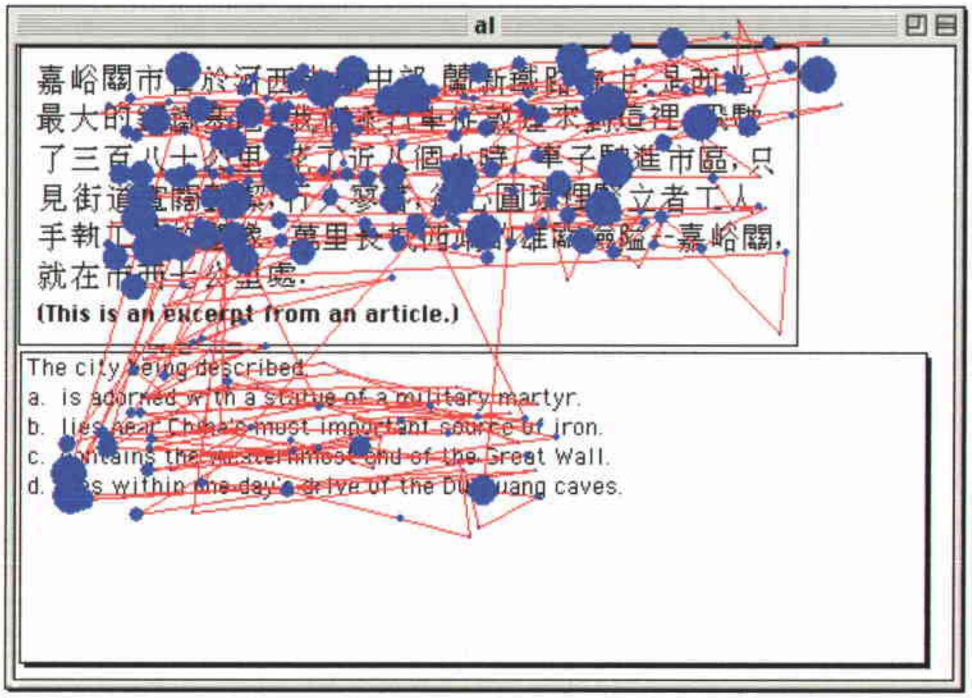


Figure 111. NNS03-Item 1 (al40).

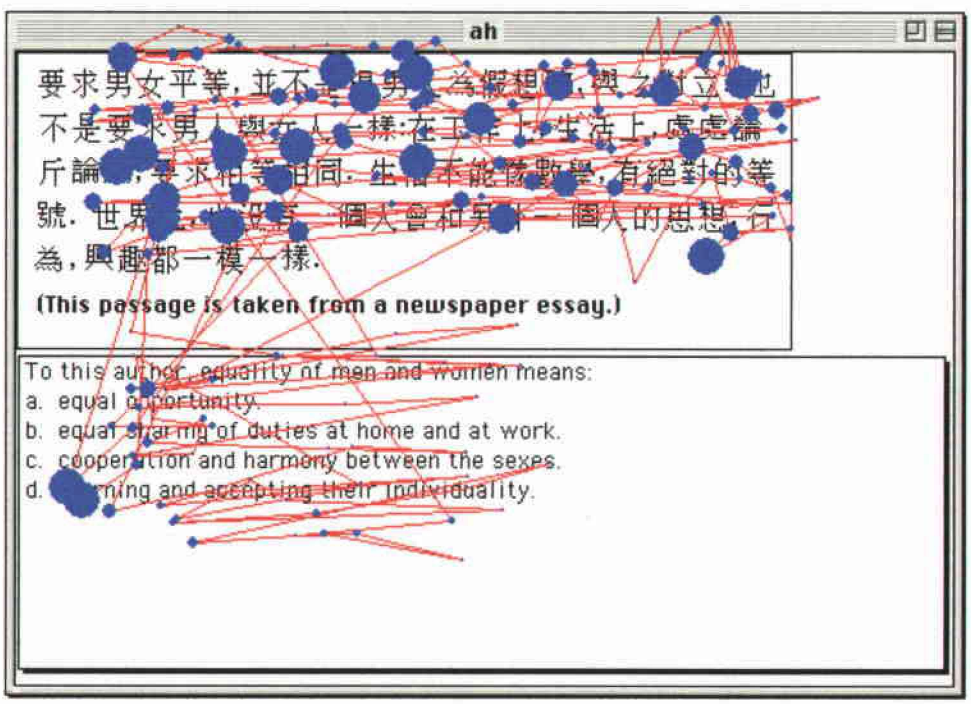


Figure 112. NNS03-Item 2 (ah33).

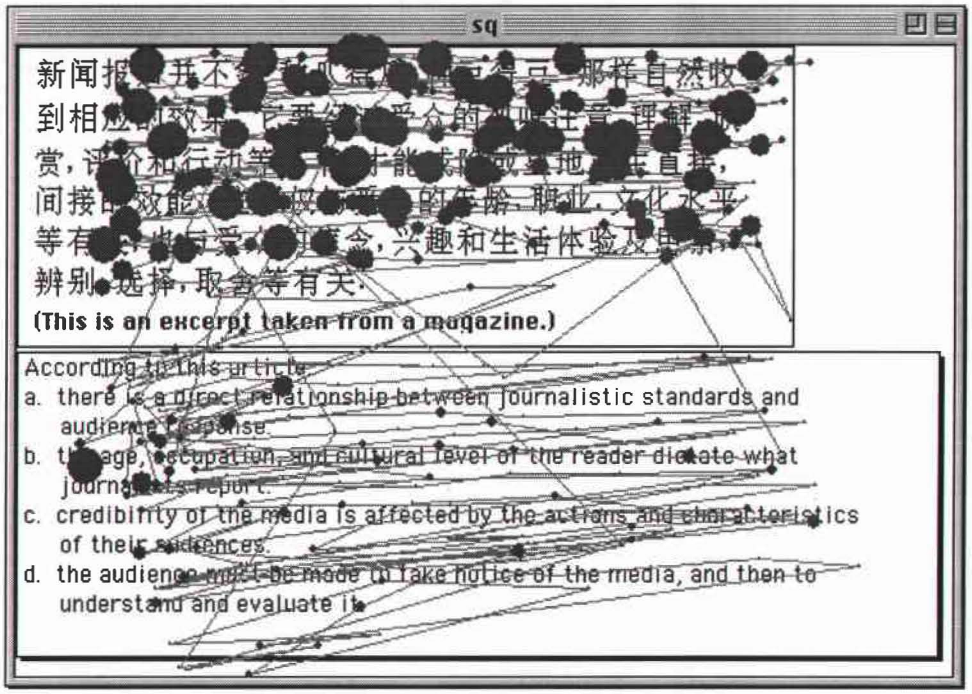


Figure 113. NNS03-Item 3 (sq03).

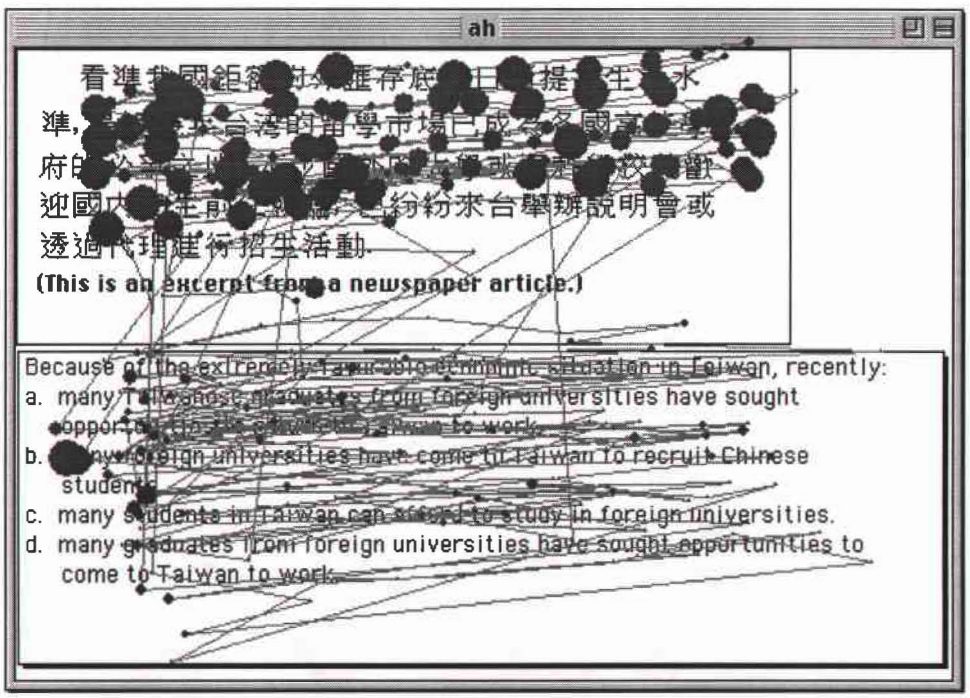


Figure 114. NNS03-Item 4 (ah08).

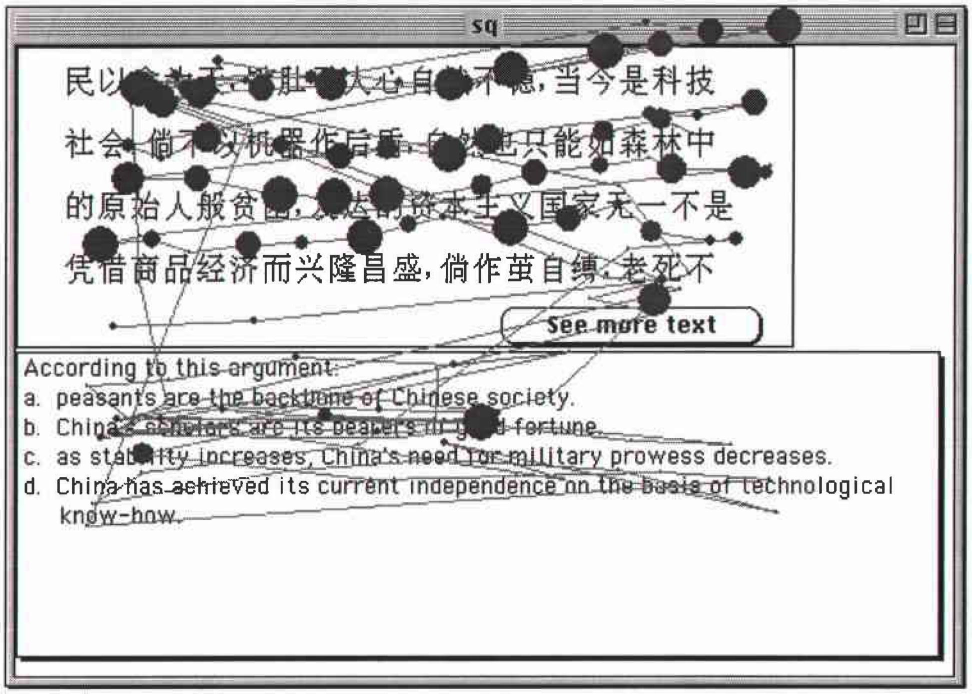


Figure 115a. NNS03-Item 5 (sq16, card 1).

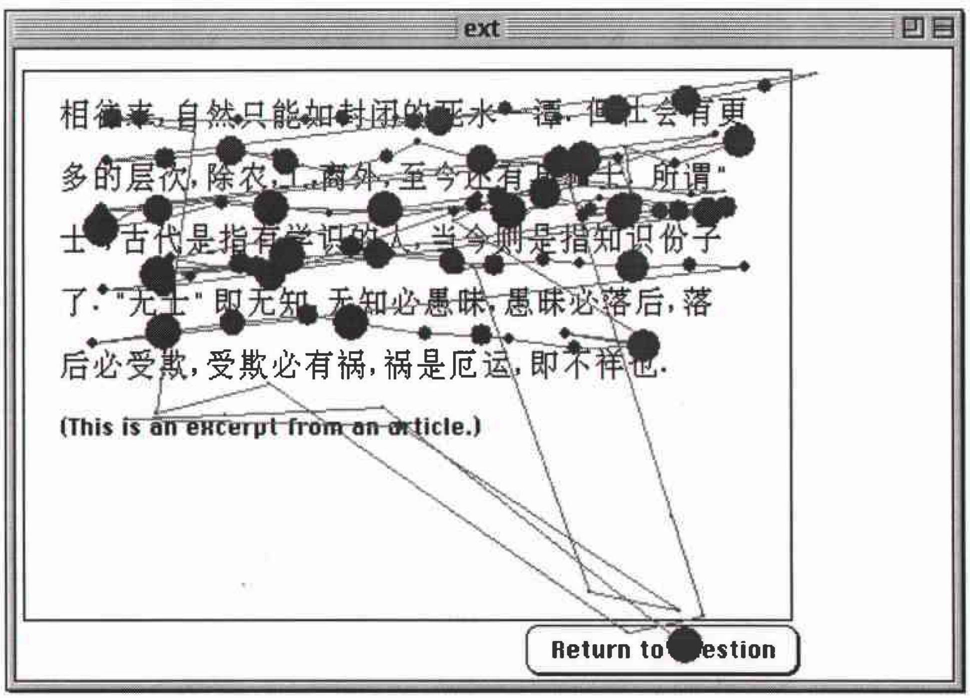


Figure 115b. NNS03-Item 5 (sq16, card 2).

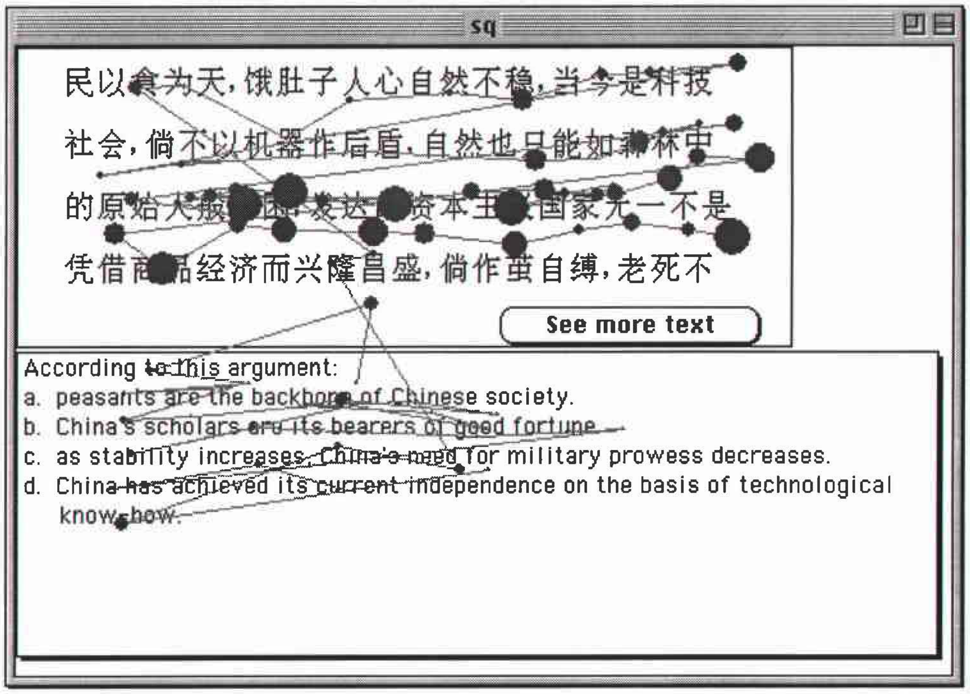


Figure 115c. NNS03-Item 5 (sq16, card 1).

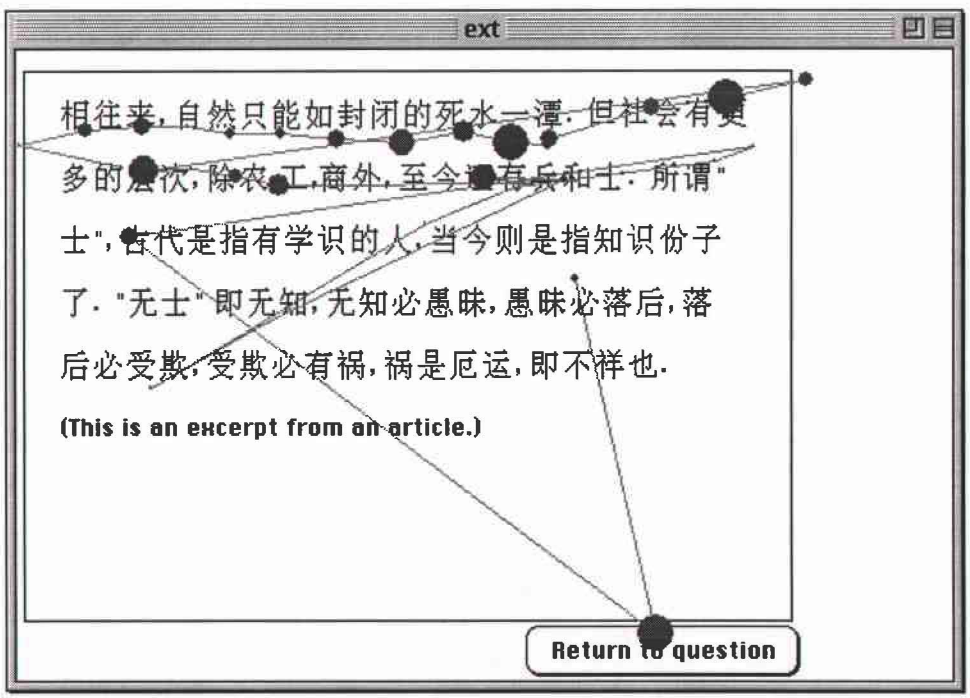


Figure 115d. NNS03-Item 5 (sq16, card 2).

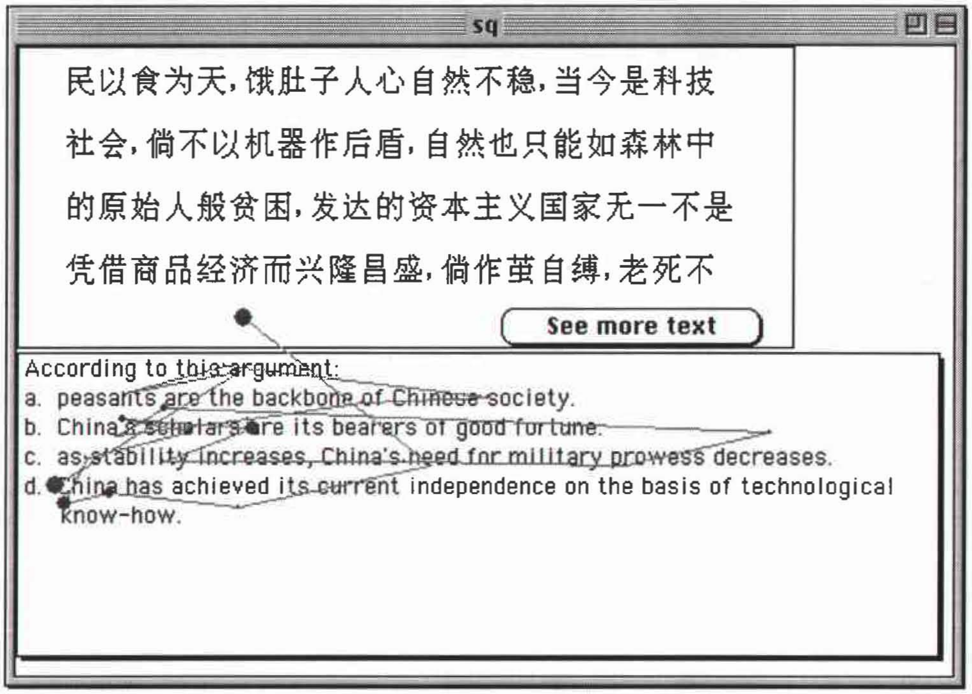


Figure 115e. NNS03-Item 5 (sq16, card 1).

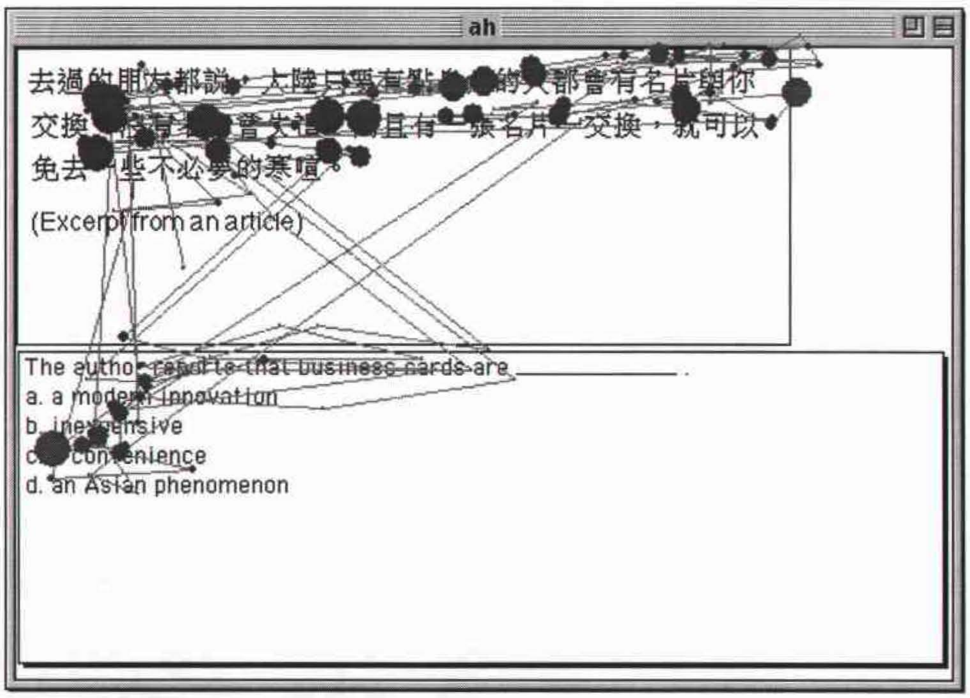


Figure 116. NNS03-Item 6 (ah48).

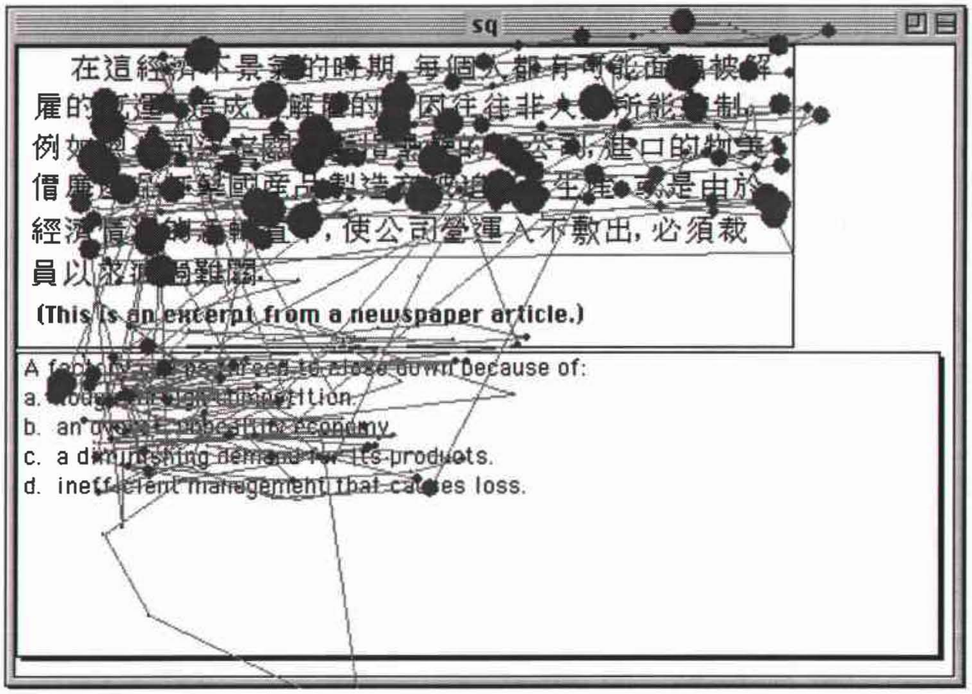


Figure 117. NNS03-Item 7 (sq07).

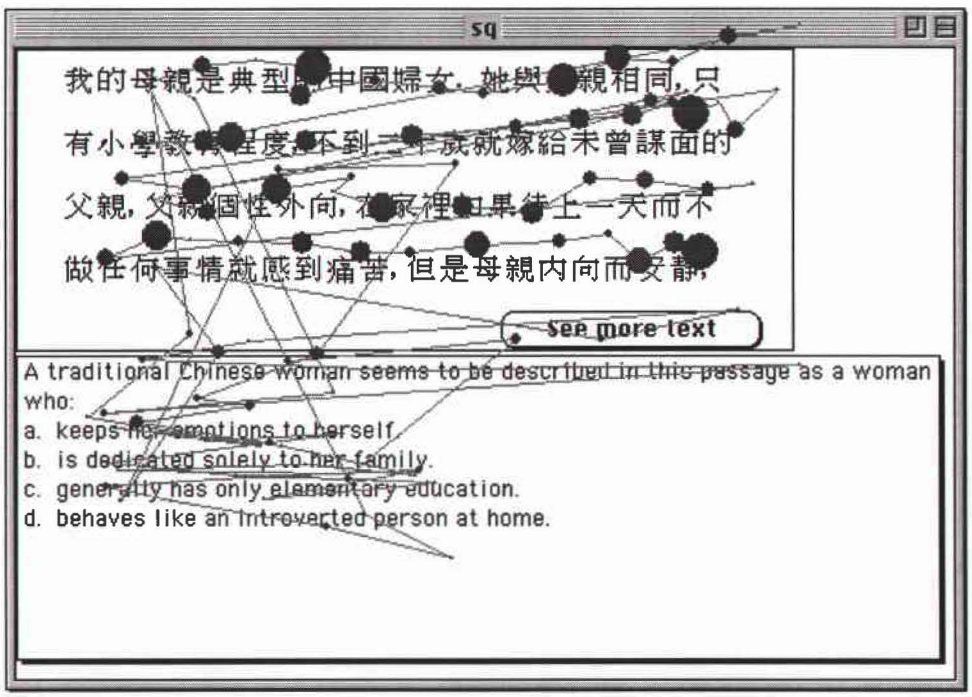


Figure 118a. NNS03-Item 8 (sq10, card 1).

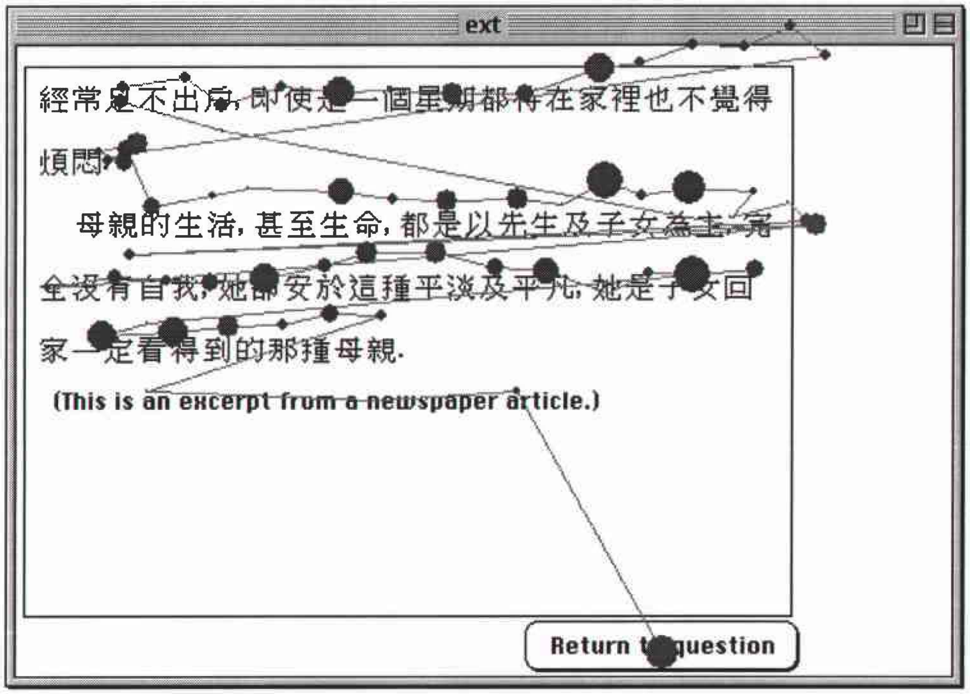


Figure 118b. NNS03-Item 8 (sq10, card 2).

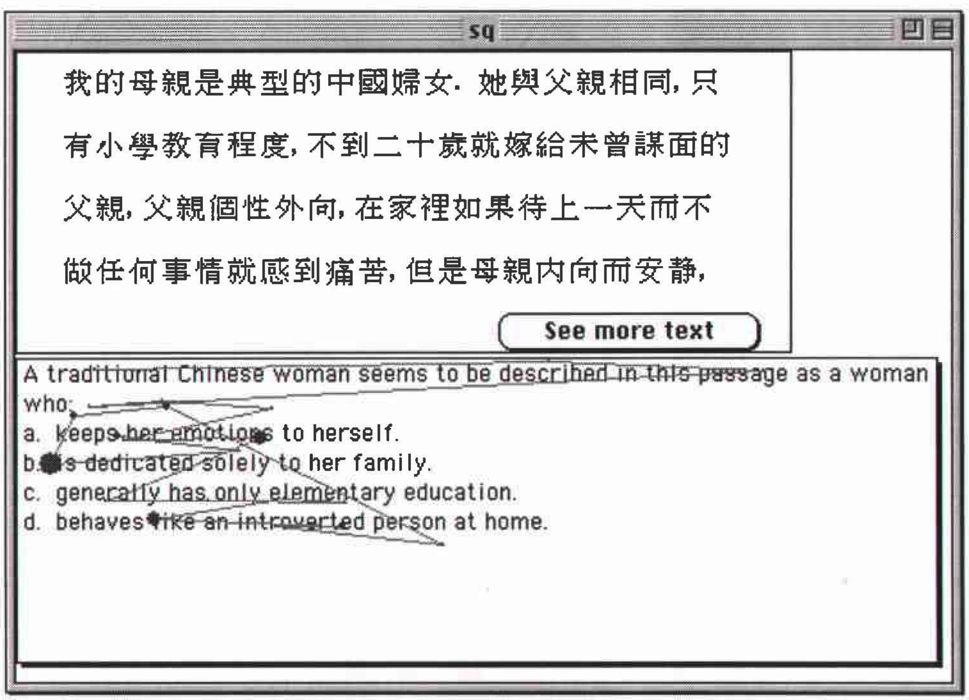


Figure 118c. NNS03-Item 8 (sq10, card 1).

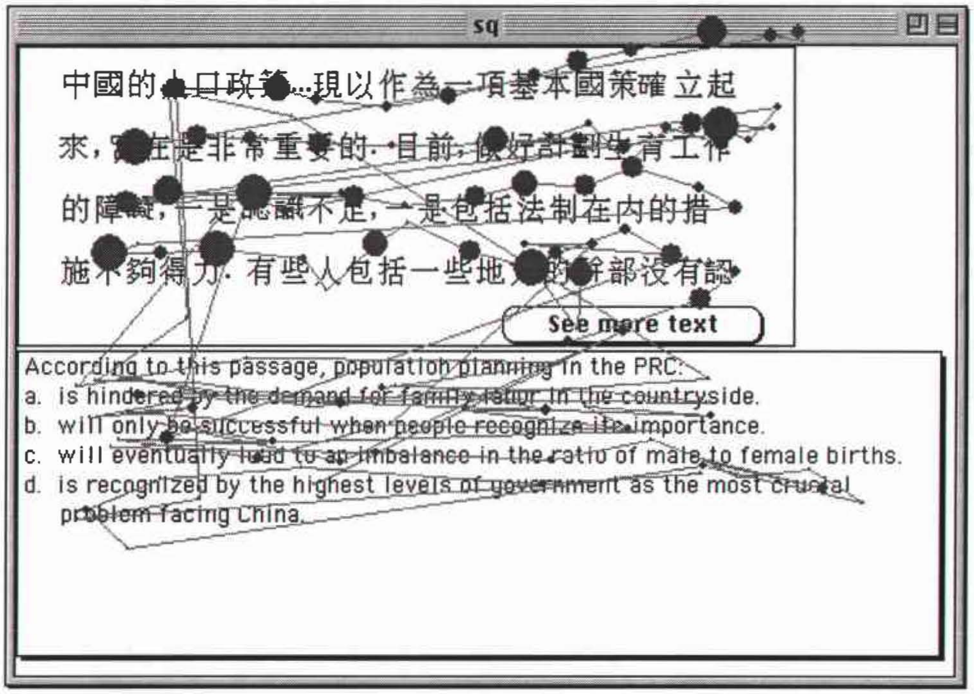


Figure 119a. NNS03-Item 9 (sq20, card 1).

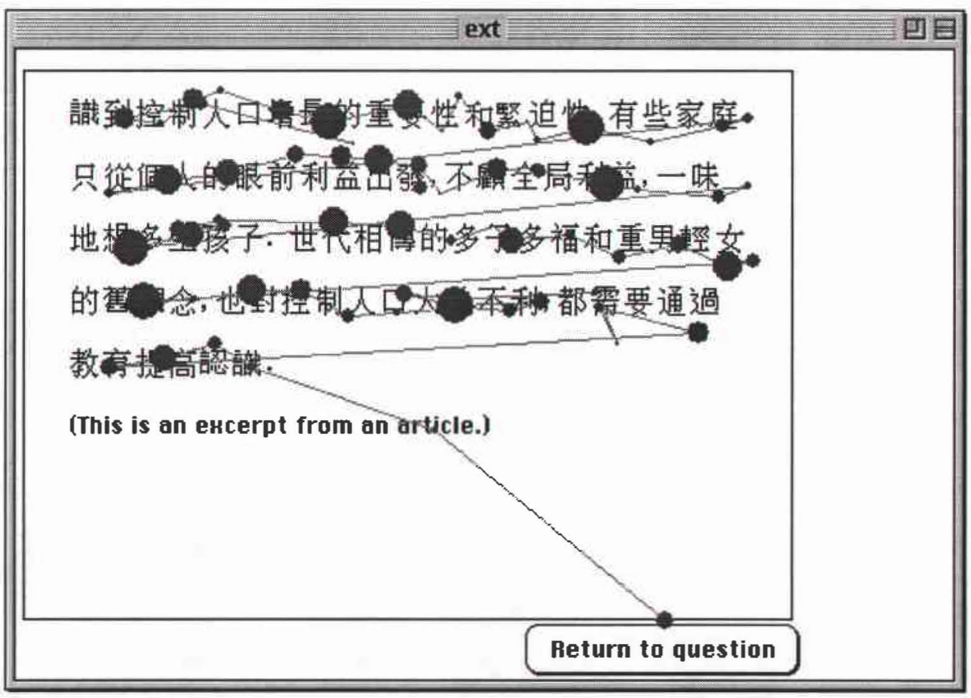


Figure 119b. NNS03-Item 9 (sq20, card 2).

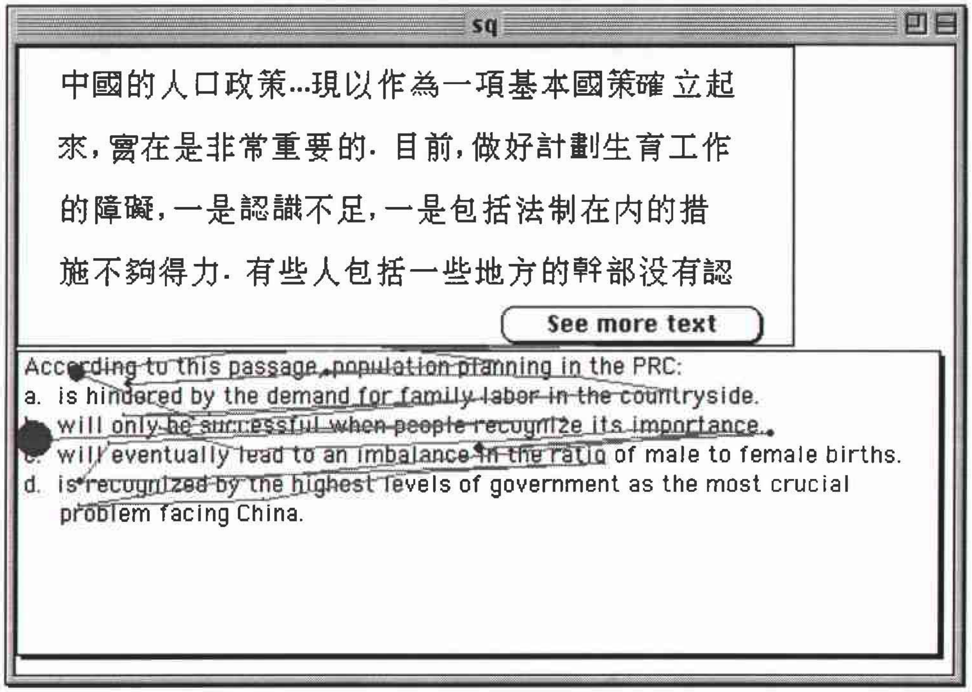


Figure 119c. NNS03-Item 9 (sq20, card 1).

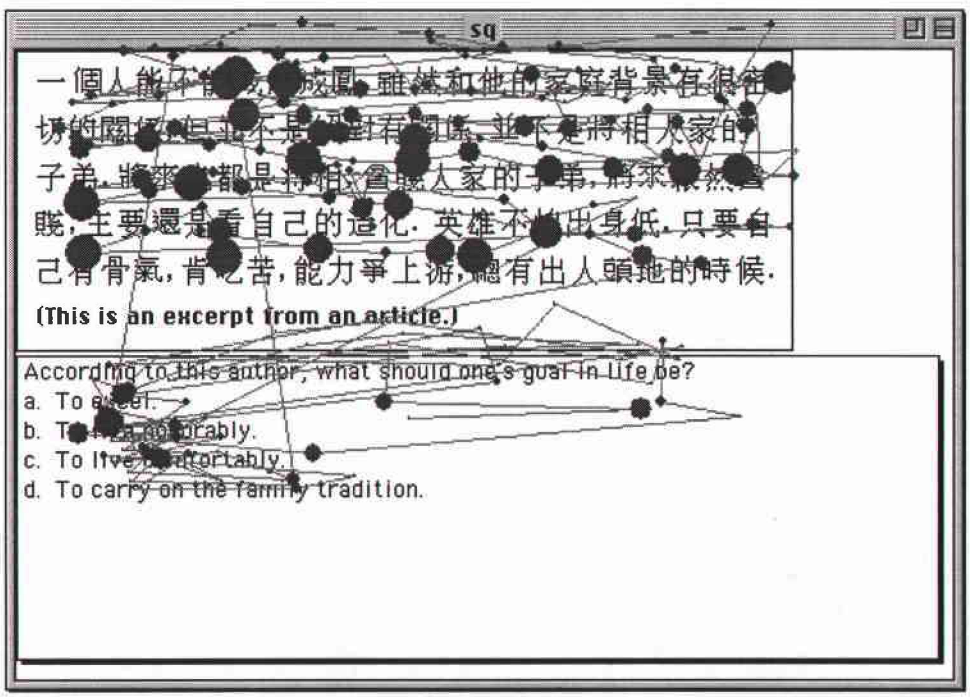


Figure 120. NNS03-Item 10 (sq14).

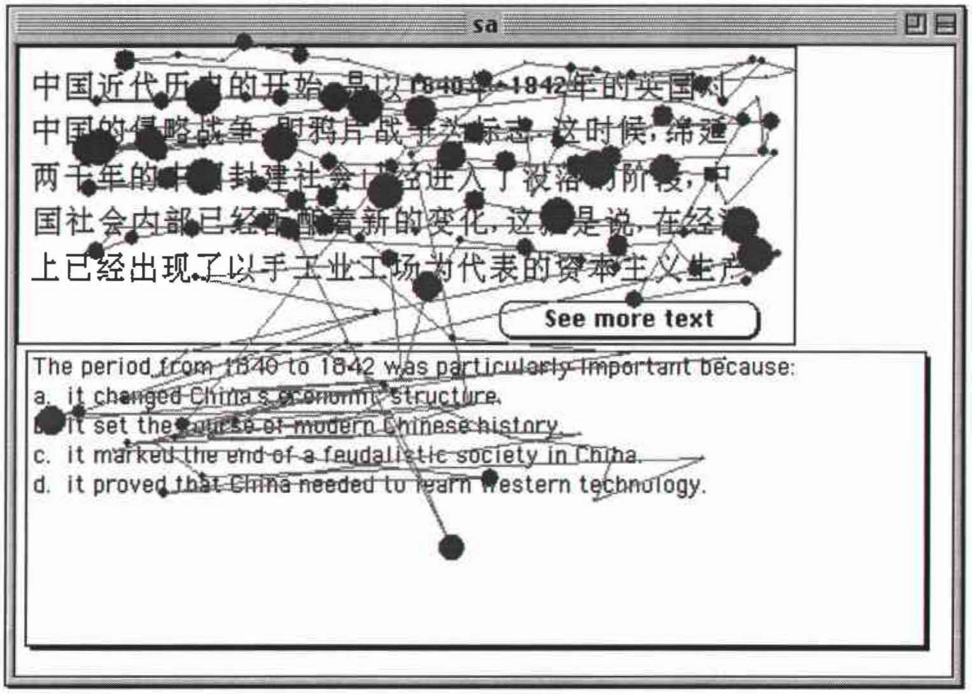


Figure 121a. NNS03-Item 11 (sa02, card 1).

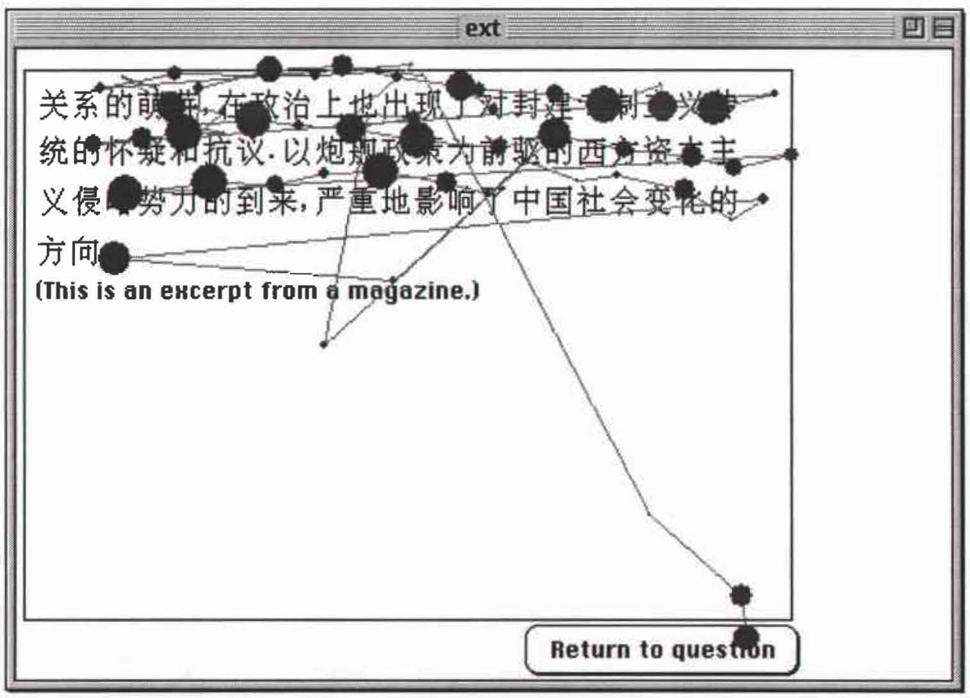


Figure 121b. NNS03-Item 11 (sa02, card 2).

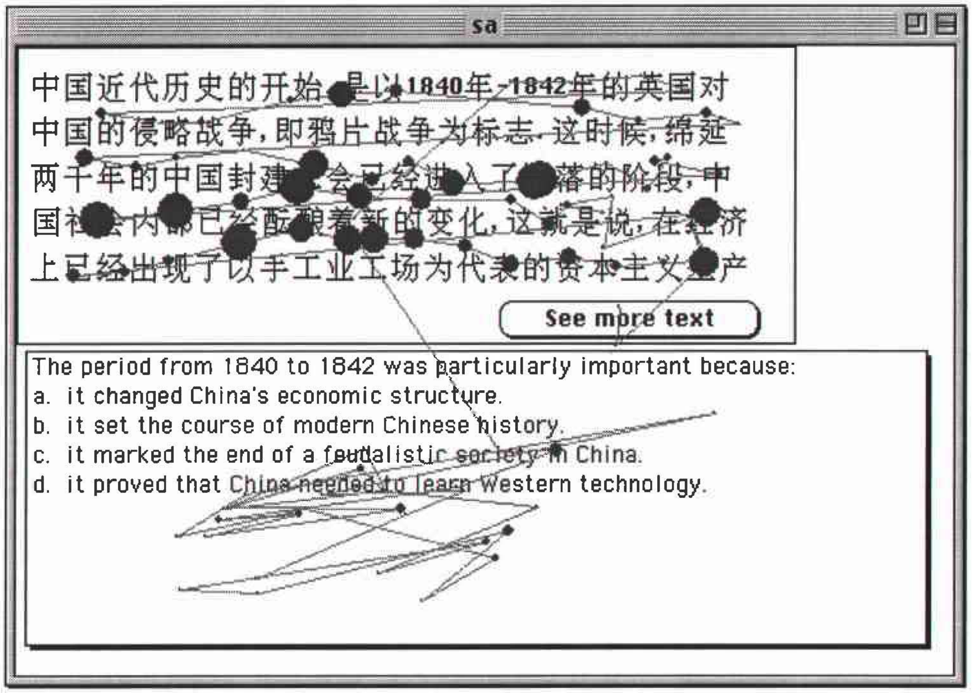


Figure 121c. NNS03-Item 11 (sa02, card 1).

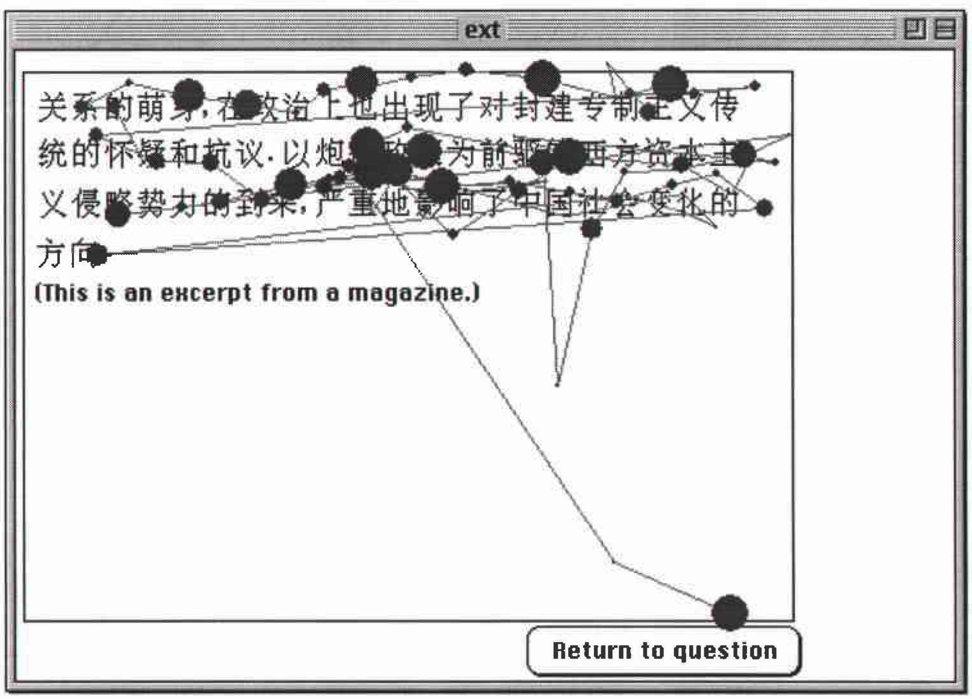


Figure 121d. NNS03-Item 11 (sa02, card 2).

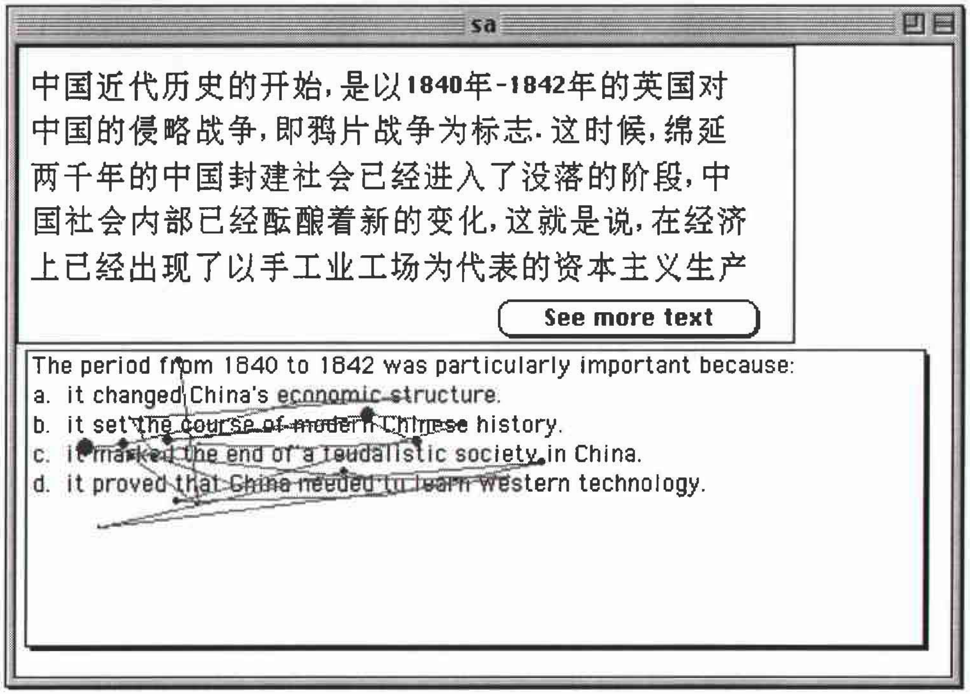


Figure 121e. NNS03-Item 11 (sa02, card 1).

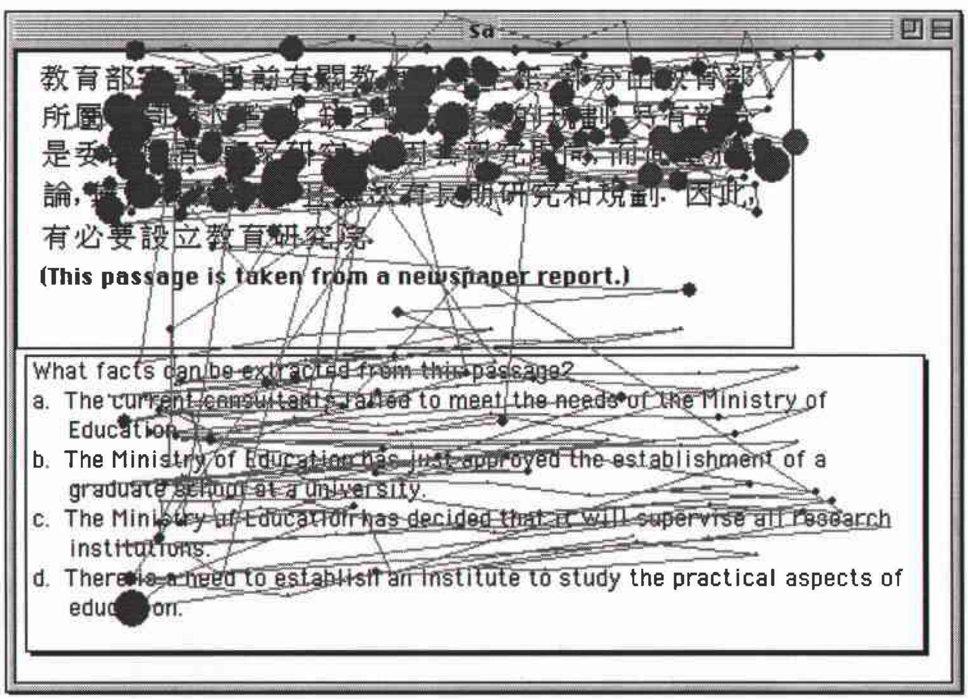


Figure 122. NNS03-Item 12 (sa14).

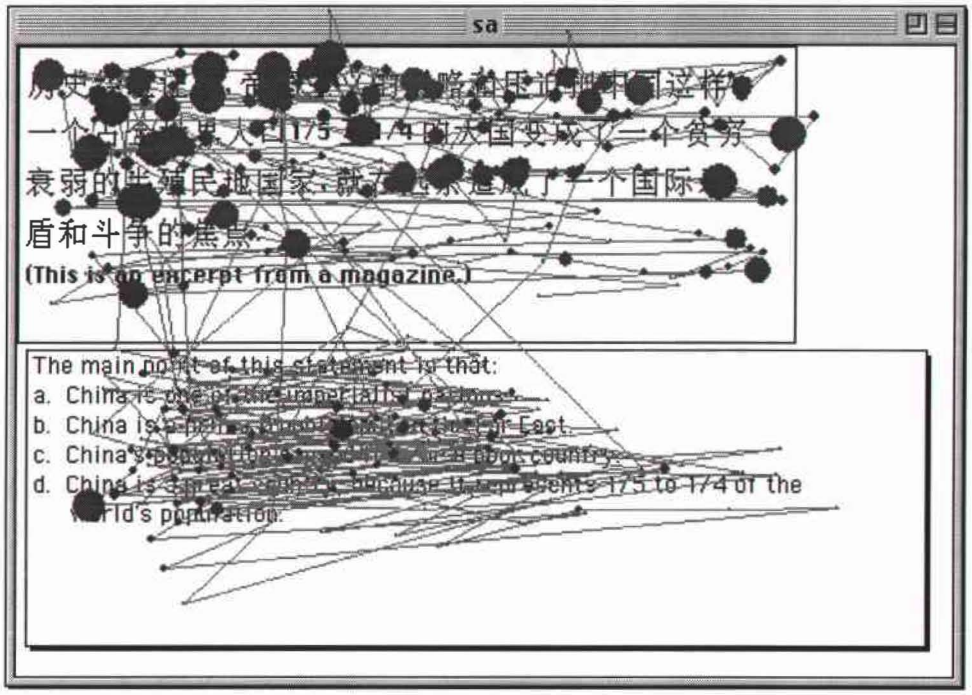


Figure 123. NNS03-Item 13 (sa01).

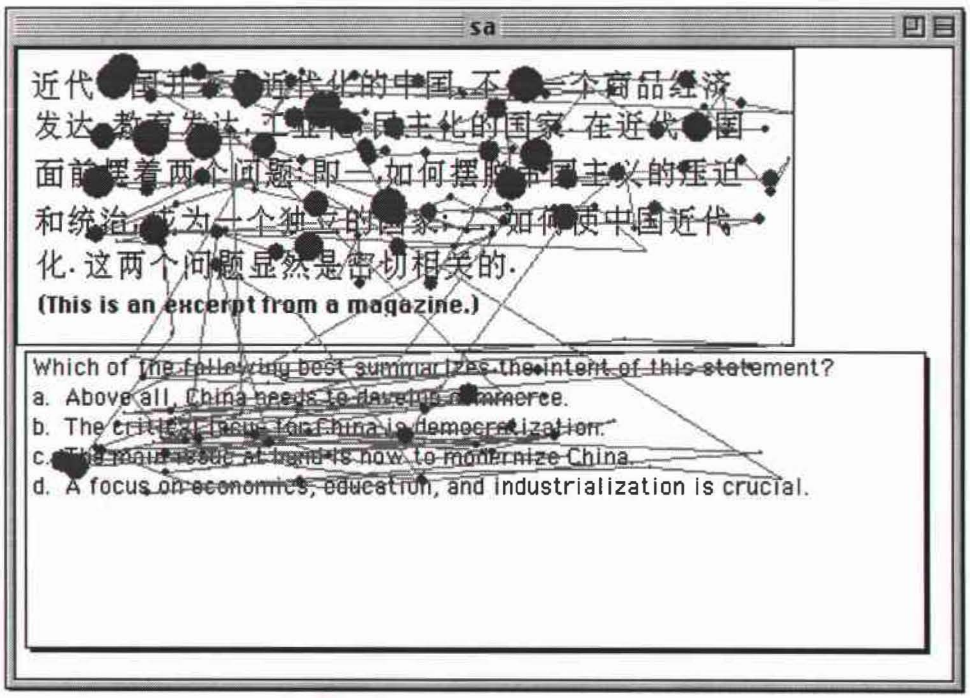


Figure 124. NNS03-Item 14 (sa03).

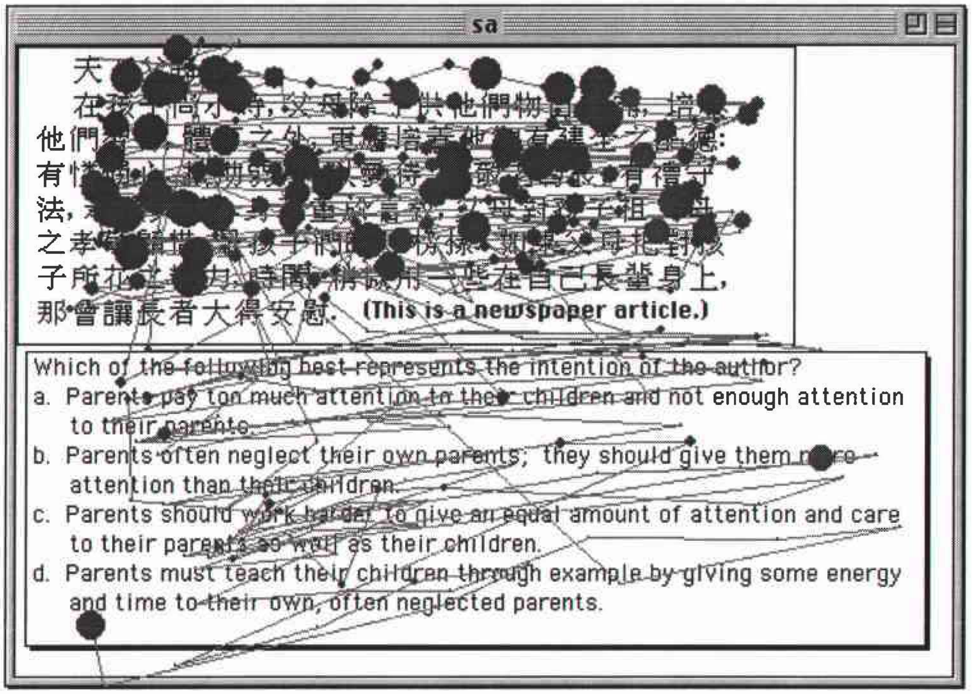


Figure 125. NNS03-Item 15 (sa04).

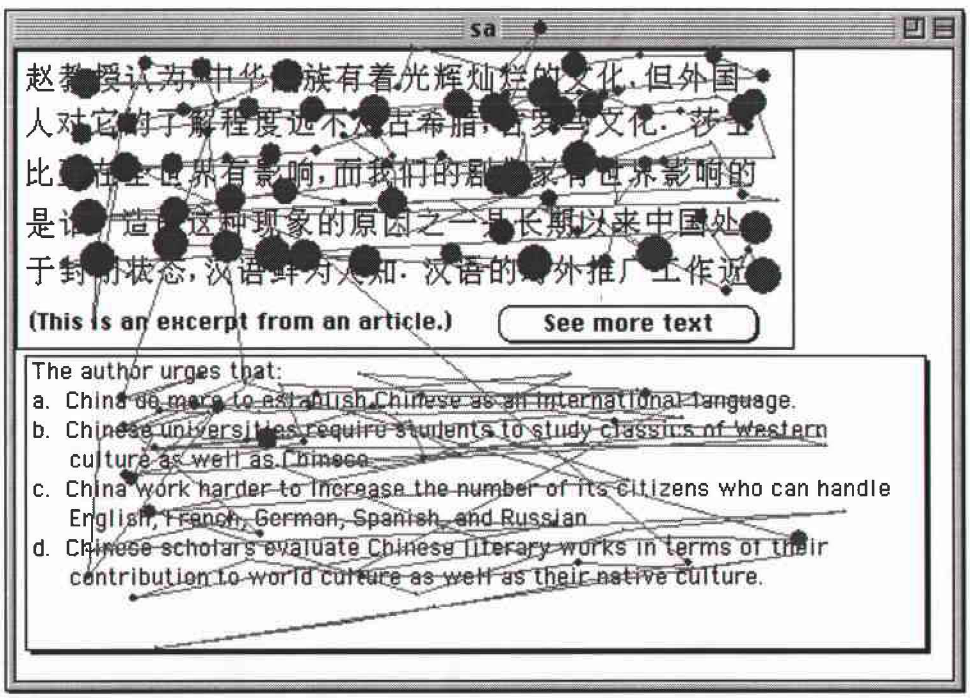


Figure 126a. NNS03-Item 16 (sa05, card 1).

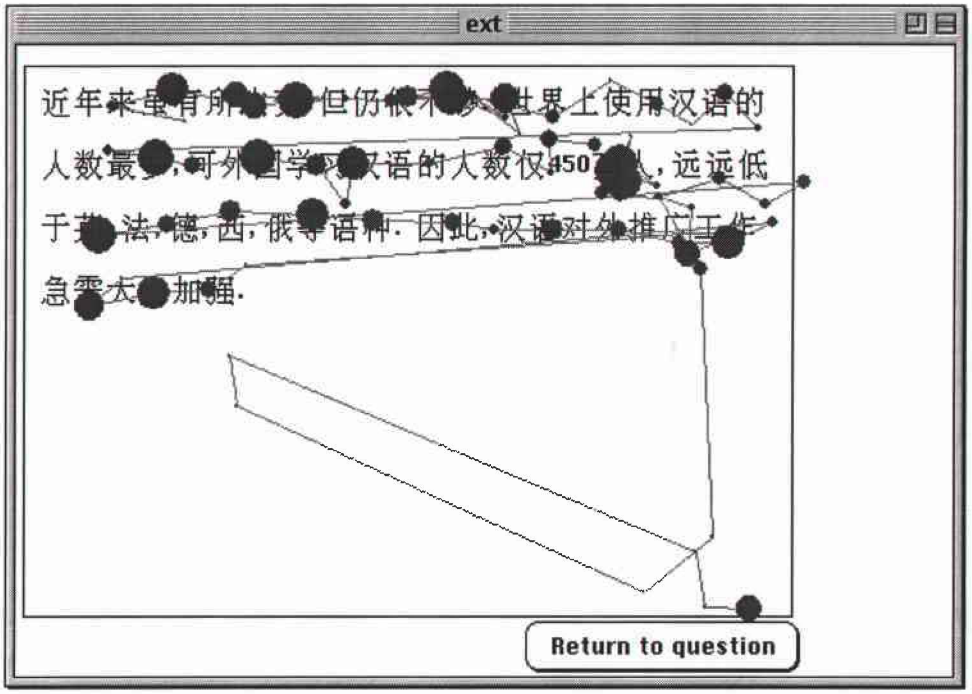


Figure 126b. NNS03-Item 16 (sa05, card 2).

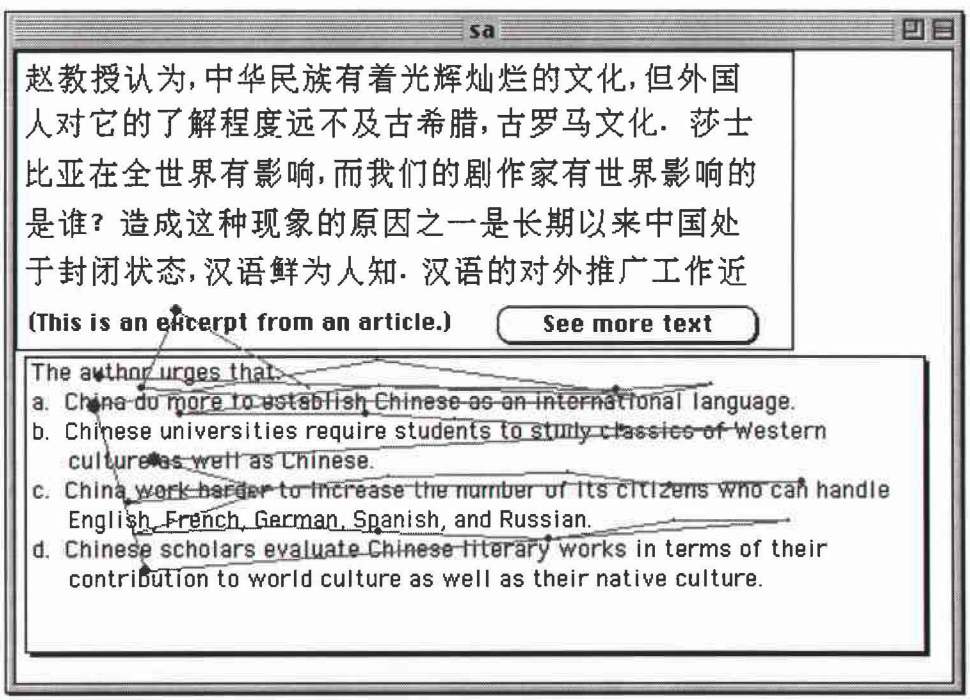


Figure 126c. NNS03-Item 16 (sa05, card 1).

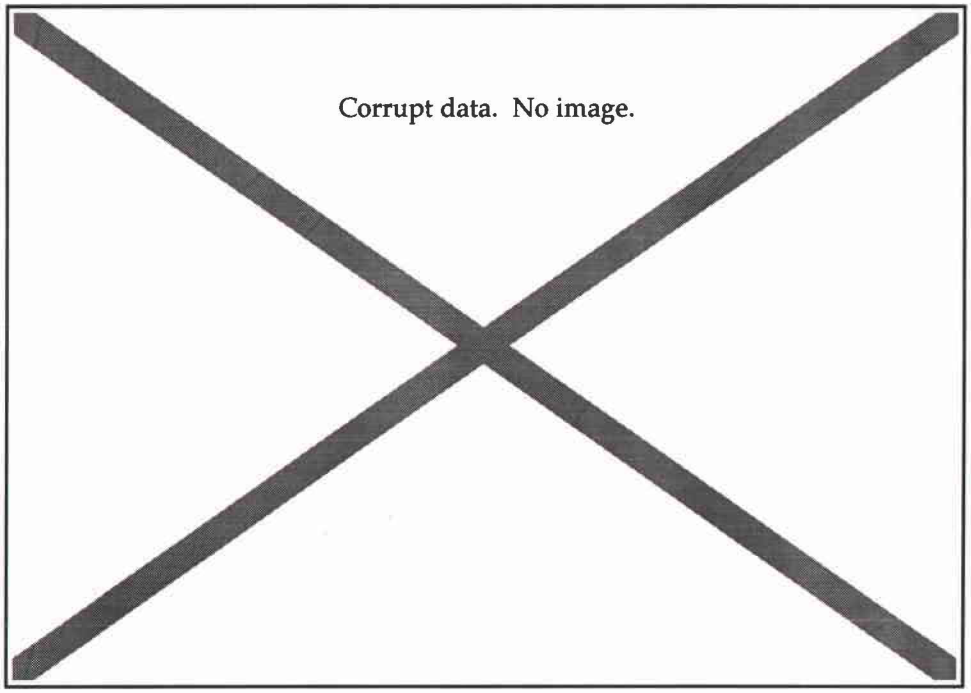


Figure 127. NNS03-Item 17 (se17).

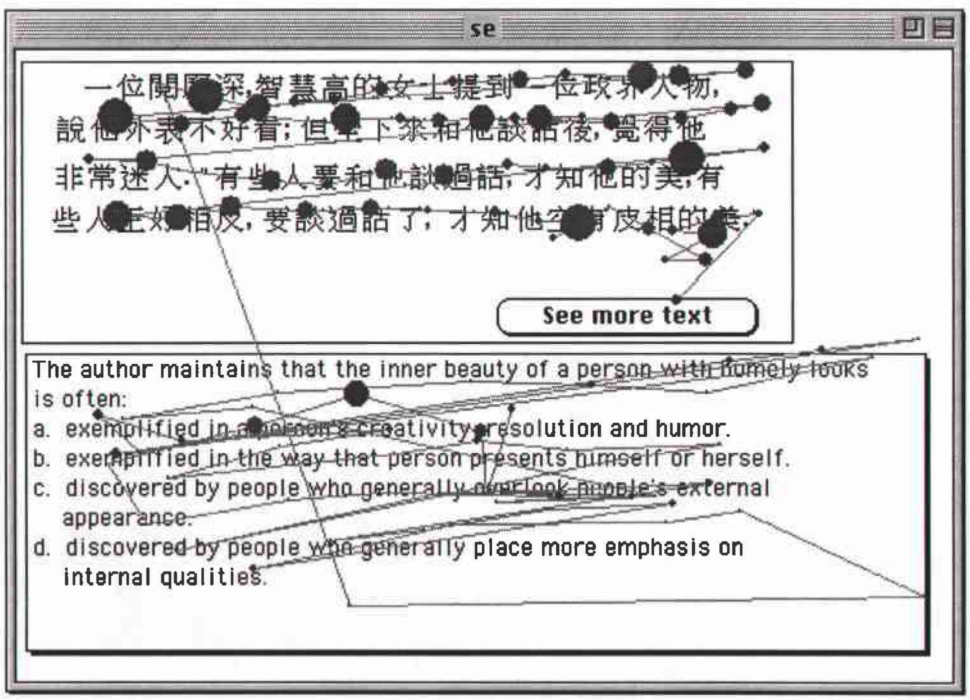


Figure 128a. NNS03-Item 18 (se04, card 1).

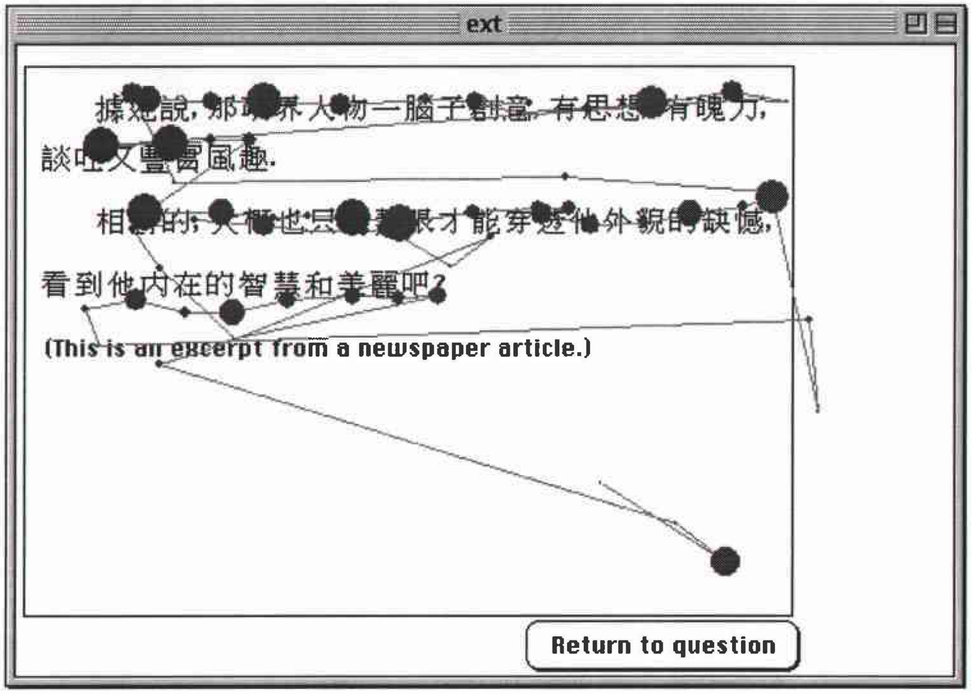


Figure 128b. NNS03-Item 18 (se04, card 2).

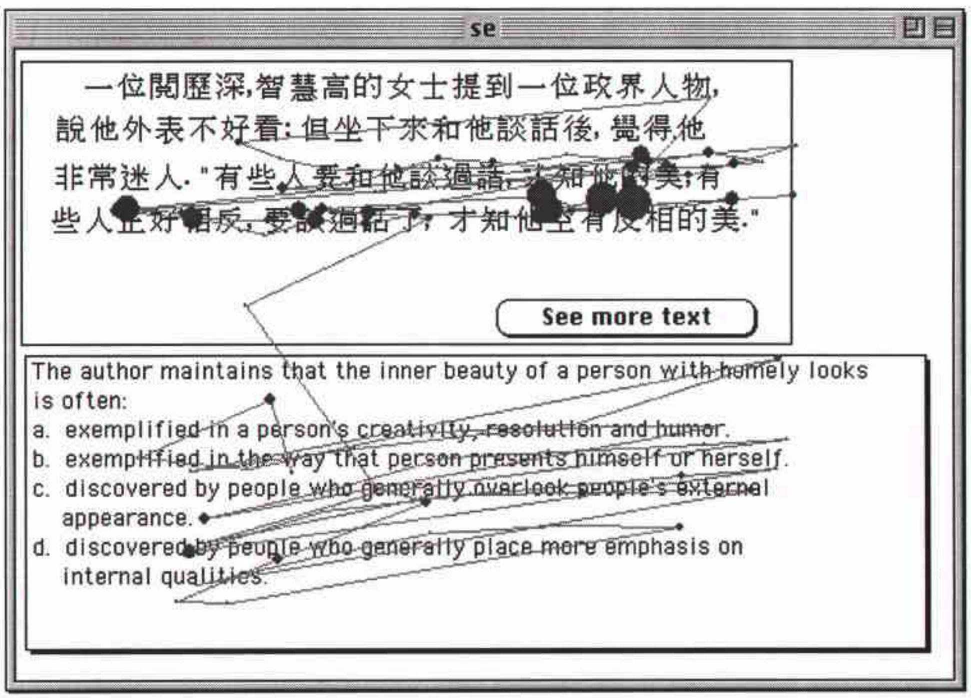


Figure 128c. NNS03-Item 18 (se04, card 1).

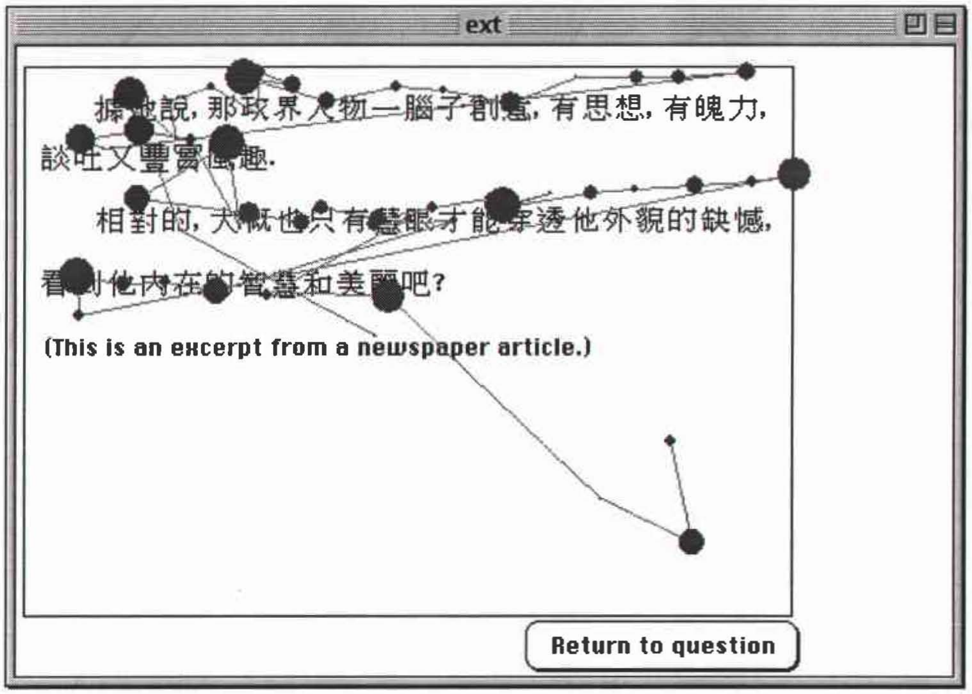


Figure 128d. NNS03-Item 18 (se04, card 2).

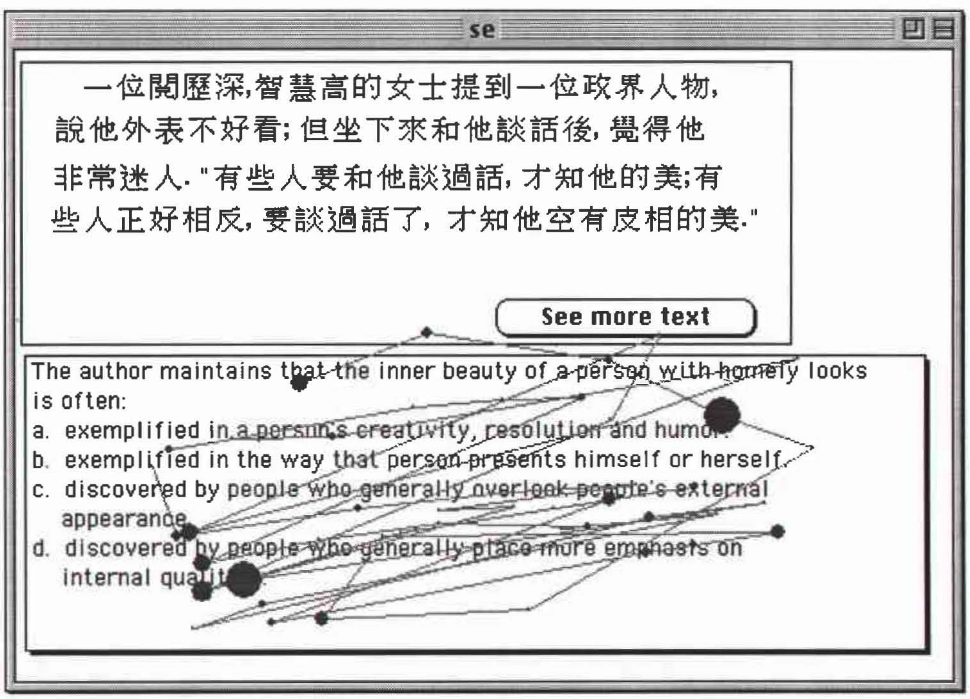


Figure 128e. NNS03-Item 18 (se04, card 1).

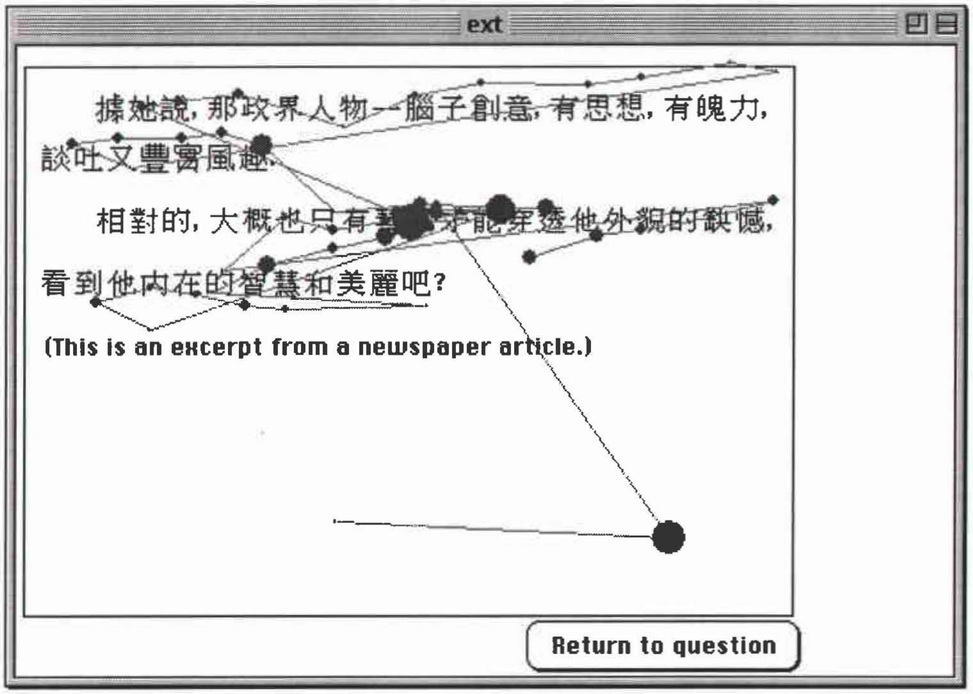


Figure 128f. NNS03-Item 18 (se04, card 2).

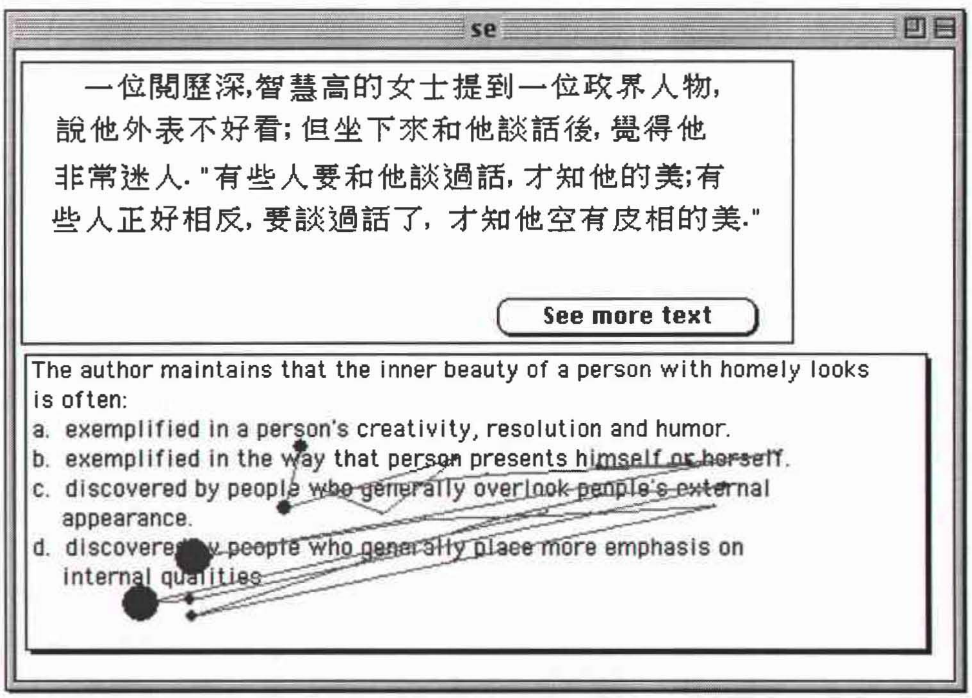


Figure 128g. NNS03-Item 18 (se04, card 1).

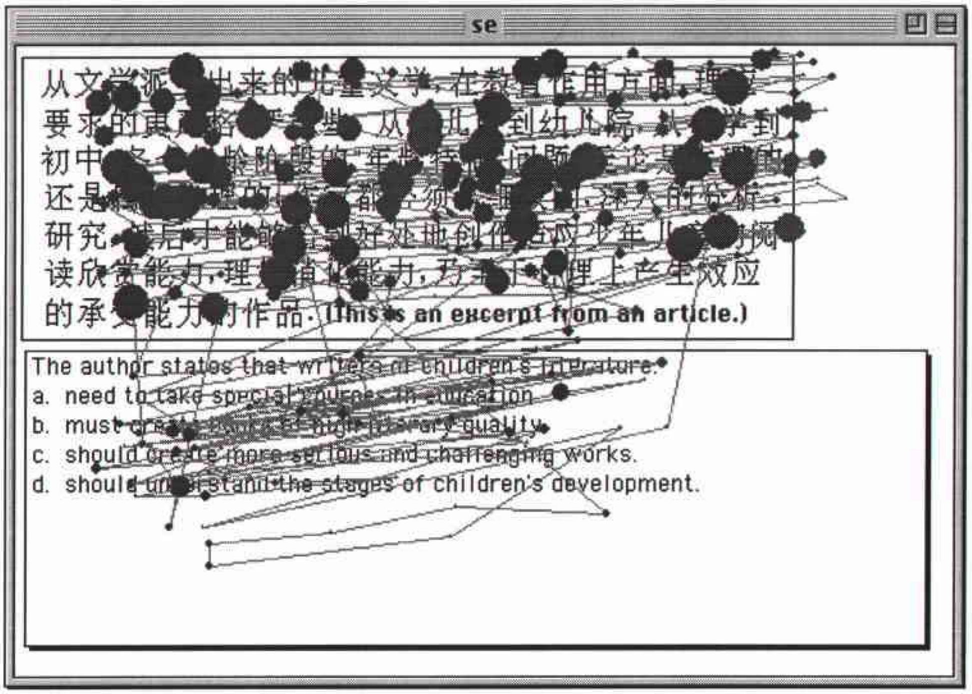


Figure 129. NNS03-Item 19 (se20).

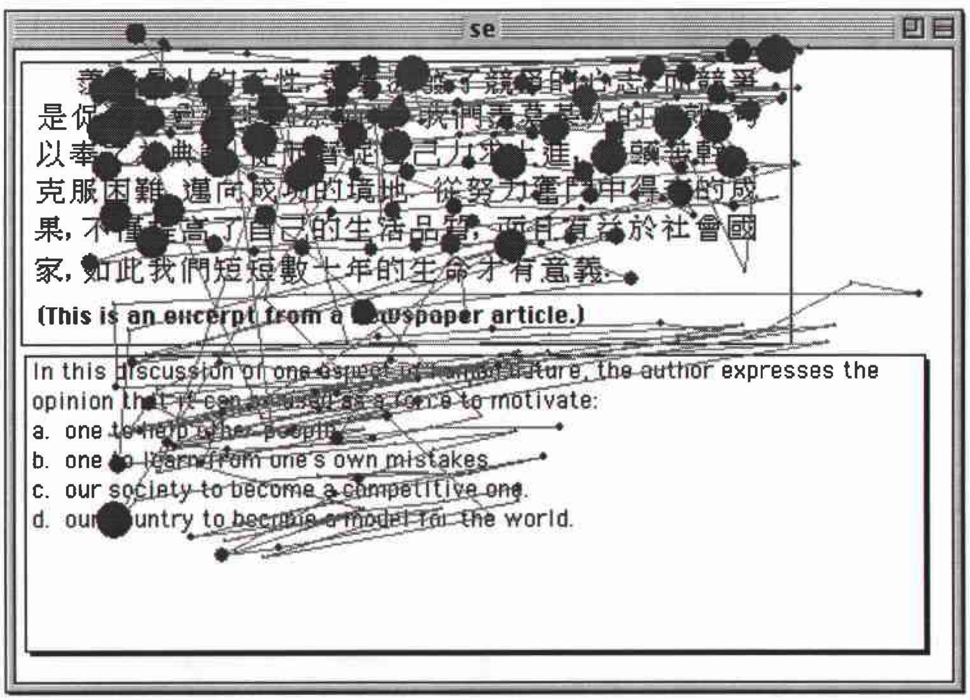


Figure 130. NNS03-Item 20 (sq02).

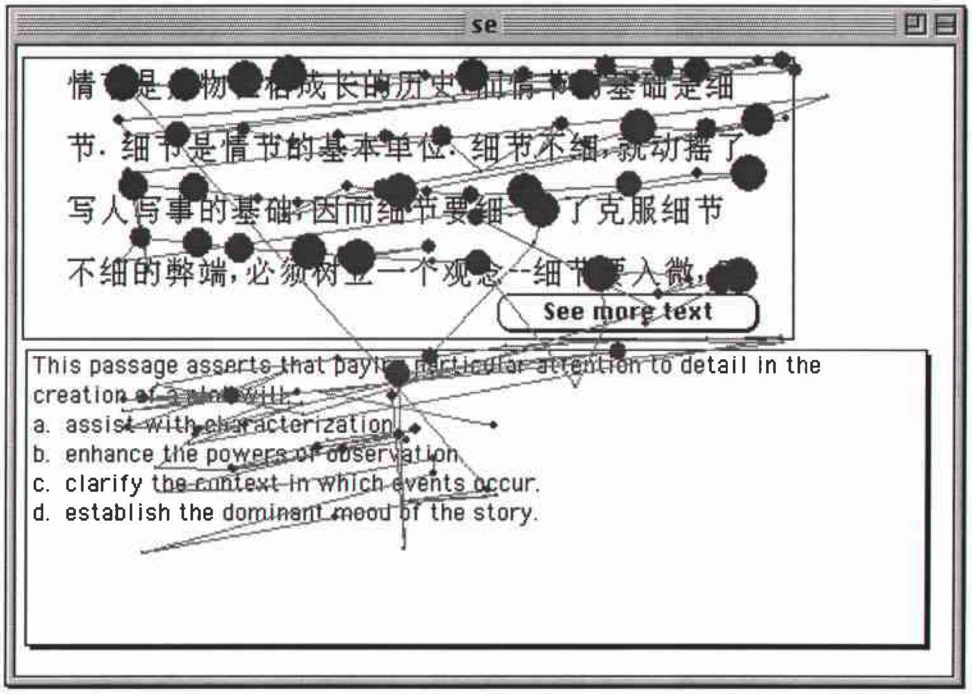


Figure 131a. NNS03-Item 21 (se16, card 1).

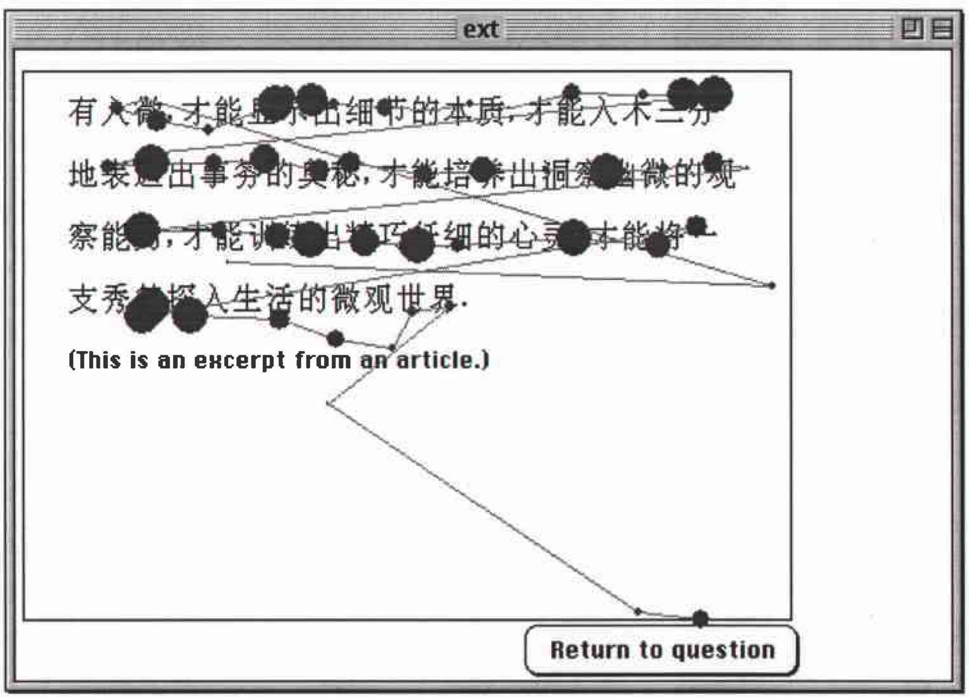


Figure 131b. NNS03-Item 21 (se16, card 2).

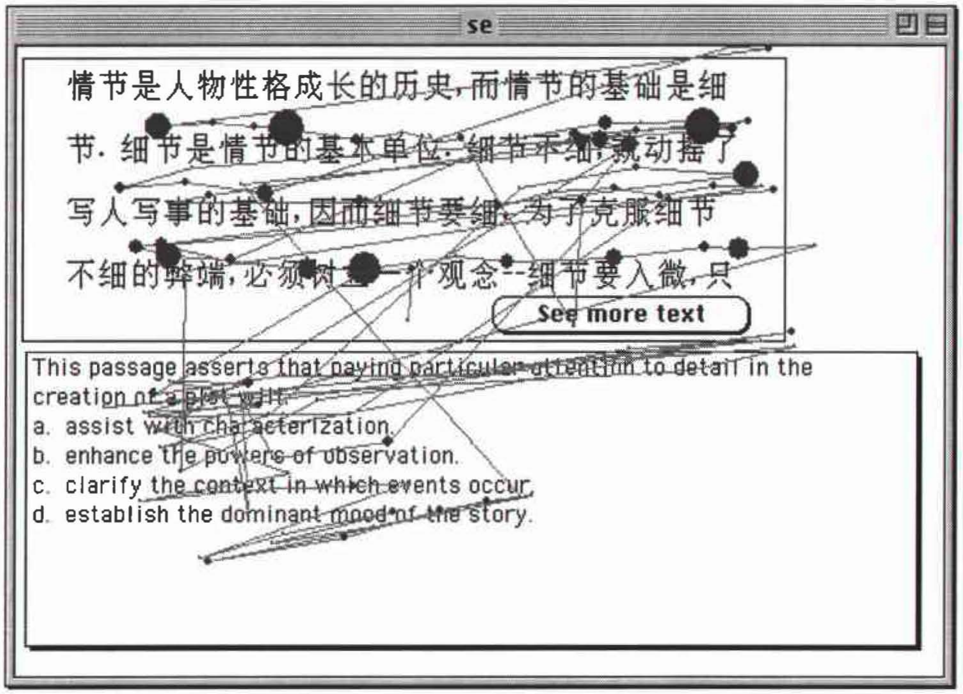


Figure 131c. NNS03-Item 21 (se16, card 1).

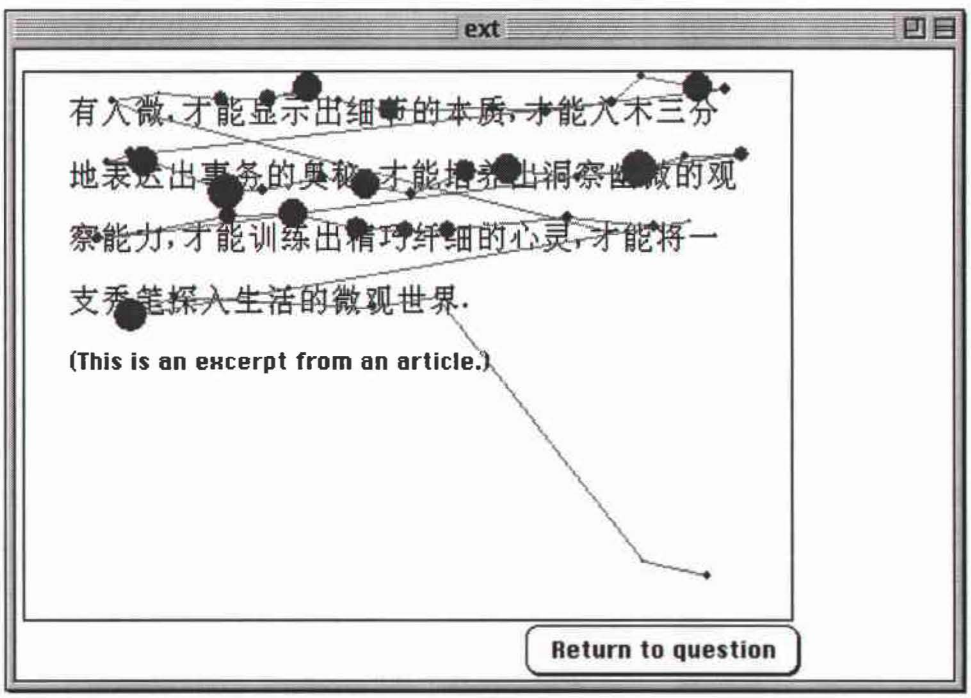


Figure 131d. NNS03-Item 21 (se16, card 2).

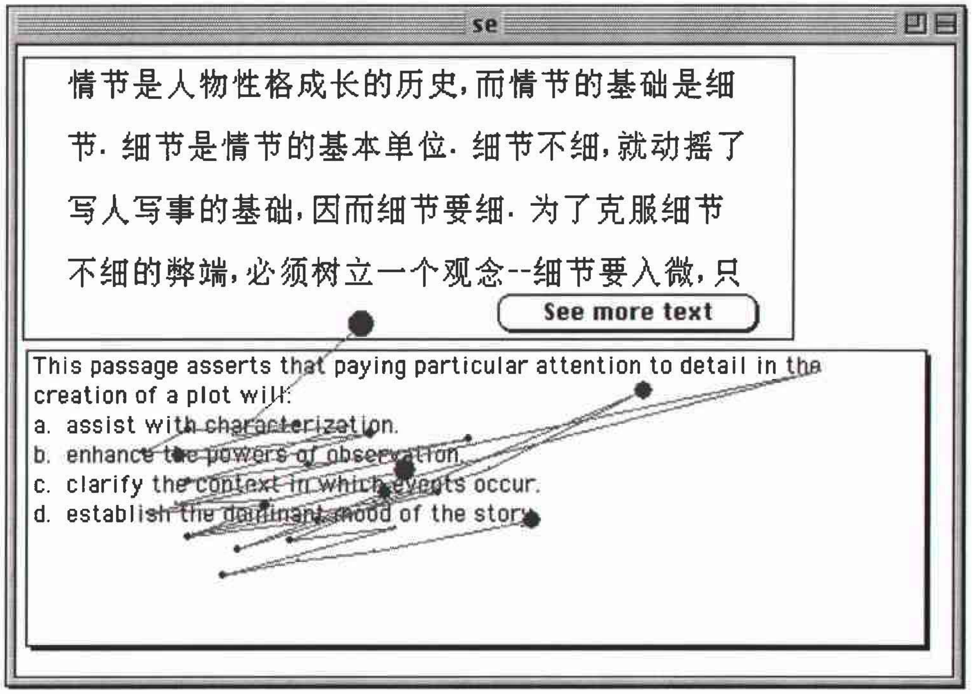


Figure 131e. NNS03-Item 21 (se16, card 1).

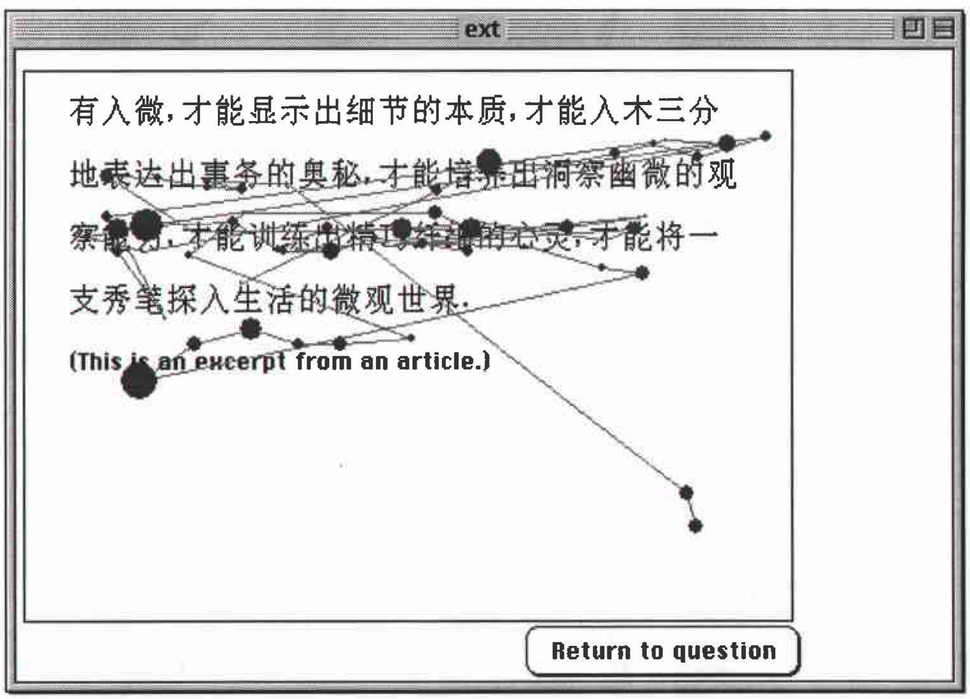


Figure 131f. NNS03-Item 21 (se16, card 2).

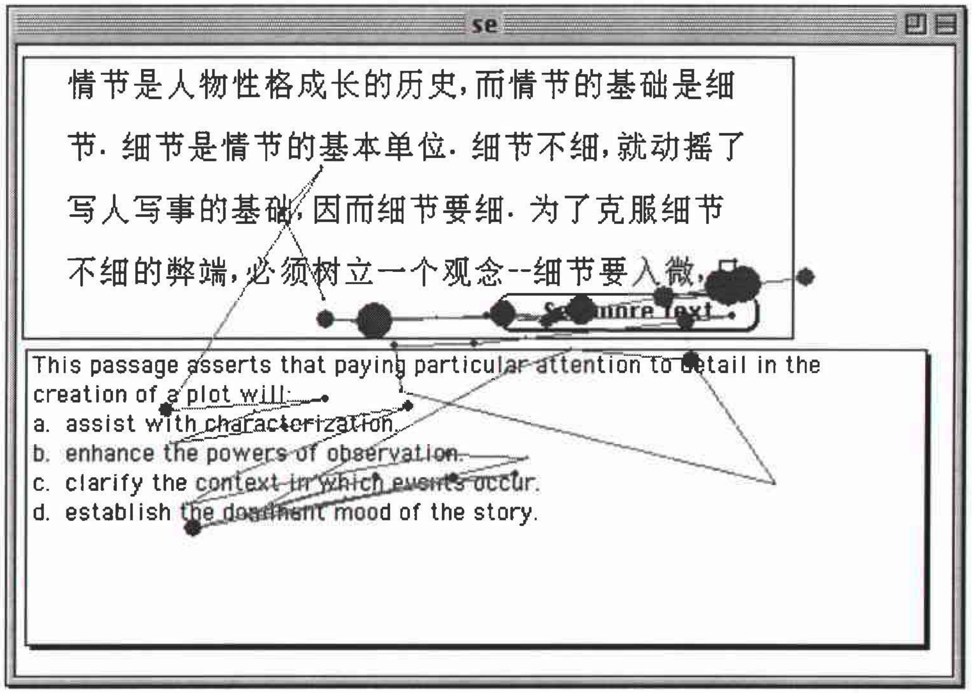


Figure 131g. NNS03-Item 21 (se16, card 1).

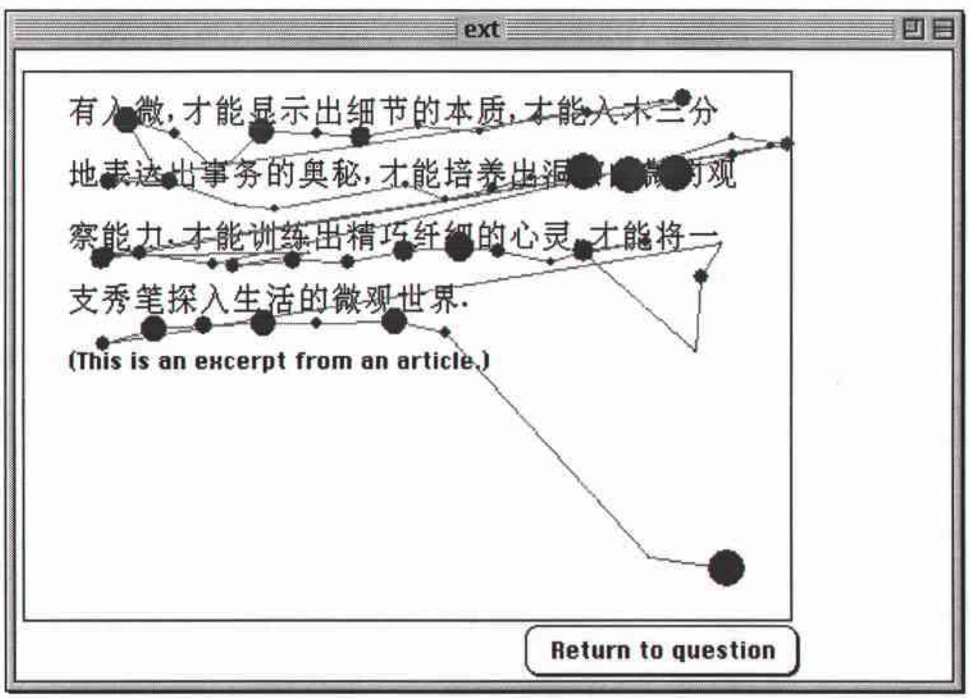


Figure 131h. NNS03-Item 21 (se16, card 2).

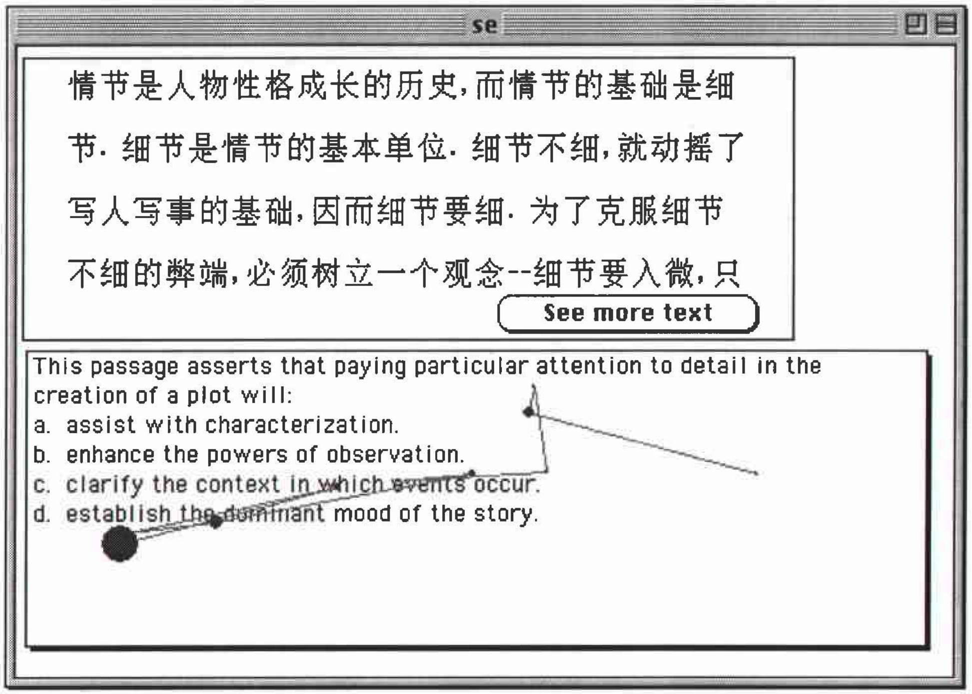


Figure 131i. NNS03-Item 21 (se16, card 1).

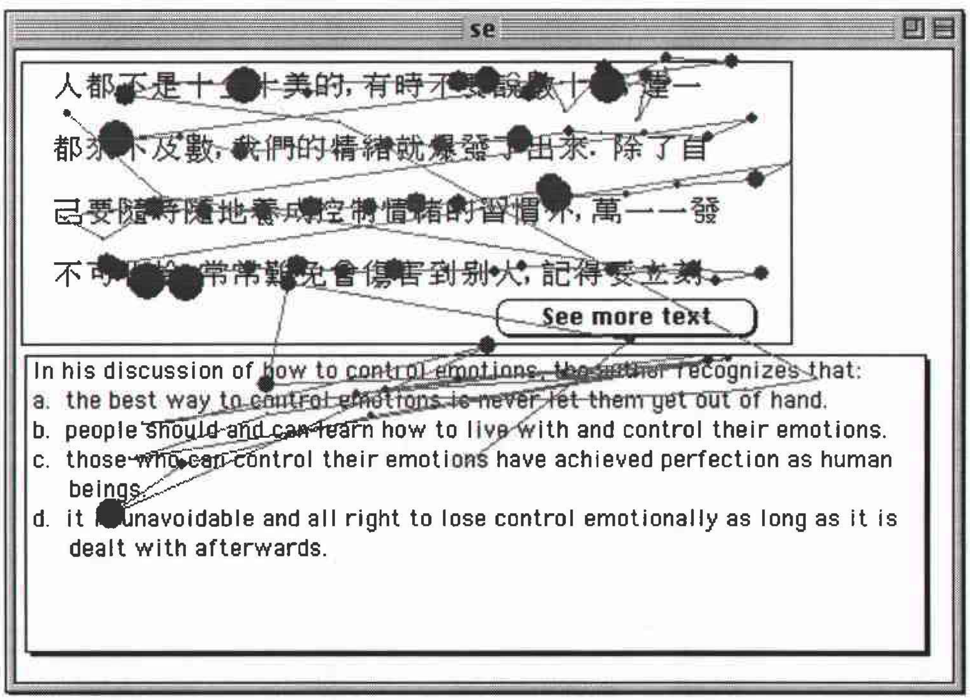


Figure 132a. NNS03-Item 22 (se06, card 1).

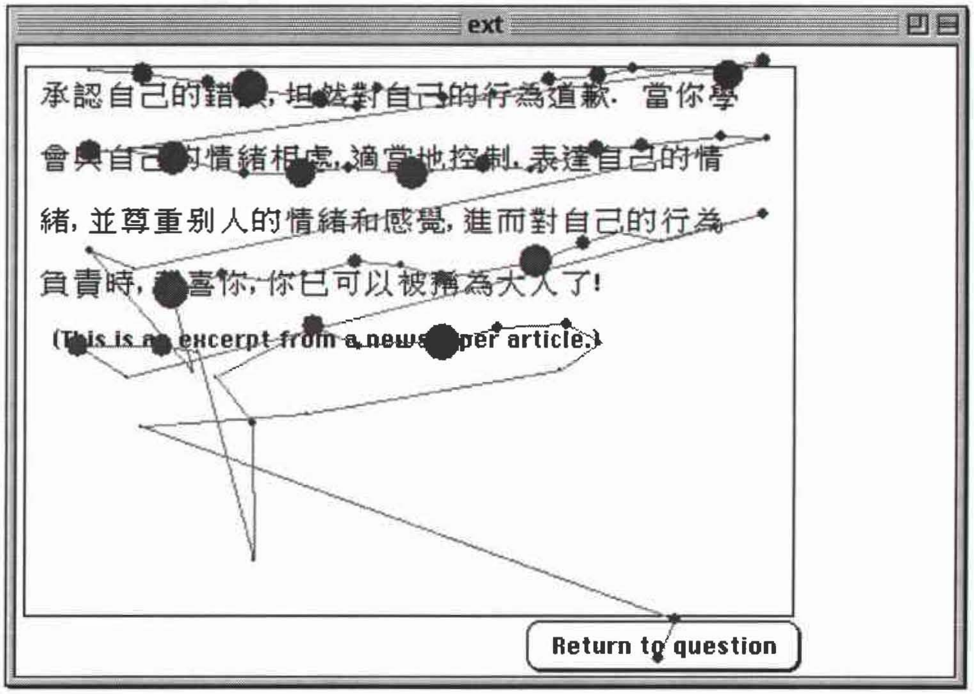


Figure 132b. NNS03-Item 22 (se06, card 2).

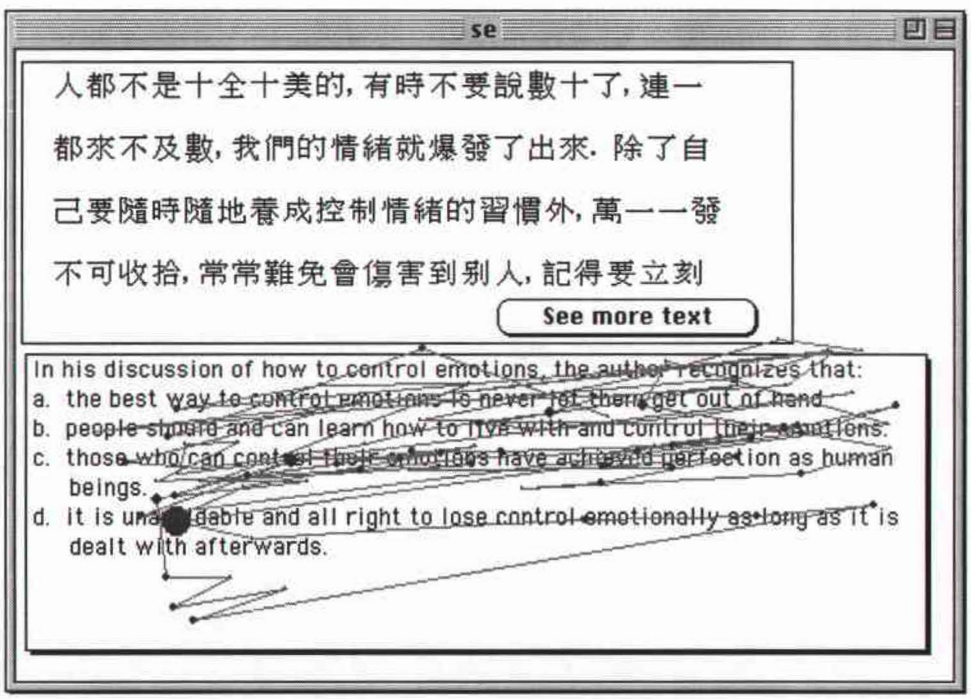


Figure 132c. NNS03-Item 22 (se06, card 1).

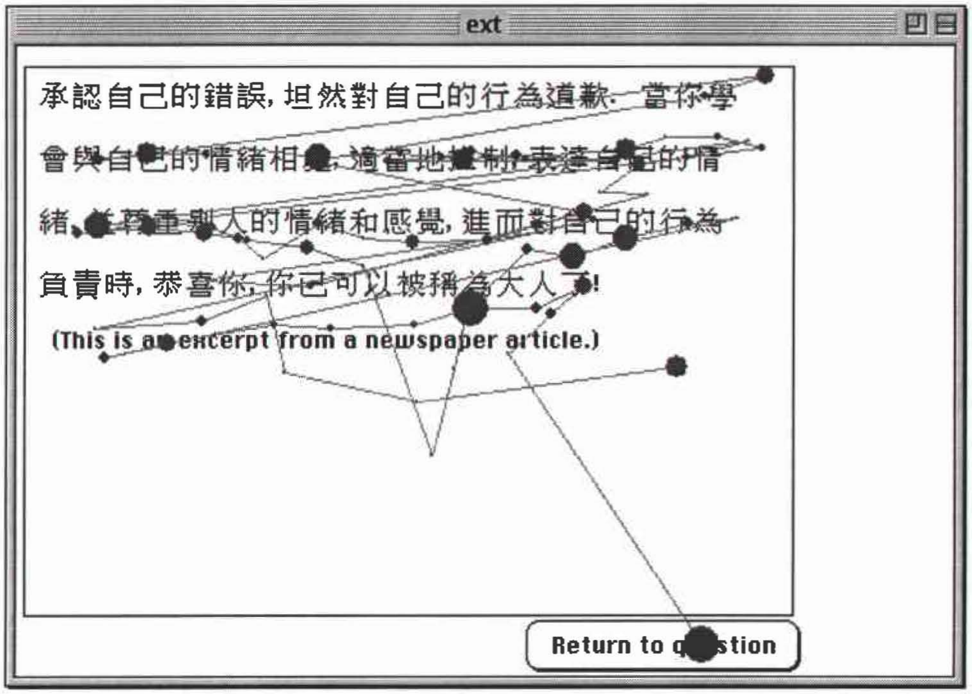


Figure 132d. NNS03-Item 22 (se06, card 2).

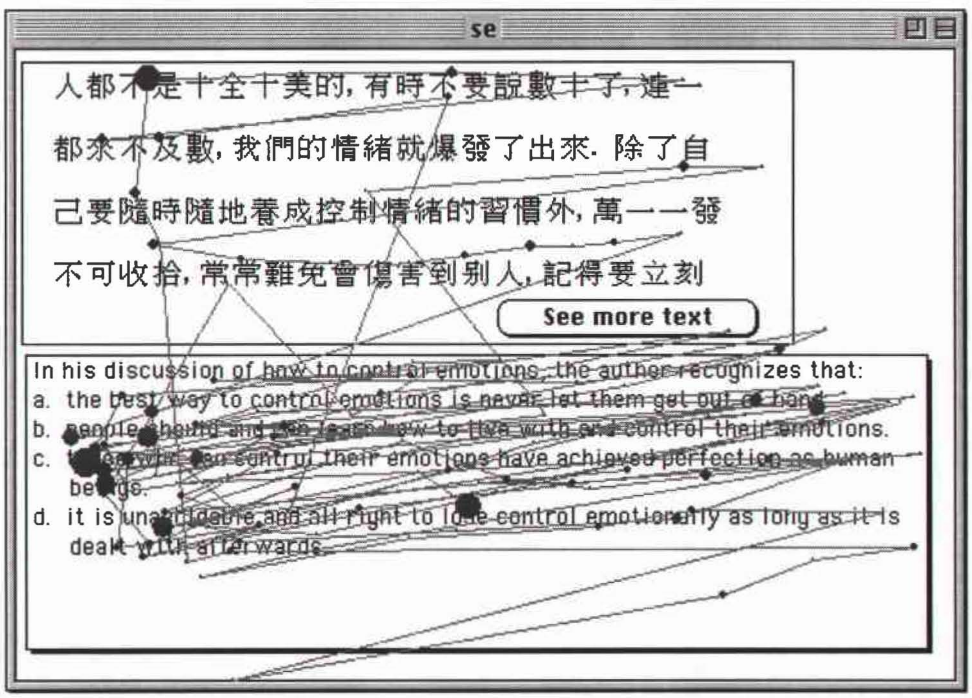


Figure 132e. NNS03-Item 22 (se06, card 1).

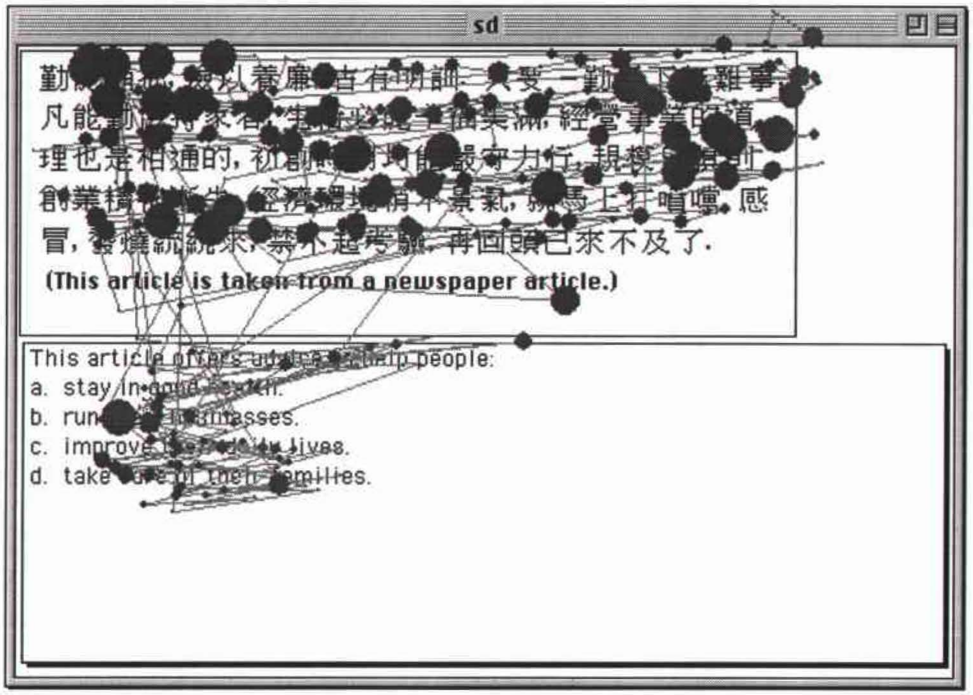


Figure 133. NNS03-Item 23 (sd18).

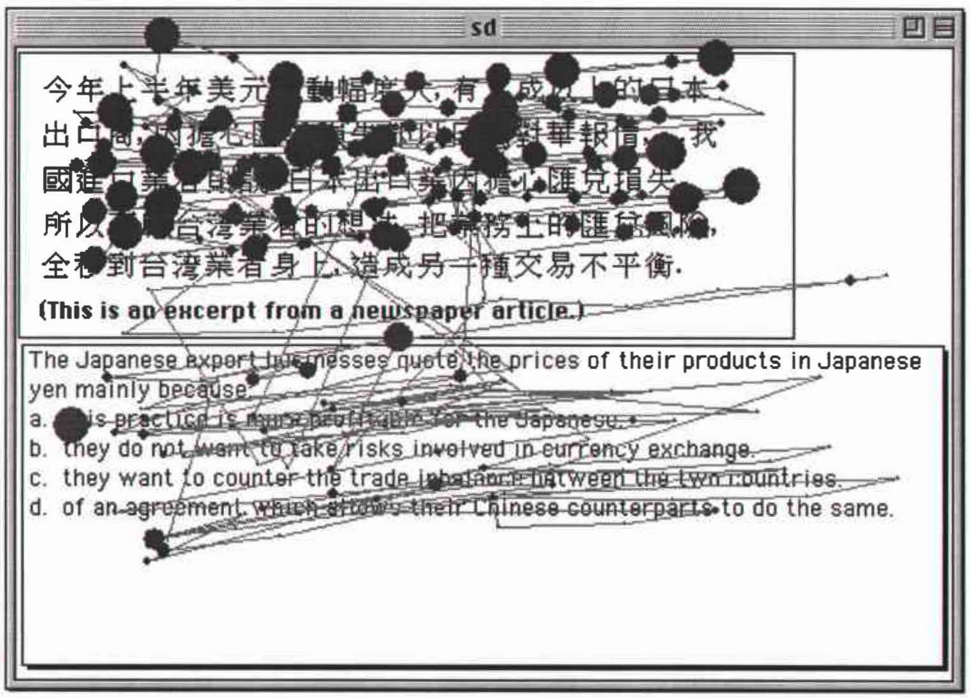


Figure 134. NNS03-Item 24 (sd15).

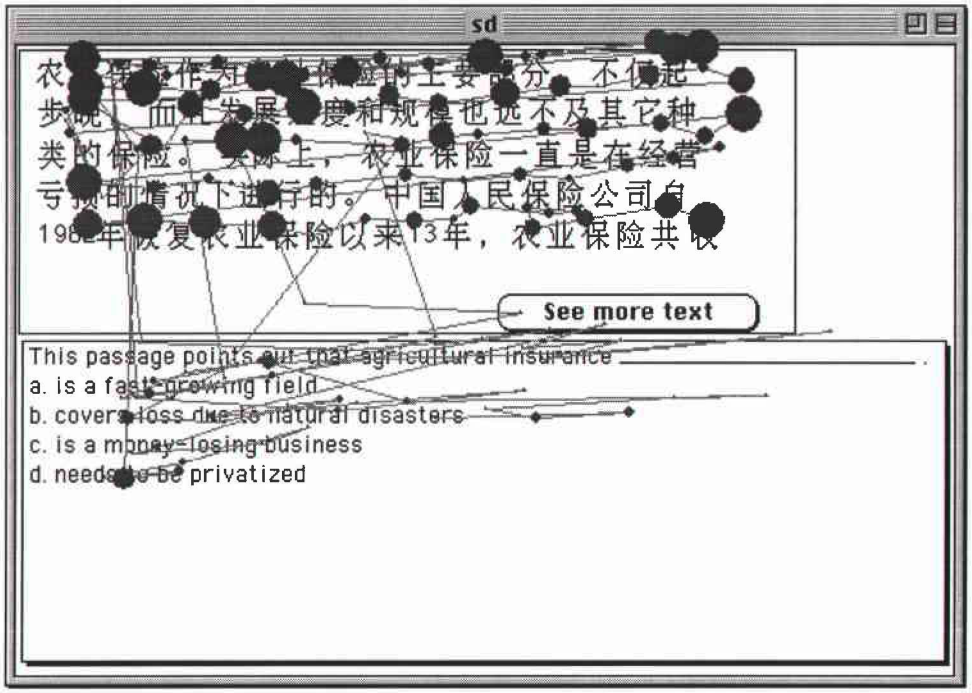


Figure 135a. NNS03-Item 25 (sd22, card 1).

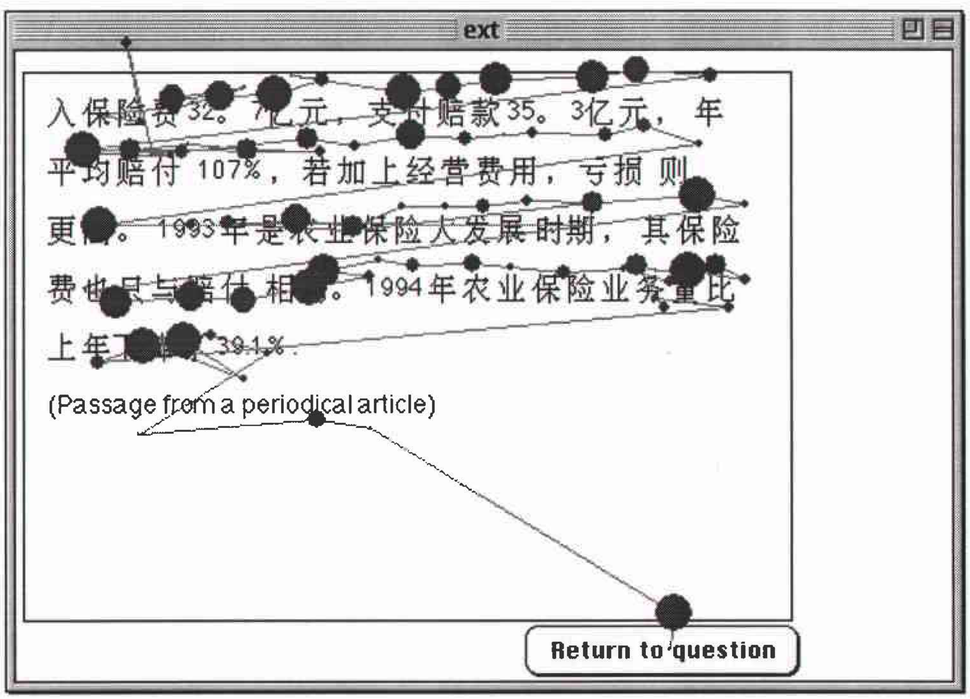


Figure 135b. NNS03-Item 25 (sd22, card 2).

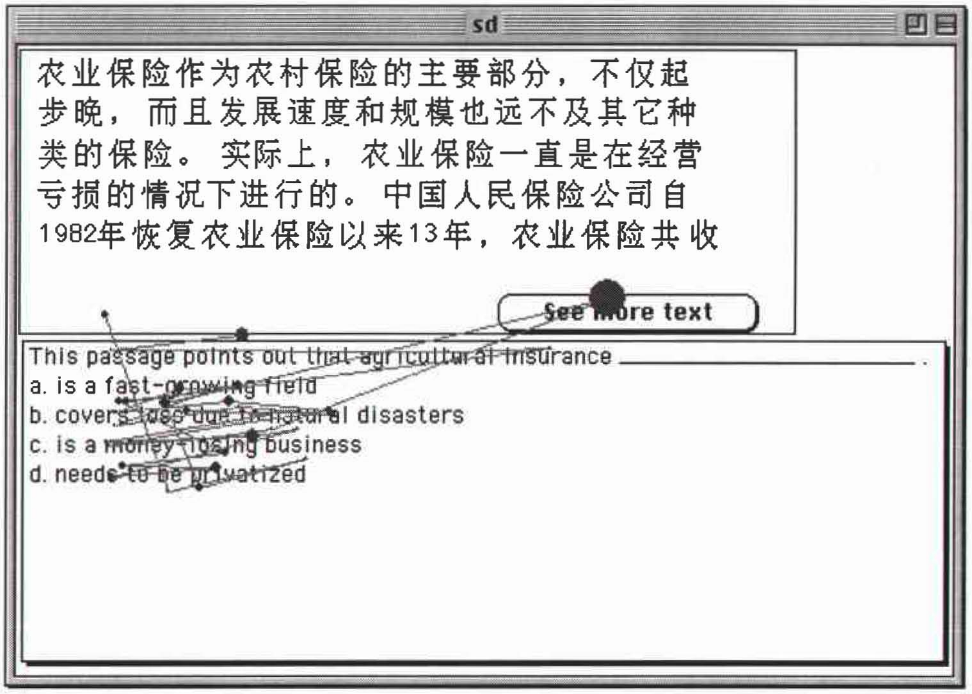


Figure 135c. NNS03-Item 25 (sd22, card 1).

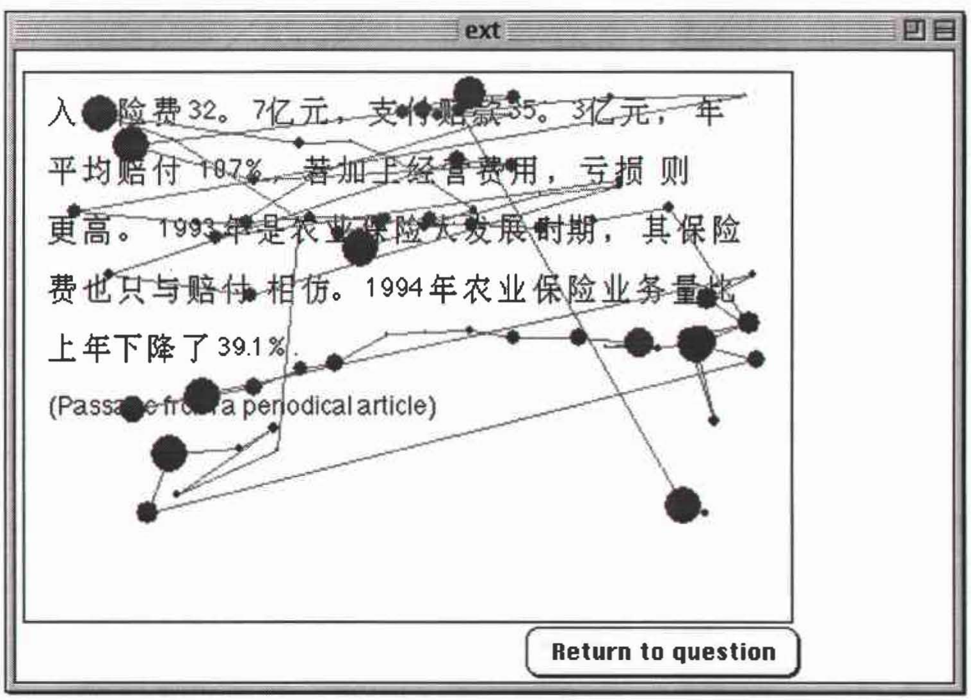


Figure 135d. NNS03-Item 25 (sd22, card 2).

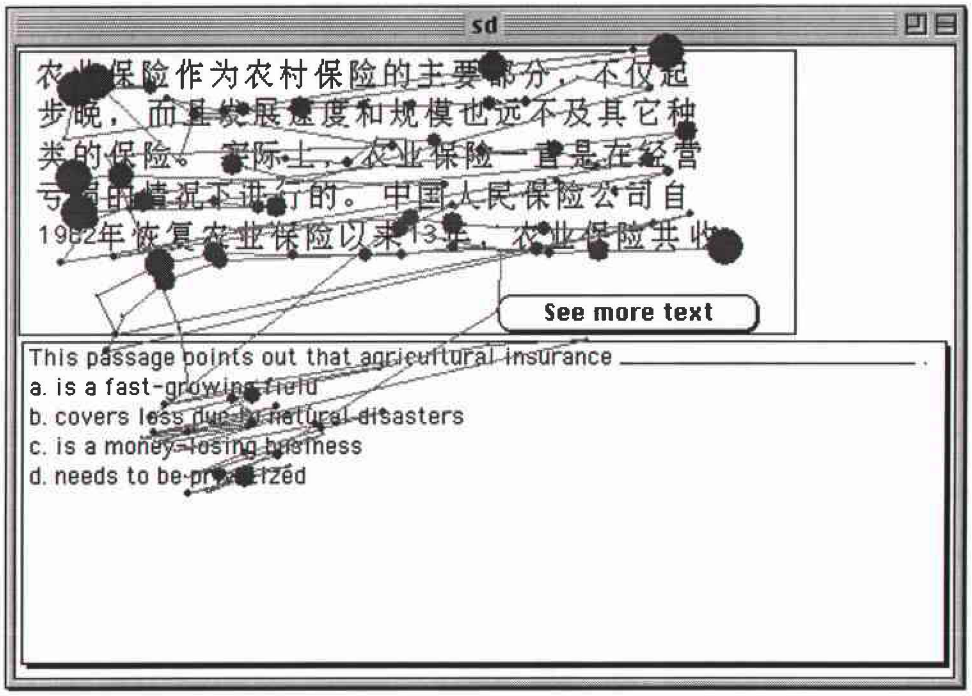


Figure 135e. NNS03-Item 25 (sd22, card 1).

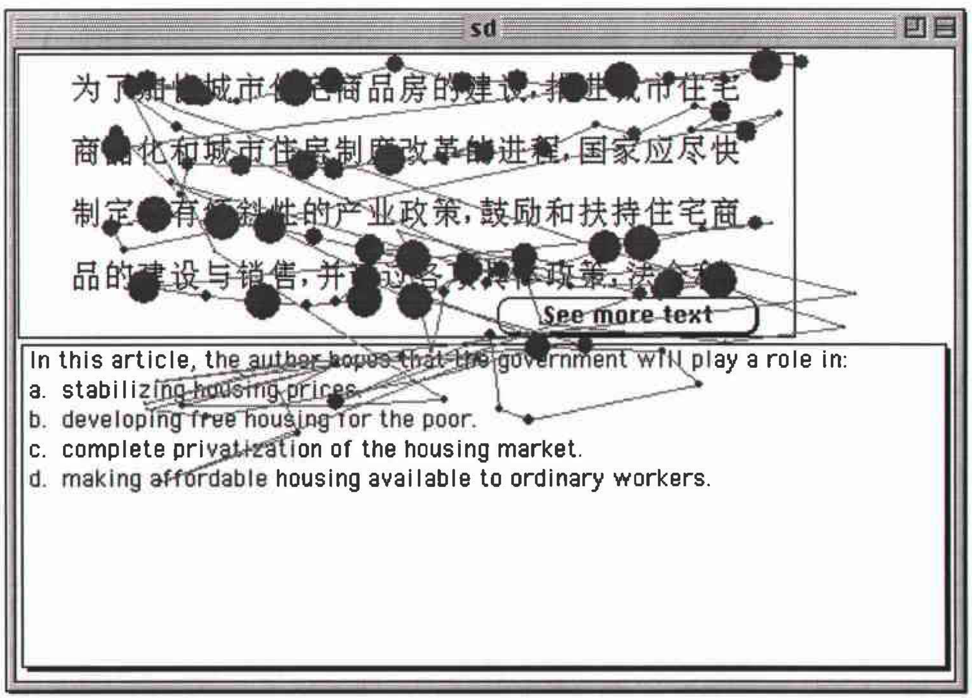


Figure 136a. NNS03-Item 26 (sd12, card 1).

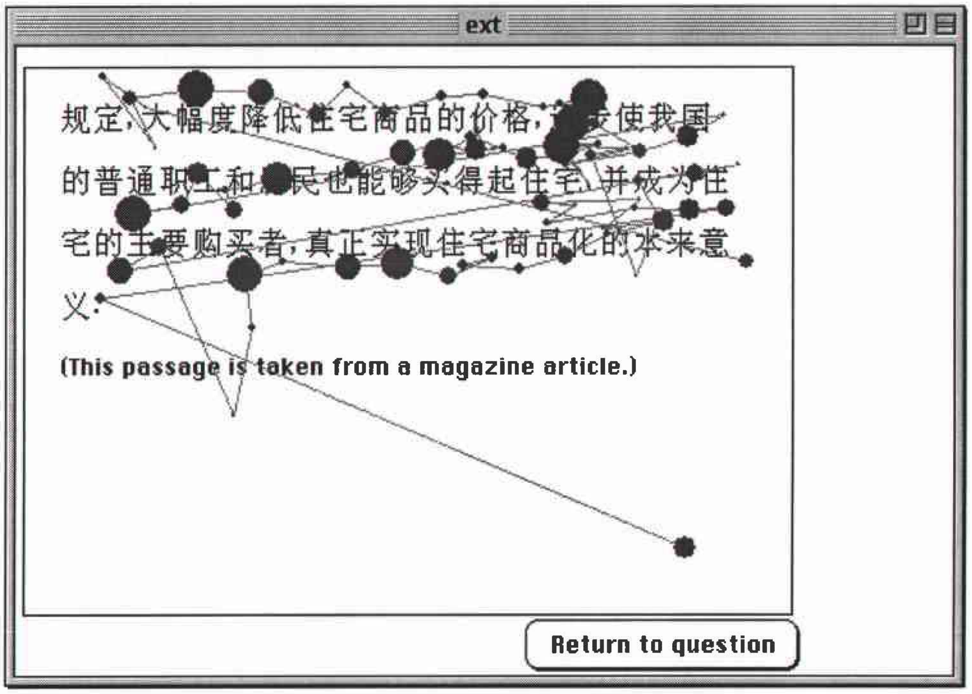


Figure 136b. NNS03-Item 26 (sd12, card 2).

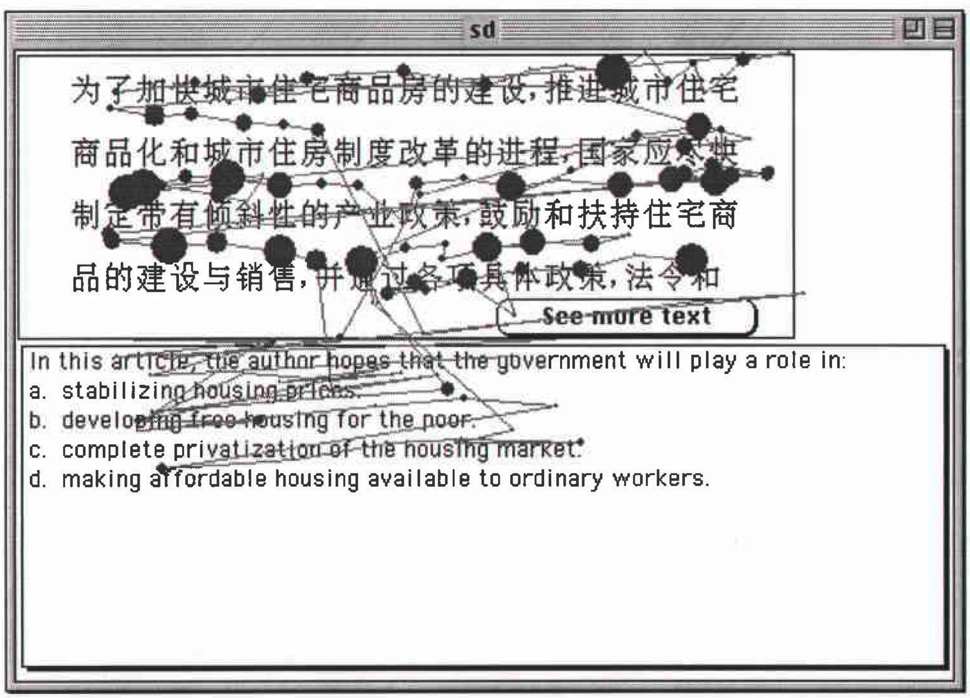


Figure 136c. NNS03-Item 26 (sd12, card 1).

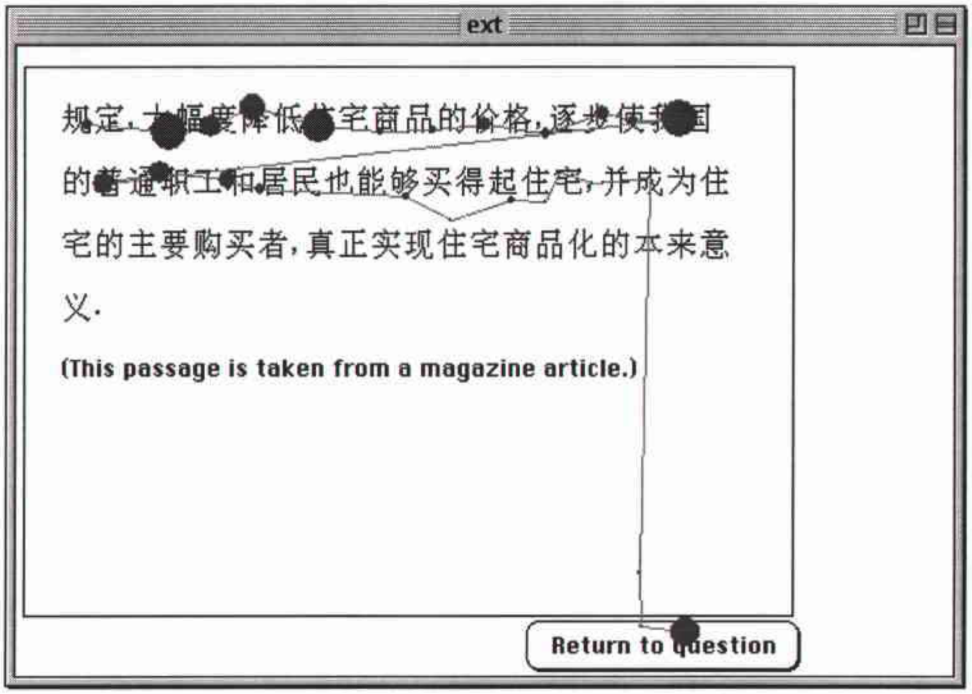


Figure 136d. NNS03-Item 26 (sd12, card 2).

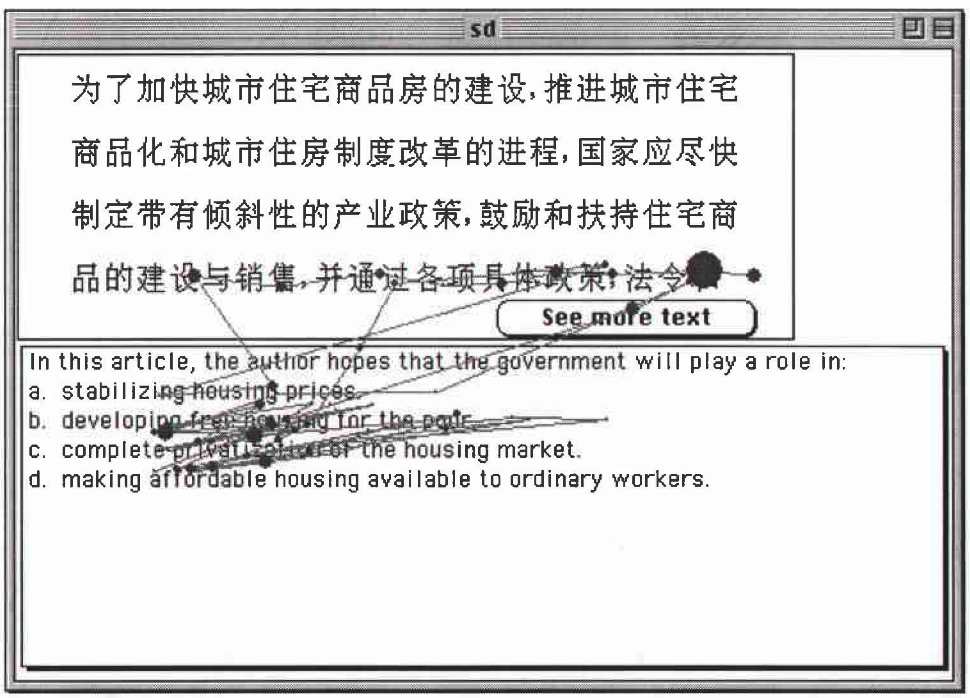


Figure 136e. NNS03-Item 26 (sd12, card 1).

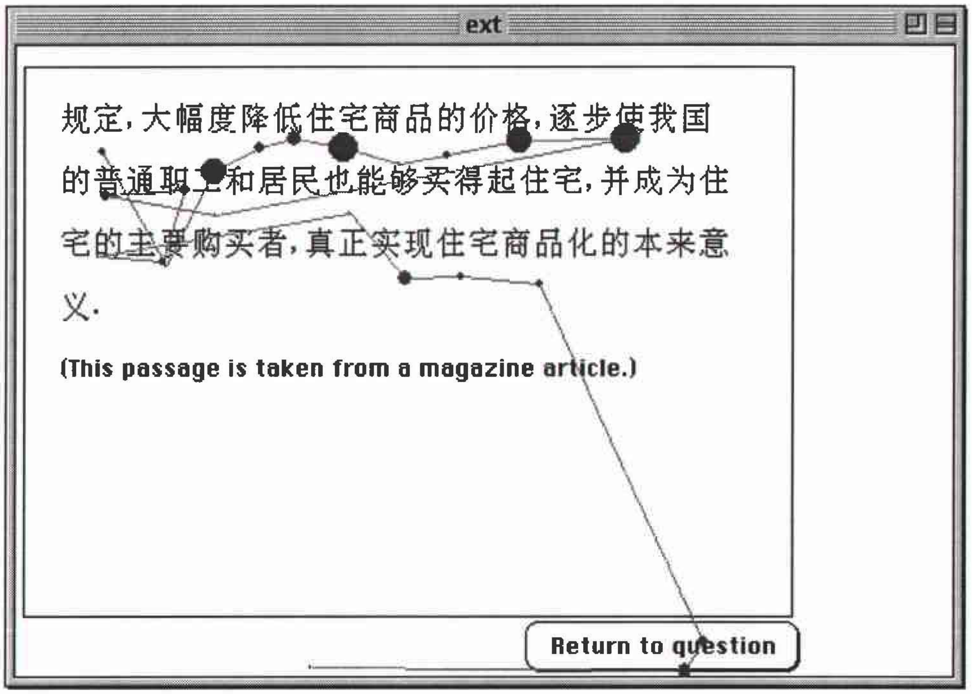


Figure 136f. NNS03-Item 26 (sd12, card 2).

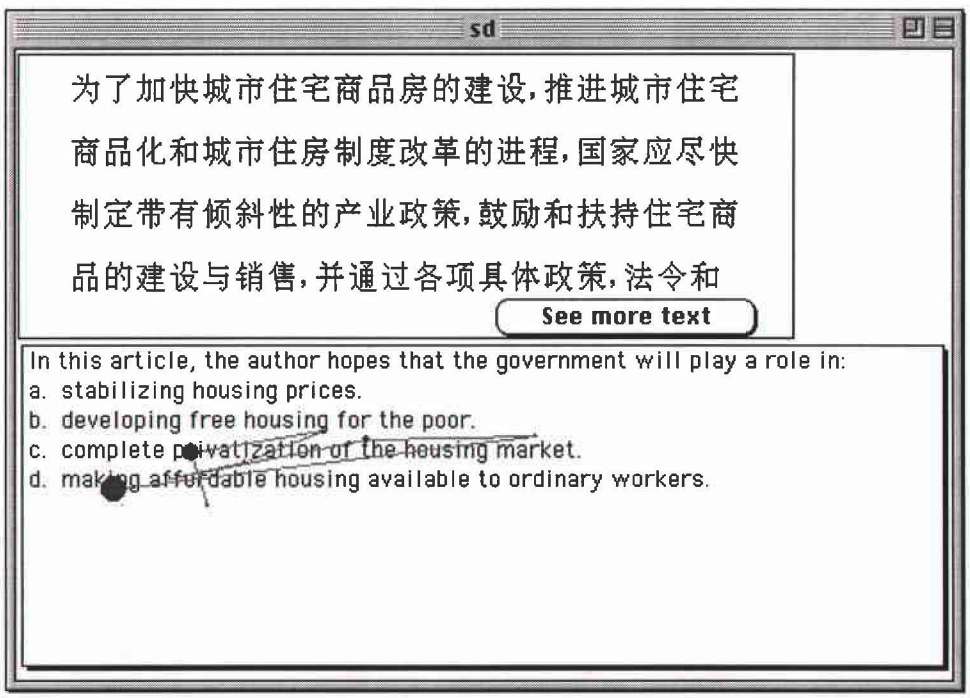


Figure 136g. NNS03-Item 26 (sd12, card 1).

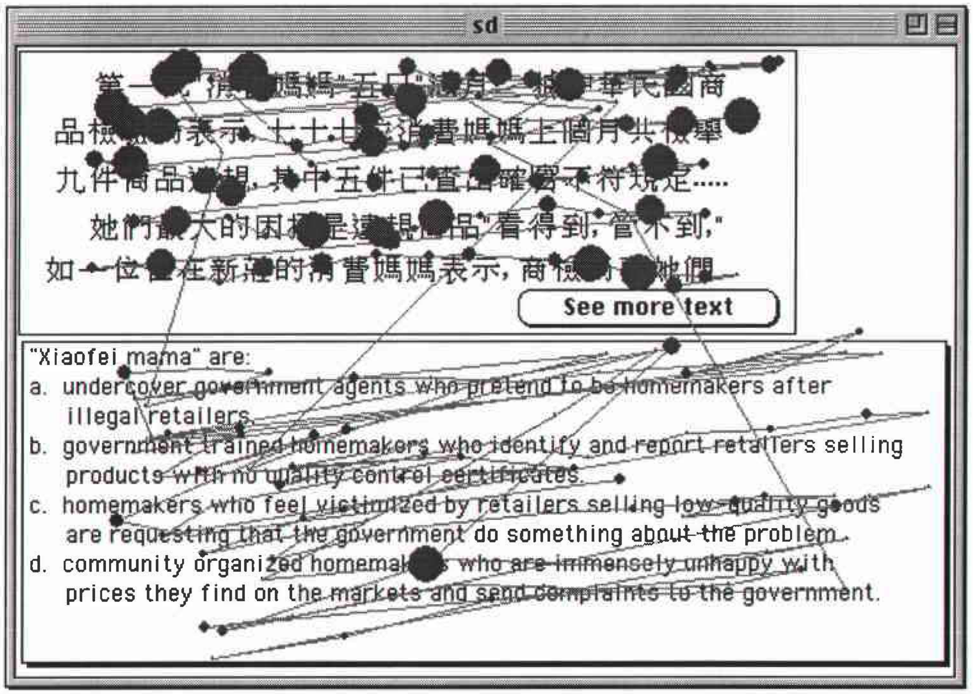


Figure 137a. NNS03-Item 27 (sd01, card 1).

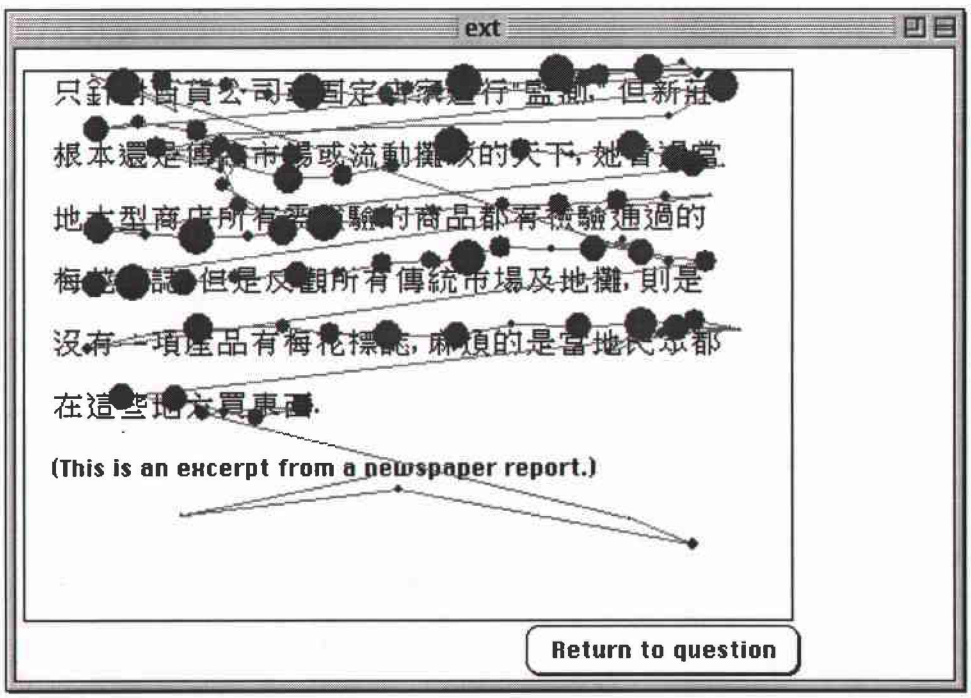


Figure 137b. NNS03-Item 27 (sd01, card 2).

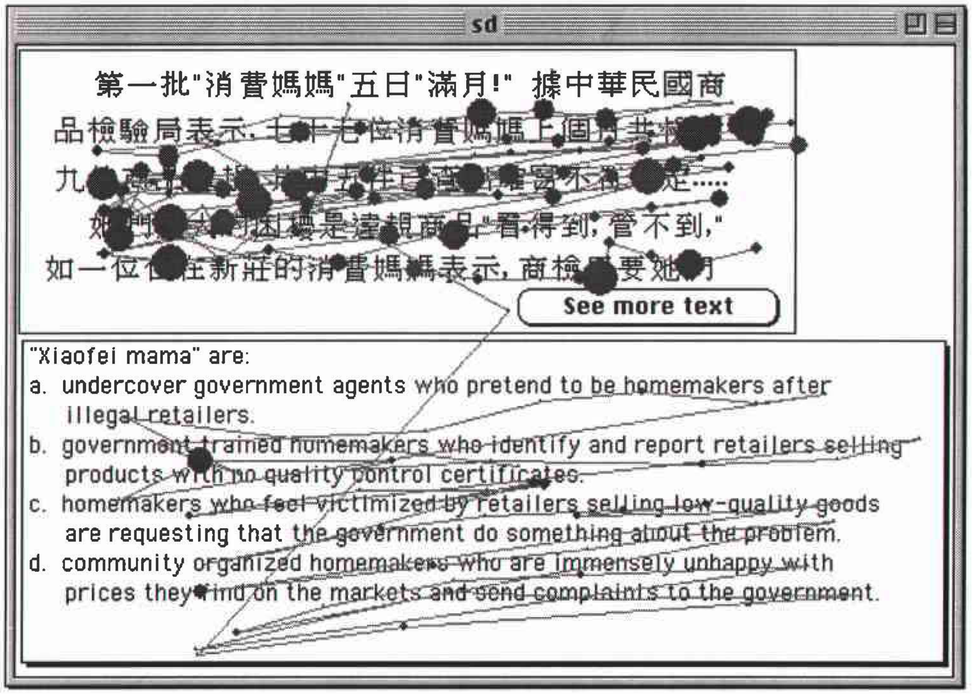


Figure 137c. NNS03-Item 27 (sd01, card 1).

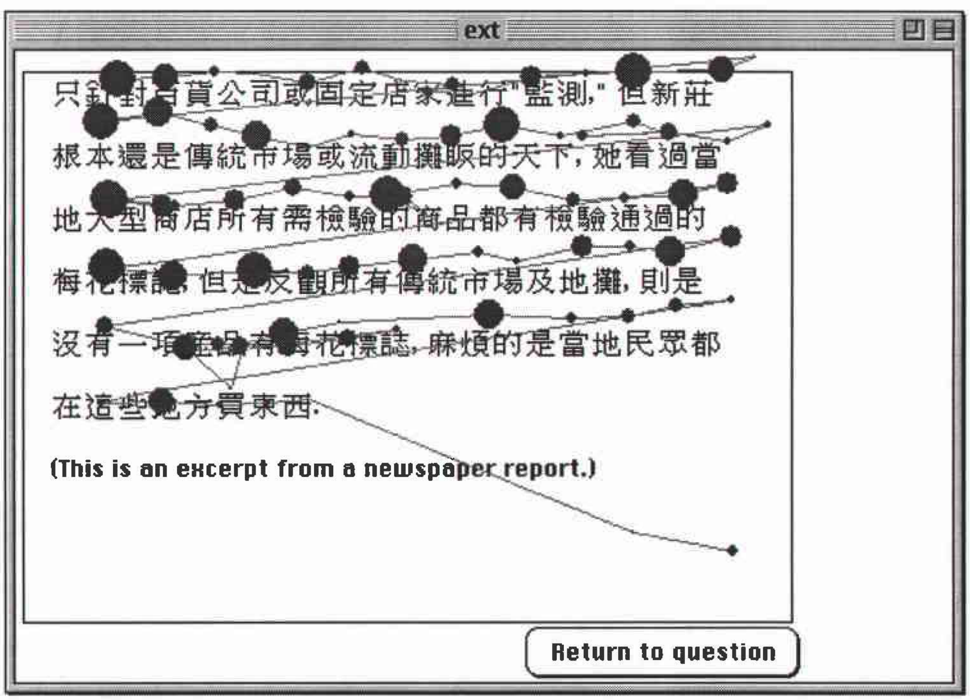


Figure 137d. NNS03-Item 27 (sd01, card 2).

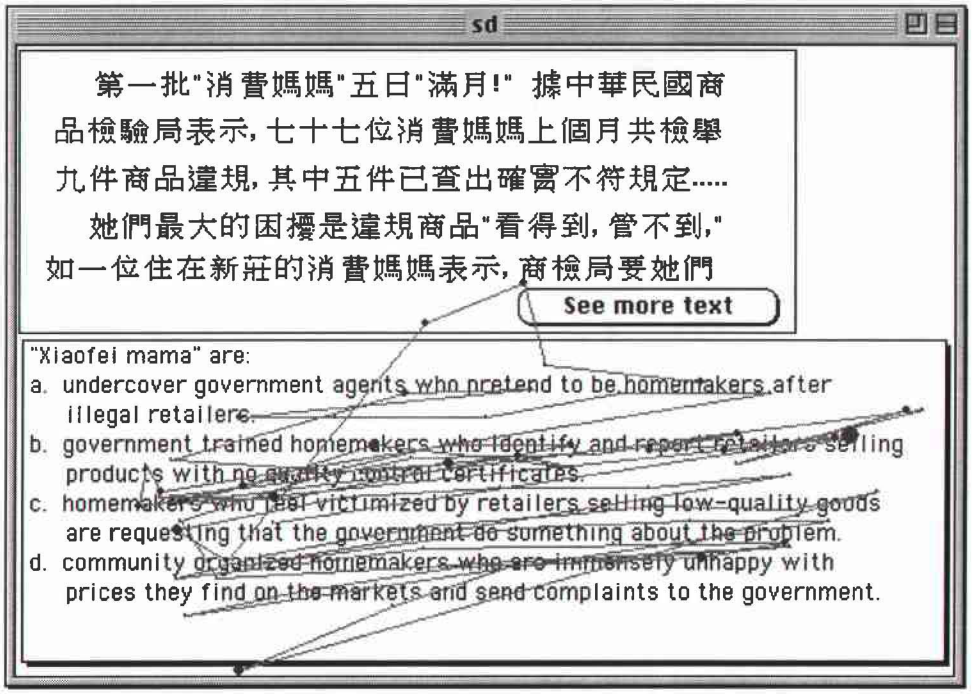


Figure 137e. NNS03-Item 27 (sd01, card 1).

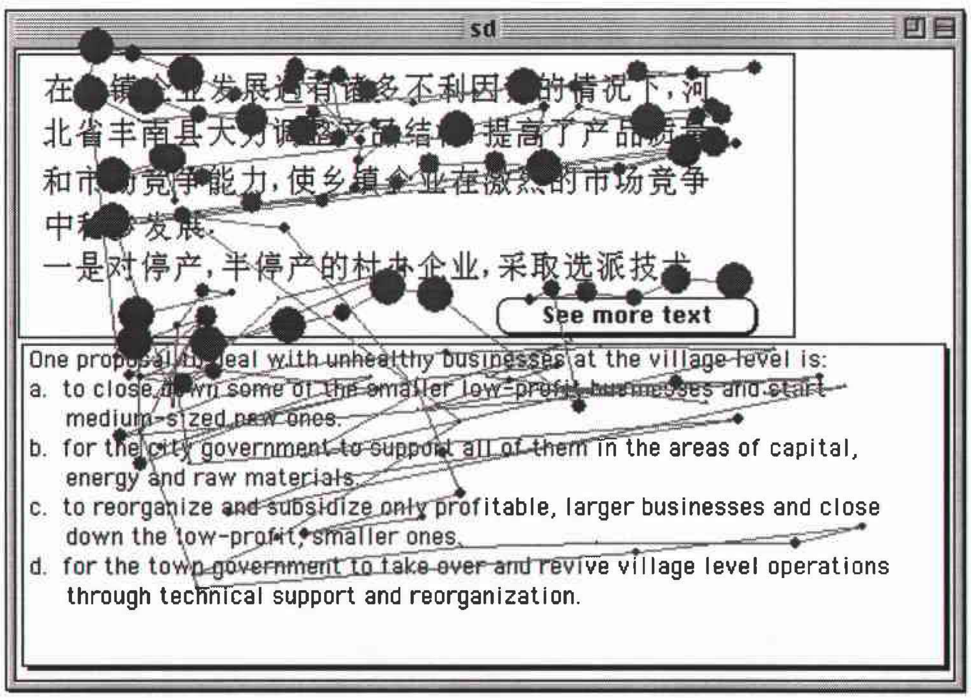


Figure 138a. NNS03-Item 28 (sd07, card 1).

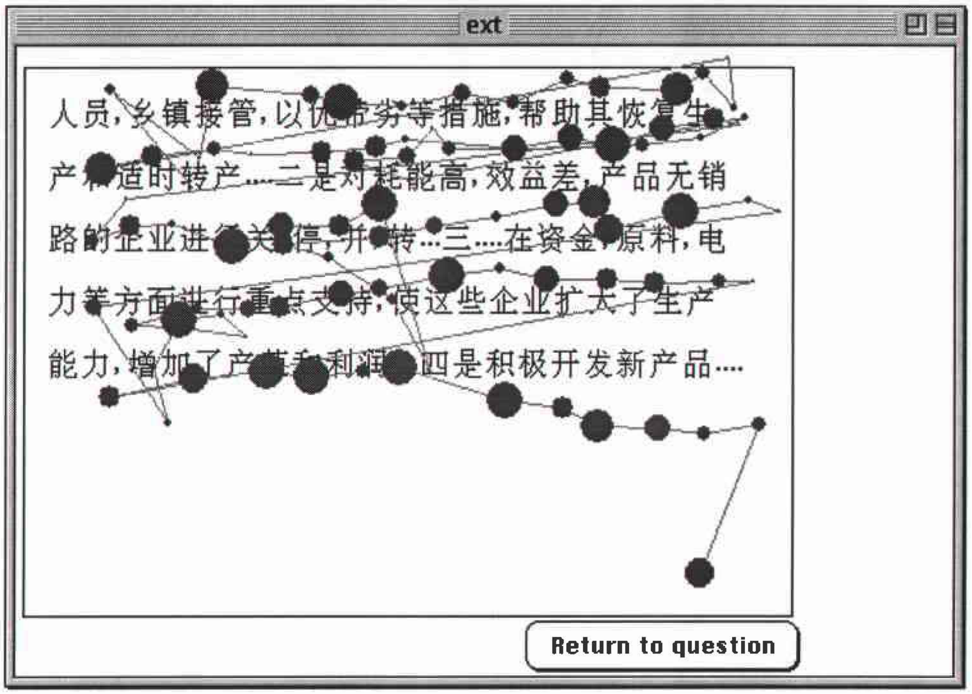


Figure 138b. NNS03-Item 28 (sd07, card 2).

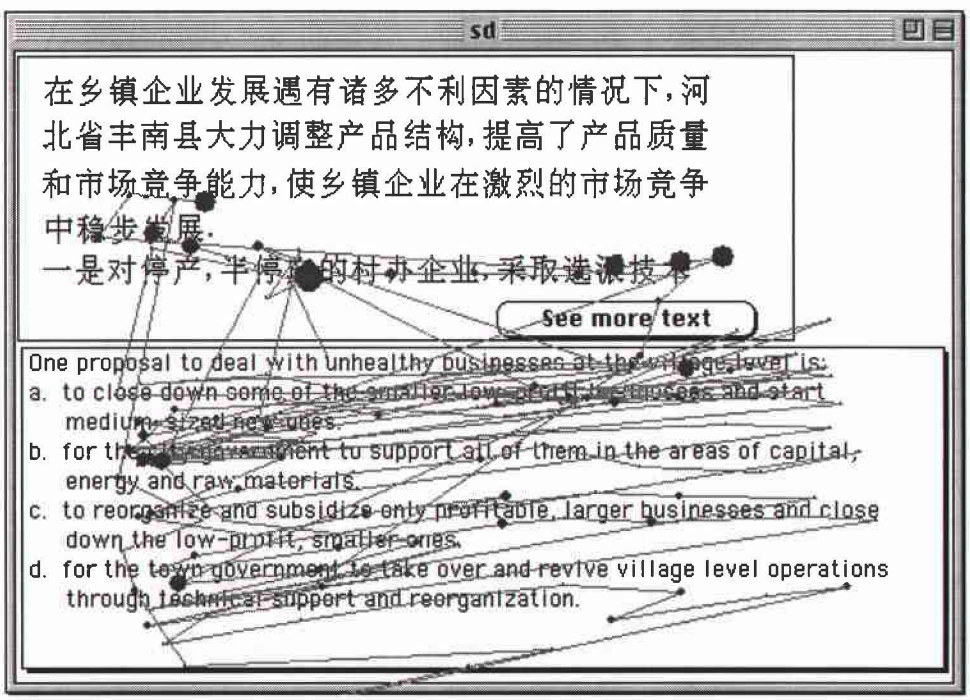


Figure 138c. NNS03-Item 28 (sd07, card 1).

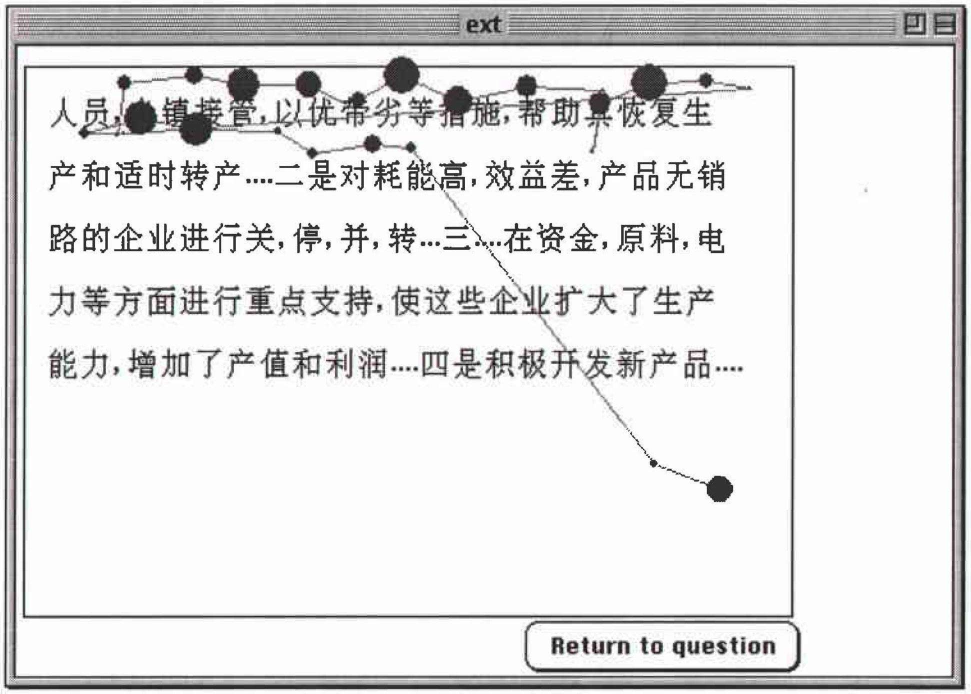


Figure 138d. NNS03-Item 28 (sd07, card 2).

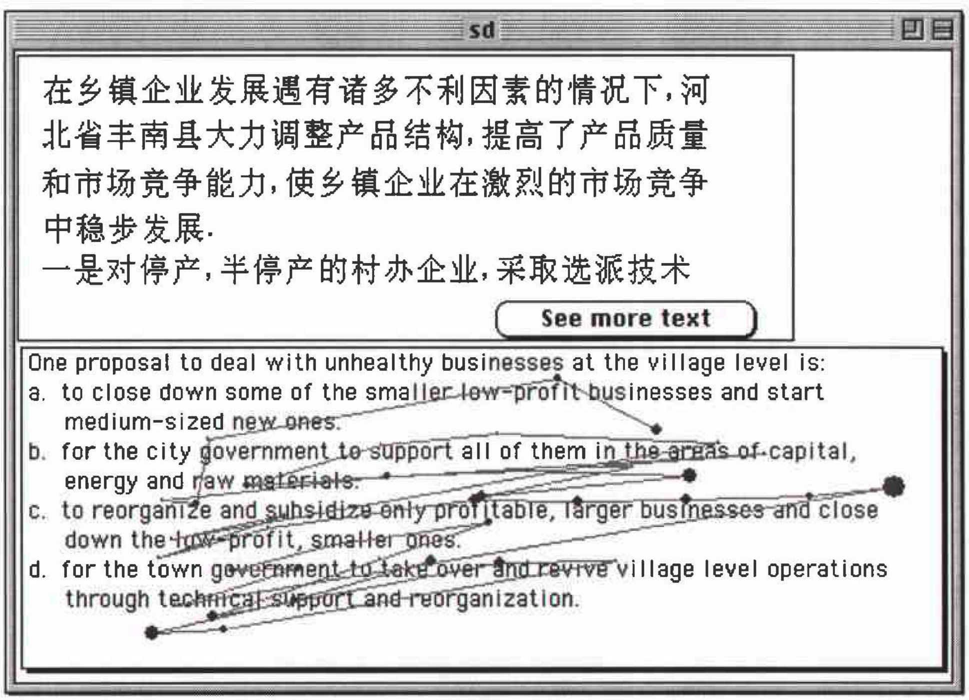


Figure 138e. NNS03-Item 28 (sd07, card 1).

APPENDIX K

EYE-TRACKING SAMPLE SET: NATIVE SUPERIOR (NS07)

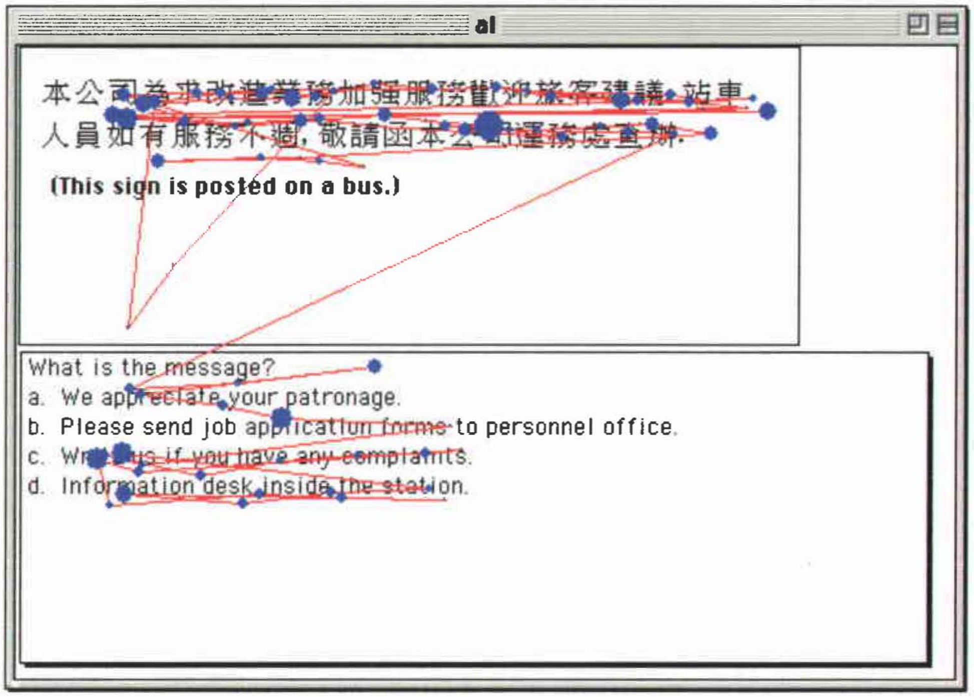


Figure 139. NS07-Item 1 (al13).

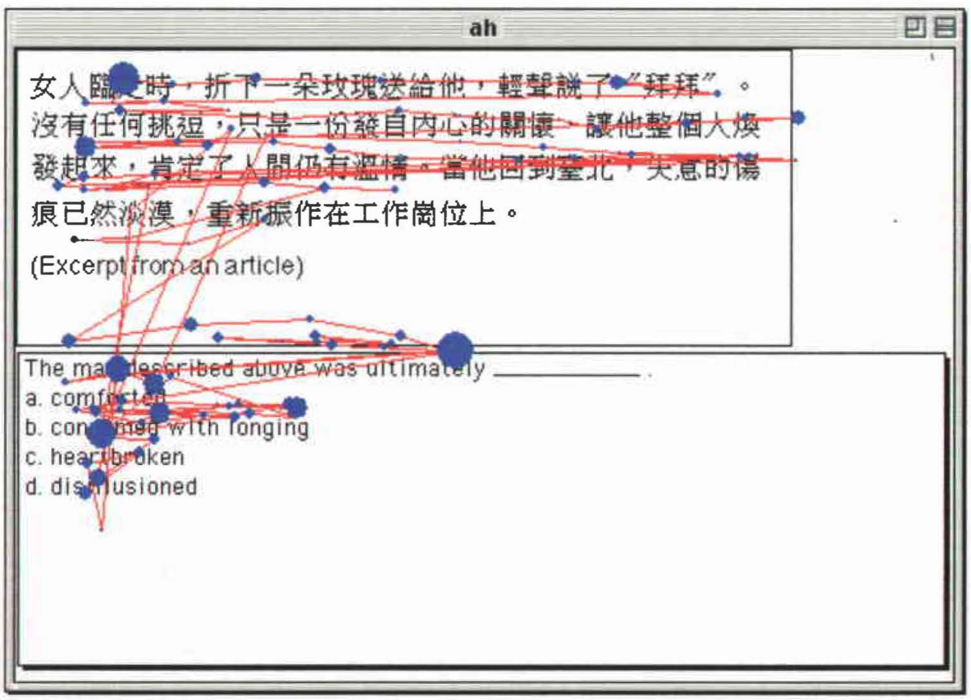


Figure 140. NS07-Item 2 (ah51).

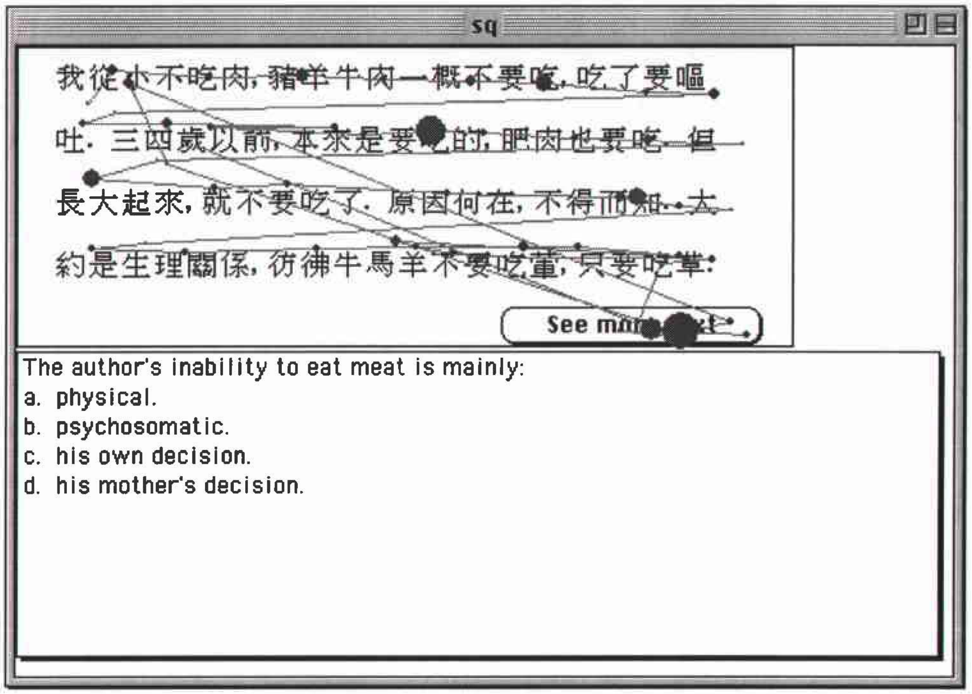


Figure 141a. NS07-Item 3 (sq08, card 1).

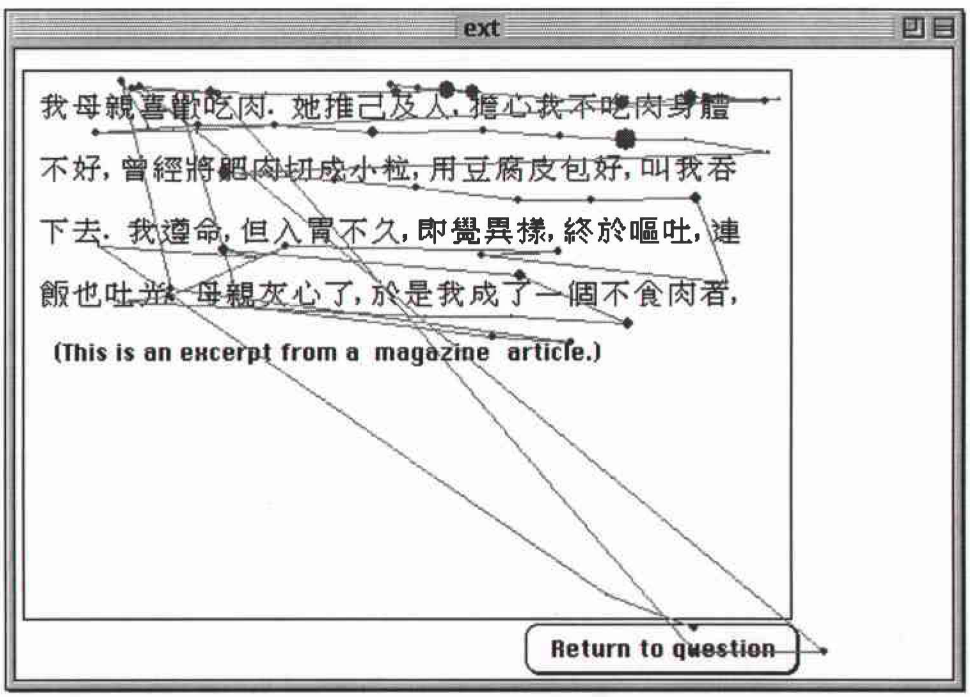


Figure 141b. NS07-Item 3 (sq08, card 2).

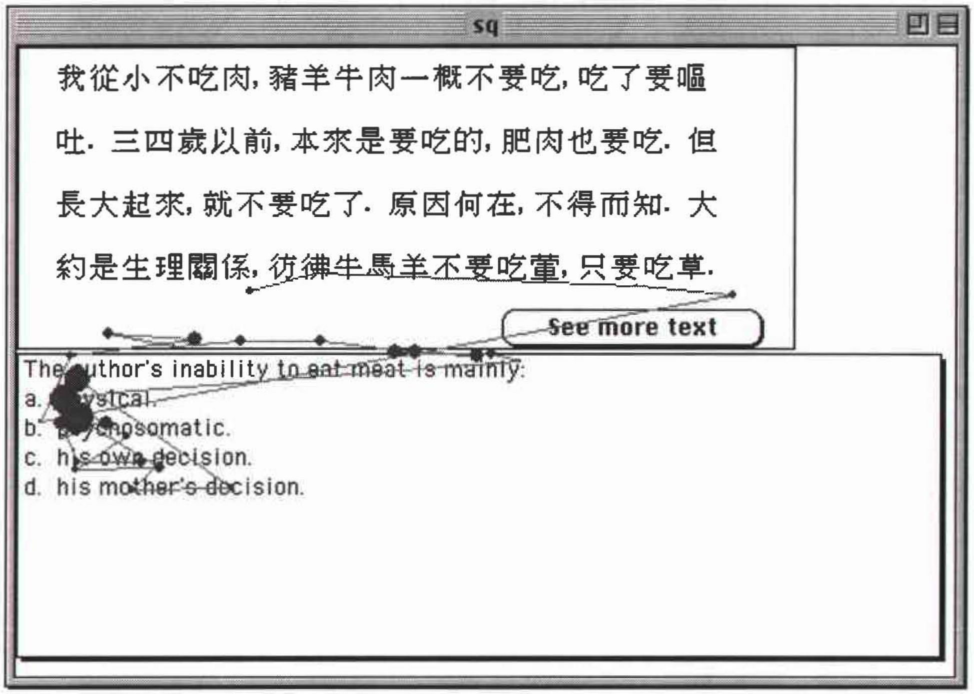


Figure 141c. NS07-Item 3 (sq08, card 1).

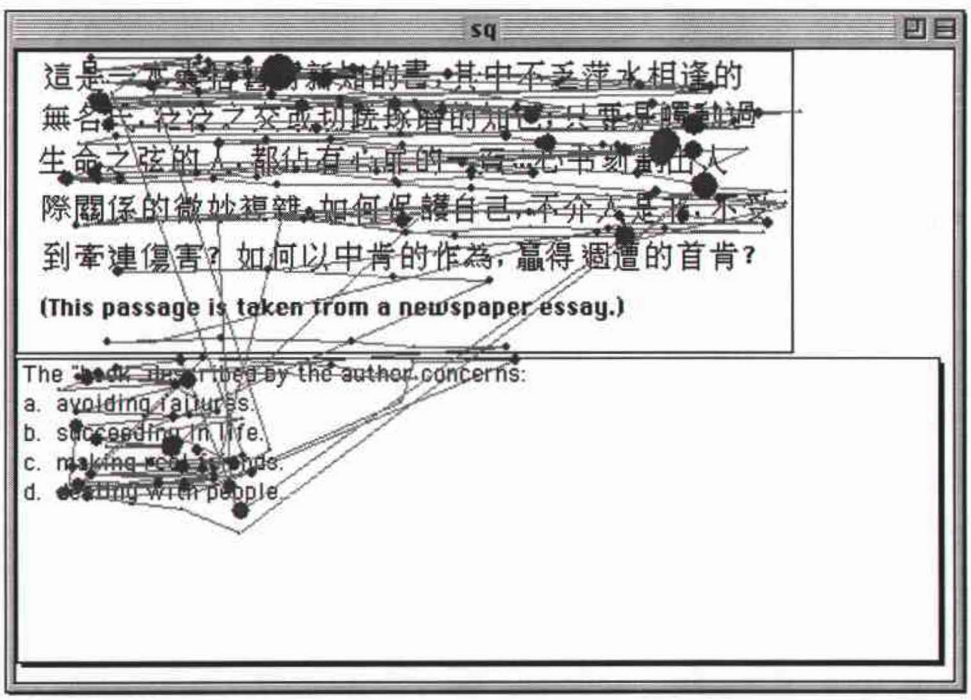


Figure 142. NS07-Item 4 (sq18).

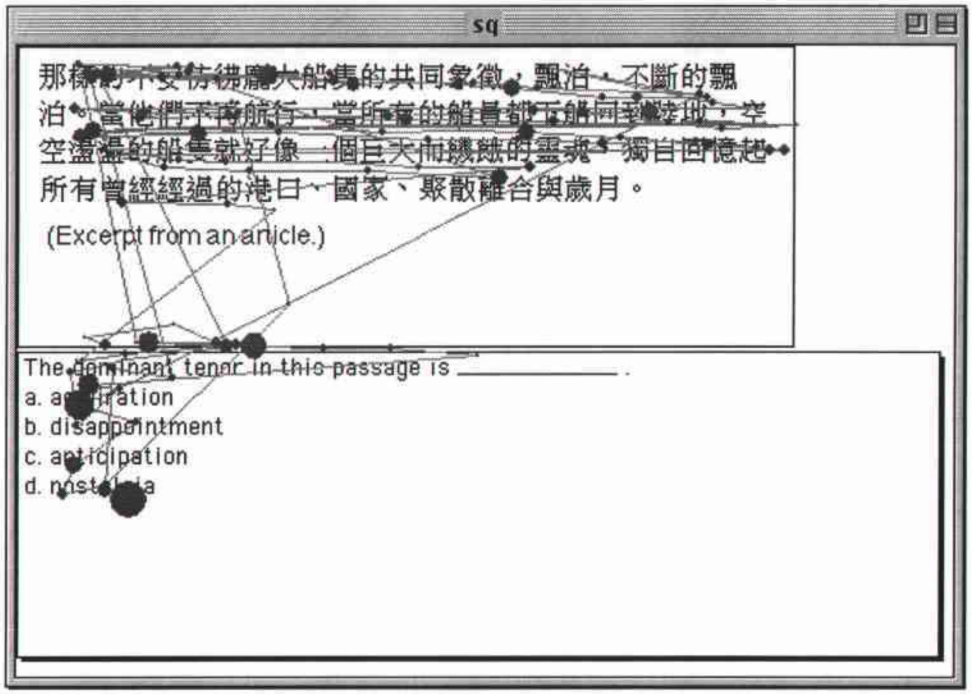


Figure 143. NS07-Item 5 (sq29).

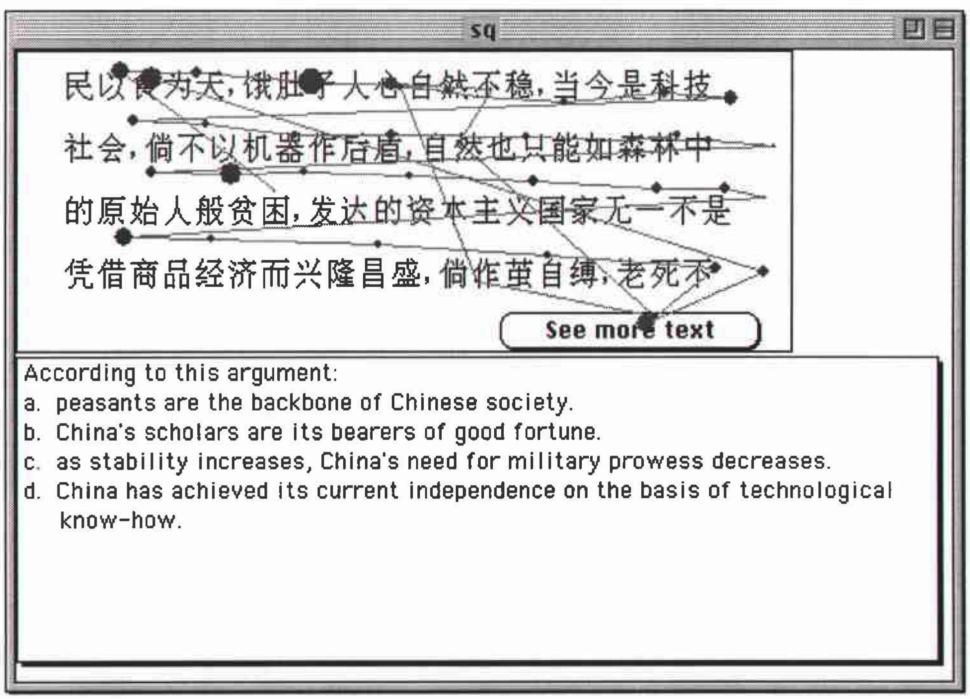


Figure 144a. NS07-Item 6 (sq16, card 1).

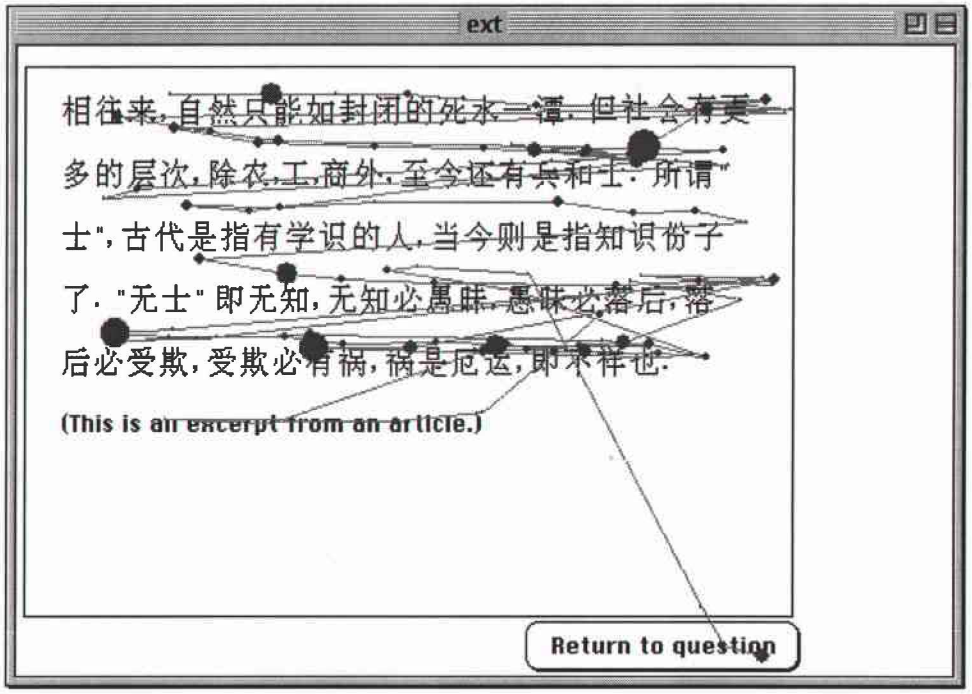


Figure 144b. NS07-Item 6 (sq16, card 2).

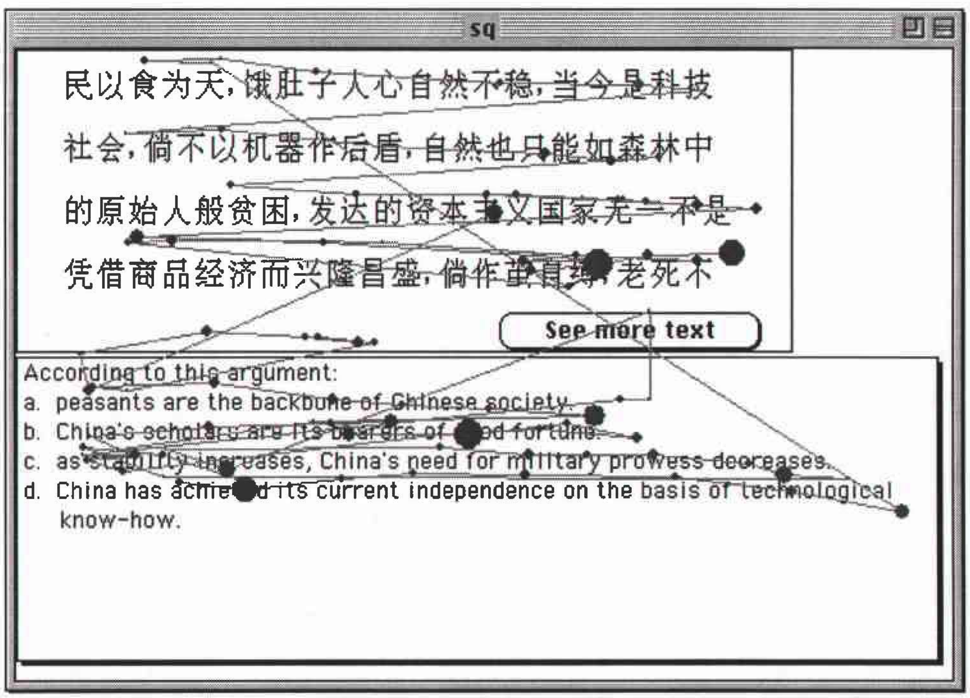


Figure 144c. NS07-Item 6 (sq16, card 1).

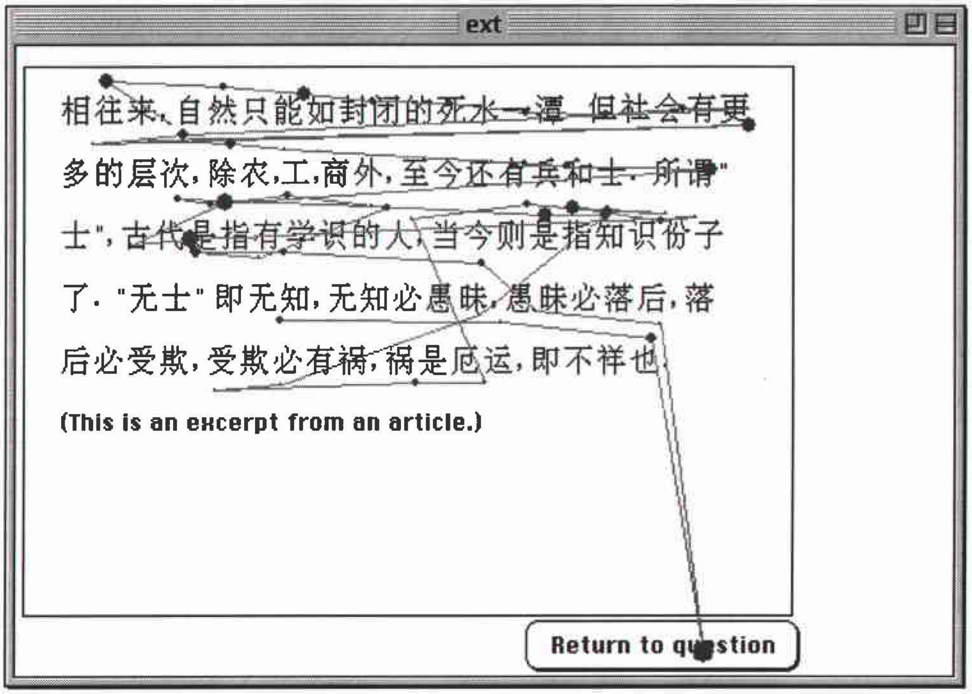


Figure 144d. NS07-Item 6 (sq16, card 2).

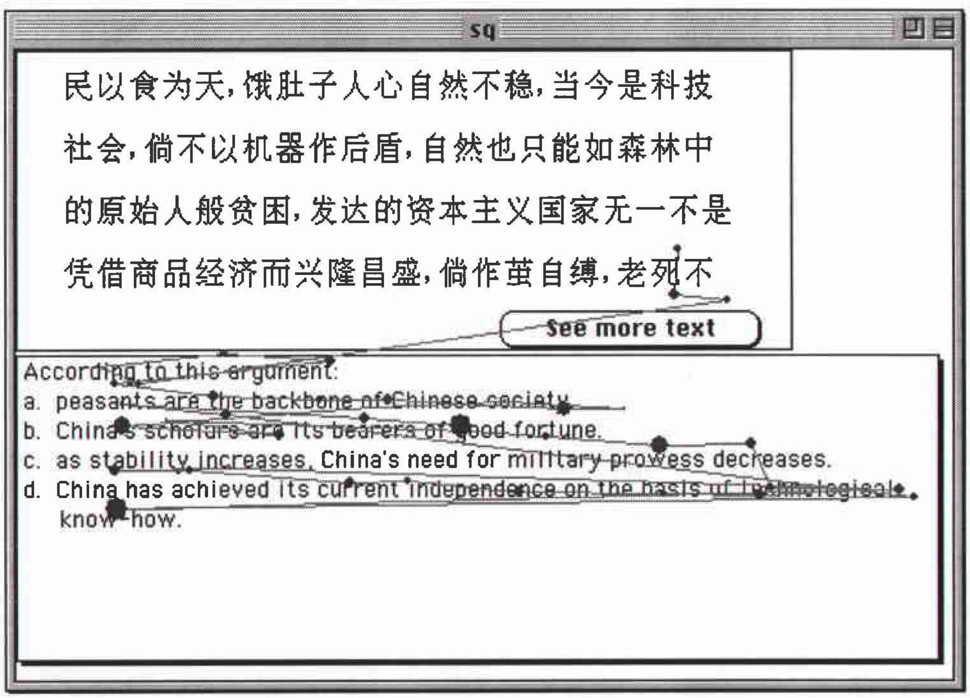


Figure 144e. NS07-Item 6 (sq16, card 1).

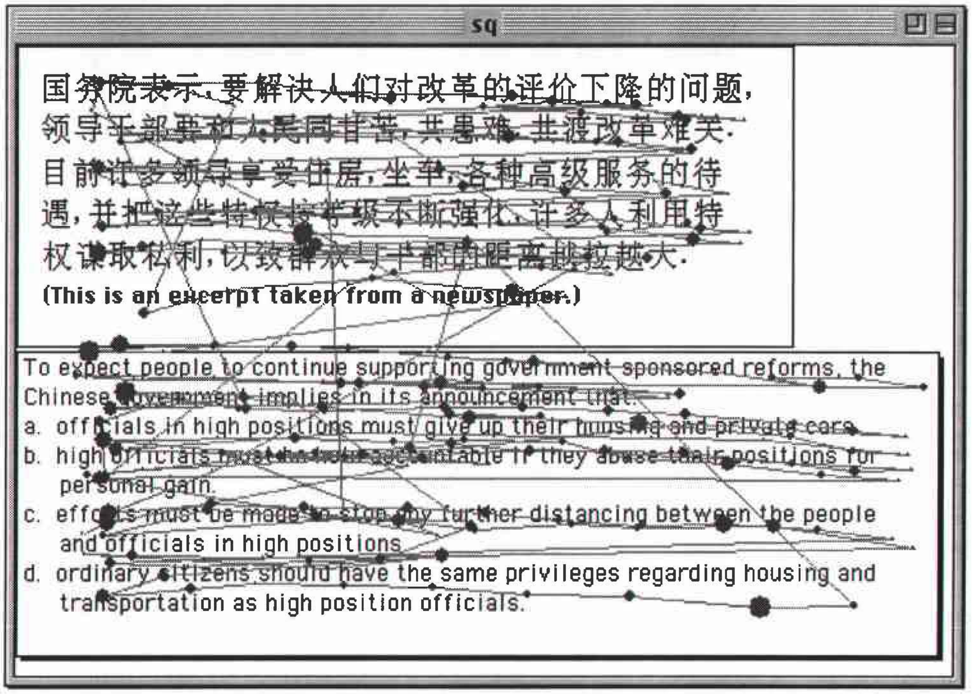


Figure 145. NS07-Item 7 (sq22).

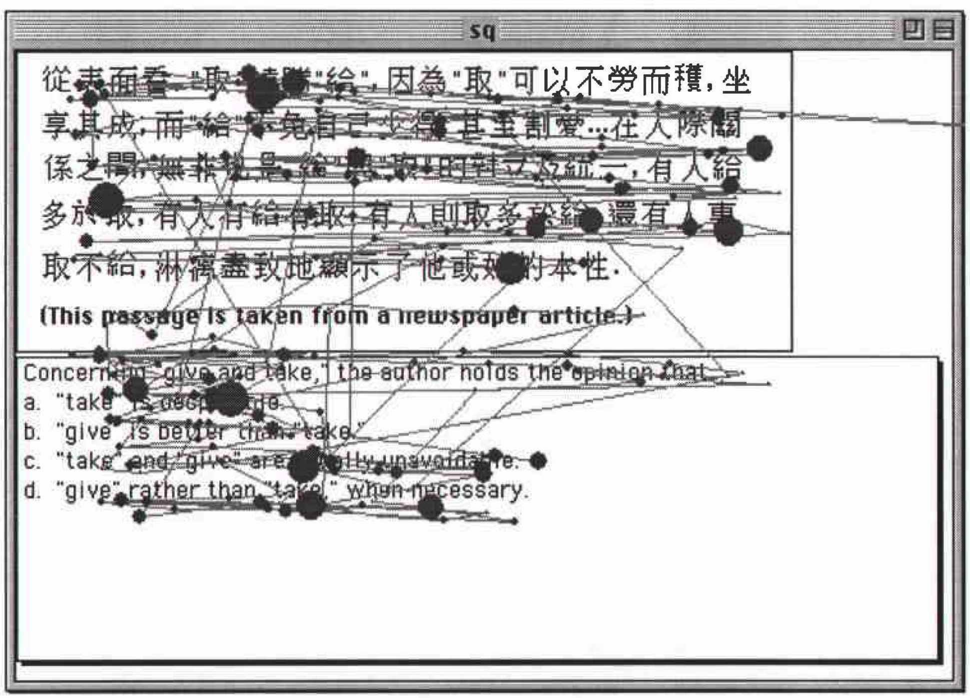


Figure 146. NS07-Item 8 (sq17).

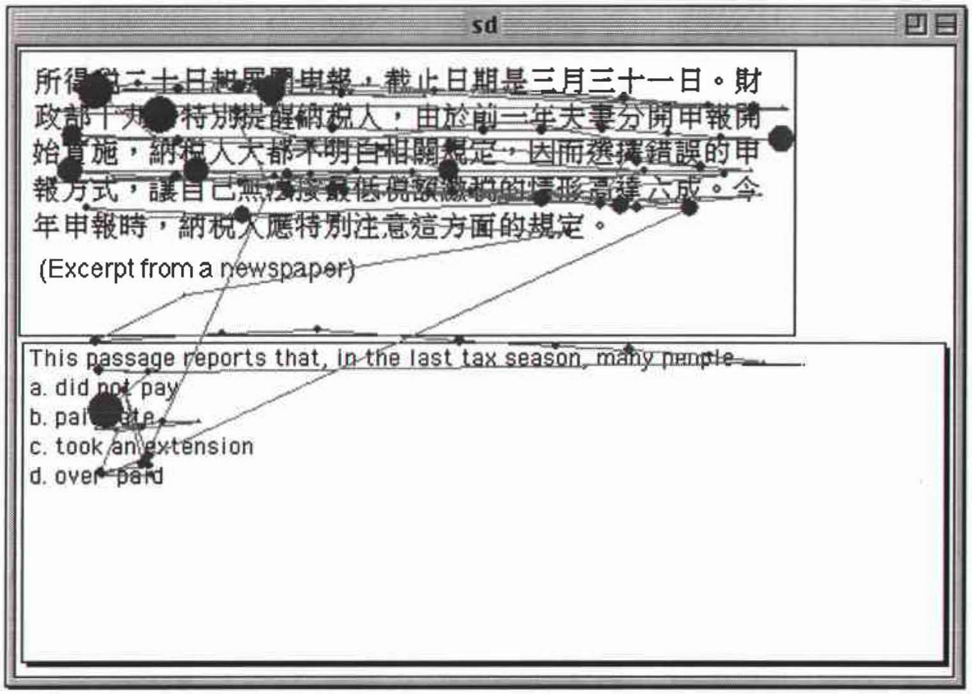


Figure 147. NS07-Item 9 (sd23).

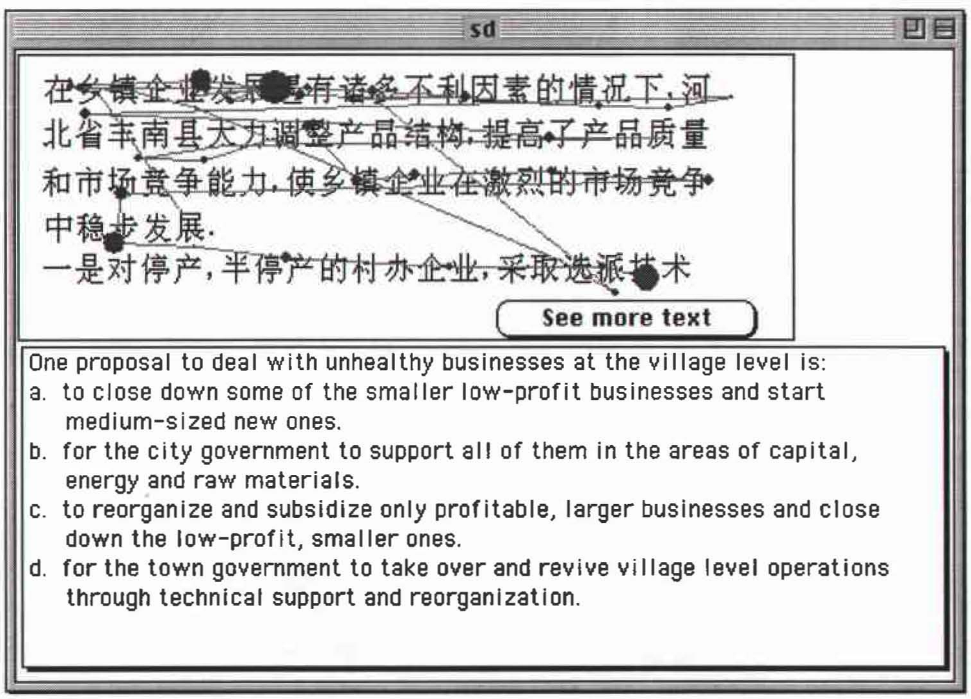


Figure 148a. NS07-Item 10 (sd07, card 1).

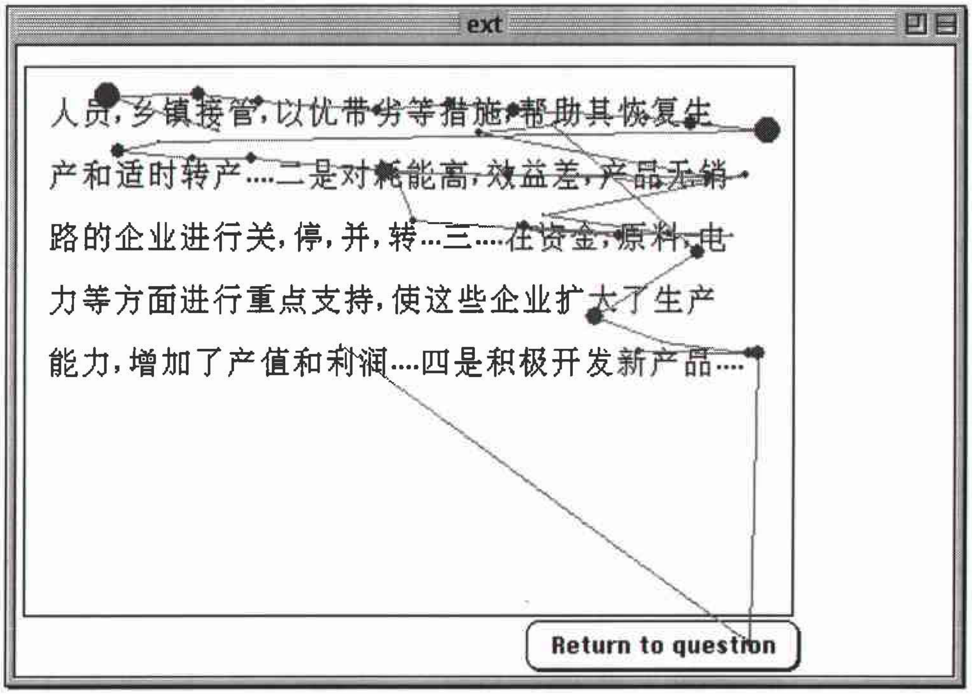


Figure 148b. NS07-Item 10 (sd07, card 2).

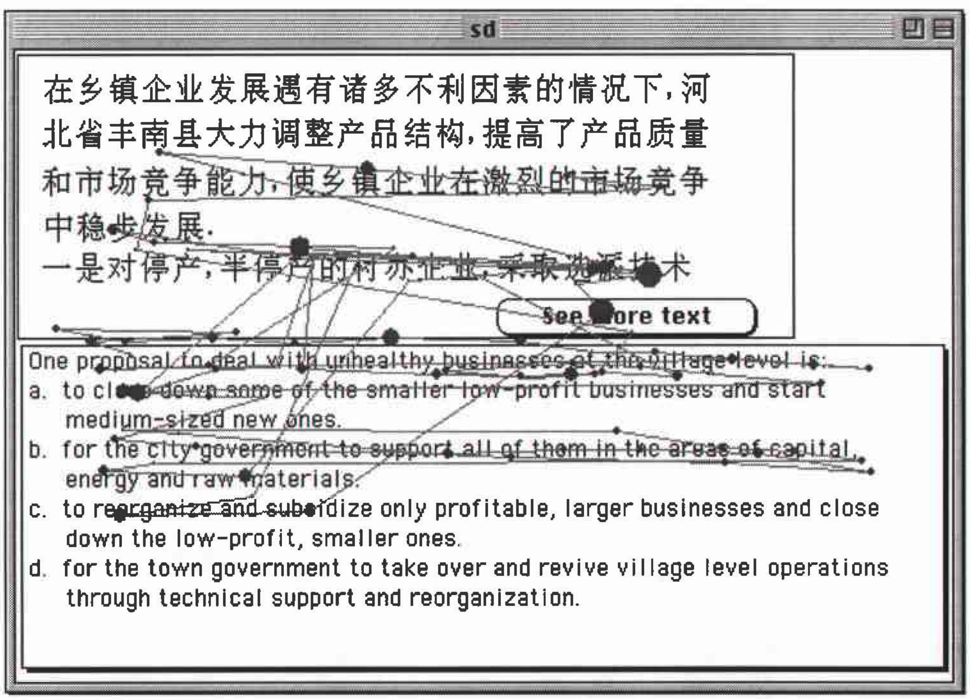


Figure 148c. NS07-Item 10 (sd07, card 1).

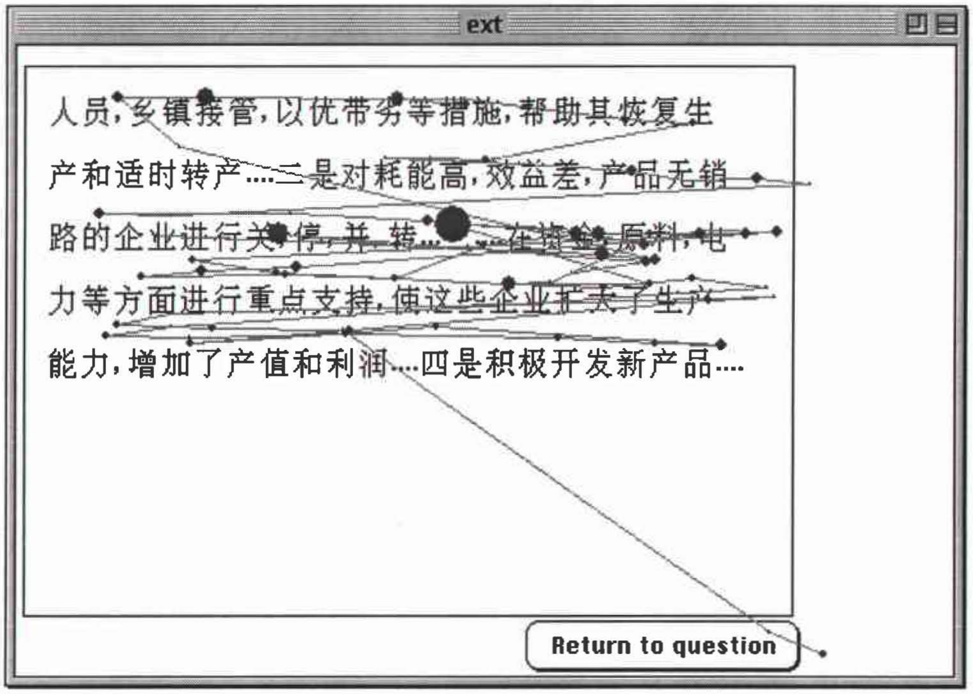


Figure 148d. NS07-Item 10 (sd07, card 2).

Figure 148e. NS07-Item 10 (sd07, card 1).

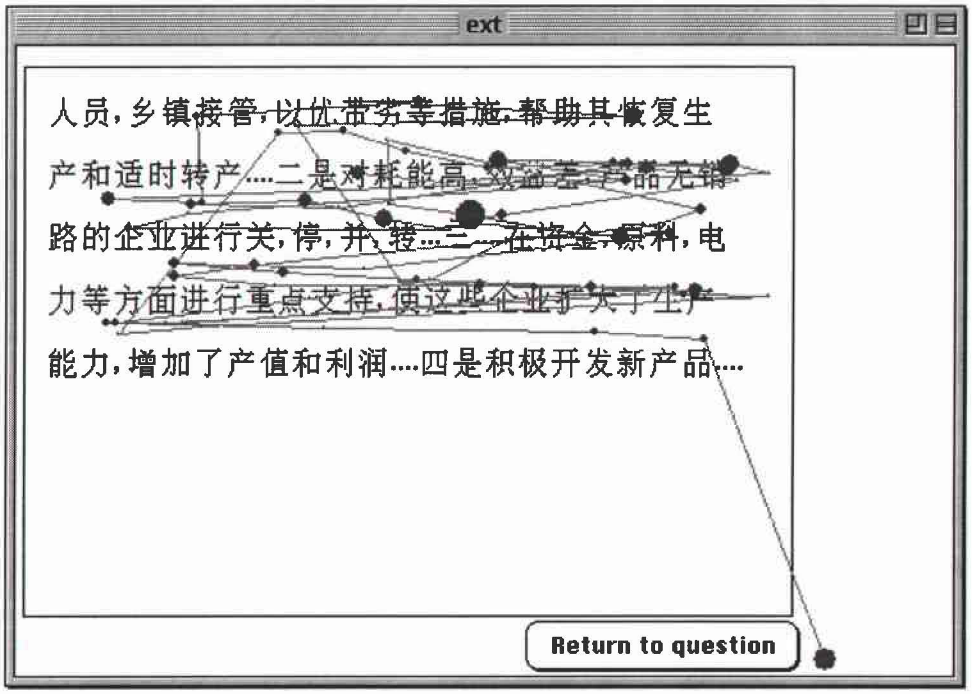


Figure 148f. NS07-Item 10 (sd07, card 2).

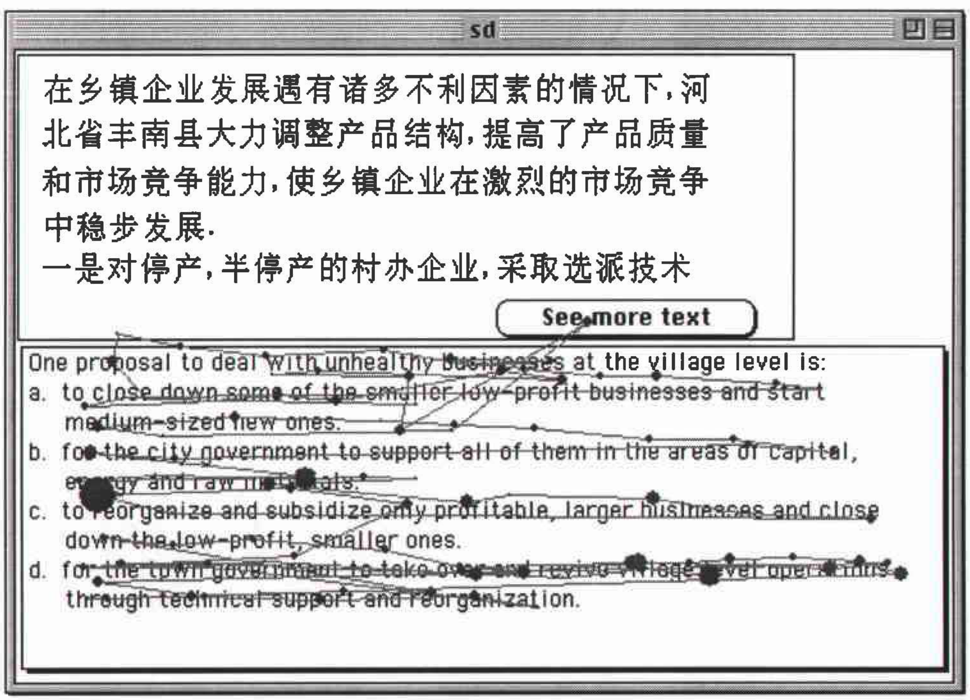


Figure 148g. NS07-Item 10 (sd07, card 1).

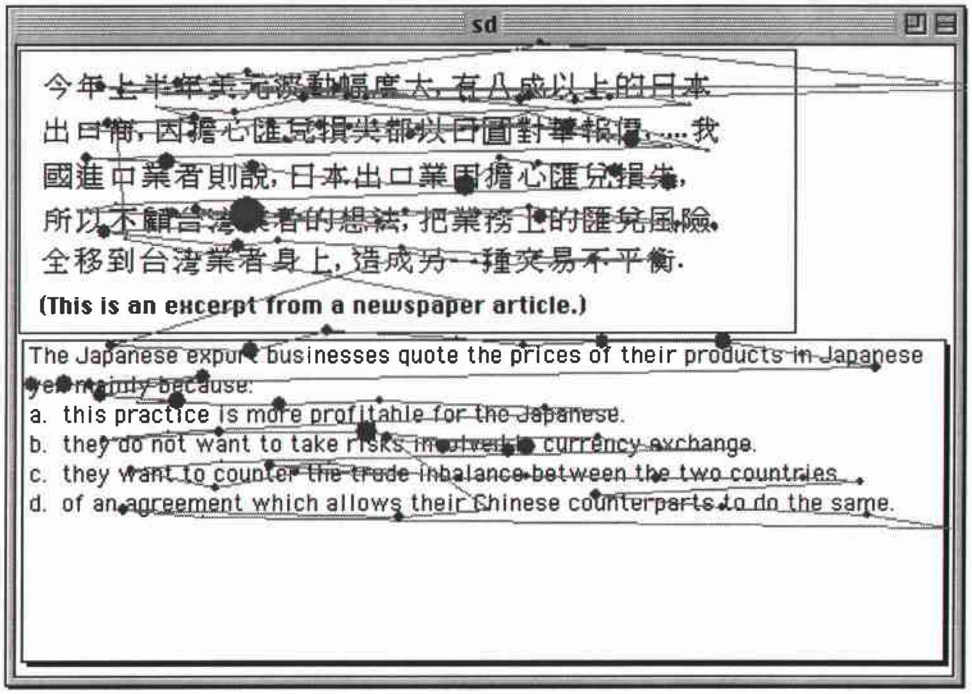


Figure 149. NS07-Item 11 (sd15).

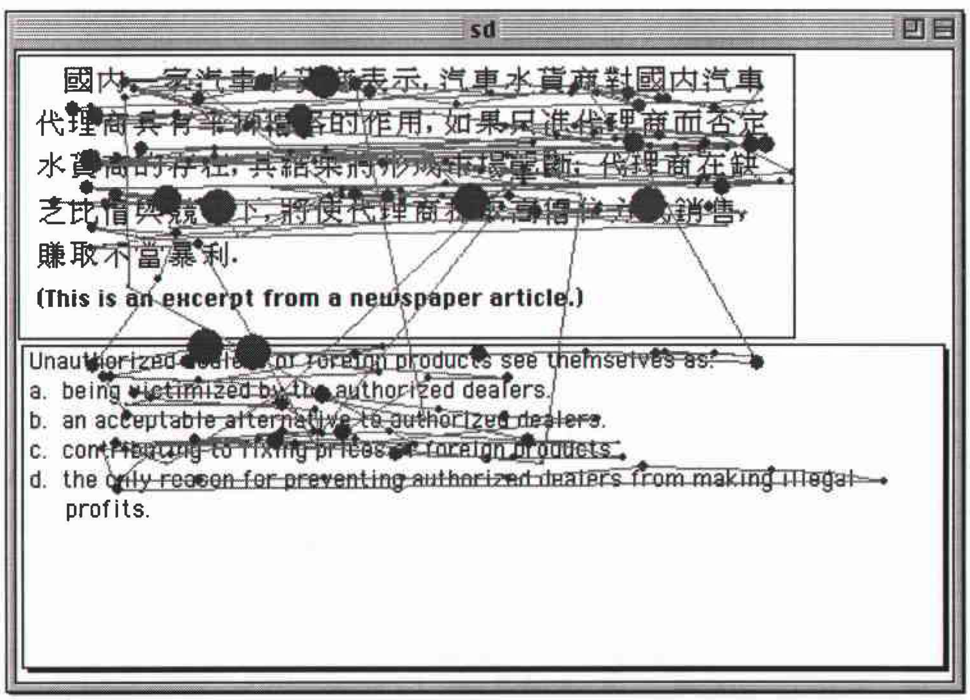


Figure 150. NS07-Item 12 (sd04).

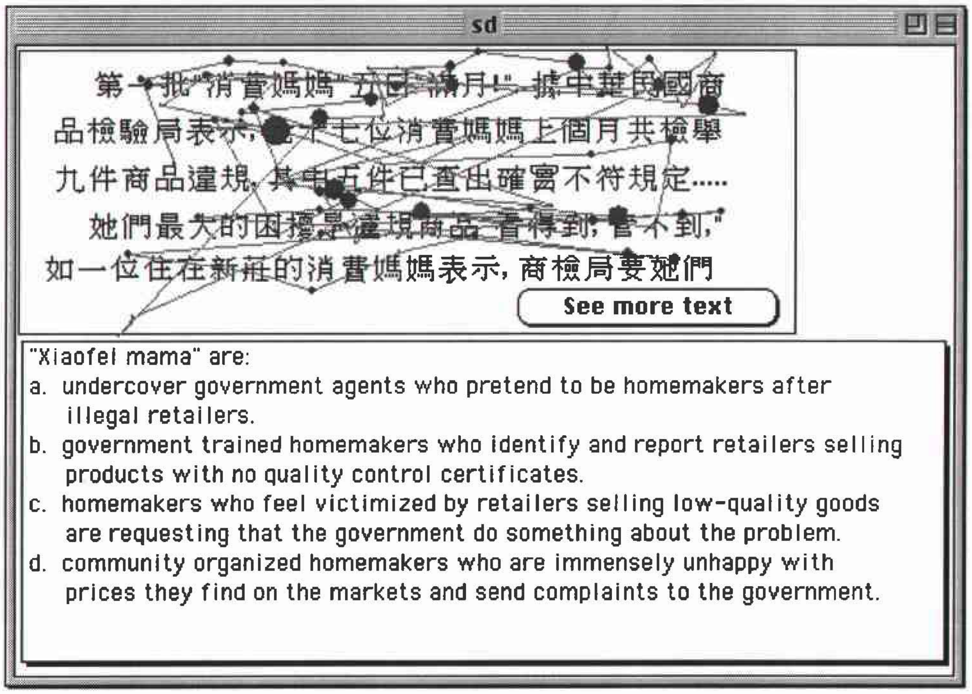


Figure 151a. NS07-Item 13 (sd01, card 1).

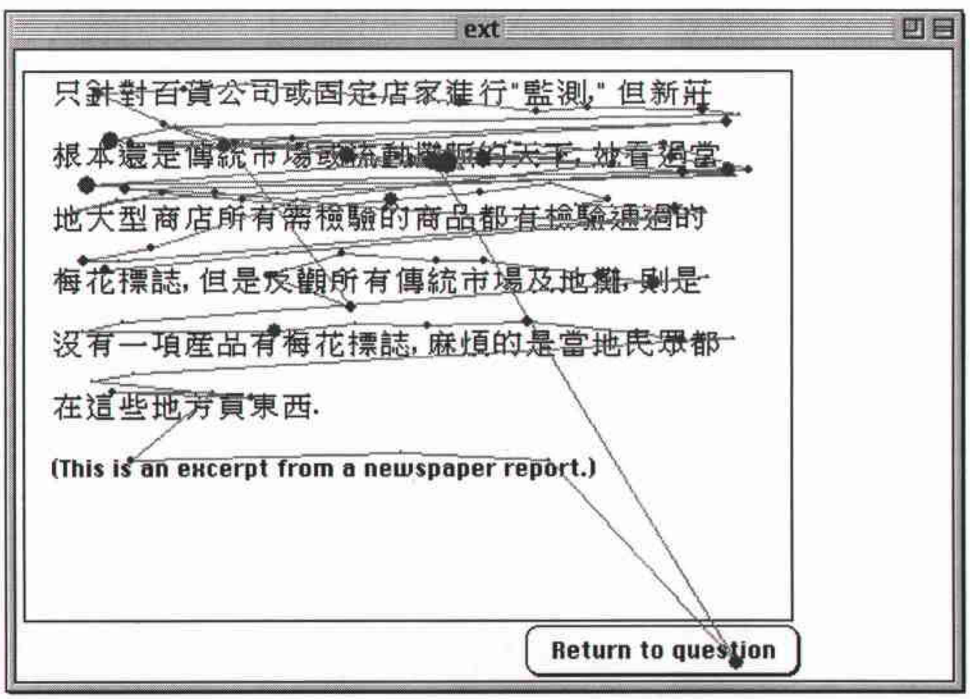


Figure 151b. NS07-Item 13 (sd01, card 2).

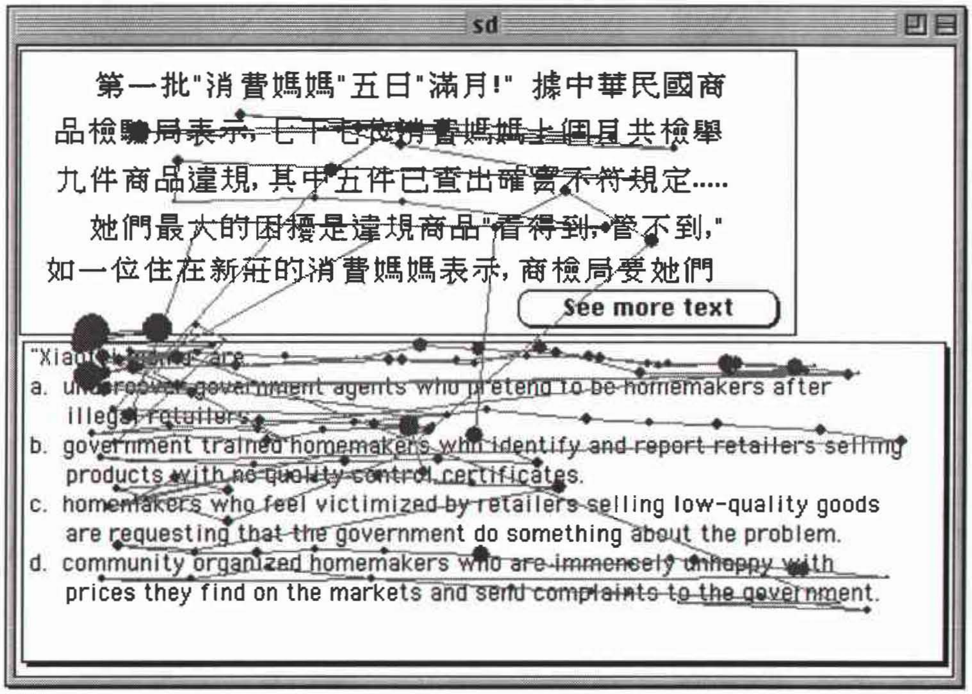


Figure 151c. NS07-Item 13 (sd01, card 1).

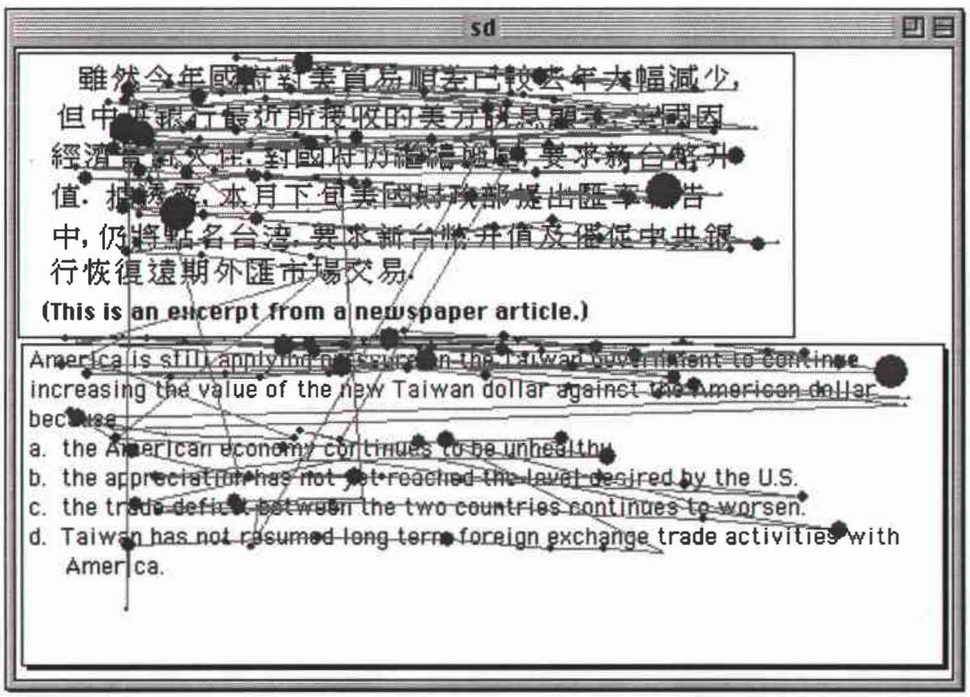


Figure 152. NS07-Item 14 (sd08).

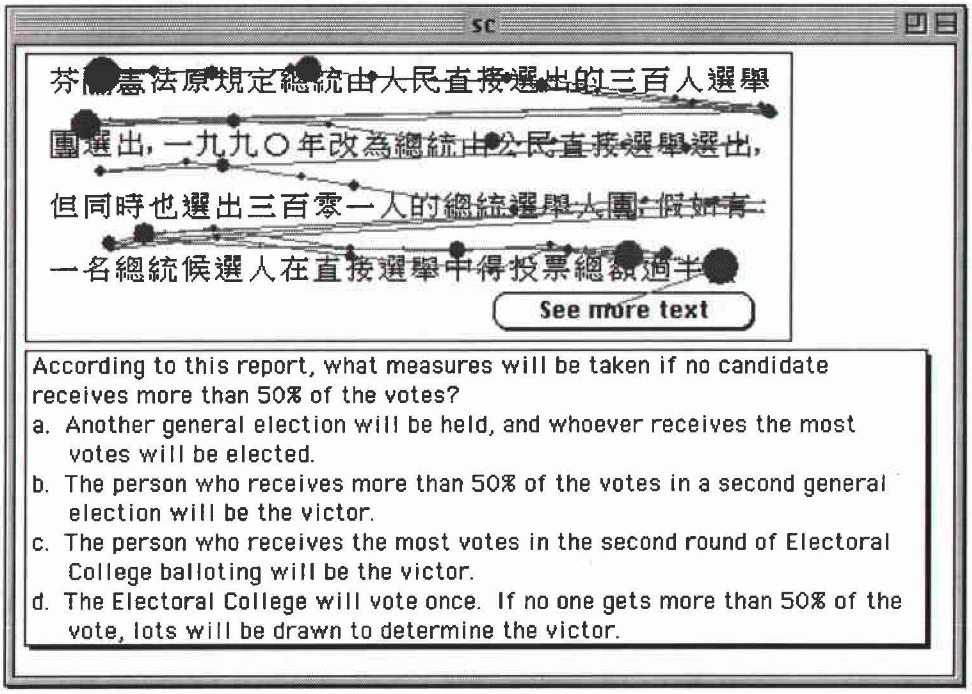


Figure 153a. NS07-Item 15 (sc17, card 1).

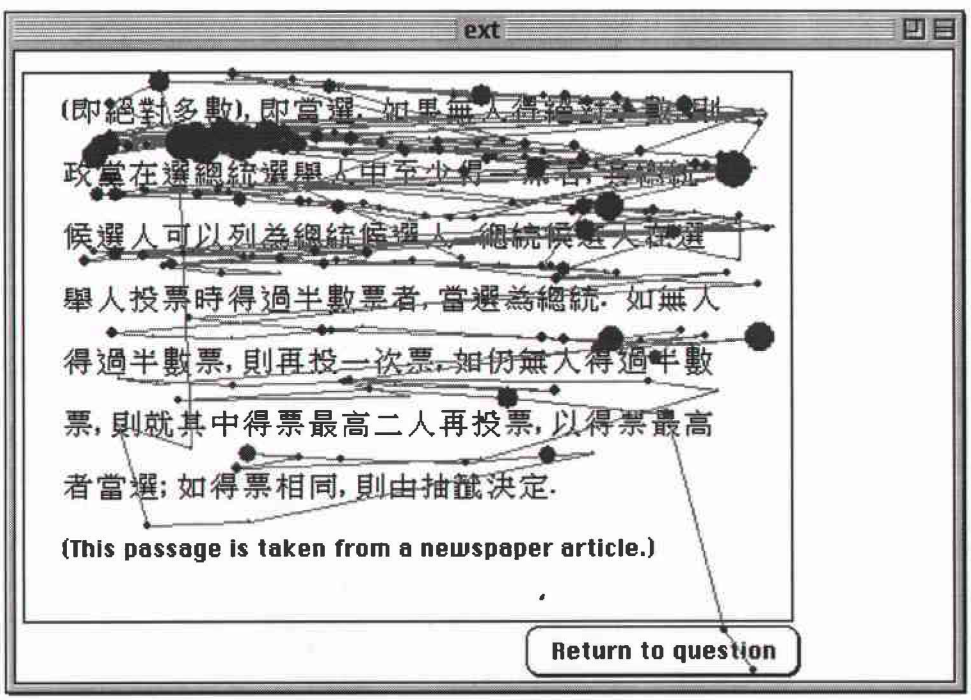


Figure 153b. NS07-Item 15 (sc17, card 2).

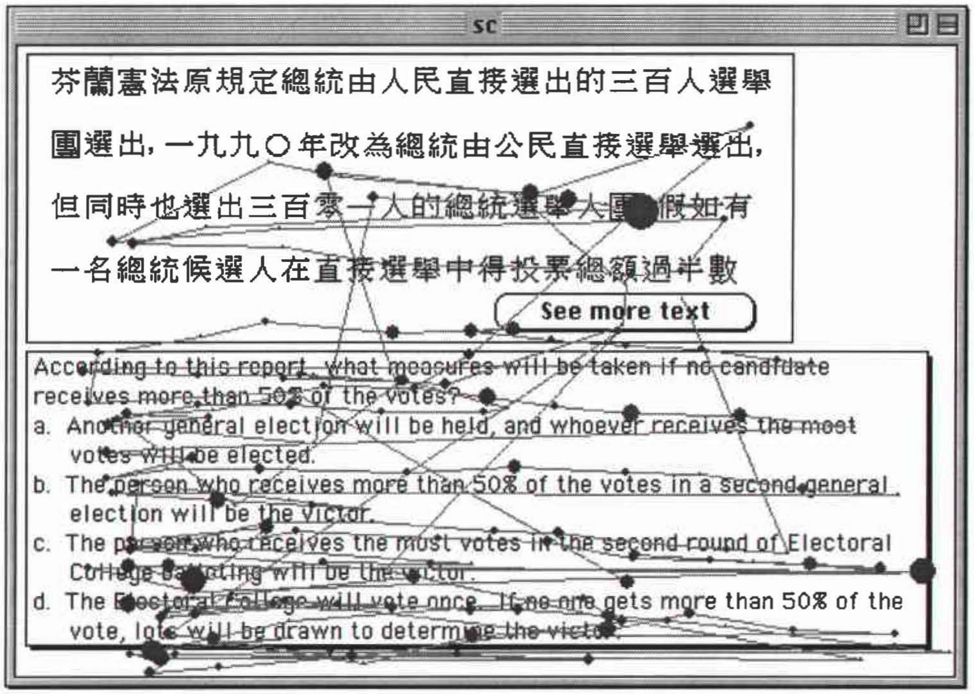


Figure 153c. NS07-Item 15 (sc17, card 1).

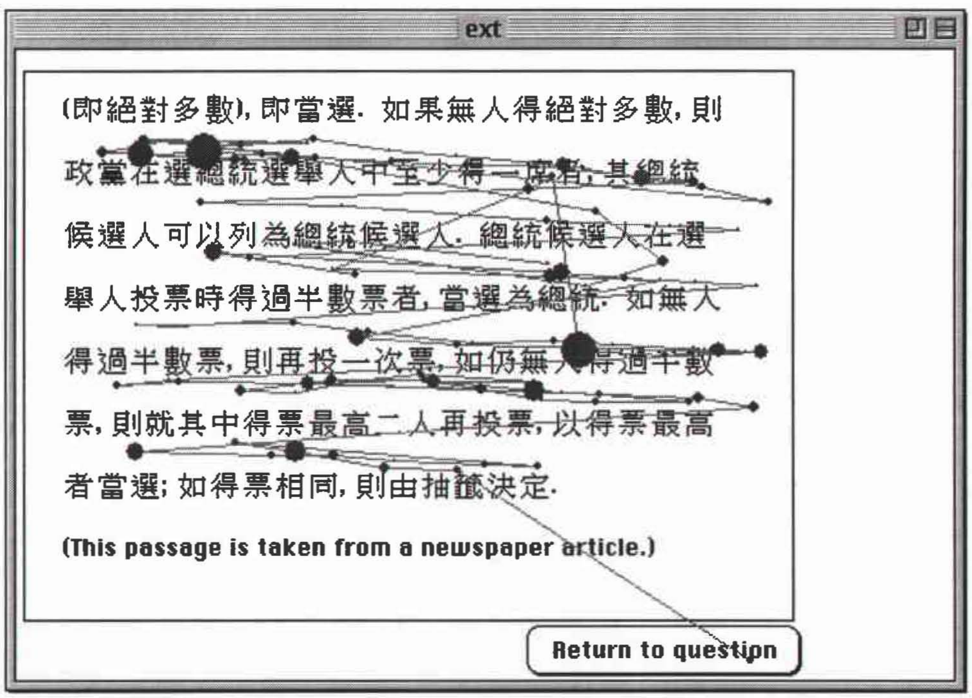


Figure 153d. NS07-Item 15 (sc17, card 2).

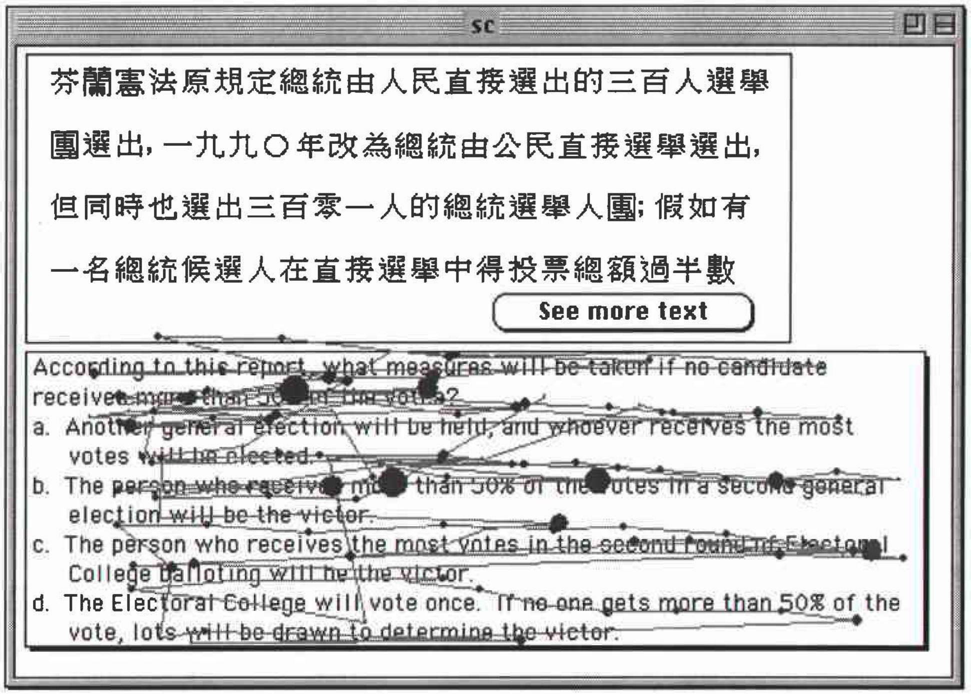


Figure 153e. NS07-Item 15 (sc17, card 1).

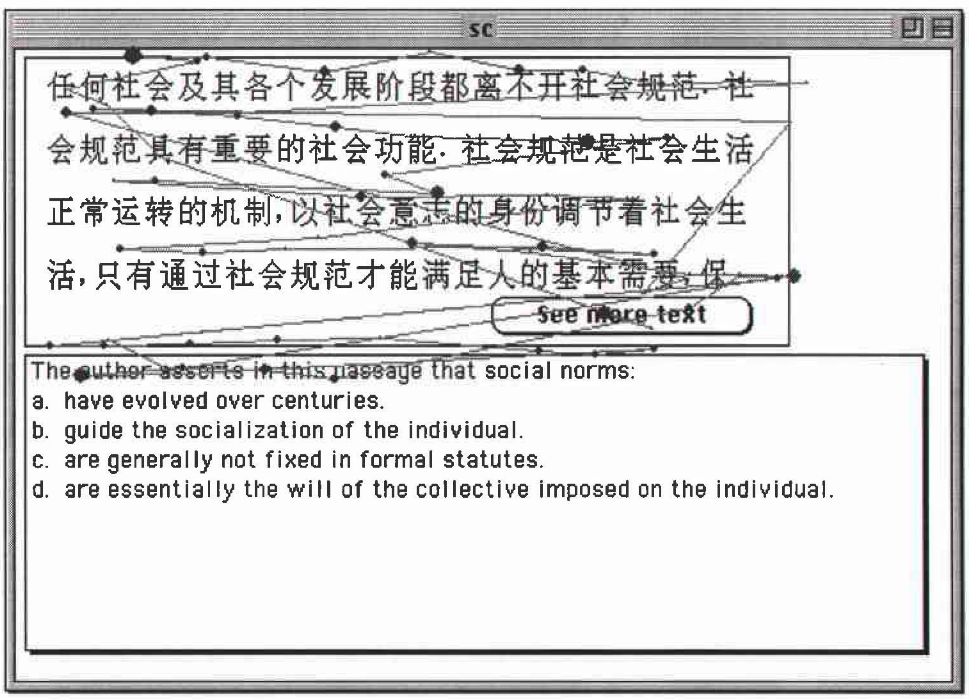


Figure 154a. NS07-Item 16 (sc13, card 1).

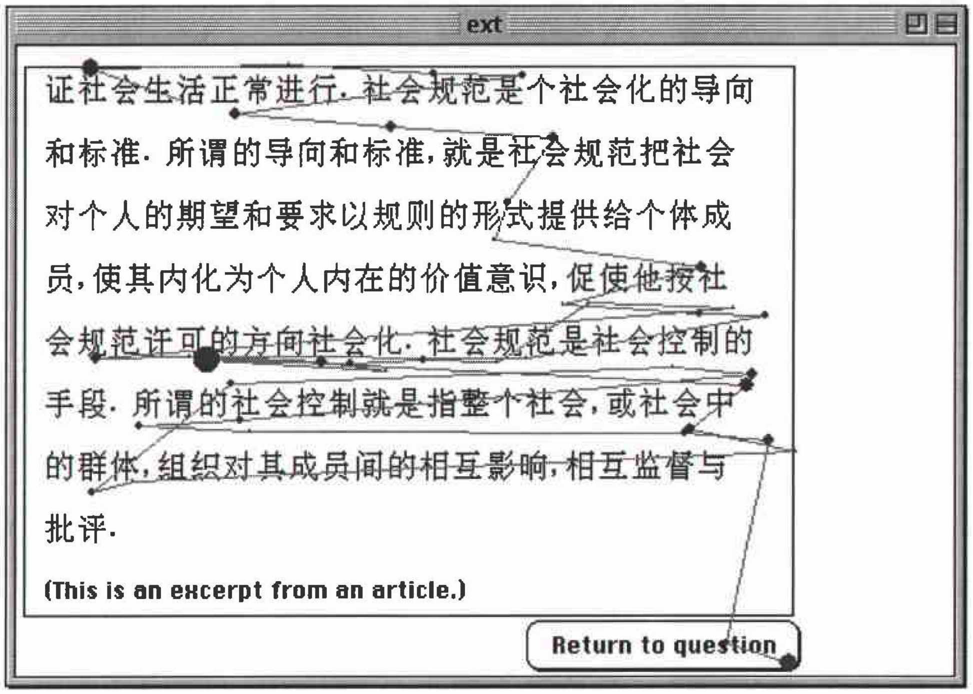


Figure 154b. NS07-Item 16 (sc13, card 2).

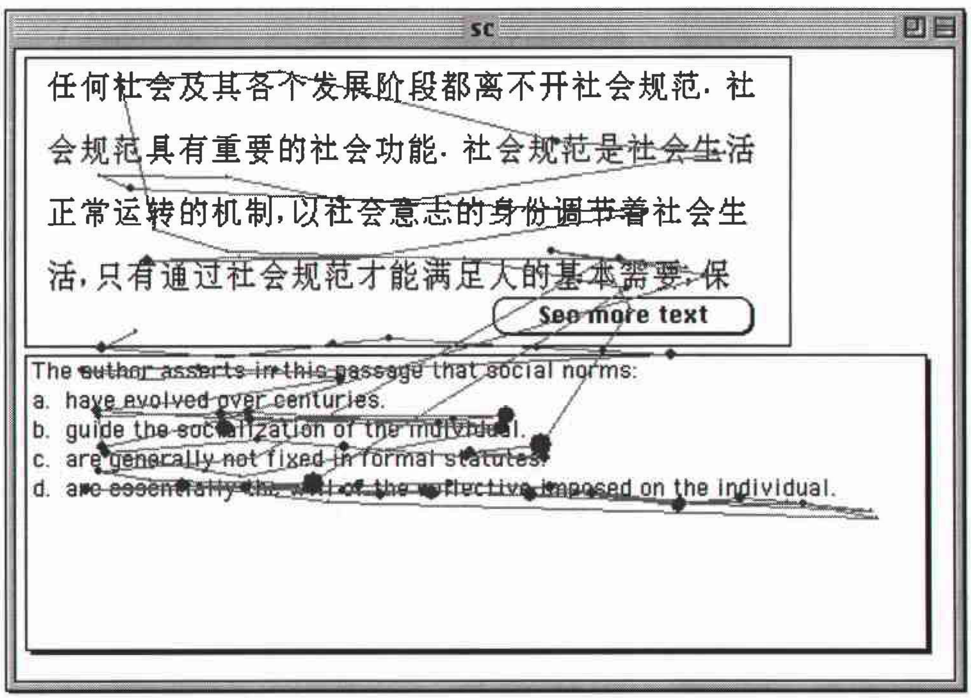


Figure 154c. NS07-Item 16 (sc13, card 1).

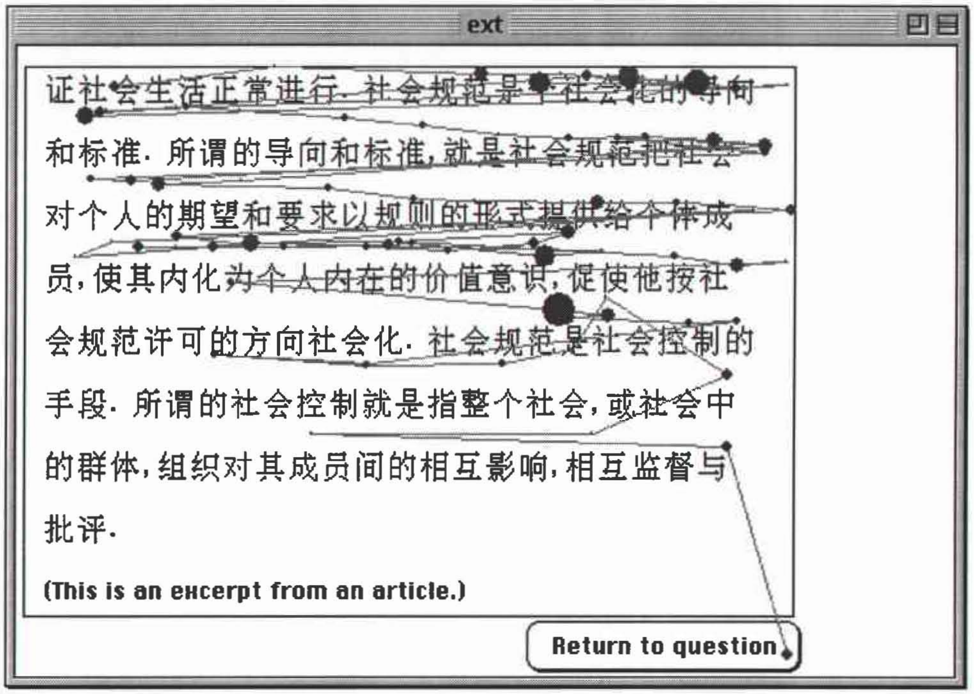


Figure 154d. NS07-Item 16 (sc13, card 2).

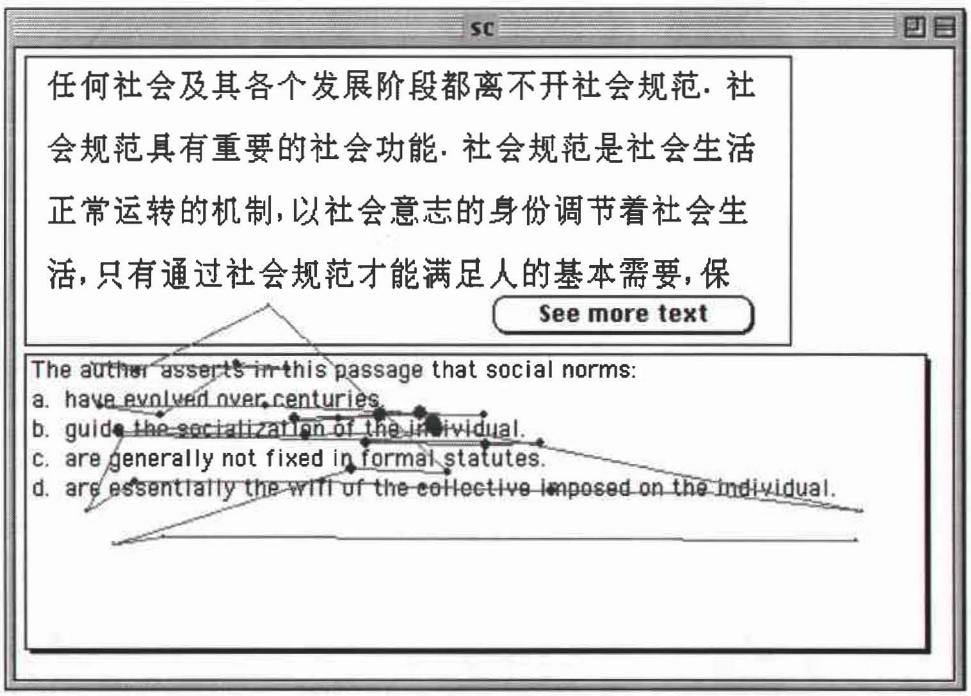


Figure 154e. NS07-Item 16 (sc13, card 1).

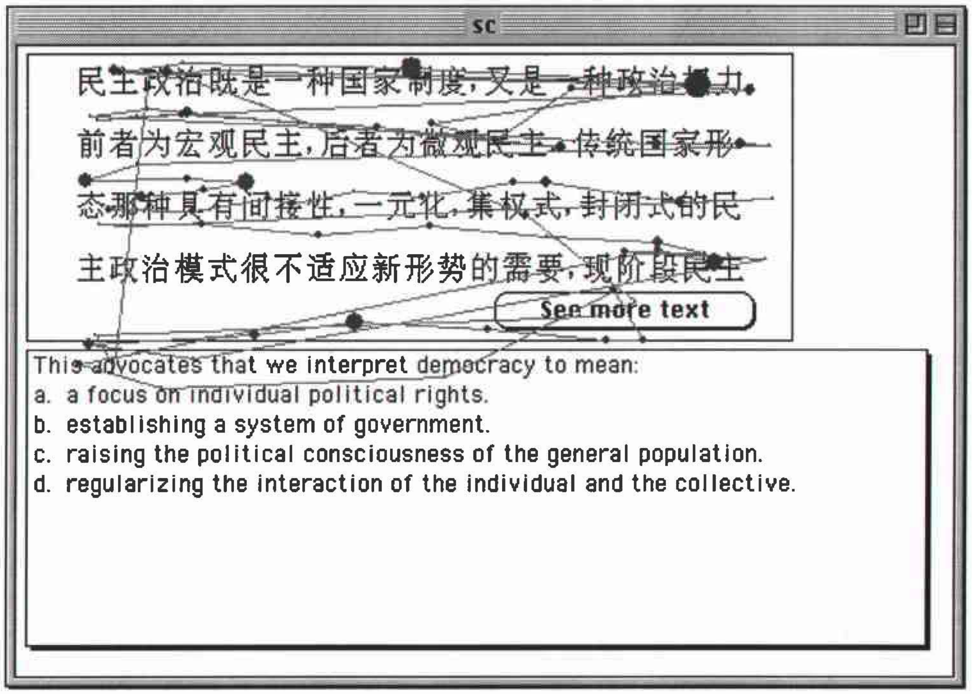


Figure 155a. NS07-Item 17 (sc12, card 1).

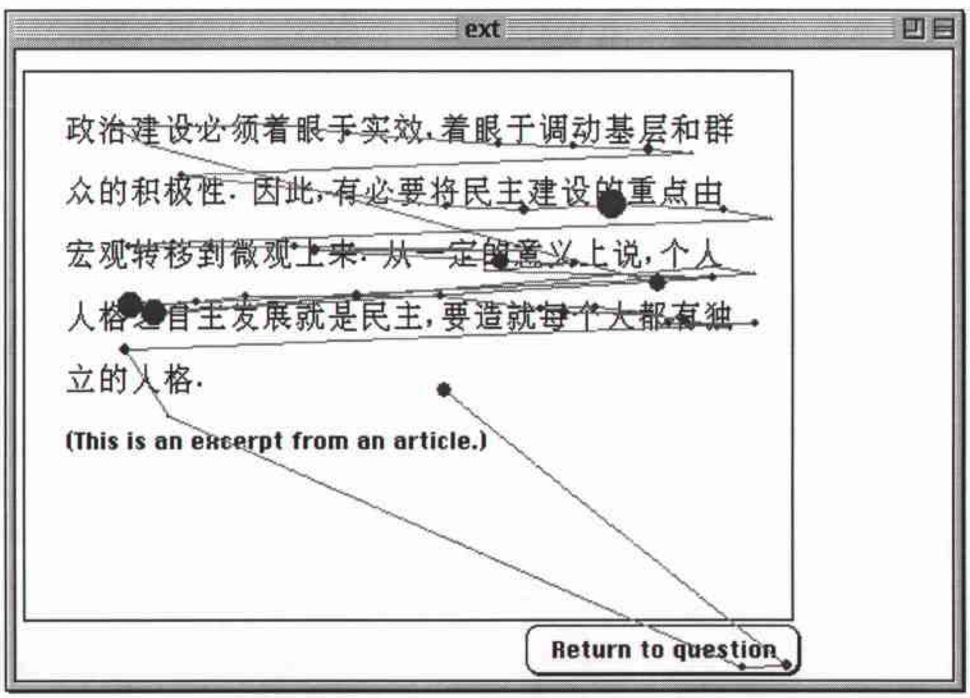


Figure 155b. NS07-Item 17 (sc12, card 2).

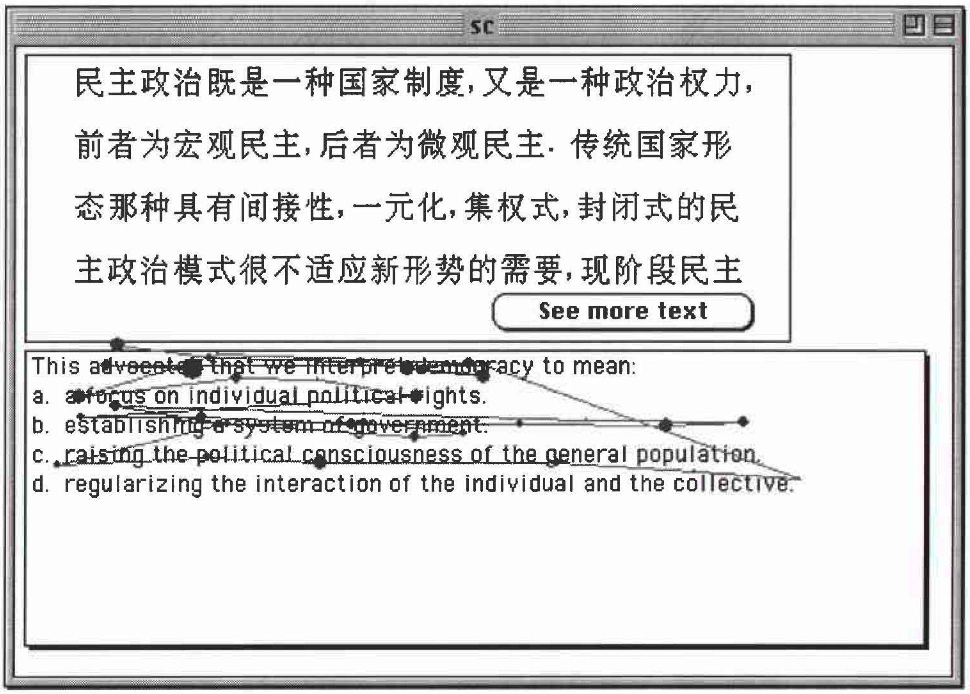


Figure 155c. NS07-Item 17 (sc12, card 1).

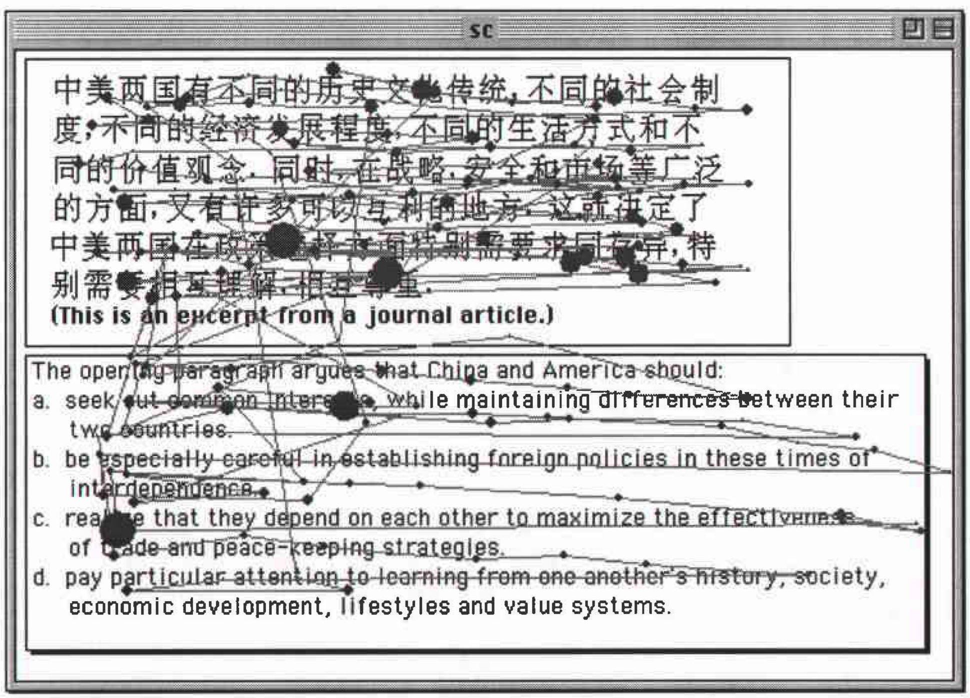


Figure 156. NS07-Item 18 (sc04).

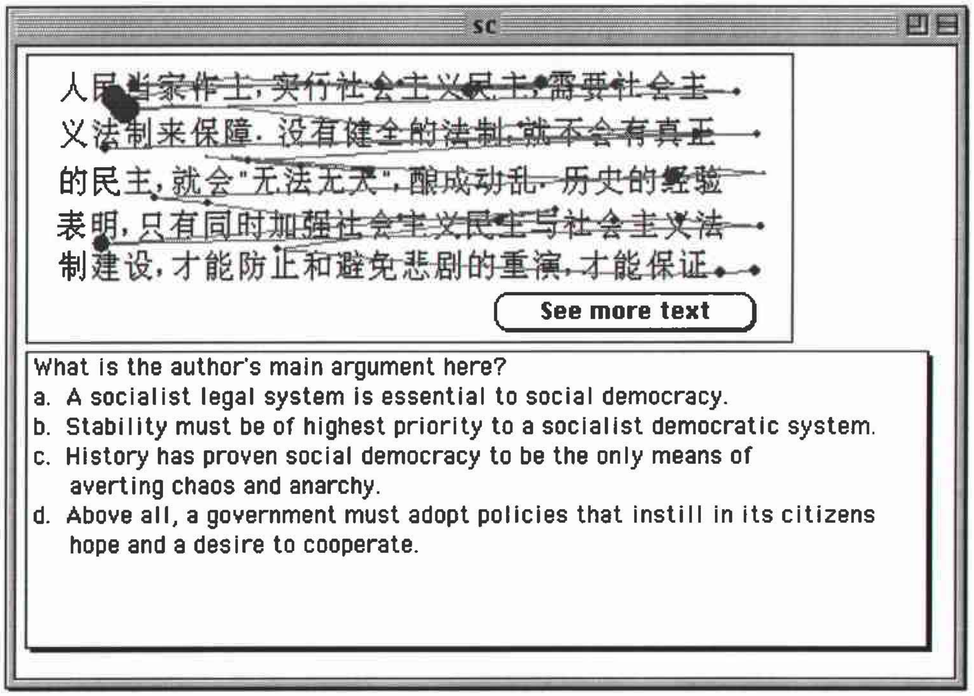


Figure 157a. NS07-Item 19 (sc05, card 1).

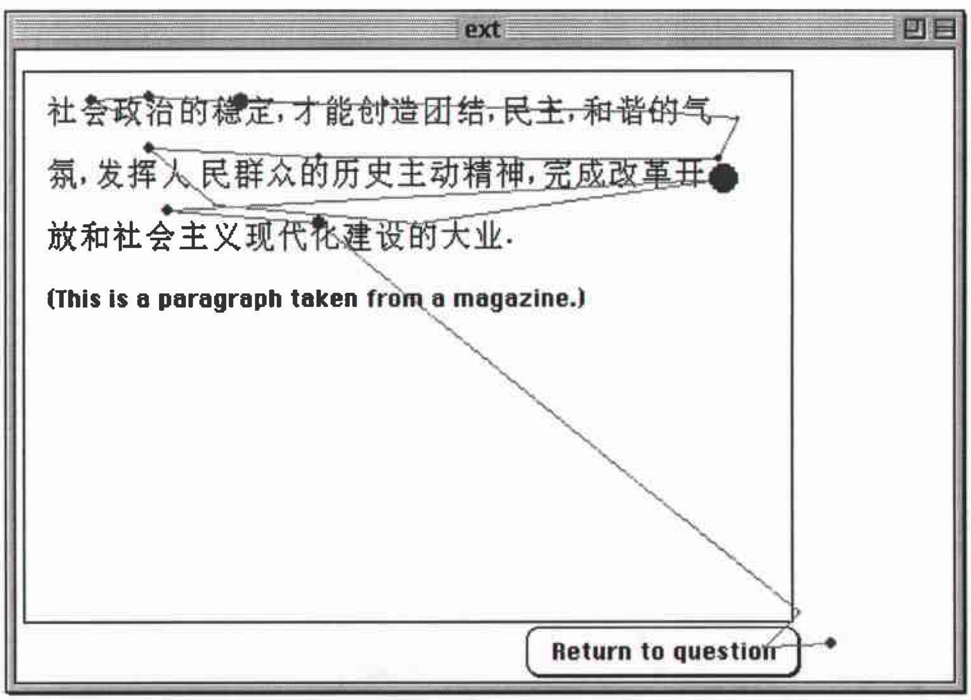


Figure 157b. NS07-Item 19 (sc05, card 2).

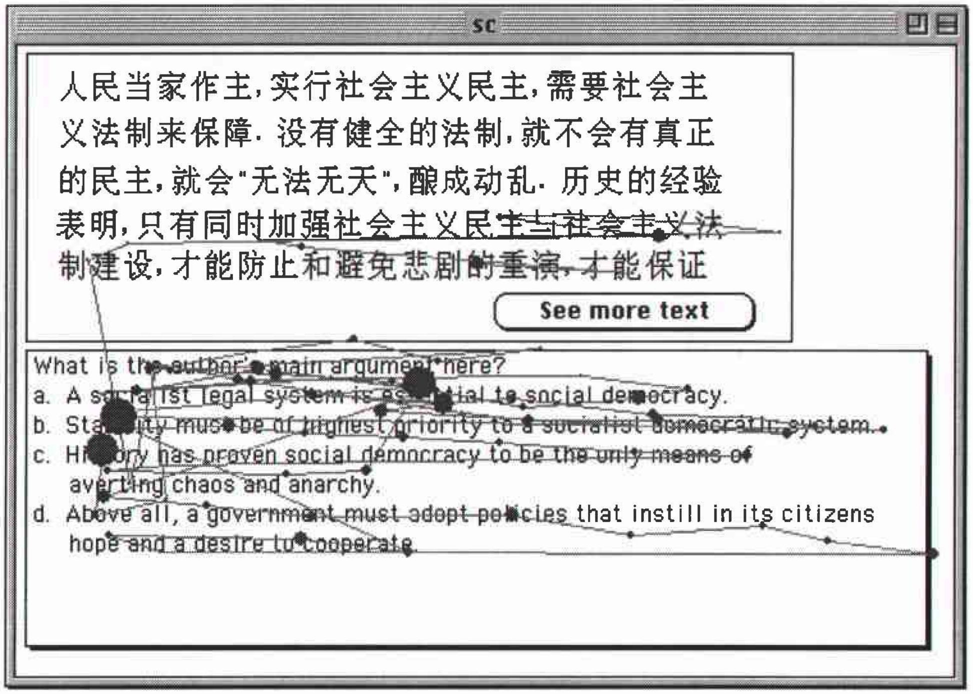


Figure 157c. NS07-Item 19 (sc05, card 1).

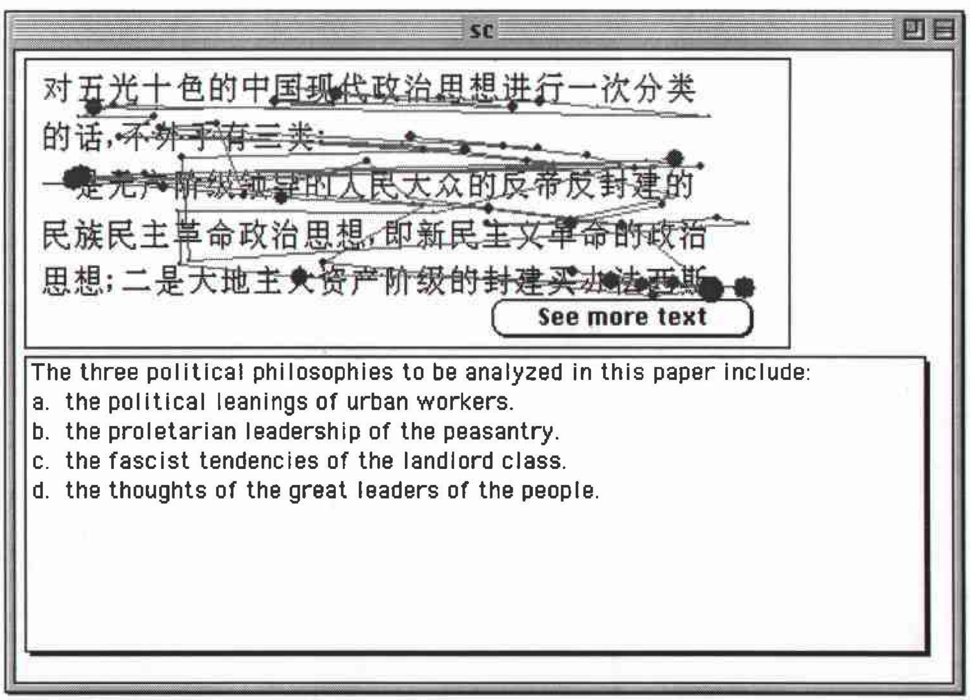


Figure 158a. NS07-Item 20 (sc06, card 1).

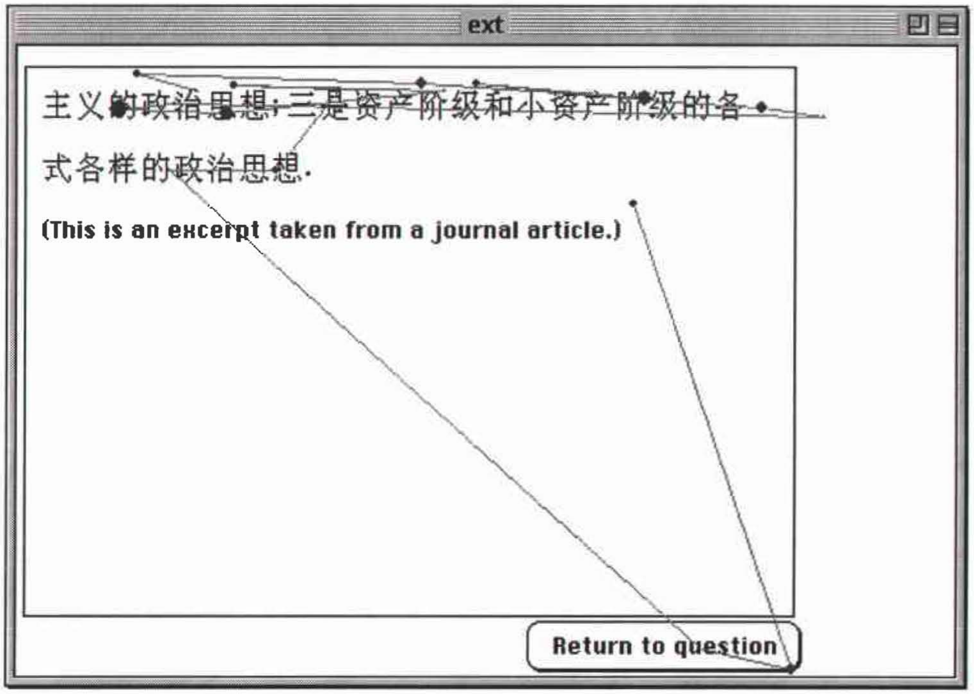


Figure 158b. NS07-Item 20 (sc06, card 2).

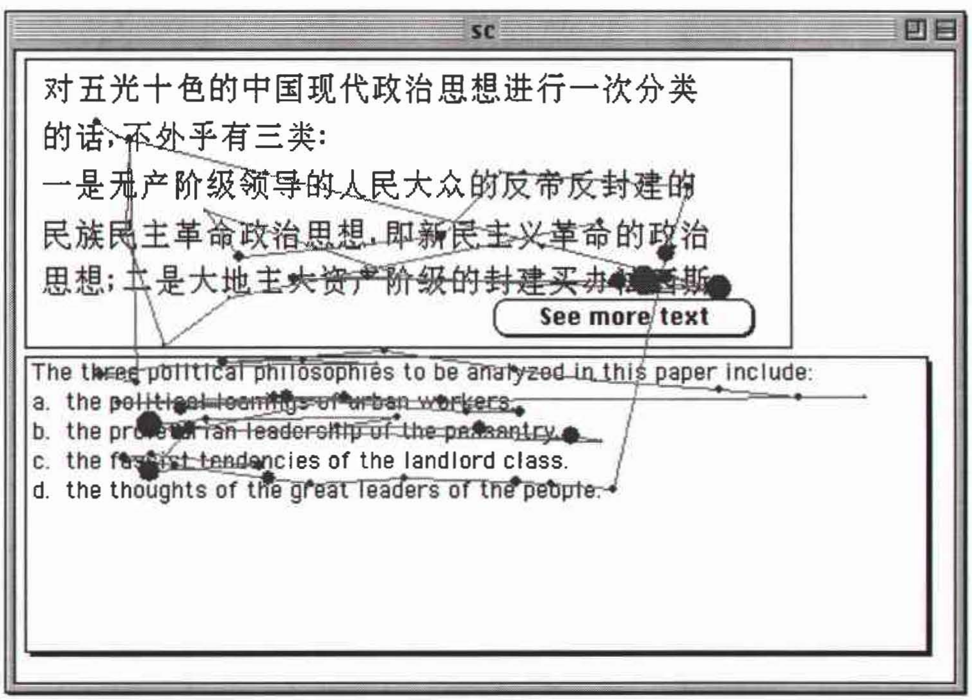


Figure 158c. NS07-Item 20 (sc06, card 1).

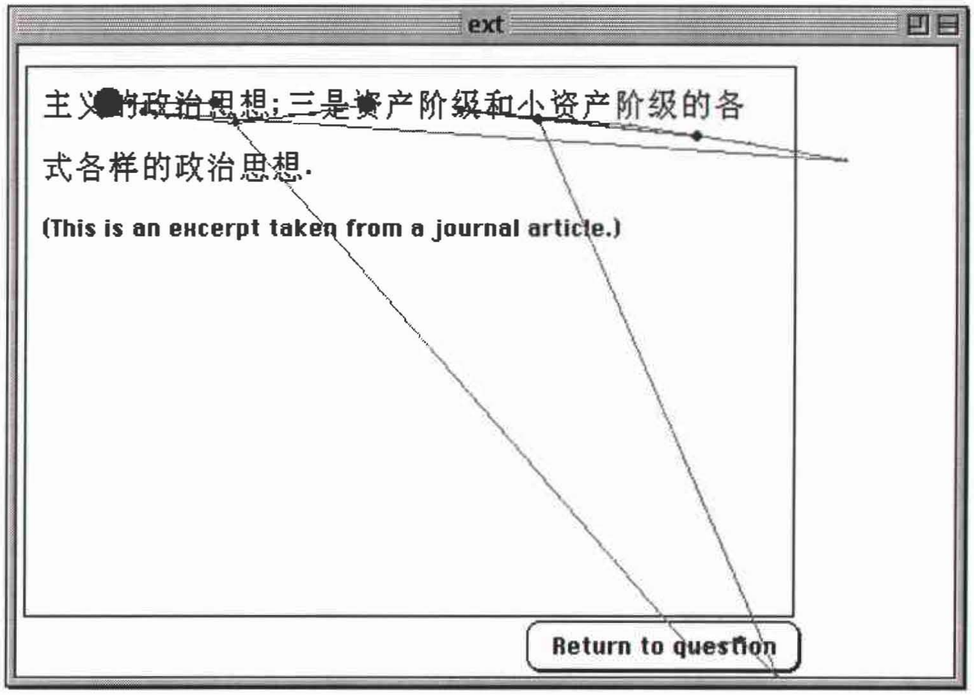


Figure 158d. NS07-Item 20 (sc06, card 2).

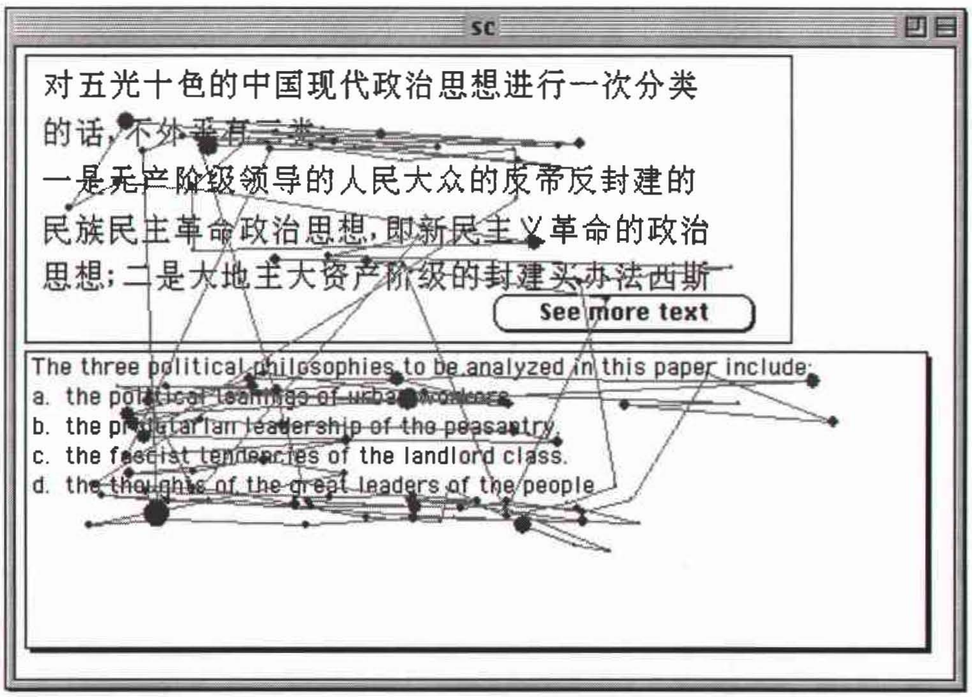


Figure 158e. NS07-Item 20 (sc06, card 1).

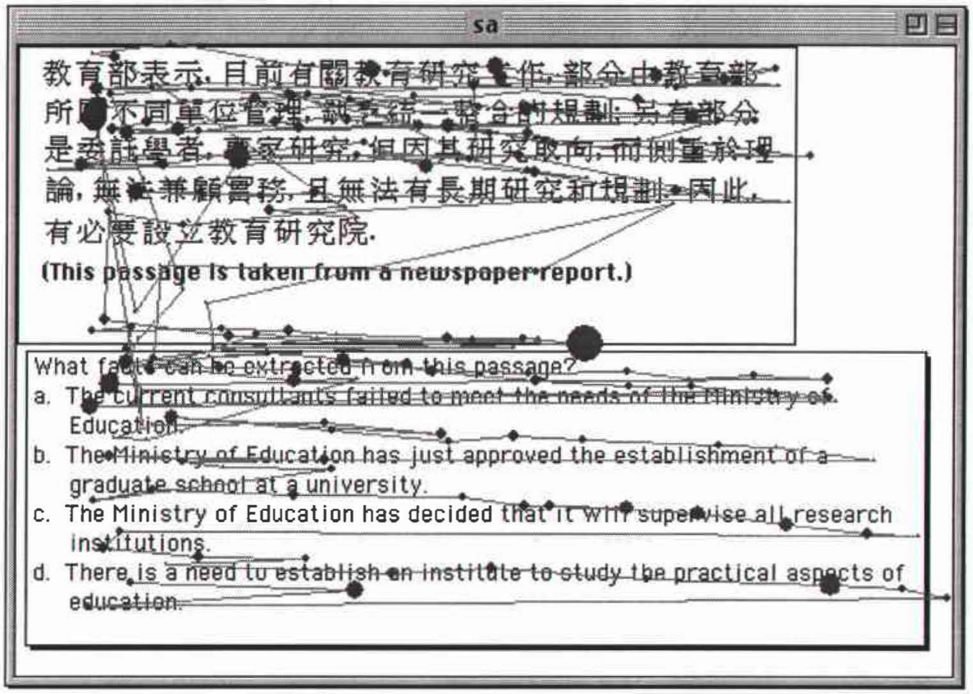


Figure 159. NS07-Item 21 (sa14).

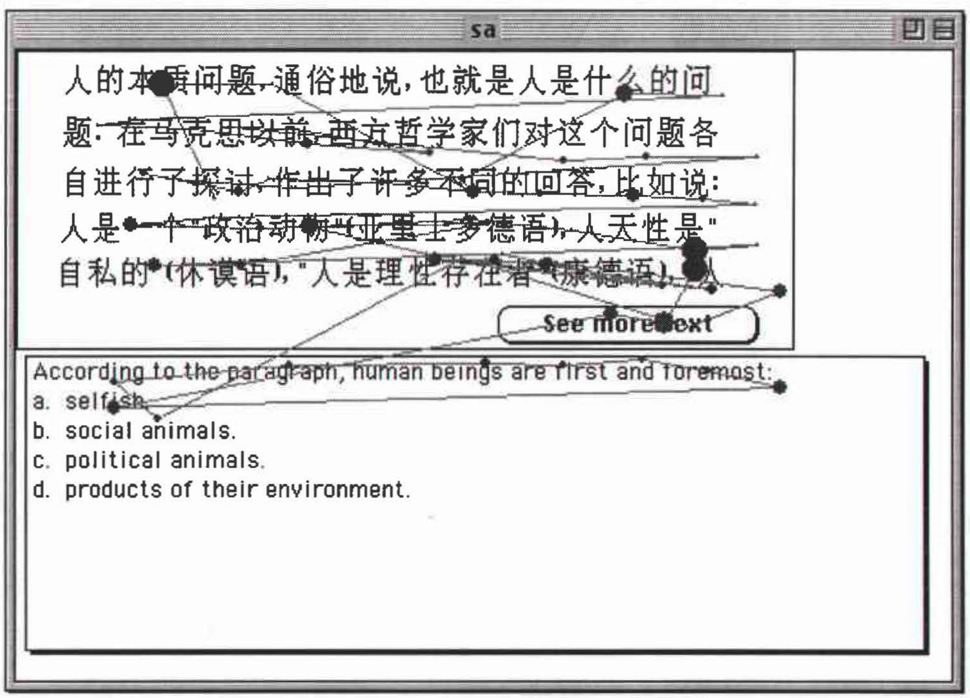


Figure 160a. NS07-Item 22 (sa06, card 1).

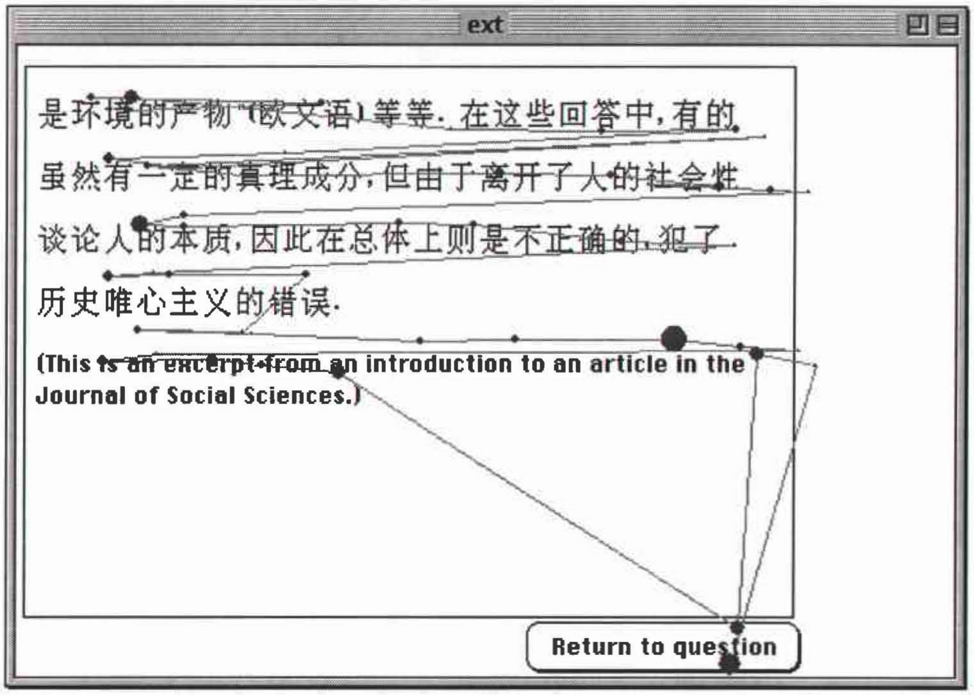


Figure 160b. NS07-Item 22 (sa06, card 2).

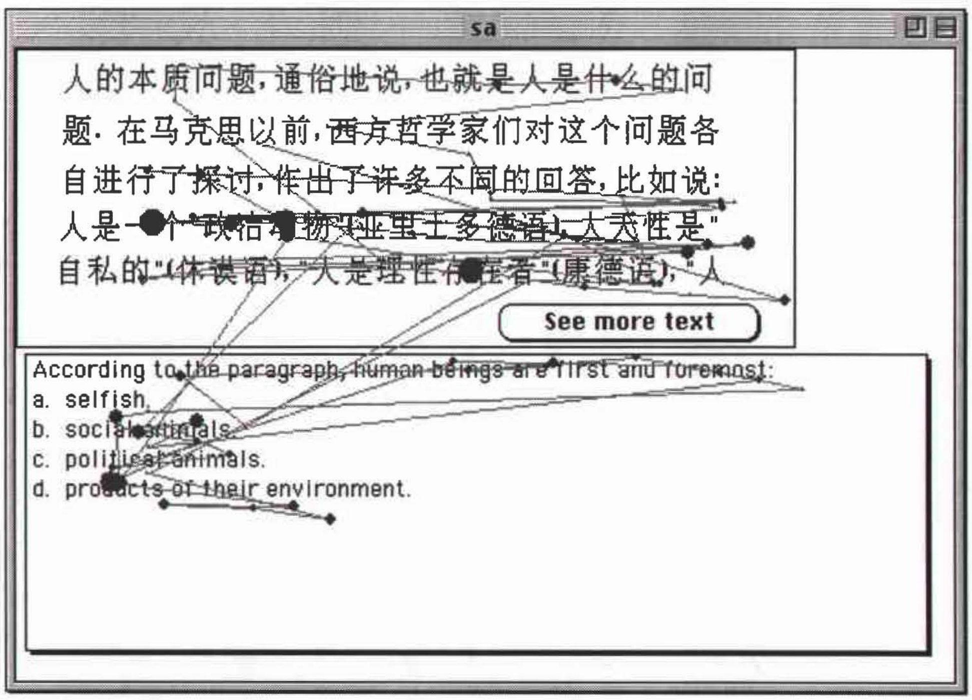


Figure 160c. NS07-Item 22 (sa06, card 1).

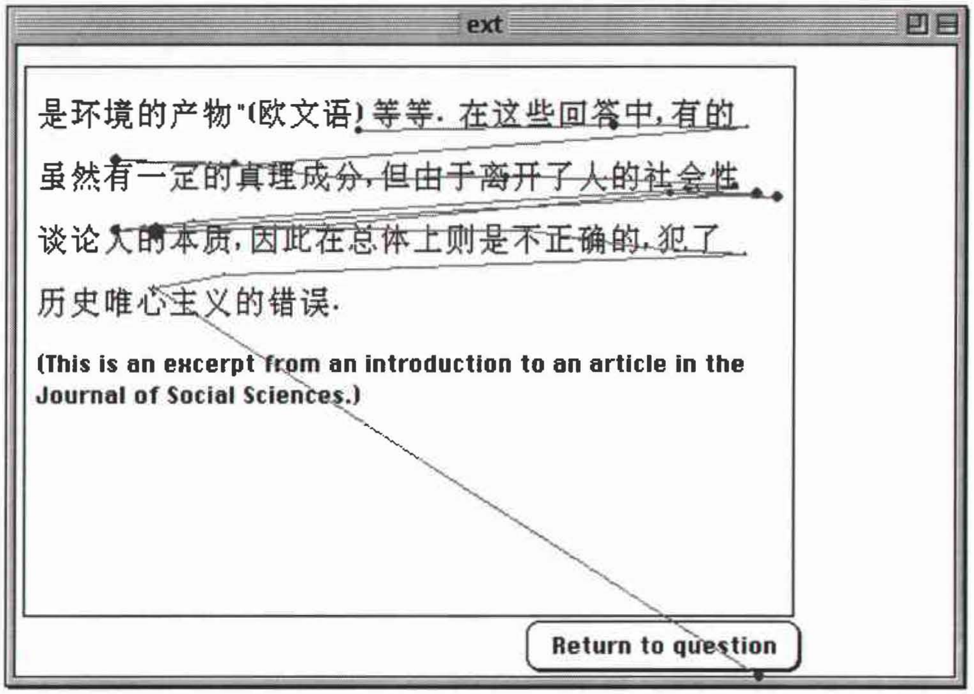


Figure 160d. NS07-Item 22 (sa06, card 2).

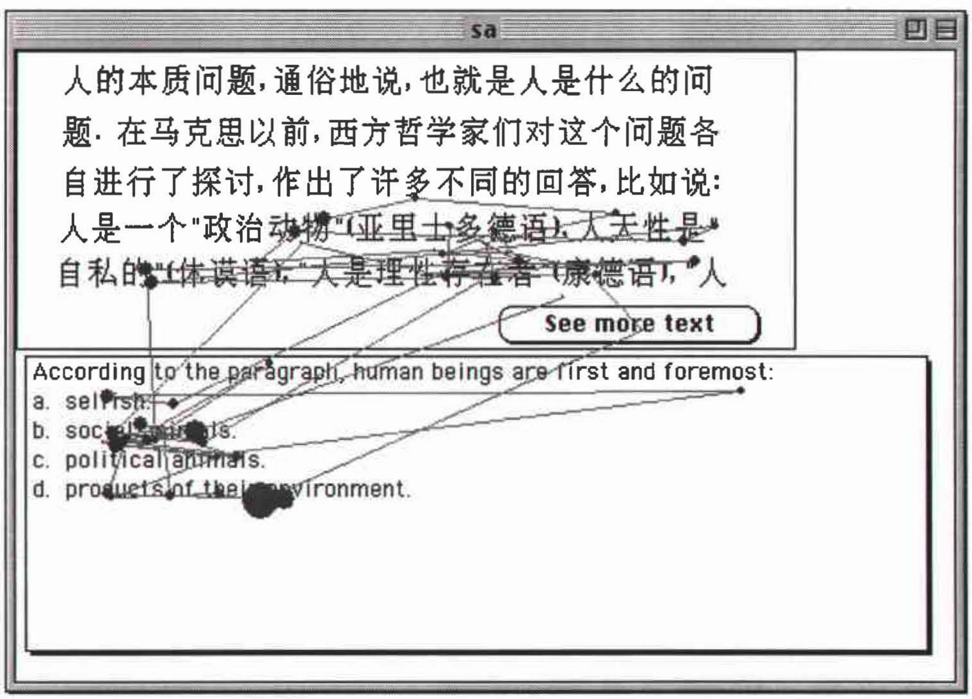


Figure 160e. NS07-Item 22 (sa06, card 1).

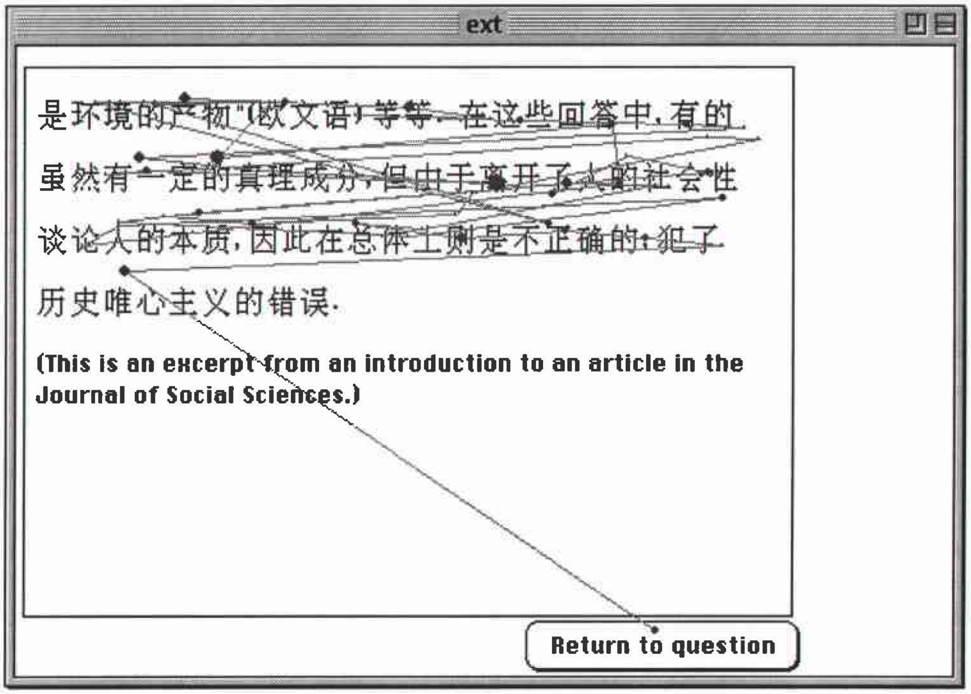


Figure 160f. NS07-Item 22 (sa06, card 2).

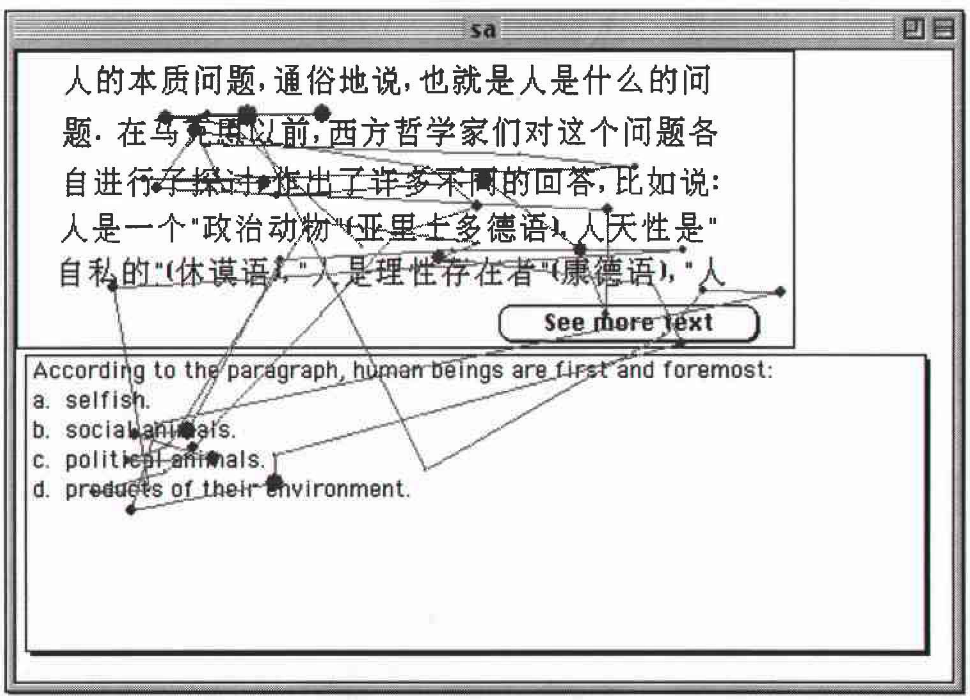


Figure 160g. NS07-Item 22 (sa06, card 1).

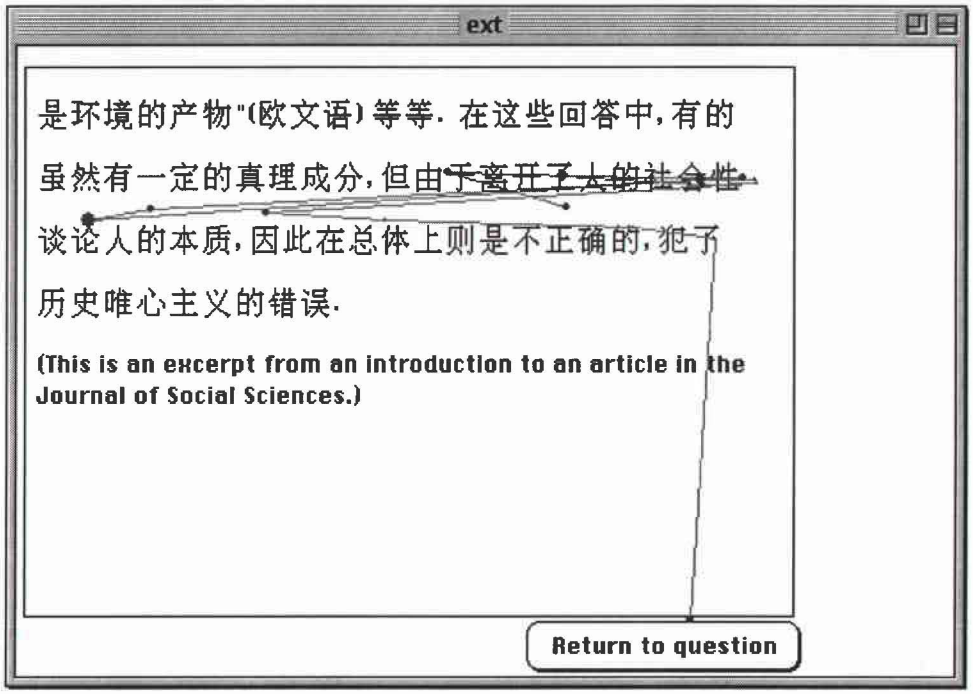


Figure 160h. NS07-Item 22 (sa06, card 2).

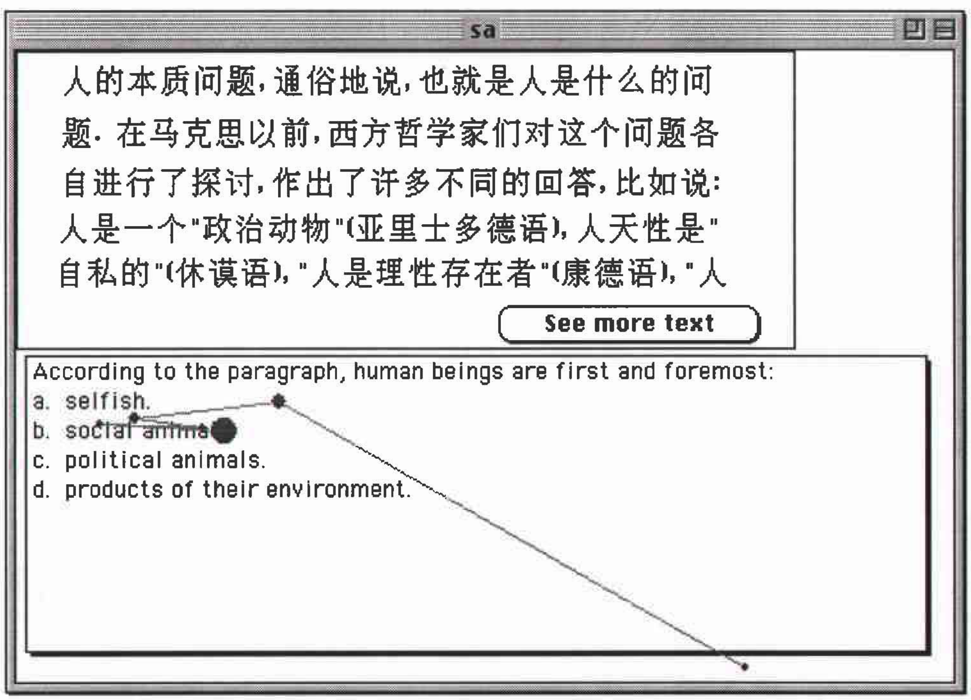


Figure 160i. NS07-Item 22 (sa06, card 1).

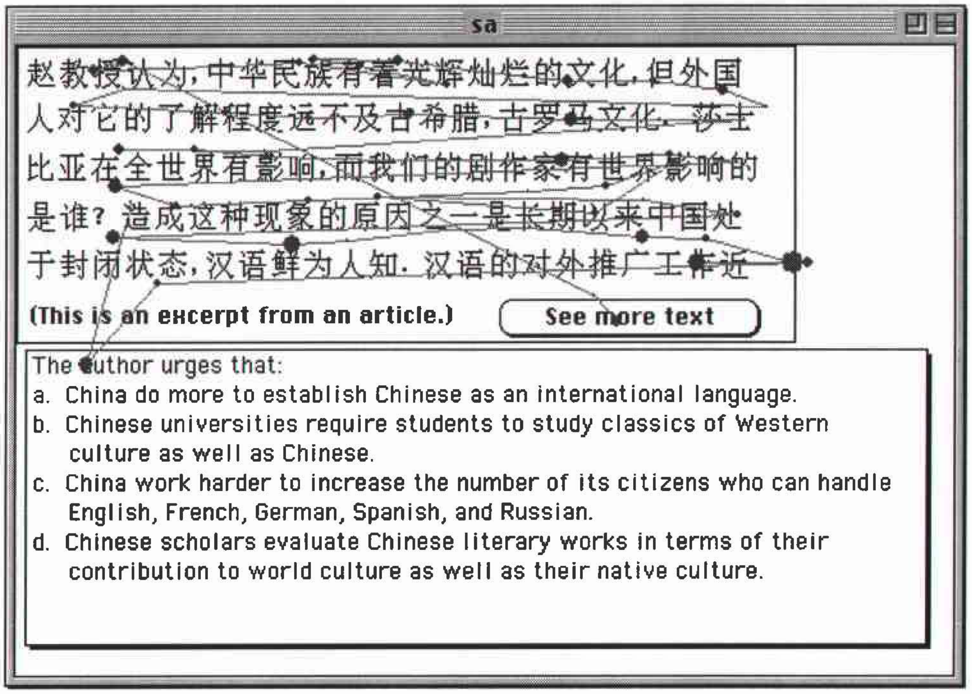


Figure 161a. NS07-Item 23 (sa05, card 1).

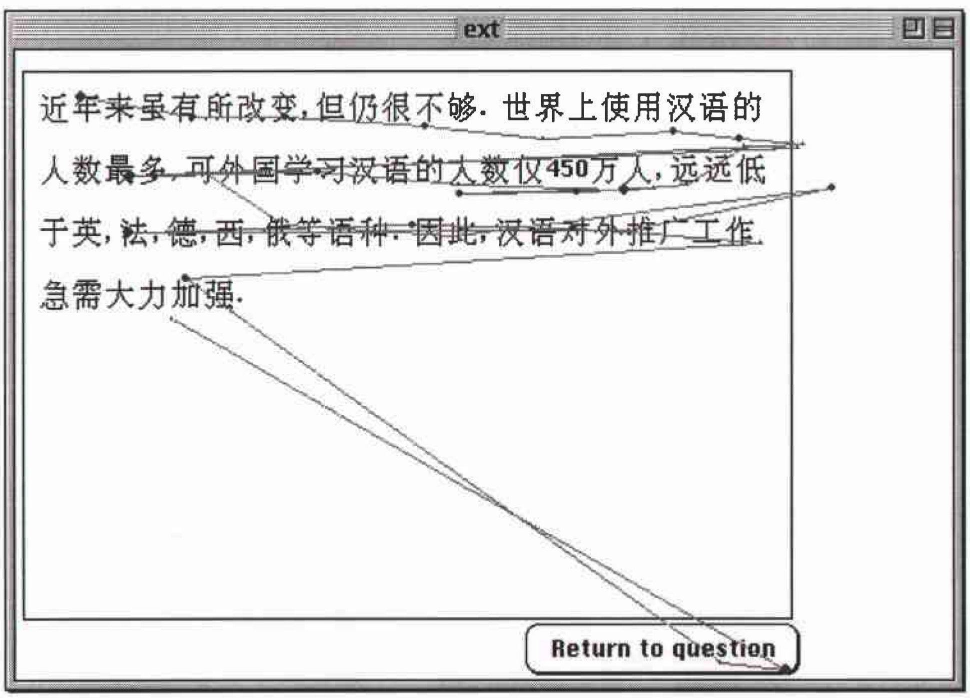


Figure 161b. NS07-Item 23 (sa05, card 2).

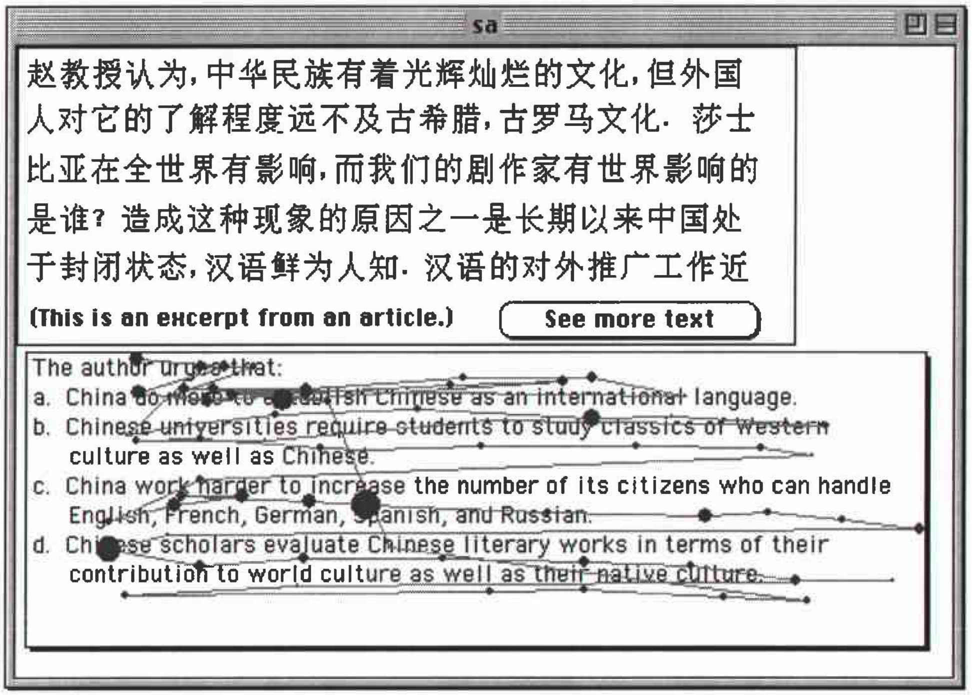


Figure 161c. NS07-Item 23 (sa05, card 1).

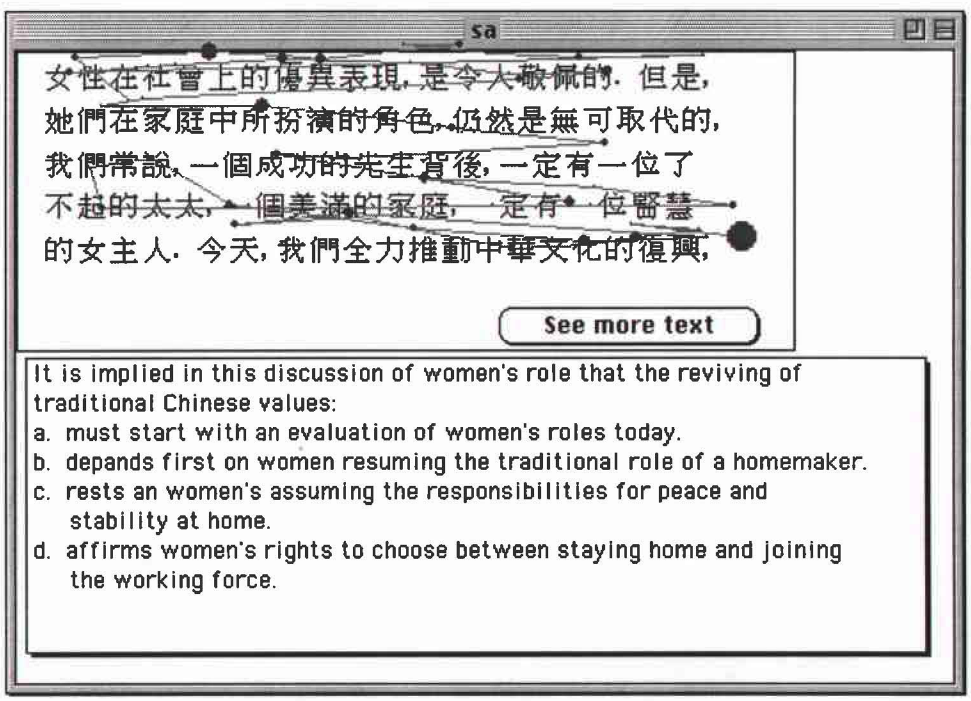


Figure 162a. NS07-Item 24 (sa09, card 1).

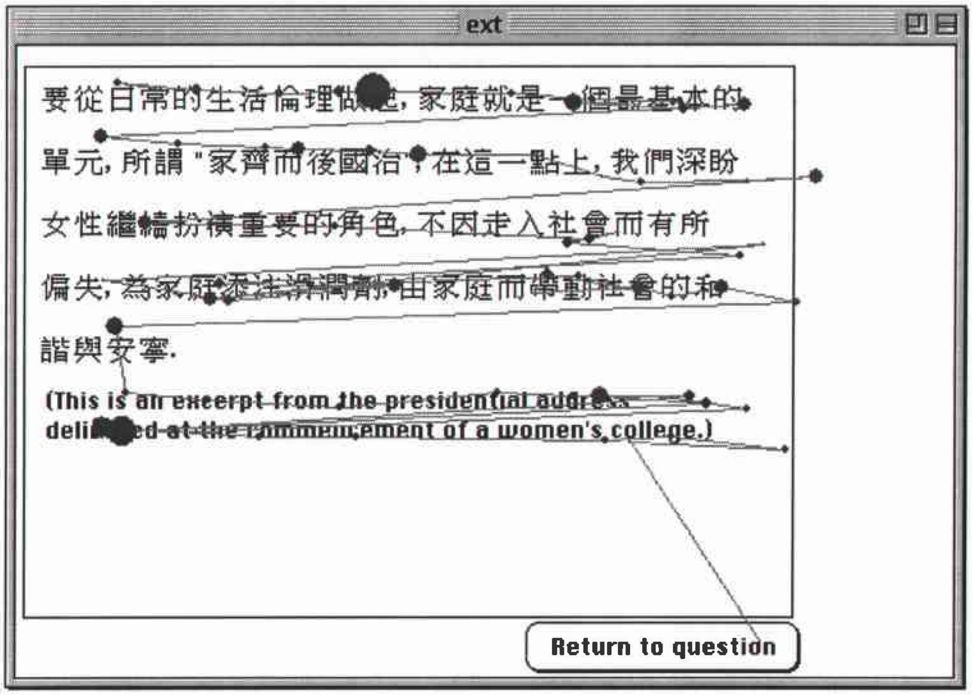


Figure 162b. NS07-Item 24 (sa09, card 2).

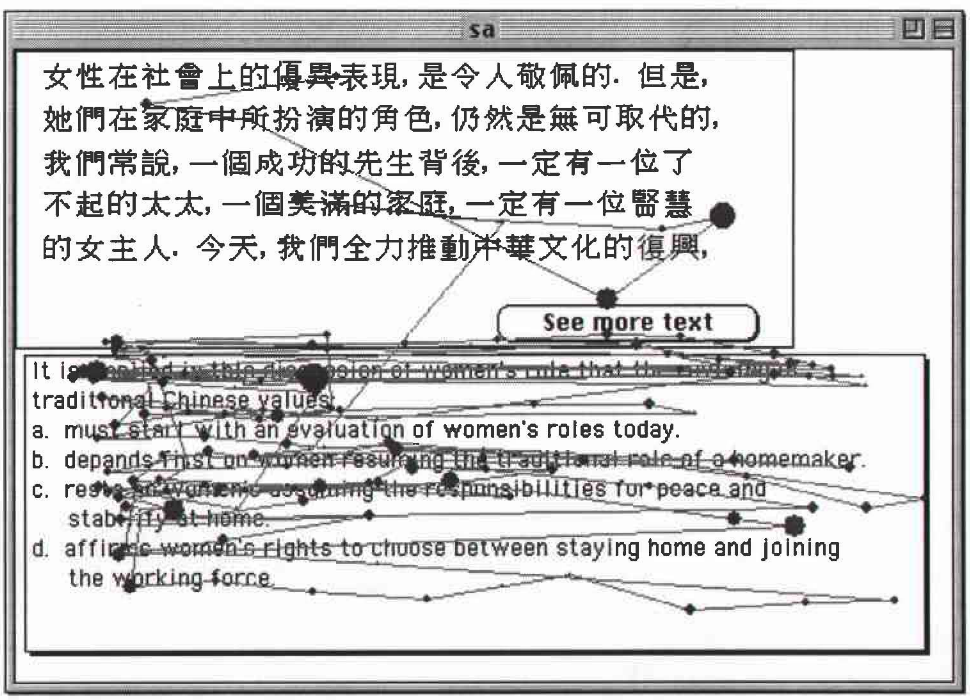


Figure 162c. NS07-Item 24 (sa09, card 1).

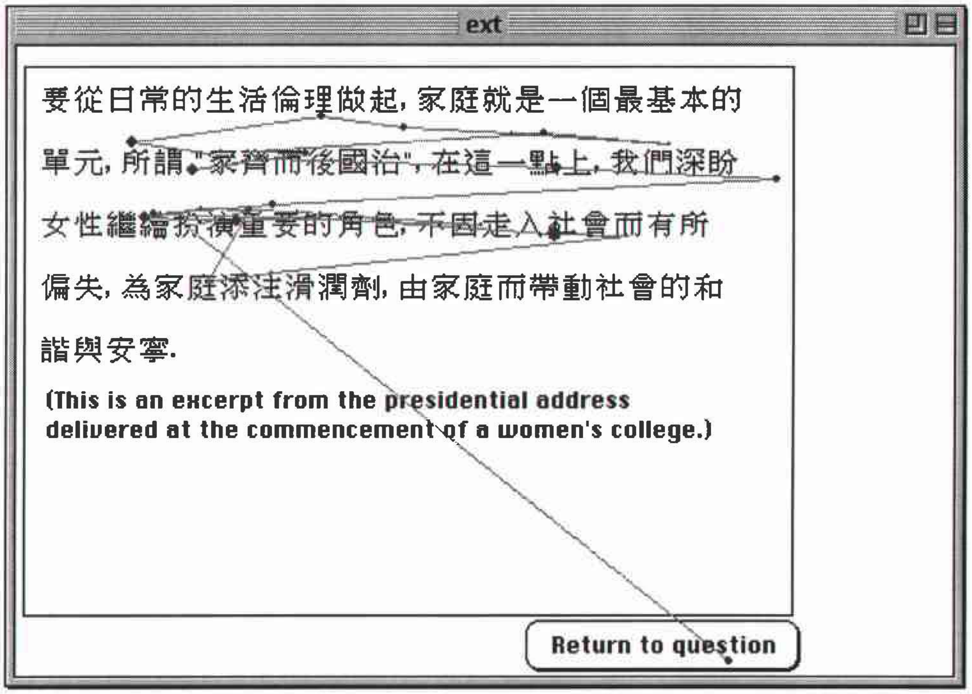


Figure 162d. NS07-Item 24 (sa09, card 2).

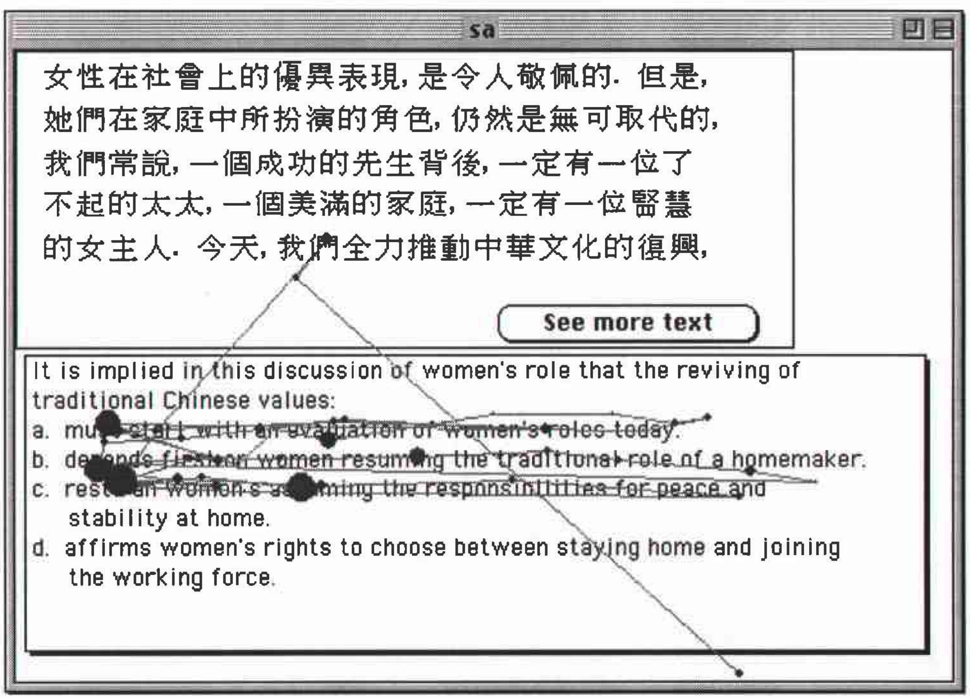


Figure 162e. NS07-Item 24 (sa09, card 1).

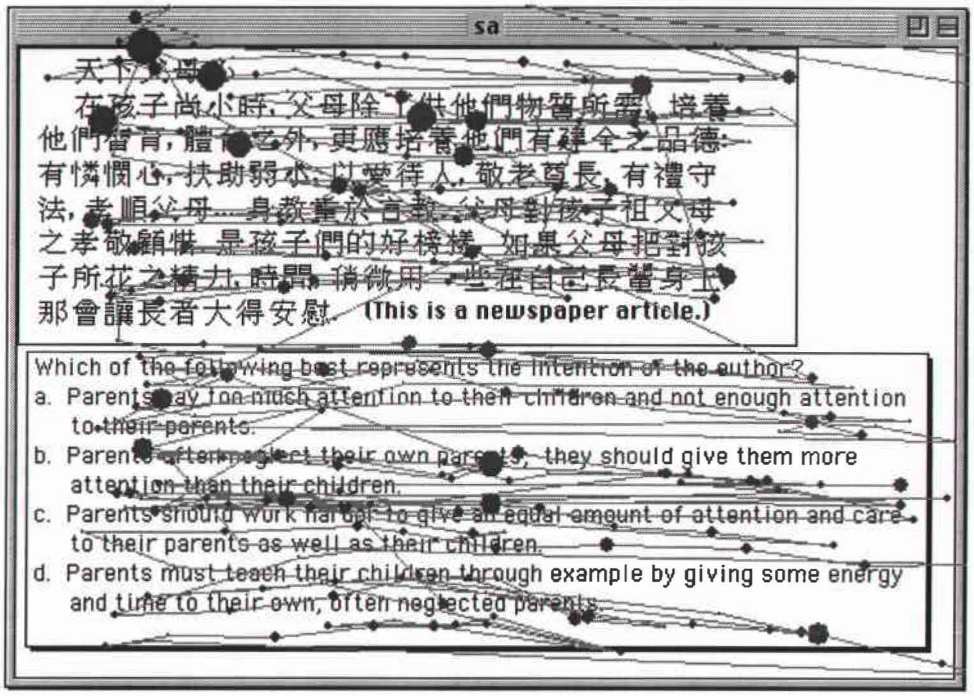


Figure 163. NS07-Item 25 (sa04).

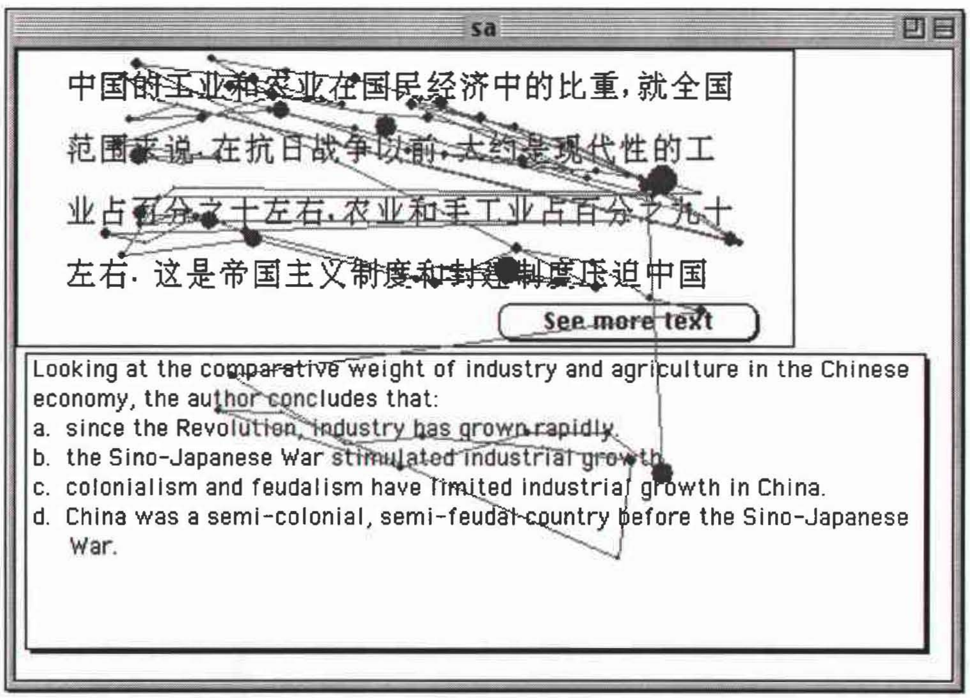


Figure 164a. NS07-Item 26 (sa13, card 1).

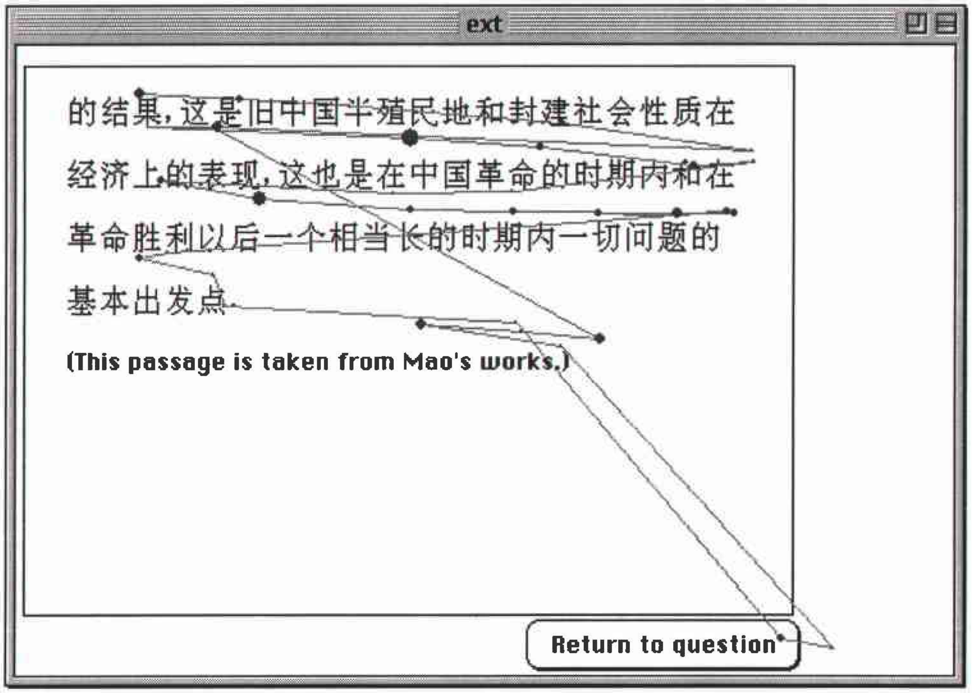


Figure 164b. NS07-Item 26 (sa13, card 2).

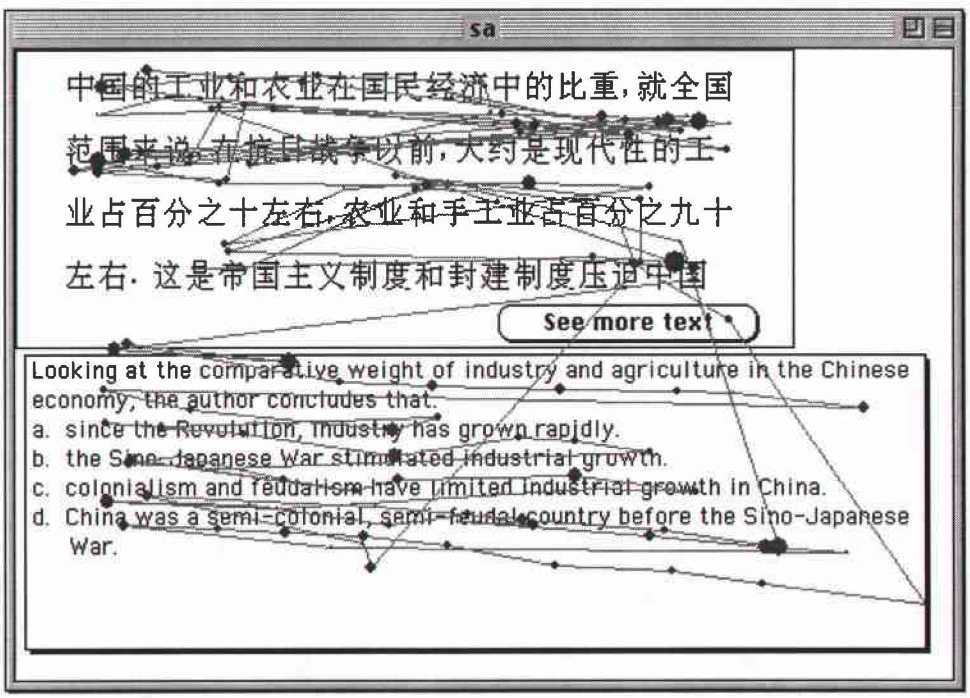


Figure 164c. NS07-Item 26 (sa13, card 1).

RECEIVED
AS
FOLLOWS

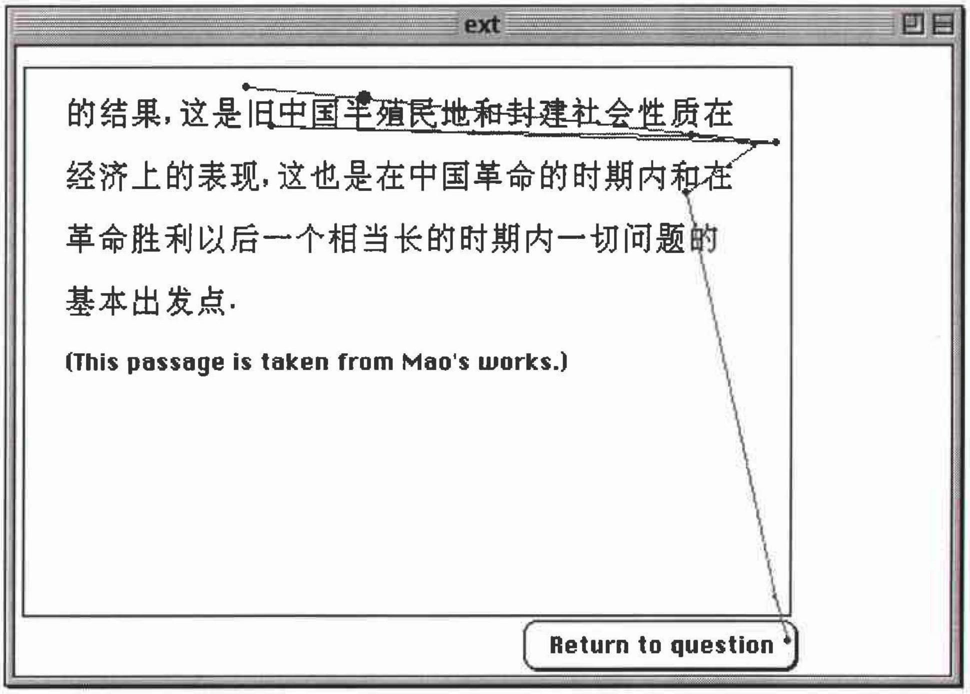


Figure 164d. NS07-Item 26 (sa13, card 2).

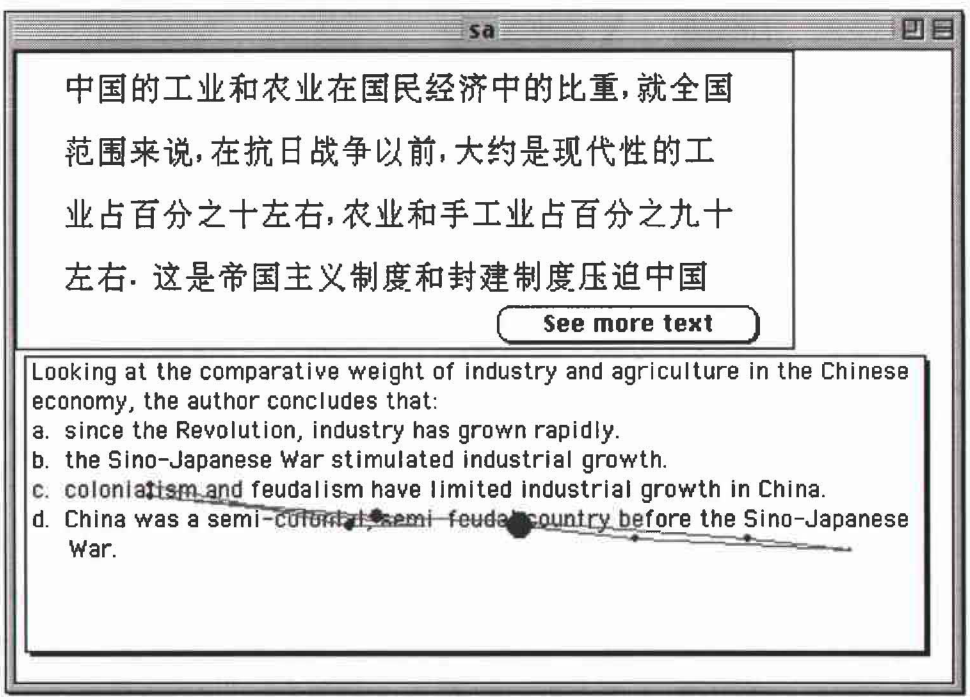


Figure 164e. NS07-Item 26 (sa13, card 1).

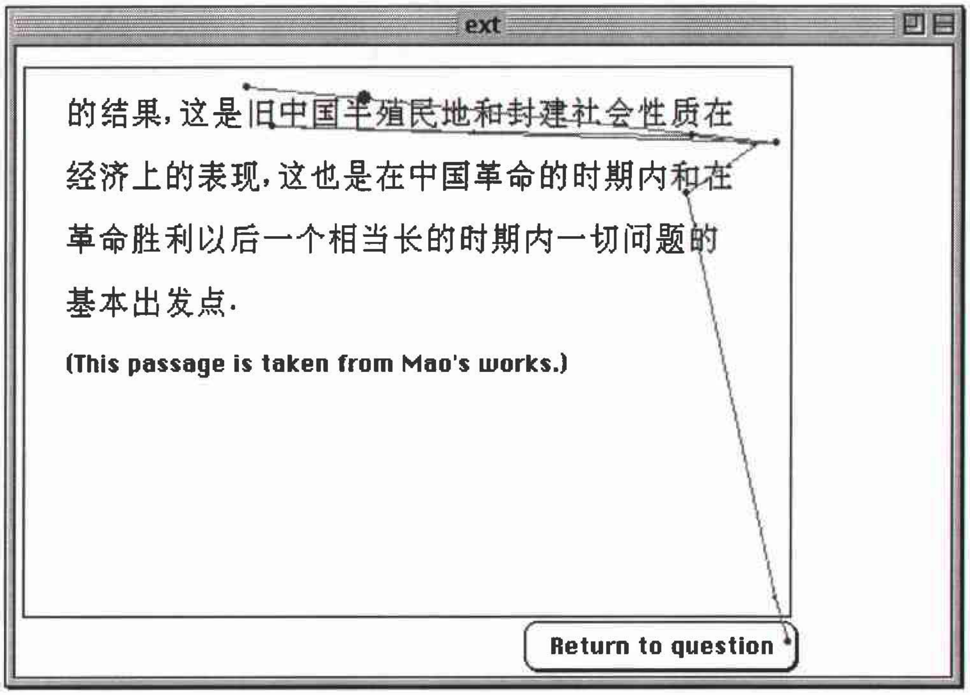


Figure 164d. NS07-Item 26 (sa13, card 2).

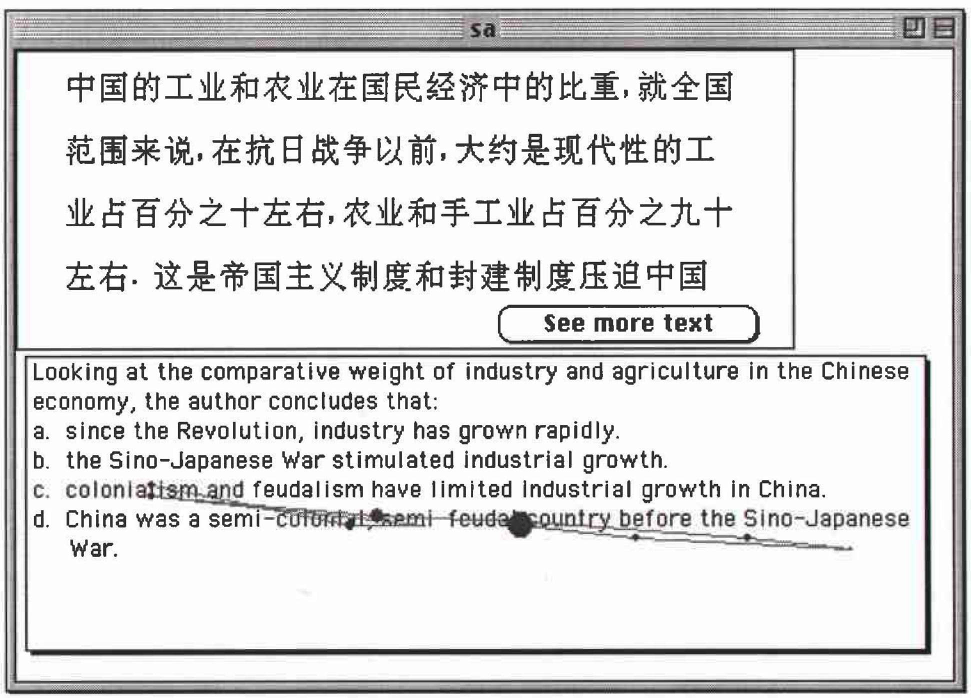


Figure 164e. NS07-Item 26 (sa13, card 1).

GLOSSARY

acuity	The sharpness and clarity of vision, best detected by the fovea, secondarily by the parafovea, tertiarily by the peripheral field. Also, spatial resolution.
Chinese character	The smallest morphological unit in the Chinese written language. Also, *Hànzì, Kanji*.
complexity	In terms of characters, the number of strokes needed to form a Chinese character. In terms of text, the measure of number of characters and level of grammar. Also, "structural complexity".
density	The proportion of space filled in the formation of a Chinese character.
effective visual field	The very narrow visual field associated with the images of objects focused onto the fovea. Abbreviated as EVF.
eye movement control	The visual skills needed to smoothly and accurately move the eyes while following or locating an object. These skills include eye-tracking, eye-jumps and near-to-far move-ments.
fixation	The extremely critical ability to point the eyes at an object and to voluntarily keep them on the target.
fovea	The area in the retina consisting of a small depression containing cones and where vision is most acute. Approxi-mately 1° of arc.
frequency	Compare "frequency of exposure" and "frequency of occurance".
frequency of exposure	How often a Chinese character or word is repeated in a language textbook, etc. (Sergent & Everson, 1992)
frequency of occurance	How often a character is used in the written language in general based on statistics gatered from corpus studies. (Sergent & Everson, 1992)
logograph	The category of written orthography to which Chinese characters are said to belong.
long-term memory	The term long-term memory refers to the unlimited capacity memory store that can hold information over lengthy periods of time.
metalinguistic awareness	The ability to reflect on and manipulate the structural features of language and treat language itself as an object, as opposed to using language to comprehend and produce utterances. (Wong & Weber, 1986)

morphosyllabic	A term used to define Chinese characters, identifying them as single syllable morphemes. (see DeFrancis, 1989)
parafovea	The area of the retina immediately surrounding the fovea, between 1° and 5° of arc
perceptual span	The portion of a text from which useful information is obtained during a single fixation. (Kess & Miyamoto, 1999)
reading strategies	The problem-solving techniques readers employ to get meaning from a text. (Barnett, 1989, p. 36)
saccade	The precise locating movements used, for example, when jumping from one word to another along a line of print. Saccadic movements are done with the eyes only, the head must remain stationary. They are critical for maintaining one's place on a printed page. Also, "saccadic movement".
scanning	Smooth eye movements enabling a person to maintain inspection of the words on a page or the lines in a design.
short-term memory	Short Term Memory is the part of the memory system where information is stored for roughly 30 seconds. Information can be maintained longer with the use of such techniques as rehearsal.
stimulus onset asynchrony	The time from the start of previous stimulus to the start of next one. Abbreviated as SOA.
subtend	To enclose in an angle. Each side of a triangle subtends the opposite angle. The angular extent of the field of vision.
working memory	Working memory is those mechanisms or processes that are involved in the control, regulation, and active maintenance of task-relevant information in the service of complex cognition, including novel as well as familiar, skilled tasks. (Miyake & Shah, 1999, p. 450)

Sources (unless otherwise indicated):

AlleyDog.com Psychology Glossary, http://www.alleydog.com/glossary/

Glossary of Vision Terms. (2001). Parents Active for Vision Education (PAVE). http://www.pavevision.org/glossary.htm

Sloan, G. D. (1999). *Risk Reduction*, http://www.gdsloan.com/RiskRed.htm

The Free Dictionary Online, http://www.thefreedictionary.com/

Webster's Online Dictionary, http://www.webster-dictionary.org

REFERENCES

Abu-Rabia, S. (1997). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers in reading paragraphs, sentences, and isolated words. *Journal of Psycholinguistic Research, 26*(4), 465-482.

ACTFL Chinese Proficiency Guidelines. (1987). *Foreign Language Annals 20*: 471-487.

Al-Fahid, J. M. (2000). *The Goodman psycholinguistic model of English reading and its applicability to Semitic languages*. Unpublished doctoral dissertation, University of Arizona.

American Council on the Teaching of Foreign Languages. (1986). *ACTFL proficiency guidelines*. New York: American Council on the Teaching of Foreign Languages.

Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press (Oxford University Press).

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A*, 5-28.

Bai, J. (1997). Teaching text structure: Why and how? *Journal of the Chinese Language Teachers Association 32*(3), 31-40.

Barnett, M. A. (1989). *More than meets the eye: Foreign language reading: Theory and practice*. Englewood Cliffs, NJ: Prentice Hall Regents.

Bernhardt, E. B. (1983). *Text processing strategies of native, non-native experienced, and non-native inexperienced readers of German: Findings and implications for the instruction of German as a foreign language*. Unpublished doctoral dissertation, University of Minnesota.

Bernhardt, E. B. (1984). Toward an information processing perspective in foreign language reading. *Modern Language Journal, 68*: 322-381.

Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, empirical, & classroom perspectives*. Norwood, NJ: Ablex Publishing Corp.

Bernhardt, E. B. (1999). If reading is reader-based, can there be a computer-adaptive test of reading? In M. Chalhoub-Deville (Ed.), *Studies in language testing 10: Issues in computer-adaptive testing of reading proficiency* (pp. 1-10). Cambridge: University of Cambridge Press.

Bernhardt, E. B. (2001). Progress and procrastination in second-language reading research [On-line].
Available: http://language.stanford.edu/conferencepapers/ (last accessed 2004.08.01)

Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between first and second language reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics, 16,* 15-34.

Bringjord, E. L. (2001). *Computerized-adaptive versus paper-and-pencil testing environments: An experimental analysis of examinee experience.* Unpublished doctoral dissertation, Rutgers University.

Brown, J. D. (1996). *Understanding research in second language learning.* New York: Cambridge University Press.

Brown, J. D. (1996). *Testing in Language Programs.* Upper Saddle River, NJ: Prentice Hall Regents.

Brown, J. D. (1997). Computers in language testing: Present research and some future directions. *Language Learning & Technology, 1*(1), 44-59.

Cantor, J. & Engle, R. W. (1993). Working memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1101-1114.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275-307). New York: Academic Press.

Carrell, P. L. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics, 12,* 159-179.

Chao, Y. R. (1968). *A grammar of spoken Chinese.* Berkeley: University of California Press.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 16* (pp. 1-58). New York: Academic Press.

Chee, M. W. L., Tan, E. W. L., & Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *Journal of Neuroscience, 19*(8), 3050-3056.

Chen, D. W. (2002). Chinese newspapers discard up-down for left-right reading. *St. Louis Chinese American News* [On-line], 607.
Available: http://www.scanews.com/spot/2002/april/s607/news/news.html

Chen, H-C. (1986). Component detection in reading Chinese characters. In H. S. R. Kao & R. Hoosain (Eds.), *Linguistics, psychology, and the Chinese language* (pp. 1-10). Hong Kong: Centre of Asian Studies.

Chen, H-C. (1992). Reading comprehension in Chinese: Implications from character reading times. In H-C. Chen & O. J. L. Tzeng (Eds.), *Advances in psychology 90: Language processing in Chinese* (pp. 175-205). New York: North-Holland.

Chen, H-C. (Ed.). (1997). *Cognitive processing of Chinese and related Asian languages*. Hong Kong: Chinese University Press.

Chen, H-C., & Chen, M. J. (1988). Directional scanning in Chinese reading. In I-M. Liu, H-C. Chen, & M. J. Chen (Eds.), *Cognitive aspects of the Chinese language, Volume I* (pp. 15-26). Hong Kong: The Chinese University of Hong Kong.

Chen, H-C., & Shu, H. (2001). Lexical activation during the recognition of Chinese characters: Evidence against early phonological activation. *Psychonomic Bulletin & Review, 8*(3), 511-518.

Chen, H-C., & Tang, C-K. (1998). The effective visual field in reading Chinese. *Reading and Writing: An Interdisciplinary Journal, 10*, 245-254.

Chen, H-C., & Tzeng, O. J. L. (Eds.). (1992). *Advances in psychology 90: Language processing in Chinese*. New York: North-Holland.

Chen, L. K., & Carr, H. A. (1926). The ability of Chinese students to read in vertical and horizontal directions. *Journal of Experimental Psychology, 9*, 110-117.

Cheng, C-M. (1992). Lexical access in Chinese: Evidence from automatic activation of phonological information. In H-C. Chen & O. J. L. Tzeng (Eds.), *Advances in psychology 90: Language processing in Chinese* (pp. 67-91). New York: North-Holland.

Chu, M. (Ed.). (1999). *Chinese Language Teachers Association Monograph Series, Volume III: Mapping the course of the Chinese language field*. Kalamazoo, MI: Chinese Language Teachers Association.

Chun, D. M., & Plass, J. L. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology, 1*(1), 60-81.

Clarke, M. A. (1979). Reading in Spanish and English: Evidence from adult ESL students. *Language Learning, 29*(1), 121-150.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments and Computers, 25,* 252-271.

Da, J. (2004). The studies on the theory and methodology of the digitalized Chinese teaching to foreigners. In P. Zhang, T. Xie, and J. Xu (Eds.), *Proceedings of the Fourth International Conference on New Technologies in Teaching and Learning Chinese* (pp. 501-511). Beijing: Tsinghua University Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Daneman, M., & Reingold, E. (1993). What eye fixations tell us about phonological recoding during reading. *Canadian Journal of Experimental Psychology, 47*(2), 153-178.

DeFrancis, J. (1966). Why Johnny can't read Chinese. *Journal of the Chinese Language Teachers Association, 1*(1), 1-20.

DeFrancis, J. (1984). *The Chinese language: Fact and fantasy*. Honolulu, HI: University of Hawaii Press.

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*. Honolulu, HI: University of Hawaii Press.

Dew, J. E. (1994). Back to the basics: Let's not lose sight of what's really important. *Journal of the Chinese Language Teachers Association, 29*(2), 31-46.

Dunkel, P. A. (1999). Considerations in developing or using second/foreign language proficiency computer-adaptive tests. *Language Learning & Technology,* 2(2), 77-93.

Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake and P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 257-297), Cambridge, UK: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211-245.

Everson, M. E. (1986). *The effect of word-unit spacing upon the reading strategies of native and non-native readers of Chinese: An eye-tracking study*. Unpublished doctoral dissertation, The Ohio State University.

Everson, M. E. (1988). Speed and comprehension in reading Chinese: Romanization vs. characters revisited. *Journal of the Chinese Language Teachers Association, 23*(2), 1-15.

Everson, M. E. (1994). Toward a process view of teaching reading in the second language Chinese curriculum. *Theory into Practice, 33*(1), 481-507.

Everson, M. E. (1996). Exploiting background knowledge in the development of Chinese pedagogical reading materials. In S. McGinnis (Ed.), *Pathways to advanced skills, Vol. II: Chinese pedagogy: An emerging field* (pp. 93-105). Columbus, OH: Ohio State University.

Everson, M. E. (1998). Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *The Modern Language Journal, 82*(2), 194-204.

Everson, M. E., & Ke, C. (1997). An inquiry into the reading strategies of intermediate and advanced learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association, 32*(1), 1-20.

Feldman, L. B., & Siok, W. W. T. (1999). Semantic radicals in phonetic compounds: Implications for visual character recognition in Chinese. In J. Wang, A. W. Inhoff, & H-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 19-35). Mahwah, NJ: Lawrence Erlbaum Associates.

Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye-movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 555-568.

Fisher, D.F. (1975). Reading and visual search. *Memory and Cognition, 3*, 188-196.

Gale, A. G., & Johnson, F. (1984). *Theoretical and applied aspects of eye movement research.* New York: North-Holland.

Gao, T. (2001). *Processing numeric table data in electronic formats: Eye movement patterns and searching efficiency.* Unpublished master's thesis, University of North Carolina, Chapel Hill.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6*, 126-135.

Goodman, K. S. (Ed.). (1968). *The psycholinguistic nature of the reading process.* Detroit: Wayne State University Press.

Goodman, K. S. (1994). Reading, writing and written texts: A transactional sociopsycholinguistic view. In R. B. Ruddell, M. R. Ruddell & H. Singer (Eds.), *Theoretical models and processes of reading*. (pp. 1093-1130). Newark, DE: International Reading Association.

Gough, P. B. (1972). One second of reading. In J. F. Kavenagh & I. G. Mattingly (Eds.), *Language by ear and by eye*. (pp. 331-338). Cambridge: MIT Press.

Hannas, W. C. (1997). *Asia's orthographic dilemma*. Honolulu: University of Hawai'i Press.

Harrington, M., & Sawyer, M. (1992). Second language working memory capacity and second language reading skills. *Studies in Second Language Acquisition*, 14, 25-38.

Hayden, J. J. (1998). The influence of a semester abroad on reading proficiency: A descriptive study. *Journal of the Chinese Language Teachers Association 33*(3), 13-24.

Hayden, J. J. (2003). Shocking our students to the next level: Language loss and some implications for teaching Chinese as a foreign language. *Journal of the Chinese Language Teachers Association 38*(3), 1-20.

Hayes, E. B. (1987). The relationship between character complexity and character recognition. *Journal of the Chinese Language Teachers Association, 22*(2), 45-58.

Hayes, E. B. (1988). Encoding strategies used by native and non-native readers of Chinese Mandarin. *Modern Language Journal, 72*(2), 188-194.

Hayes, E. B. (1990). The relationship between 'word length' and memorability among non-native readers of Chinese Mandarin. *Journal of the Chinese Language Teachers Association, 25*, 31-41.

Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47*(2), 201-221.

Henderson, L. (1982). *Orthography and word recognition in reading*. New York: Academic Press.

Ho, S. H. (1975). Comments on teaching reading Chinese. *Collection of Papers presented at Asian Studies on the Pacific Coast: Panel on Chinese Linguistics and Language Teaching*. June 20, 1975, Honolulu, HI. University of Hawai'i.

Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Horiba, Y. (1996a). Comprehension processes in L2 reading: Language competence, textual coherence, and interferences. *Studies in Second Language Acquisition, 18*, 433-473.

Horiba, Y. (1996b). The role of elaboration in L2 text memory: The effect of encoding task on recall of causally related sentences. *The Modern Language Journal, 80*(2), 151-164.

Horiba, Y., van den Broek, P. W., & Fletcher, C. R. (1993). Second language readers' memory for narrative texts: Evidence for structure-preserving top-down processing. Language Learning, 43(3), 345 372.

Horodeck, R. A. (1987). *The role of sound in reading and writing kanji*. Unpublished doctoral dissertation, Cornell University.

Hsueh-chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430.

Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition, 54*, 73-98.

Hudson, T. (1982). The effects of induced schemata on the "short circuit" in L2 reading: Non-decoding factors in L2 reading performance. *Language Learning, 32*(1), 1-31.

Hue, C-w. (1992). Recognition processes in character naming. In H-C. Chen & O. J. L. Tzeng (Eds.), *Advances in psychology 90: Language processing in Chinese* (pp. 93-107). New York: North-Holland.

Hue, C-w., & Erickson, J. R. (1988). Short-term memory for Chinese characters and radicals. *Memory & Cognition, 16*(3), 196-205.

Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.

Hulme, C., Maughan, S., & Brown, D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*, 685-701.

Hulstijn, J. H. (2000). The use of computer technology in experimental studies of second language acquisition: A survey of some techniques and some ongoing studies. *Language Learning & Technology, 3*(2), 32-43.

Hyönä, J. (1995). Do irregular letter combinations attract readers' attention? Evidence from fixation locations in words. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 68-81.

Hyönä, J., Niemi, P., & Underwood, G. (Eds.). (1989). Reading long words embedded in sentences: Informativeness of word halves affects eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 15*(1), 142-152.

Hyönä, J., Tommola, J., & Alaja, A-M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *Quarterly Journal of Experimental Psychology, 48A*, 598-612.

Ihara, L. C. (1983). *Kanji recognition ability: Related to English whole-word ability?* Unpublished master's thesis, Honolulu, University of Hawai'i at Mānoa.

Inhoff, A. W., & Liu, W. (1997). The range of effective vision during the reading of Chinese sentences. In H-C. Chen (Ed.), *Cognitive processing of Chinese and related Asian languages* (pp. 243-265). Hong Kong: Chinese University Press.

Inhoff, A. W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 20-34.

Inhoff, A. W., Liu, W., & Tsang, Z. (1999). Use of prelexical and lexical information during Chinese sentence reading: Evidence from eye-movement studies. In J. Wang, A. W. Inhoff, & H-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 223-238). Mahwah, NJ: Lawrence Erlbaum Associates.

Jackson, N. E., Lu, W-H., & Ju, D. (1994). Reading Chinese and reading English: Similarities, differences, and second language reading. In V. W. Berninger (Ed.), *The varieties of orthographic knowledge* (vol I, pp. 73-110). Dordrecht: Kluwer Academic Publisher.

Jennings, T. A. (2001). *Differential cognitive processing of text and graphically presented items in an on-line testing environment*. Unpublished doctoral dissertation, Arizona State University.

Jiang, S. (2004, April). *A dynamic approach to advanced Chinese: Course design and materials development of a fourth year Chinese language course*. Paper presented at the University of Hawai'i, Honolulu, HI.

Jorden, E. (1975). Reading Japanese: A suggested route for the foreigner. *Journal of the Association of Teachers of Japanese, 10*(2/3), 111-118.

Jorden, E. (1995). Teaching Johnny to read Japanese: Some thoughts on Chinese characters. In M. S. Erbaugh (Ed.), *Pathways to advanced skills, Vol. VI: Difficult characters: Interdisciplinary studies of Chinese and Japanese writing* (pp. 92-104). Columbus, OH: National Foreign Languages Resource Center.

Just, M. A., & Carpenter, P. A. (1978). Inference processes during reading: Reflections from eye fixations. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 157-174). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329-354.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension.* Boston: Allyn and Bacon.

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

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology, 47*(2), 310-339.

Kao, H. S. R., & Hoosain, R. (Eds.). (1986). *Linguistics, psychology, and the Chinese language.* Hong Kong: Centre of Asian Studies.

Ke, C. (1996). An empirical study on the relationship between Chinese and character recognition and production. *Modern Language Journal, 80,* 340-349.

Ke, C. (1998). Effects of language background on the learning of Chinese characters among foreign language students. *Foreign Language Annals, 31*(1), 91-101.

Kennedy, G. A. (1937). A minimum vocabulary in Modern Chinese. *Modern Language Journal, 21(8),* 587-592.

Kess, J. F., & Miyamoto, T. (1999). *The Japanese mental lexicon: Psycholinguistic studies of Kana and Kanji processing.* Philadelphia, PA: John Benjamins Publishing Company.

Khaldieh, S. A. (2001). The relationship between knowledge of iᶜraab, lexical knowledge, and reading comprehension of nonnative readers of Arabic. *Modern Language Journal,* 85(3), 416-431.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49,* 294-303.

Kintsch, W., Patel, V. L., & Ericsson, K. A. (1999). The role of long-term working memory in text comprehension. *Psychologia 42*(4), 186-198.

Kliegl, R. (1981). Automated and interactive analysis of eye fixation data in reading. *Behavior Research Methods & Instrumentation, 13*(2), 115-120.

Ko, S-l. (2003). Western writing style to debut. *Taipei Times* [On-line]. Available: http://www.taipeitimes.com/News/taiwan/archives/2003/08/12/2003063299

Kobrin, J. L. (2000). *The cognitive equivalence of reading comprehension test items via computerized and paper-and-pencil administration.* Unpublished doctoral dissertation, Rutgers University.

Koda, K. (1992). The effects of lower-level processing skills on FL reading performance: Implications for instruction. *The Modern Language Journal, 76*, 502-512.

Koda, K. (1996). L2 word recognition research: A critical review. *The Modern Language Journal, 80*(4), 450-460.

Kolers, P. A. (1970). Three stages in reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading* (pp. 90-118). New York: Plenum Press.

Kowler, E. (Ed.). (1990). *Eye movements and their role in visual and cognitive processes.* New York: Elsevier.

Krashen, S. (1993). *The Power of Reading: Insights from the Research.* Englewood, CO: Libraries Unlimited.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity? *Intelligence, 14*, 389-433.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.

Lee, L-C. C. (1998). *The application of reading strategies by American learners of Chinese as a foreign language when processing narrative and argumentative texts.* Unpublished doctoral dissertation, Indiana University.

Li, Y. (1998). *Teaching and learning L2 reading: The Chinese case.* Unpublished doctoral dissertation, Purdue University.

Light, T. (1975). Controlled composition and reading. *Journal of the Chinese Language Teachers Association, 10*(2), 70-79.

Light, T. (1976). Comparative reading speeds with romanized and character texts. *Journal of the Chinese Language Teachers Association, 11*(1), 1-10.

Liu, I-M. (1988). Context effects on word/character naming: Alphabetic versus logographic languages. In I-M. Liu, H-C. Chen, & M. J. Chen (Eds.), *Cognitive aspects of the Chinese language, Volume I.* (pp. 81-91). Hong Kong: The Chinese University of Hong Kong.

Liu, I-M. (1999). Reading comprehension and reading strategies. In M. Chu (Ed.), *Chinese Language Teachers Association Monograph Series, Volume III: Mapping the Course of the Chinese Language Field* (pp. 39-49). Kalamazoo, MI: Chinese Language Teachers Association.

Liu, I-M., Chen, H-C., & Chen, M. J. (Eds.). (1988). *Cognitive aspects of the Chinese language, Volume I.* Hong Kong: The Chinese University of Hong Kong.

Lüer, G. (1987). *European conference on eye movements.* Göttingen, Germany: C. J. Hogrefe, Inc.

Magloire, J. G. (2002). *Eye movements by good and poor readers during reading of regular and phrase-segmented texts.* Unpublished doctoral dissertation, University of Arizona.

Marian, V. (2000). *Bilingual language processing: Evidence from eye-tracking and functional neuroimaging.* Unpublished doctoral dissertation, Cornell University.

Matsunaga, S. (1994). *The linguistic and psycholinguistic nature of Kanji: Do Kanji represent and trigger only meanings?* Unpublished doctoral dissertation, University of Hawai'i at Mānoa.

Matsunaga, S. (1999). Universality in reading processes: Evidence from an eye-tracking study. *Psychologia, 42*(4), 281-289.

McGinnis, S. (1995). *Does Johnny even want to read?: Student attitudes and approaches in the learning of written Chinese.* Paper presented at the Association for Asian Studies (AAS) Annual Meeting, Washington, DC, April 6-9, 1995.

McGinnis, S. (1997). Tonal spelling versus diacritics for teaching pronunciation of Mandarin Chinese. *Modern Language Journal, 81*(2), 228-236.

McGinnis, S. (1999). Student goals and approaches. In M. Chu (Ed.),*Chinese Language Teachers Association Monograph Series, Volume III: Mapping the Course of the Chinese Language Field* (pp. 151-188). Kalamazoo, MI: Chinese Language Teachers Association.

Mecartty, F. H. (1998 Winter). The effects of proficiency level and passage content on reading skills assessment. *Foreign Language Annals, 31*(4), 521-534.

Miles, W. R., & Shen, E. (1925). Photographic recording of eye movements in the reading of Chinese in vertical and horizontal axes: Method and preliminary results. *Journal of Experimental Psychology, 8*, 344-362.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* Cambridge: Cambridge University Press.

Mori, Y. (1999). Beliefs about language learning and their relationship to the ability to integrate information from word parts and context in interpreting novel *Kanji* words. *Modern Language Journal, 83*(4), 534-547.

Moser, D. (1991). Why Chinese is so damn hard. In V. H. Mair (Ed.), *Schriftfestschrift: Essays in Honor of John DeFrancis, Sino-Platonic Papers, 27* (pp. 59-70). Philadelphia, PA: University of Pennsylvania.

Muller, G. H. (1972). Testing Chinese language proficiency through multiple choice items. *Journal of the Chinese Language Teachers Association, 7*(1), 25-36.

Okita, Y. (1996). *Kanji learning strategies and verbal and visual learning style preferences.* Unpublished doctoral dissertation, University of Hawai'i at Mānoa.

Okurowski, M. E. (1987). Teaching Chinese at the text level through textual cohesion. *Journal of the Chinese Language Teachers Association, 22*(2), 59-80.

Osaka, N. (1993). Reading vertically without a fovea. In S. F. Wright & R. Groner (Eds.), *Facets of Dyslexia and its Remediation* (pp. 257-266). Amsterdam: North-Holland.

Packard, J. L. (1990). Effects of time lag in the introduction of characters into the Chinese language curriculum. *The Modern Language Journal, 74,* 167-175.

Packard, J. L. (2000). *The morphology of Chinese.* Cambridge: Cambridge University Press.

Park, S. (1999). Testing the EFL skills and text hierarchy of the ACTFL reading guidelines. Unpublished master's thesis, Honolulu, University of Hawai'i at Mānoa.

Peng, D. L, & Orchard, L. N., & Stern, J. A. (1983). Evaluation of eye movement variables of Chinese and American readers. *Pavlovian: Journal of Biological Science, 18*(2),94-102.

Peng, 彭聃龄、舒华、陈烜之。（编）。（1997）。汉语认知研究。济南，中华人民共和国：山东教育出版社。

[Péng, D. L., Shū, H., & Chén, H-C. (Biān.). (1997). Hànyǔ Rènzhī Yánjiū. Jìnán, Zhōnghuá Rénmín Gònghéguó: Shāndōng Jiàoyù Chūbǎnshè.]

[Peng, D. L., Shu, H., & Chen, H-C. (Eds.). (1997). *Cognitive Research on Chinese Language*. Jinan, People's Republic of China: Shandong Education Publishers.]

Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies in Reading, 7*(1), 3-24.

Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1), 101-118.

Pollatsek, A., Raney, G. E., Lagasse, L., & Rayner, K. (1993). The use of information below fixation in reading and in visual search. *Canadian Journal of Experimental Psychology, 47*(2), 179-200.

Rayner, K. (Ed.) (1983). *Eye movements in reading: Perceptual and language processes*. New York: Academic Press.

Rayner, K. (Ed.) (1992). *Eye movements and visual cognition: Scene perception and reading*. New York: Springer-Verlag.

Rothkopf, E. Z. (1978). Analyzing eye movements to infer processing styles during learning from text. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 209-223). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Saito, Y., Garza, T. J., & Horwitz, E. K. (1999). Foreign language reading anxiety. *Modern Language Journal, 83*(2), 202-218.

Samuels, S. J., & Kamil, M. L. (1984). Models of the Reading Process. In P. David Pearson (Ed.) *Handbook of Reading Research*. New York: Longman, 185-224.

Senders, J. W., Fisher, D. F., & Monty, R. A. (1978). *Eye movements and their higher psychological functions*. New York: Lawrence Erlbaum Associates, Publishers.

Seo, H., & Lee, C. (2002). Head-free reading of horizontally and vertically arranged texts. *Vision Research, 42*, 1325-1337.

Sergent, W. K., Jr. (1990). A study of the oral reading strategies of advanced and highly advanced second language readers of Chinese. *Dissertation Abstracts International, 51*(03), 772.

Sergent, W. K., and Everson, M. E. (1992). The effects of frequency and density on character recognition speed and accuracy by elementary and advanced L2 readers of Chinese. *Journal of the Chinese Language Teachers Association* 27(1/2), 29- 44.

Shen, D., & Forster, K. I. (1999). Masked phonological priming in reading Chinese words depends on the task. *Language and Cognitive Processes, 14*(5/6), 429-459.

Shen, E. (1927). An analysis of eye movements in the reading of Chinese. *Journal of Experimental Psychology, 10*, 158-183.

Shen, H. (2000). The interconnections of reading text based writing and reading comprehension among college intermediate learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association, 35*(3), 29-48.

Shu, H., & Anderson, R. C. (1999). Learning to read Chinese: The development of metalinguistic awareness. In J. Wang, A. W. Inhoff, & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 1-18). Mahwah, NJ: Lawrence Erlbaum Associates.

Smith, F. (1973). *Psycholinguistics and reading*. New York: Holt, Rinehart and Winston.

Smith, F. (1988). *Understanding reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Smith, F. (1994). *Understanding reading: A psycholinguistic analysis of reading and learning to read*, 5th edition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Spring, M. (1999). Improving reading instruction in upper-level Chinese courses: Challenges and possibilities. In M. Chu (Ed.), *Chinese Language Teachers Association Monograph Series, Volume III: Mapping the Course of the Chinese Language Field* (pp. 51-69). Kalamazoo, MI: Chinese Language Teachers Association.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*, 32-71.

Stern, J. A. (1978). Eye movements, reading, and cognition. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 145-155). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Su, I-R. (2001). Transfer of sentence processing strategies: A comparison of L2 learners of Chinese and English. *Applied Psycholinguistics, 22*, 83-112.

Sun, F. (1993). Eye movements in reading Chinese: Paragraphs, single paragraphs, and Pinyin. In S. F. Wright & R. Groner (Eds.), *Facets of Dyslexia and its Remediation* (pp. 245-256). Amsterdam: North-Holland.

Sun, F., & Feng, D. (1999). Eye movements in reading Chinese and English text. In J. Wang, A. W. Inhoff, & H-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 189-205). Mahwah, NJ: Lawrence Erlbaum Associates.

Sun, F. C., Morita, M., & Stark, L. W. (1985). Comparative patterns of reading eye movement in Chinese and English. *Perception & Psychophysics, 37*, 502-506.

Sun, Y. (1992). *First and second language word recognition processes in Chinese*. Unpublished doctoral dissertation, University of Toronto.

Swanson, H. L., & Berninger, V. (1995). The role of working memory in skilled and less skilled readers comprehension. *Intelligence, 21*, 83-108.

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language, 37*, 41-57.

Tan, L. H., & Perfetti, C. A. (1998). Phonological codes as early sources of constraint in Chinese word identification: A review of current discoveries and theoretical accounts. *Reading and Writing, 10*(3/5), 165-200.

Tan, L. H., Liu, H-L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J-H. (2001). The neural system underlying Chinese logograph reading. *NeuroImage, 13*, 836-846.

Tang, C-k., Yeung, L-h. A., & Chen, H-C. (1997). The effective visual field in Chinese reading. In In H-C. Chen (Ed.), *Cognitive processing of Chinese and related Asian languages* (pp. 267-286). Hong Kong: Chinese University Press.

Tao, L., & Healy, A. F. (2002). The unitization effect in reading Chinese and English text. *Scientific Studies in Reading, 6*(2), 167-197.

Taylor, C., Jamieson, J., Eignor, D., & Kirsch, I. (1998). *The relationship between computer familiarity and performance on computer-based TOEFL test tasks*. (ETS Research Rep. No. 61). Princeton, NJ: Educational Testing Service.

Taylor, I., & Taylor, M. M. (1995). *Studies in written language and literacy 3: Writing and literacy in Chinese, Korean and Japanese*. Amsterdam: John Benjamins.

Tien, S. O. L. (1983). *Chinese adult readers: A psycholinguistic and transactional study of the reading process in Chinese, with comparison to English*. Unpublished doctoral dissertation, Michigan State University, East Lansing.

Tsai, C-h. (1996). Tsai's list of Chinese words. [On-line].
Available: http://technology.chtsai.org/wordlist/ (Last accessed 2004.09.14)

Tsai, C-h. (2001). *Word identification and eye movements in reading Chinese: A modeling approach*. University of Illinois at Urbana-Champaign.

Tzeng, O. J. L., Hung, D. L., & Garro, L. (1978). Reading the Chinese characters: An information processing view. *Journal of Chinese Linguistics, 6*, 287-305.

Tzeng, O. J. L., Hung, D. L., & Wang, W. S-Y. (1977). Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Memory and Learning, 3*, 621-630.

Tu, H. T. C. (1930). The effects of different arrangements of the Chinese language upon speed and comprehension of silent reading. *The Pedagogical Seminary and Journal of Genetic Psychology: Child Behavior, Animal Behavior, and Comparative Psychology, 38(1-4)*, 321-337.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127-145.

Underwood, G., & Everatt, J. (1992). The role of eye movements in reading: Some limitations of the Eye-Mind Assumption. In E. Chekaluk & K. Llewellyn (Eds.), *Advances in Psychology 88: The role of eye movements in perceptual processes* (pp. 111-169). Amsterdam: North Holland.

Unger, J. M. (2004). Ideogram: Chinese characters and the myth of disembodied meaning. Honolulu: University of Hawai'i Press.

Walczyk, J., Kelly, K., Meche, S., & Braud, H. (1999). Time limitations enhance reading comprehension. *Contemporary Educational Psychology, 24,* 156-165.

Walker, G. (1984). 'Literacy' and 'reading' in a Chinese language program. *Journal of the Chinese Language Teachers Association, 19*(1), 67-84.

Walter, C. (2000). *The involvement of working memory in reading in a foreign language.* Unpublished doctoral dissertation, Cambridge University.

Wang, A., & Thomas, M. (1992). The effect of imagery-based mnemonics on the long-term retention of Chinese characters. *Language Learning 42*(3), 359-376.

Wang, F. C. (1934). *An experimental study of eye movements in the silent reading of Chinese.* Unpublished doctoral dissertation, University of Chicago.

Wang, F. C. (1935). An experimental study of eye movements in the silent reading of Chinese. *Elementary School Journal, 35,* 527-539.

Wang, J., Inhoff, A. W., & Chen, H-C. (1999). *Reading Chinese script: A cognitive analysis.* Mahwah, NJ: Lawrence Erlbaum Associates.

Wang, M., Perfetti, C. A., Liu, Y. (2003). Alphabetic readers quickly acquire orthographic structure in learning to read Chinese. *Scientific Studies in Reading, 7*(2), 183-208.

Wang, W. S-Y. (1973). The Chinese language. *Scientific American, 228*(2), 50-63.

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition 19,* 287-307.

Watanabe, Y., & Yao, T-c. (1998). Examining the validity of the *Computer-Adaptive Test for Reading Chinese.* Paper presented at the Second Annual College-wide Conference for Students in Languages, Linguistics, and Literatures, University of Hawai'i, Honolulu, HI, March 1998.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology 49A*, 51-79.

Wolf, D. F. (1993). A comparison of assessment tasks used to measure FL reading comprehension. *The Modern Language Journal, 77*(4), 473-488.

Wong, C. H., & Weber, R-M. (1986). Metalinguistic awareness, Chinese orthography, and learning to read. In H. S. R. Kao & R. Hoosain (Eds.), *Linguistics, psychology, and the Chinese language* (pp. 39-45). Hong Kong: Centre of Asian Studies.

Wong, K. F. E. (2000). *The time course of semantic activation in reading Chinese two-character words*. Unpublished doctoral dissertation, Chinese University of Hong Kong (People's Republic of China).

Wong, K. F. E., & Chen, H-C. (1999). Orthographic and phonological processing in reading Chinese text: Evidence from eye fixations. *Language and Cognitive Processes, 14*(5/6), 461-480.

Wright, S. F., & Groner, R. (Eds.). (1993). *Facets of Dyslexia and its Remediation*. Amsterdam: North-Holland.

Xiao, Y. (2002). The effect of character density on learning Chinese as a foreign language. *Journal of the Chinese Language Teachers Association 37*(3), 71-83.

Xu, J. (1998). *A study of the reading process in Chinese through detecting errors in a meaningful text*. Unpublished doctoral dissertation, University of Arizona.

Yamamoto, K. (1980). *Relative recognition speed of hiragana, katakana and kanji by adult native and non-native speakers*. Unpublished master's thesis, Honolulu, University of Hawai'i at Mānoa.

Yang, H-M. (1994). *Word perception and eye movements in Chinese reading*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Yang, H-M., & McConkie, G. W. (1999). Reading Chinese: Some basic eye-movement characteristics. In J. Wang, A. W. Inhoff, & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 207-222). Mahwah, NJ: Lawrence Erlbaum Associates.

Yang, J. (2000). Orthographic effect on word recognition by learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association, 35*(2), 1-18.

Yano, Y., Long, M. H., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning 44*(2), 189-219.

Yao, T-c. (1994). Computer-Adaptive Test for Reading Chinese [Computer software]. Honolulu: University of Hawai'i at Mānoa.

Yao, T-c. (1995). A computer-adaptive test for reading Chinese (CATRC): A preliminary report. *Journal of the Chinese Language Teachers Association 30*(1), 1-20.

Yarbus, A. L. (1967). *Eye movements and vision.* New York: Plenum Press.

Yim-Ng, Y-Y., Varley, R., & Andrade, J. (2000). Contribution of finger tracing to the recognition of Chinese characters. *International Journal of Language & Communication Disorders, 35*(4), 561-571.

Yu B., Zhang, W., Jing, Q., Peng, R., Zhang, G., & Simon, H. H. (1985). STM capacity for Chinese and English language materials. *Memory & Cognition, 13*(3), 202-207.

Zhang, G, & Simon, H. H. (1985). STM capacity for Chinese words and idioms: Chunking and acoustic loop hypotheses. *Memory & Cognition, 13*(3), 193-201.

Zhang, S., & Perfetti, C. A. (1993). The tongue-twister effect in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(5), 1082-1093.

Zheng, Y., Liu, C., & Ding, X. (2002). Single character type identification. *Proceedings of SPIE Document Recognition and Retrieval IX, Vol. 4670,* Jan. 21-22, 2002: 49-56. [On-line]. Available: http://www.ece.umd.edu/~zhengyf/SingleCharacterTypeIdentification_SPIE2002.pdf (last accessed 2004.08.15)

Zhou, X., & Marslen-Wilson, W. (1999). Sublexical processing in reading Chinese. In J. Wang, A. W. Inhoff, & H-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 37-63). Mahwah, NJ: Lawrence Erlbaum Associates.

Zhou, X., Marslen-Wilson, W., Taft, W., & Shu, H. (1999). Morphology, orthography, and phonology in reading Chinese compound words. *Language and Cognitive Processes, 14*(5/6), 525-565.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin 123,* 162-185.