



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2009년12월16일
(11) 등록번호 10-0932141
(24) 등록일자 2009년12월08일

(51) Int. Cl.

G09B 3/06 (2006.01) G09B 3/00 (2006.01)

- (21) 출원번호 10-2008-7000607
- (22) 출원일자 2006년05월24일
심사청구일자 2008년01월09일
- (85) 번역문제출일자 2008년01월09일
- (65) 공개번호 10-2008-0021771
- (43) 공개일자 2008년03월07일
- (86) 국제출원번호 PCT/JP2006/310372
- (87) 국제공개번호 WO 2006/134759
국제공개일자 2006년12월21일
- (30) 우선권주장
JP-P-2005-174675 2005년06월15일 일본(JP)
- (56) 선행기술조사문헌
JP2001056634 A
JP2001318792 A
JP2003203071 A

(73) 특허권자

각코호진 와세다다이가쿠

일본국 도쿄도 신주쿠구 도쓰카마치 1쥬메 104번지

가부시카이가이사 교우이쿠 소쿠테이 켄큐쇼

일본 도쿄도 신주쿠구 요코테라쥬 55

(72) 발명자

안토니 로렌스

일본 도쿄도 신주쿠구 오쿠보 3쥬메 4반 1고 각코호진와세다다이가쿠 리코가쿠주쓰인내

하야시 노리오

일본 도쿄도 신주쿠구 요코테라쥬 55 가부시카이가이사 교우이쿠소쿠테이 켄큐쇼내

(뒷면에 계속)

(74) 대리인

유미특허법인

전체 청구항 수 : 총 12 항

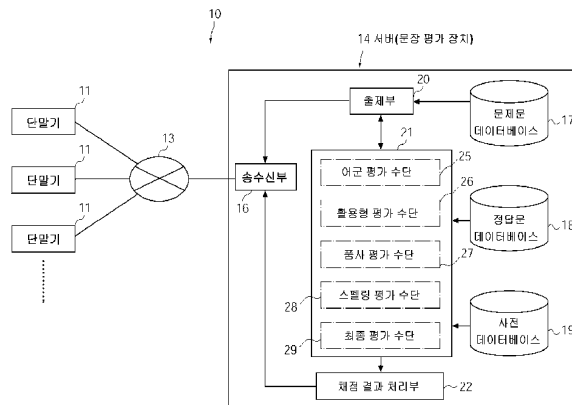
심사관 : 류제준

(54) 문장 평가 장치 및 문장 평가 프로그램

(57) 요약

문장 평가 장치(14)는, 이용자가 작성한 해답문과, 해당 해답문에 대응한 정답문 데이터베이스(18)의 각 정답문을 각각 대비하고, 이들 정답문마다, 전술한 기준에 의한 감점법으로 채점을 행하고, 그 중에서 최고 득점을 상기 해답문의 득점으로 하는 채점부(21)를 구비하고 있다. 이 채점부(21)는, 정답문 및 해답문 내의 어군의 상이에 착안하여 채점하는 어군 평가 수단(25)과, 상기 정답문 및 상기 해답문 내의 단어의 품사에 착안하여 채점하는 품사 평가 수단(27)과, 이들 각 평가 수단(25 ~ 28)에 의한 채점을 행한 후에 최종적인 채점을 행하는 최종 평가 수단(29)을 구비하고 있다.

도 1



(72) 발명자

소네 가즈오

일본 도쿄도 신쥬쿠구 요코테라쵸 55 가부시키가이
샤 교우이쿠소쿠테이 켄큐쇼내

야마나시 도시오

일본 도쿄도 신쥬쿠구 요코테라쵸 55 가부시키가이
샤 교우이쿠소쿠테이 켄큐쇼내

마에다 가오리

일본 도쿄도 신쥬쿠구 요코테라쵸 55 가부시키가이
샤 교우이쿠소쿠테이 켄큐쇼내

야마키시 야스코

일본 도쿄도 신쥬쿠구 요코테라쵸 55 가부시키가이
샤 교우이쿠소쿠테이 켄큐쇼내

쓰바키모토 야요이

일본 도쿄도 신쥬쿠구 요코테라쵸 55 가부시키가이
샤 교우이쿠소쿠테이 켄큐쇼내

특허청구의 범위

청구항 1

원문이 되는 문제문에 대하여 소정 언어로 번역된 해당문의 평가를 행하는 문장 평가 장치에 있어서,
상기 문제문에 대한 정답 번역문이 되는 정답문이 기억된 정답문 데이터 베이스와, 상기 정답문 데이터 베이스에 기억된 정답문과, 사용자에게 의해 작성된 상기 해당문의 대비에 의해, 상기 해당문에 대한 채점 처리를 행하는 채점부를 구비하고,

상기 채점부는, 상기 정답문 데이터 베이스에 기억된 정답문 및 상기 사용자에게 의해 작성된 해당문 내의 어군(語群)의 상이에 착안하여 채점하는 어군 평가 수단을 구비하고,

상기 어군 평가 수단은, 상기 정답문 데이터 베이스에 기억된 정답문 및 상기 사용자에게 의해 작성된 해당문의 대비에 의해, 서로 동일한 어군을 결정하거나, 또는 상기 동일한 어군에 이웃하는 전(前) 또는 후(後)의 단어에 대해, 단어의 활용형이 기억된 사전 데이터 베이스에 기초하여 활용형이 상이한지 여부를 판단하거나, 미리 기억된 방법에 기초하여 스펠링 미스인지 여부를 판단하여, 상기 단어가 활용형의 상이 또는 스펠링 미스인 경우에 상기 단어를 상기 동일한 어군에 포함시켜 근사한 어군으로 결정한 다음, 상기 동일 또는 근사한 어군을 평가 대상으로 하고, 상기 정답문 및 상기 해당문 사이에서의 각 평가 대상어의 위치의 상이 또는 상기 평가 대상어의 배열 순서의 상이에 따라 점수에 차이를 부여하는, 문장 평가 장치.

청구항 2

제1항에 있어서,

상기 어군 평가 수단은, 상기 동일한 어군과 근사한 어군으로 점수 차이를 부여하는, 문장 평가 장치.

청구항 3

삭제

청구항 4

제1항에 있어서,

소정의 단어가 그 활용형과 함께 기억된 사전 데이터 베이스를 구비하고,

상기 채점부는, 상기 사전 데이터 베이스를 기초로 하여, 상기 정답문 데이터 베이스에 기억된 정답문 및 상기 해당문 내의 단어의 활용형의 상이에 착안하여 채점하는 활용형 평가 수단을 구비하고,

상기 활용형 평가 수단은, 이미 평가된 평가 대상어를 제외한 단어 중에서, 상기 정답문 및 상기 해당문의 대비에 의해, 상기 사전 데이터 베이스를 기초로 활용형의 상이에 불과한 단어를 새로운 평가 대상으로서 추출하고, 상기 새로운 평가 대상어에 대하여, 상기 정답문 및 상기 해당문 사이에서의 위치의 상이에 따라 점수에 차이를 부여하는, 문장 평가 장치.

청구항 5

제1항 또는 제4항에 있어서,

상기 채점부는, 상기 해당문 내의 단어의 스펠링 미스에 착안하여 채점하는 스펠링 평가 수단을 구비하고,

상기 스펠링 평가 수단은, 이미 평가된 평가 대상어를 제외한 단어 중에서, 상기 정답문 데이터 베이스에 기억된 정답문 및 상기 해당문 사이에서, 미리 기억된 방법에 기초하여 철자가 근사한 단어를 새로운 평가 대상으로서 추출하고, 해당되는 새로운 평가 대상어에 대하여, 상기 정답문 및 상기 해당문 사이에서의 위치의 상이에 따라 점수에 차이를 부여하는, 문장 평가 장치.

청구항 6

제1항, 제2항 및 제4항 중 어느 한 항에 있어서,

소정의 단어가 그 품사의 종류와 함께 기억된 사전 데이터 베이스를 구비하고,

상기 채점부는, 상기 사전 데이터 베이스를 기초로 하여, 상기 정답문 데이터 베이스에 기억된 정답문 및 상기 해답문 내의 단어의 품사의 동일성에 착안하여 채점하는 품사 평가 수단을 구비하고,

상기 품사 평가 수단은, 이미 평가된 평가 대상어를 제외한 단어에 대해, 상기 사전 데이터 베이스에 기억된 품사의 종류를 기초로 품사의 판별을 행하고, 상기 정답문 및 상기 해답문의 대비에 의해, 이들 사이에서 품사의 종류가 일치하는 단어를 새로운 평가 대상어로서 추출하고, 상기 새로운 평가 대상어에 대하여, 상기 정답문 및 상기 해답문 사이에서의 위치의 상이 및 품사의 종류 중 하나 이상에 따라 점수에 차이를 부여하는, 문장 평가 장치

청구항 7

제1항, 제2항 및 제4항 중 어느 한 항에 있어서,

상기 채점부는, 상기 정답문 데이터 베이스에 기억된 정답문 및 상기 해답문 내에서, 상기 평가 대상어를 제외한 나머지 단어의 개수에 따라 점수에 차이를 부여하는 최종 평가 수단을 구비하는, 문장 평가 장치.

청구항 8

제7항에 있어서,

상기 최종 평가 수단은, 상기 나머지 단어가, 상기 정답문 데이터 베이스에 기억된 정답문 및 상기 해답문 사이에 대응하는 위치에 존재하는지의 여부에 의해 점수에 차이를 두는, 문장 평가 장치.

청구항 9

제7항에 있어서,

소정의 단어가 그 품사의 종류와 함께 기억된 사전 데이터 베이스를 구비하고,

상기 최종 평가 수단은, 상기 나머지 단어에 대해, 상기 사전 데이터 베이스를 기초로 품사의 판별을 행하고, 판별된 품사의 종류에 따라 점수에 차이를 부여하는, 문장 평가 장치.

청구항 10

제1항, 제2항 및 제4항 중 어느 한 항에 있어서,

상기 정답문 데이터 베이스에는, 정답문이 복수개 기억되고,

상기 채점부에서는, 각 정답문에 대하여, 각 평가 수단에 의한 상기 해답문의 채점 처리가 각각 행해지고, 정답문마다 구해진 채점 결과 중에서 가장 양호한 평가를 상기 해답문의 평가로 하는, 문장 평가 장치.

청구항 11

원문이 되는 문제문에 대한 정답 번역문이 되는 정답문과, 상기 문제문에 대하여 소정 언어로 번역된 해답문을 대비함으로써, 상기 해답문의 평가를 컴퓨터에 실행시키는 문장 평가 프로그램이 기억된, 컴퓨터로 판독 가능한 기억 매체로서,

미리 정답문 데이터 베이스에 기억된 상기 정답문 및 사용자에게 의해 작성된 해답문의 대비에 의해, 서로 동일한 어근을 결정하거나, 또는 상기 동일한 어근에 이웃하는 전(前) 또는 후(後)의 단어에 대해, 단어의 활용형이 기억된 사전 데이터 베이스에 기초하여 활용형이 상이한지 여부를 판단하거나, 미리 기억된 방법에 기초하여 스펠링 미스인지 여부를 판단하여, 상기 단어가 활용형의 상이 또는 스펠링 미스인 경우에 상기 단어를 상기 동일한 어근에 포함시켜 근사한 어근으로 결정한 다음, 상기 동일 또는 근사한 어근을 평가 대상어로서 추출하고, 상기 정답문 및 상기 해답문 사이에서의 각 평가 대상어의 위치의 상이 또는 상기 평가 대상어의 배열 순서의 상이에 따라 점수에 차이를 부여하는 어근 평가 수단으로서 상기 컴퓨터를 기능시키기 위한 문장 평가 프로그램이 기억된, 컴퓨터로 판독 가능한 기억 매체.

청구항 12

원문이 되는 문제문에 대한 정답 번역문이 되는 정답문과, 상기 문제문에 대하여 소정 언어로 번역된 해답문을 대비함으로써, 상기 해답문의 평가를 컴퓨터에 실행시키는 문장 평가 프로그램이 기억된, 컴퓨터로 판독 가능한

기억 매체로서,

미리 정답문 데이터 베이스에 기억된 상기 정답문 및 사용자에게 의해 작성된 해답문의 대비에 의해, 서로 동일한 어군을 결정하거나, 또는 상기 동일한 어군에 이웃하는 전(前) 또는 후(後)의 단어에 대해, 단어의 활용형이 기억된 사전 데이터 베이스에 기초하여 활용형이 상이한지 여부를 판단하거나, 미리 기억된 방법에 기초하여 스펠링 미스인지 여부를 판단하여, 상기 단어가 활용형의 상이 또는 스펠링 미스인 경우에 상기 단어를 상기 동일한 어군에 포함시켜 근사한 어군으로 결정한 다음, 상기 동일 또는 근사한 어군을 평가 대상어로서 추출하고, 상기 정답문 및 상기 해답문 사이에서의 각 평가 대상어의 위치의 상이 또는 상기 평가 대상어의 배열 순서의 상이에 따라 점수에 차이를 부여하는 어군 평가 수단과,

상기 어군 평가 수단으로 평가된 평가 대상어를 제외한 단어에 대해, 소정의 단어가 그 품사의 종류와 함께 기억된 사전 데이터 베이스를 기초로 품사의 판별을 행하고, 상기 정답문과 상기 해답문의 대비에 의해, 이들 사이에서 품사의 종류가 일치하는 단어를 새로운 평가 대상어로서 추출하고, 상기 새로운 평가 대상어에 대하여, 상기 정답문 및 상기 해답문 사이에서의 위치의 상이 및 품사의 종류 중 하나 이상에 따라 점수에 차이를 부여하는 품사 평가 수단과,

각각의 상기 어군 평가 수단 및 상기 품사 평가 수단에서의 평가 대상어를 제외한 나머지 단어의 개수에 따라 점수에 차이를 두는 최종 평가 수단으로서 상기 컴퓨터를 기능시키기 위한 문장 평가 프로그램이 기억된, 컴퓨터로 판독 가능한 기억 매체.

청구항 13

제12항에 있어서,

상기 문장 평가 프로그램은 추가로, 각각의 상기 어군 평가 수단, 상기 품사 평가 수단, 및 상기 최종 평가 수단을, 상기 정답문 데이터 베이스에 기억된 복수개의 정답문 각각에 대하여 실행하고, 이들 정답문마다 구해진 채점 결과 중에서 가장 양호한 평가를 상기 해답문의 평가로 결정하는 수단으로서 상기 컴퓨터를 기능시키기 위한 프로그램을 포함하는, 컴퓨터로 판독 가능한 기억 매체.

명세서

기술분야

<1> 본 발명은 문장 평가 장치 및 문장 평가 프로그램에 관한 것으로서, 보다 상세하게는, 원문이 되는 문제문에 대하여 소정 언어로 번역된 해답문의 평가를 객관적이면서 간단하게 행할 수 있는 문장 평가 장치 및 문장 평가 프로그램에 관한 것이다.

배경기술

<2> 중학교나 고등학교 등의 영어 교육 현장에서, 학생에게 출제된 영작문의 테스트 문제의 채점은, 교사의 독자적인 채점 기준으로 행해지는 경우가 많다. 즉, 학생이 작성한 해답문을 교사가 채점할 때는, 해답문과 정답문을 보고 비교하여, 이들의 상이의 정도에 따라 주관적인 판단으로 행해진다. 그런데, 문제문의 정답이 되는 영문은, 여러 개 존재하므로, 이들 정답문을 모두 학생의 해답문과 비교하는 것은 막대한 시간이 걸리고, 실제 교육 현장에서는 실행 불가능하다. 따라서, 채점 교사는, 한 문장 또는 수개의 문장의 정답문을 모범 해답으로 하고, 한정된 모범 해답을 학생의 해답문과 비교하여 채점하는 것이 일반적이며, 난이도가 높은 문제문일수록, 다수의 학생의 해답문 중에서, 모범 해답과 완전하게 일치하는 것은 적어지게 된다. 또한, 모범 해답문과 해답문의 상이는, 단어의 상이, 단어 위치의 상이, 단어 스펠링의 상이, 단어의 누락, 단어의 활용형의 상이, 시제의 상이 등, 다양하므로, 패턴화하기 곤란하다. 이상과 같은 이유로 인하여, 여러 명의 교사가 테스트 문제의 채점을 분담하도록 한 경우, 모범 해답과 해답문의 상이점에 관한 평가는, 각 교사의 주관에 맡기고, 각 교사 사이에서 채점 기준이 완전히 일치하지는 않게 되기 쉽다. 또한, 한 사람의 교사가 모든 학생의 채점을 행한 경우에도, 채점의 처음과 마지막에서는 채점 기준이 바뀌는 경우도 많이 있다. 그러므로, 대체로, 사람에 의한 영작문의 채점에서는, 객관성을 담보하는 것이 곤란한 것이 현재 상황이다.

<3> 그런데, 소정의 번역자나 자동 번역 장치가 작성한 번역문에 대한 평가를 행하는 번역 능력 평가 시스템이 알려져 있고(특히 문헌 1 참조), 이 시스템을 영작문의 채점에 응용하여 적용하는 것도 생각할 수 있다. 이 번역 능력 평가 시스템은, 일본어 테스트문에 대한 정답 후보가 되는 영문(정답문)이 다수 기억되어 있고, 각 정답문

과, 일본어 테스트문에 대응하여 작성된 번역문이 각각 대비되고, 정답문 각각에 대하여, 번역 정답율로 불리는 지표가 구해지고, 그 중 최고의 번역 정답율이 채택된다. 이 번역 정답율은, 번역문을 구성하는 단어와 동일한 단어가, 정답문과 동일한 위치에 존재하는지의 여부를 기준으로 산출된다.

<4> [특허 문헌 1]: 일본국 특개 2002-140326호 공보

발명의 상세한 설명

- <5> 그러나, 상기 번역 능력 평가 시스템에서는, 정답문과 번역문 사이에서, 서로 상대(相對)하는 단어의 동일성만으로 번역문의 평가가 행해지므로, 실정에 맞지 않는 번역문의 평가가 행해질 우려가 있다. 예를 들면, 정답문과 번역문 사이에서, 서로 동일한 어군(語群)이 존재하지만, 이들 어군의 위치가 상이한 경우, 동일한 어군을 도출한 능력이 전혀 고려되지 않고, 전혀 다른 어군을 배치한 경우와 동등하게 취급되는 문제가 있다.
- <6> 본 발명은, 이와 같은 문제에 착안하여 고안된 것이며, 그 목적은, 원문이 되는 문제문에 대하여 소정 언어로 번역된 해당문의 평가를 실정에 따라 객관적으로 행할 수 있는 문장 평가 장치 및 문장 평가 프로그램을 제공하는 것에 있다.
- <7> (1) 상기 목적을 달성하기 위하여, 원문이 되는 문제문에 대하여 소정 언어로 번역된 해당문의 평가를 행하는 문장 평가 장치에 있어서,
- <8> 상기 문제문에 대한 정답 번역문이 되는 정답문이 기억된 정답문 데이터 베이스와, 상기 정답문과 상기 해당문의 대비에 의해, 상기 해당문에 대한 채점 처리를 행하는 채점부를 구비하고,
- <9> 상기 채점부는, 상기 정답문 및 상기 해당문 내의 어군의 상이에 착안하여 채점하는 어군 평가 수단을 구비하고,
- <10> 상기 어군 평가 수단은, 상기 정답문 및 해당문 사이에 동일 또는 근사한 어군을 추출하여 해당되는 동일 또는 근사한 어군을 평가 대상어로 하고, 상기 정답문 및 상기 해당문 사이에서의 각 평가 대상어의 위치의 상이 또는 상기 평가 대상어의 배열 순서의 상이에 따라 점수에 차이를 두는 구성을 채용하고 있다.
- <11> (2) 또한, 상기 어군 평가 수단은, 상기 동일한 어군과 근사한 어군으로 점수에 차이를 두는 구성을 채용하고 있다.
- <12> (3) 또한, 상기 어군 평가 수단은, 정답문 및 해당문 사이에서의 어군의 상이가, 해당 어군을 구성하는 단어 사이에서의 활용형의 상이 및/또는 스펠링 미스에 지나지 않는 경우에, 근사한 어군으로 하는 구성을 채용하고 있다.
- <13> (4) 또한, 소정의 단어가 그 활용형과 함께 기억된 사전 데이터 베이스를 구비하고,
- <14> 상기 채점부는, 상기 사전 데이터 베이스를 기초로 하여, 상기 정답문 및 상기 해당문 내의 단어의 활용형의 상이에 착안하여 채점하는 활용형 평가 수단을 구비하고,
- <15> 상기 활용형 평가 수단은, 이미 평가된 평가 대상어를 제외한 단어 중에서, 상기 정답문 및 해당문 사이에서 활용형이 상이한 단어를 새로운 평가 대상어로서 추출하고, 해당되는 새로운 평가 대상어에 대하여, 상기 정답문 및 상기 해당문 사이에서의 위치의 상이에 따라 점수에 차이를 두는 구성을 병행하여 채용할 수 있다.
- <16> (5) 또한, 상기 채점부는, 상기 해당문 내의 단어의 스펠링 미스에 착안하여 채점하는 스펠링 평가 수단을 구비하고,
- <17> 상기 스펠링 평가 수단은, 이미 평가된 평가 대상어를 제외한 단어 중에서, 상기 정답문 및 상기 해당문 사이에 철자가 근사한 단어를 새로운 평가 대상어로서 추출하고, 해당되는 새로운 평가 대상어에 대하여, 상기 정답문 및 상기 해당문 사이에서의 위치의 상이에 따라 점수에 차이를 두는 구성을 병행하여 채용할 수 있다.
- <18> (6) 또한, 소정의 단어가 그 품사의 종류와 함께 기억된 사전 데이터 베이스를 구비하고,
- <19> 상기 채점부는, 상기 사전 데이터 베이스를 기초로 하여, 상기 정답문 및 상기 해당문 내의 단어의 품사의 동일성에 착안하여 채점하는 품사 평가 수단을 구비하고,
- <20> 상기 품사 평가 수단은, 이미 평가된 평가 대상어를 제외한 단어 중에서, 상기 정답문 및 해당문 사이에 품사의 종류가 일치하는 단어를 새로운 평가 대상어로서 추출하고, 해당되는 새로운 평가 대상어에 대하여, 상기 정답문 및 상기 해당문 사이에서의 위치의 상이 및/또는 품사의 종류에 따라 점수에 차이를 두는 구성을 병행하여

채용할 수 있다.

- <21> (7) 또한, 상기 채점부는, 상기 정답문 및 상기 해답문 내에서, 상기 평가 대상어를 제외한 나머지 단어의 개수에 따라 점수에 차이를 두는 최종 평가 수단을 구비하는 구성을 병행하여 채용할 수 있다.
- <22> (8) 여기서, 상기 최종 평가 수단은, 상기 나머지 단어가, 상기 정답문 및 상기 해답문 사이에 대응하는 위치에 존재하는지의 여부에 의해 점수에 차이를 두는 구성을 채용하는 것이 바람직하다.
- <23> (9) 또한, 상기 최종 평가 수단은, 상기 나머지 단어의 품사의 종류에 따라 점수에 차이를 두는 구성을 채용할 수도 있다.
- <24> (10) 이상에서, 상기 정답문 데이터 베이스에는, 상기 정답문이 복수개 기억되고,
- <25> 상기 채점부에서는, 각 정답문에 대하여, 각 평가 수단에 의한 상기 해답문의 채점 처리가 각각 행해지고, 정답문마다 구해진 채점 결과 중에서 가장 양호한 평가를 상기 해답문의 평가로 하는 구성을 채용하면 된다.
- <26> (11) 또한, 본 발명은, 원문이 되는 문제문에 대한 정답 번역문이 되는 정답문과, 상기 문제문에 대하여 소정 언어로 번역된 해답문을 대비함으로써, 상기 해답문의 평가를 컴퓨터에 실행시키는 문장 평가 프로그램으로서,
- <27> 상기 정답문 및 해답문 사이에 동일 또는 근사한 어군을 평가 대상어로서 추출하고, 상기 정답문 및 상기 해답문 사이에서의 각 평가 대상어의 위치의 상이 또는 상기 평가 대상어의 배열 순서의 상이에 따라 점수에 차이를 두는 처리를 상기 컴퓨터에 실행시키는 구성을 채용하고 있다.
- <28> (12) 또한, 본 발명은, 원문이 되는 문제문에 대한 정답 번역문이 되는 정답문과,
- <29> 상기 문제문에 대하여 소정 언어로 번역된 해답문을 대비함으로써, 해당 해답문의 평가를 컴퓨터에 실행시키는 문장 평가 프로그램으로서,
- <30> 상기 정답문 및 해답문 사이에 동일 또는 근사한 어군을 평가 대상어로서 추출하고, 상기 정답문 및 상기 해답문 사이에서의 각 평가 대상어의 위치의 상이 또는 상기 평가 대상어의 배열 순서의 상이에 따라 점수에 차이를 두는 어군 평가 단계와,
- <31> 상기 어군 평가 단계에서의 평가 대상어를 제외한 단어 중에서, 상기 정답문과 상기 해답문 사이에서 품사의 종류가 일치하는 단어를 새로운 평가 대상어로서 추출하고, 해당되는 새로운 평가 대상어에 대하여, 상기 정답문 및 상기 해답문 사이에서의 위치의 상이 및/또는 품사의 종류에 따라 점수에 차이를 두는 품사 평가 단계와,
- <32> 각 단계에서의 평가 대상어를 제외한 나머지 단어의 개수에 따라 점수에 차이를 두는 최종 평가 단계를,
- <33> 차례로 상기 컴퓨터에 실행시키는 구성을 채용하고 있다.
- <34> (13) 여기서, 각 단계는, 복수개의 정답문 각각에 대하여 실행되고, 이들 정답문마다 구해진 채점 결과 중에서 가장 양호한 평가를 상기 해답문의 평가로 결정하도록, 상기 컴퓨터에 실행시키는 구성을 채용하면 된다.
- <35> 그리고, 「어군」은, 청크(chunk)로 불리우며, 하나의 단어, 또는 소정의 문장 내에서, 연속하는 복수개의 단어의 집합을 총칭한 의미로서 사용한다.
- <36> 또한, 「동일한 어군」은, 비교 대상문 사이에서 서로 일치하는 어군 중, 해당 어군에 인접하는 전후의 단어가 비교 대상문 사이에서 상이한 것을 가리킨다. 예를 들면, 도 5를 참조하면, 해답문과 정답문 사이에서, 서로 일치하는 어군으로서 도 5에서 사각형으로 에워싸인 3개의 어군 외에, 「eating」, 「and」, 「more」, 「than」, 「two」, 「more than」, 「than two」의 7개의 어군의 조합을 들 수 있다. 그 중, 도 5의 사각형으로 에워싸인 어군이, 해당 어군의 전후의 단어가 비교 대상문 사이에서 상이한 것으로서 「동일한 어군」이 된다.
- <37> 또한, 「근사한 어군」은, 동일한 어군에 인접하는 전후의 단어가, 소정 조건하에서 상이한 경우, 해당 단어를 동일한 어군에 포함한 것을 가리킨다.
- <38> 이상 (1) ~ (13)의 발명에 의하면, 정답문과 해답문 사이에서, 동일 위치에서의 단어의 동일성의 판단 이외의 요소를 포함한 다양한 평가가 행해짐으로써, 해답문의 평가를 실정에 따라 객관적으로 행할 수 있다.

실시예

- <54> 이하, 본 발명의 일실시예에 대하여 도면을 참조하면서 설명한다.

- <55> [실시예 1]
- <56> 도 1에는, 본 실시예에 따른 영작문 학습 시스템의 개략적인 시스템 구성도가 나타나 있다. 이 도면에서, 영작문 학습 시스템(10)은, 일본어에 의한 소정의 문제문(원문)에 대하여 사용자가 작성한 영작문을 평가하는 시스템이다. 영작문 학습 시스템(10)은, 미리 등록된 사용자가 소유하는 단말기(11)와, 단말기(11)에 대하여, 인터넷 등의 컴퓨터 네트워크(13)를 통하여 각종 정보의 송수신을 행하고, 상기 영작문의 평가를 행하는 문장 평가 장치로서의 서버(14)를 구비하는 구성으로 이루어져 있다.
- <57> 이 영작문 학습 시스템(10)은, 문제문이 서버(14)로부터 이용자의 각 단말기(11)에 분배되고, 이용자에 의해 작성된 영어의 해답문이 단말기(11)로부터 서버(14) 측에 송신되면, 서버(14)에서, 상기 해답문의 채점, 첨삭 및 학습 레벨의 판정 등이 컴퓨터에 의해 자동적으로 행해지고, 이들 결과가 이용자의 단말기(11)에 송신되도록 되어 있다.
- <58> 단말기(11)는, 특히 한정되지 않고, 퍼스널 컴퓨터, 휴대 정보 단말기, 휴대 전화기 등, 본 발명에서 필요로 하는 정보의 송수신이 가능하면, 각종 기기를 채용할 수 있다.
- <59> 서버(14)는, 소프트웨어 및/또는 하드웨어에 의해 구성되며, 프로세서 등, 복수개의 프로그램 모듈 및/또는 처리 회로에 의해 이루어지고 있다. 이 서버(14)는, 단말기(11)에 대하여 각종 정보의 송수신을 행하는 송수신부(16)와, 다수의 문제문이 기억된 문제문 데이터 베이스(17)와, 문제문 데이터 베이스(17)의 각 문제문에 대하여 정답이 되는 영문의 정답문이 기억된 정답문 데이터 베이스(18)와, 소정의 단어가 각 활용형 및 품사의 종류와 함께 기억된 사전 데이터 베이스(19)와, 송수신부(16)로부터 단말기(11)에 송신되는 임의의 문제문을 문제문 데이터 베이스(17)로부터 추출하는 출제부(20)와, 송수신부(16)에서 수신한 이용자의 해답문에 대하여 채점 처리를 행하는 채점부(21)와, 채점부(21)에서 채점된 결과를 기억 및 집계하고, 정답문 등의 관련 정보를 채점 결과와 함께 송수신부(16)로부터 단말기(11)에 송신하는 채점 결과 처리부(22)를 구비하고 있다. 그리고, 서버(14)는, 그 외에, 각종 콘텐츠를 제공하는 기능 등도 구비하고 있지만, 본 발명의 요지가 아니므로, 여기서는, 도시 및 설명을 생략한다.
- <60> 정답문 데이터 베이스(18)에는, 문제문 데이터 베이스(17)에 기억된 문제문 각각에 대하여, 생각할 수 있는 모든 정답문이 기억되어 있다.
- <61> 채점부(21)는, 이용자가 작성한 해답문과, 해당 해답문에 대응한 정답문 데이터 베이스(18)의 각 정답문을 각각 대비하고, 이들 정답문마다, 후술하는 기준에 의한 감점법으로 채점을 행하고, 그 중 최고 득점을 상기 해답문의 득점으로 하도록 기능한다.
- <62> 이 채점부(21)는, 상기 정답문 및 상기 해답문 내의 어군의 상이에 착안하여 채점하는 어군 평가 수단(25)과, 상기 정답문 및 상기 해답문 내의 단어의 활용형의 상이에 착안하여 채점하는 활용형 평가 수단(26)과, 상기 정답문 및 상기 해답문 내의 단어의 품사의 동일성에 착안하여 채점하는 품사 평가 수단(27)과, 상기 정답문 및 상기 해답문 내의 단어의 스펠링 미스에 착안하여 채점하는 스펠링 평가 수단(28)과, 이들 각 평가 수단(25 ~ 28)에 의한 채점을 행한 후 최종적인 채점을 행하는 최종 평가 수단(29)을 구비하고 있다.
- <63> 다음에, 이상과 같이 구성된 영작문 학습 시스템(10)의 채점 프로세스에 대하여, 일례를 들면서 설명한다.
- <64> 먼저, 이용자가 단말기(11)로부터 서버(14)에 요구하면, 출제부(20)로부터 소정의 문제문이 출제되고, 해당 문제문이 송수신부(16)로부터 단말기(11)에 송신된다. 그리고, 이용자는, 해당 문제문에 대하여 영작문을 행하고, 그 해답문을 단말기(11)로부터 서버(14)에 송신한다. 여기서, 이하, 일례에 의한 설명에 대해서는, 도 2에 나타난 문제문이 출제되고, 그에 대하여, 이용자에 의해, 도 2에 나타난 해답문이 작성된 것으로 가정한다.
- <65> 그리고, 서버(14)의 채점부(21)에서는, 이용자가 작성한 해답문의 채점이 행해진다. 여기서의 채점은, 이용자의 해답문이, 정답문 데이터 베이스(18)에 기억된 다수의 정답문과 각각 대비되고, 정답문 각각에 대해, 해답문의 득점이 후술하는 바와 같이 구해진다. 그리고, 정답문마다 구해진 득점 중 최고 득점이 이용자의 득점이 되고, 상기 득점의 기준이 된 정답문과 함께 채점 결과 처리부(22)에 보내지며, 이용자의 학습 레벨의 판정이 행해지고, 해당 레벨과 함께, 첨삭이나 정답문의 설명 등의 각종 콘텐츠가 단말기(11)에 보내지게 된다.
- <66> 여기서, 채점부(21)에서의 채점은, 도 3의 흐름도에 나타난 순서로 행해진다.
- <67> 먼저, 이용자의 해답문이 채점부(21)에 받아들여진다(S101). 그리고, 출제된 문제문에 대응하는 각 정답문 중에서, 하나의 정답문이 정답문 데이터 베이스(18)로부터 추출된다. 여기서, 정답문 데이터 베이스(18)에는, 예를 들면, 도 4에 나타난 바와 같이, 복수개의 정답문이 기억되어 있다. 이하에서는, 이들 정답문 중, 도 4의

「No.1」의 정답문과의 대비에 의한 채점을 예로서 설명한다.

- <68> 처음에, 어군 평가 수단(25)에서, 하나의 단어 또는 연속하는 복수개의 단어의 집합체인 어군의 상이에 착안한 채점이 행해진다(어군 평가 단계: S103). 즉, 여기서는, 해답문 및 정답문의 대비에 의해, 해답문과 정답문 사이에서 동일한 어군이 추출되고, 해당 어군이 어군 평가 수단(25)에서의 평가 대상이 된다. 전술한 예에서는, 도 5에 나타낸 바와 같이, 각 문장 내에 사각형으로 에워싸인 어군이, 해답문과 정답문 사이에서 동일한 어군이며, 이들 어군이 평가 대상어가 된다.
- <69> 다음에, 이들 평가 대상어에 대하여, 문두로부터의 단어의 위치가 해답문과 정답문 사이에서 일치하는지의 여부가 판정된다. 전술한 예에서, 평가 대상어 「Jack」은 위치가 일치하고 있지만, 다른 평가 대상어 「eating and」와 「more than two」는 위치가 불일치한다.
- <70> 그리고, 해답문과 정답문 사이에서 위치가 상이한 동일한 어군 한 조(組)에 대하여, 각 1점의 감점 포인트가 부여되고, 이들을 합산함으로써, 어군 평가 수단(25)에서의 감점 포인트가 구해진다. 전술한 예에서는, 두 조의 어군 「eating and」와 「more than two」가, 해답문과 정답문 사이에서 서로 위치가 일치하지 않기 때문에, 각 1점의 감점 포인트가 부여되고, 어군 평가 수단(25)에서의 감점 포인트는 합계 2점이 된다.
- <71> 다음에, 활용형 평가 수단(26)에서, 어군 평가 수단(25)에서 평가 대상어로 된 어군을 제외한 각 단어에 대하여, 이들 활용형에 착안한 채점이 행해진다(활용형 평가 단계: S104). 즉, 여기서는, 어군 평가 수단(25)에서 평가 대상으로 되지 않았던 나머지 단어에 대하여, 해답문 및 정답문의 대비에 의해, 사전 데이터 베이스(19)를 기초로 하여, 해답문과 정답문 사이에서 활용형의 상이에 지나지 않는 단어가 추출되고, 해당 단어가 활용형 평가 수단(26)에서의 평가 대상이 된다. 전술한 예에서는, 도 6에 나타낸 바와 같이, 사각형으로 에워싸인 단어 「have」와 단어 「has」가, 해답문과 정답문 사이에서 활용형의 상이에 지나지 않는 단어이며, 이들 단어가 여기서의 평가 대상어가 된다. 그리고, 이후의 각 도면 내에서, 「×」가 표시된 단어는, 그 전에 이미 평가된 평가 대상어를 의미한다. 여기서의 활용형의 상이는, 전술한 동사의 경우 외에, 예를 들면, 명사의 경우에는, 단수형과 복수형의 상이, 형용사나 부사의 경우에는, 원급, 비교급, 최상급의 상이 등을 들 수 있다.
- <72> 또한, 이들 평가 대상어에 대하여, 문두로부터의 위치가, 해답문과 정답문 사이에서 서로 일치하는지의 여부가 판정된다. 전술한 예에서는, 해답문 내의 단어 「have」와 정답문 내의 단어 「has」는, 이들 위치가 일치하고 있다.
- <73> 여기서, 이 활용형 평가에 대한 감점 포인트로서, 해답문과 정답문 사이에서 위치가 상이하지 않으면, 활용형의 상이에 대한 것만 1점이 부여되는 한편, 해답문과 정답문 사이에서 위치가 상이하면, 활용형 및 위치의 상이에 대한 2점이 부여된다. 이들 감점 포인트를 합산함으로써, 활용형 평가 수단(26)에서의 감점 포인트가 구해진다. 전술한 예에서는, 해답문 내의 단어 「have」와 정답문 내의 단어 「has」는, 활용형이 상이하지만 위치가 일치하므로, 1점의 감점 포인트가 부여되고, 활용형 평가 수단(26)에서의 감점 포인트는 합계 1점이 된다.
- <74> 다음에, 품사 평가 수단(27)에서, 어군 평가 수단(25) 및 활용형 평가 수단(26)에서의 평가 대상어를 제외한 각 단어에 대하여, 품사의 동일성에 주목하고 채점이 행해진다(품사 평가 단계: S105). 즉, 여기서는, 각 평가 수단(25, 26)에서 평가 대상이 되지 않았던 나머지 단어에 대하여, 사전 데이터 베이스(19)의 데이터를 기초로 하여, 품사의 판별이 행해지고, 해답문 및 정답문의 대비에 의해, 이들 사이에서 품사가 동일하게 되는 단어가 추출되고, 해당 단어가 품사 평가 수단(27)에서의 평가 대상이 된다.
- <75> 여기서의 품사의 동일은, 미리 설정된 품사 체계에 기초하여 판단된다. 이 품사 체계는, 특히 한정되지 않고, 「품사」나 「형용사」 등의 대분류에 의한 것이나, 「보통 명사」나 「고유 명사」 등, 대분류를 보다 세분화한 분류에 의한 것이 사용된다. 본 실시예에서는, 후자의 분류에 기초한 품사 체계가 사용된다.
- <76> 전술한 예에서는, 도 7에 나타낸 바와 같이, 사각형으로 에워싸인 단어 「in」 및 「for」의 조와 「minutes」 및 「hours」의 조가, 해답문과 정답문 사이에서 품사가 각각 동일한 단어의 조이며, 이들 각 단어가, 여기서의 평가 대상어가 된다.
- <77> 그리고, 전술한 예에는 없지만, 한쪽의 문장의 품사에 대하여, 동일한 품사가 다른 쪽의 문장에 복수개 있는 경우에는, 상기 다른 쪽의 문장의 문두에 가까운 단어가 선택된다.
- <78> 또한, 이들 평가 대상어가 된 각 단어의 조에 대하여, 문두로부터의 위치가, 해답문과 정답문 사이에서 서로 일치하는지의 여부가 판정된다. 전술한 예에서는, 해답문 중의 단어 「in」과 정답문 중의 단어 「for」의 조는,

이들 위치가 불일치하며, 또한, 해답문 중의 단어 「minutes」와 정답문 중의 단어 「hours」의 조도 이들 위치가 불일치한다.

- <79> 여기서, 이 품사 평가에 대한 감점 포인트로서 해답문과 정답문 사이에서 위치가 상이하지 않으면, 단어의 상이에 대한 1점이 부여되는 한편, 해답문과 정답문 사이에서 위치가 상이하면, 단어 및 위치의 상이에 대한 2점이 부여된다. 이들 감점 포인트를 합산함으로써, 품사 평가 수단(27)에서의 감점 포인트가 구해진다. 전술한 예에서는, 해답문 중의 단어 「in」과 정답문 중의 단어 「for」의 조는, 품사는 동일하지만 단어가 상이하며, 위치가 일치하고 있지 않으므로, 2점의 감점 포인트가 부여된다. 또한, 해답문 내의 단어 「minutes」와 정답문 내의 단어 「hours」의 조도 동일하게 2점의 감점 포인트가 부여되고, 품사 평가 수단(27)에서의 감점 포인트는 합계 4점이 된다.
- <80> 다음에, 스펠링 평가 수단(28)에서, 어근 평가 수단(25), 활용형 평가 수단(26) 및 품사 평가 수단(27)에서의 각 평가 대상어를 제외한 각 단어에 대하여, 해답문의 단어의 스펠링 미스에 착안한 채점이 행해진다(스펠링 평가 단계: S106). 즉, 여기서는, 각 평가 수단(25 ~ 27)에서 평가 대상이 되지 않았던 나머지 단어에 대하여, 해답문 및 정답문의 대비에 의해, 이들 사이에서 철자가 근사한 단어가 추출되고, 해당 단어가, 여기서의 평가 대상이 된다.
- <81> 여기서, 해답문 및 정답문의 각 단어 사이에서 철자가 근사한지의 여부의 판단은, 공지된 레벤슈타인 거리(Levenshtein distance)를 사용하여 행해진다. 이 레벤슈타인 거리는, 서로 비교 대상이 되는 단어에 대하여, 문자의 누락, 문자의 오기, 문자의 삽입을 고려하여, 정답에 합치시키는 단계 중 최소로 되는 단계의 개수(코스트)에 의해 정해진다.
- <82> 예를 들면, 「puzzle」이라는 정답문의 단어에 대하여, 스펠링 미스에 의해, 해답문에서는 「pzzel」이라는 단어가 사용되었을 경우,
- <83> 해답문의 「pzzel」에 대하여, 정답문의 「puzzle」로 바꾸려면,
- <84> (1) 문자 「p」와 문자 「z」의 사이에 문자 「u」를 삽입하고,
- <85> (2) 문자 「z」와 문자 「e」의 사이에 문자 「l」을 삽입하고,
- <86> (3) 마지막 문자 「l」을 삭제,
- <87> 의 3 단계가 필요해지므로, 여기서의 「코스트 3」이, 「puzzle」 및 「pzzel」 사이의 레벤슈타인 거리가 된다.
- <88> 그리고, 해답문 및 정답문 사이에서, 평가 대상어가 되는 각 단어를 대비할 때, 레벤슈타인 거리가 소정의 임계값 이하로 되는 단어의 조가, 서로 근사한 단어로 판단되고, 그 해답문의 단어가 스펠링 미스의 단어인 것으로 판정된다.
- <89> 전술한 예에서는, 도 8에 나타낸 바와 같이, 사각형으로 에워싸인 해답문 내의 단어 「dlinking」과, 정답문 내의 단어 「drinking」이, 해답문과 정답문 사이에서 스펠링이 근사한 단어이며, 이들 단어가 여기서의 평가 대상으로 되고, 해답문 내의 단어 「dlinking」이 스펠링 미스의 단어인 것으로 판단된다.
- <90> 또한, 이들 평가 대상어에 대하여, 문두로부터의 위치가, 해답문과 정답문 사이에서 서로 일치하는지의 여부가 판정된다. 전술한 예에서는, 해답문 중의 단어 「dlinking」, 정답문 중의 단어 「drinking」은, 이들 위치가 불일치하다.
- <91> 여기서의 감점 포인트로서, 해답문과 정답문 사이에서 위치가 동일하면, 스펠링 미스에 대한 1점이 부여되는 한편, 해답문과 정답문 사이에서 위치가 상이하면, 스펠링 미스 및 위치의 상이에 대한 2점이 부여된다. 이들 감점 포인트를 합산함으로써, 스펠링 평가 수단(28)에서의 감점 포인트가 구해진다. 전술한 예에서는, 해답문 중의 단어 「dlinking」과, 정답문 중의 단어 「drinking」은, 위치가 일치하고 있지 않으므로, 2점의 감점 포인트가 부여되고, 스펠링 평가 수단(28)에서의 감점 포인트는 합계 2점이 된다.
- <92> 다음에, 최종 평가 수단(29)에서는, 전술한 각 평가 수단(25 ~ 28)에서 평가 대상어가 되지 않았던 나머지 단어에 대한 채점이 행해진다(최종 평가 단계: S107). 즉, 여기서는, 나머지 단어에 대하여, 다음과 같이 감점 포인트가 구해진다.
- <93> 먼저, 아직 평가 대상이 되고 있지 않은 단어 중, 해답문과 정답문 사이에서, 동일 위치에 단어가 존재하는지의

여부가 판정된다. 전술한 예에서는, 도 9에 나타난 바와 같이, 동일 위치에 존재하는 단어가 없고, 여기서의 감점 포인트는 0 점이다.

- <94> 그리고, 더 남은 단어 중, 해답문과 정답문 사이에서, 한단어씩의 조를 작성할 수 있는지의 여부가 판별되고, 조로 작성할 수 있으면, 감점 포인트로서 단어 및 위치의 상이에 대한 2점이 한 조마다 부여된다. 전술한 예에서는, 도 10에 나타난 바와 같이, 해답문 중의 단어 「thirty」와, 정답문 중의 단어 「been」이 평가 대상이 되고, 한 조이므로, 감점 포인트로서 2점이 부여된다.
- <95> 마지막으로, 해답문 및 정답문 중 어느 한쪽에만 남은 단어에 대하여, 한 단어마다 감점 포인트 1점이 부여된다. 전술한 예에서는, 도 11에 나타난 바와 같이, 정답문 중의 단어 「and」와 단어 「a」와 단어 「half」가 남아 있으며, 감점 포인트로서 3단어분의 3점이 부여된다.
- <96> 다음에, 이상의 각 수단에 의해 구해진 감점 포인트가 합산되고, 다음 식(1)에 의해, 정답문에 대한 해답문의 근사도가 구해진다(S108).
- <97> 여기서, 근사도를 N이라 하고, 감점 포인트의 합계를 P라 하고, 해답문의 단어 개수를 W1이라 하고, 정답문의 단어 개수를 W2라 하면,
- <98>
$$N = 1 - (P/(W1 + W2)) \dots \text{식(1)}$$
- <99> 전술한 예에서는, 도 5 ~ 도 11의 결과, 감점 포인트의 총합계는 14점이 되고, 해답문의 단어 개수는 11개 단어, 정답문의 단어 개수는 14개 단어이므로, 근사도 N은,
- <100> $1 - (14/(11 + 14)) = 0.44$ 로 구해진다.
- <101> 이상의 채점 처리를 각 정답문에 대하여 행하고(S109), 그 중 가장 높은 근사도가 그 해답문의 점수가 되며(S110), 해당 근사도에 대응하는 정답문이 해답문의 침삭 대상으로 된다.
- <102> 그리고, 이상에서, 각 수단에서의 감점 포인트를 1점 또는 2점으로 하고 있지만, 각 평가 항목에 따라, 부여되는 감점 포인트의 점수를 바꾸어서, 각 항목의 중요도를 부여할 수도 있다.
- <103> 따라서, 이와 같은 실시예에 따르면, 해답문에 대하여 각종 항목으로 채점할 수 있고, 실정에 따른 객관적인 채점을 간단하게 행할 수 있는 효과를 얻는다.
- <104> 다음에, 본 발명의 다른 실시예에 대하여 설명한다. 그리고, 이하의 설명에서, 실시예 1과 동일 또는 동등한 구성 부분에 대하여는 동일 부호를 사용하기로 하고, 설명을 생략 또는 간략하게 한다.
- <105> [실시예 2]
- <106> 실시예 2는, 실시예 1에 대하여 채점부(21)의 구성을 바꾸고, 해답문의 채점 방법, 즉 상기 감점 포인트 및 근사도 N의 산출 방법을 바꾼 것이 특징이다.
- <107> 즉, 본 실시예의 채점부(21)는, 도 12에 나타난 바와 같이, 상기 정답문 및 상기 해답문 내의 어군의 상이에 착안하여 채점하는 어군 평가 수단(25)과, 상기 정답문 및 상기 해답문 내의 단어의 품사의 동일성에 착안하여 채점하는 품사 평가 수단(27)과, 이들 각 평가 수단(25, 27)에 의한 채점을 행한 후 최종적인 채점을 행하는 최종 평가 수단(29)을 구비하고 있고, 어군 평가 수단(25), 품사 평가 수단(27), 최종 평가 수단(29)의 순서로 채점 처리가 행해진다.
- <108> 본 실시예의 채점부(21)에서의 채점 방법에 대하여, 일례를 들어 이하에 설명한다.
- <109> 본 실시예에서도, 실시예 1과 동일한 문제문, 정답문 및 해답문을 사용하고, 해당 해답문에 대한 채점 프로세스를 설명한다.
- <110> 먼저, 어군 평가 수단(25)에서, 하나의 단어 또는 연속하는 복수개의 단어의 집합체인 어군의 상이에 착안한 채점이 행해진다(어군 평가 단계). 즉, 여기서는, 해답문 및 정답문의 대비에 의해, 해답문과 정답문 사이에서 동일 또는 근사한 어군이 추출되고, 해당 어군이 어군 평가 수단(25)에서의 평가 대상이 된다. 구체적으로는, 정답문 및 해답문 사이에 동일한 어군이 상호 매칭 등에 의해 결정된다. 그리고, 해당되는 동일한 어군에 인접하는 전후의 단어에 대하여, 정답문 및 해답문 사이에서, 활용형의 상이 또는 스펠링 미스인지의 여부가 판단되고, 활용형의 상이 또는 스펠링 미스로 판단된 경우에, 이들 전후의 단어를 상기 동일한 어군에 포함하여 근사한 어군이 된다. 한편, 동일한 어군에 인접하는 전후의 단어가 활용형의 상이 또는 스펠링 미스로 판단되지 않

은 경우에는, 해당 단어를 포함하지 않고 그대로 동일한 어군이 된다.

- <111> 전술한 예에서는, 도 13의 각 문장 내에 사각형으로 에워싸인 어군이, 해답문과 정답문 사이에서 동일 또는 근사한 것이며, 이들 어군이 평가 대상으로 된다.
- <112> 즉, 해답문 내의 어군 A 「Jack have」 및 정답문 내의 어군 A 「Jack has」는, 「Jack」 부분이 동일하며, 그 뒤의 단어인 해답문의 「have」가 정답문의 「has」에 대하여 활용형이 잘못되어 있는 것으로 판단되어, 근사한 어군이 된다.
- <113> 또한, 해답문 내의 어군 B 「eating and dlinking」 및 정답문 내의 어군 B 「eating and drinking」은, 「eating and」 부분이 동일하며, 그 뒤의 단어인 해답문의 「dlinking」이 정답문의 「drinking」에 대하여 스펠링 미스로 판단되어, 근사한 어군이 된다.
- <114> 또한, 해답문 내 및 정답문 내의 어군 C 「more than two」는, 서로 완전히 일치하므로, 동일한 어군이 된다.
- <115> 그리고, 여기서의 활용형의 에러 및 스펠링 미스의 판단은, 실시예 1에서 설명한 활용형 평가 수단(26) 및 스펠링 평가 수단(28)과 마찬가지로의 방법에 의해 행해진다.
- <116> 다음에, 이들 동일 또는 근사한 어군으로 이루어지는 평가 대상어에 대하여, 정답문에 대하여 활용형이 잘못된 단어, 스펠링 미스의 단어에 대하여, 각각 한 단어당 1점의 감점 포인트가 부여된다. 전술한 예에서, 해답문의 어군 A 「Jack have」는, 정답문의 어군 A에 대하여 활용형이 잘못되어 있는 단어 「have」를 한 단어 포함하고, 또한, 해답문의 어군 B 「eating and dlinking」은, 정답문의 어군 B에 대하여 철자가 잘못되어 있는 단어 「dlinking」을 한 단어 포함하고 있으므로, 이들 어군 A 및 어군 B에 대한 감점 포인트는, 각각 1점이 된다. 또한, 해답문의 어군 C 「more than two」는, 정답문의 어군 C에 대하여 활용형의 상이나 스펠링 미스가 없으므로, 감점 포인트는 0점이 된다.
- <117> 다음에, 동일 또는 근사한 어군에 대하여, 그 배열 순서의 상이에 주목하고, 감점 포인트를 부여할지의 여부가 결정된다. 전술한 예에서, 해답문은, 어군 A, 어군 B 및 어군 C의 순서로 배열되어 있고, 정답문도 어군 A, 어군 B 및 어군 C의 순서로 배열되어 있다. 이와 같은 경우에는, 배열 순서가 동일하게 되어 감점 없음, 즉 감점 포인트 0점이 된다.
- <118> 한편, 만일, 정답문이, 어군 A, 어군 B 및 어군 C의 순서에 대하여, 해답문이, 어군 C, 어군 A 및 어군 B의 순서로 배열되어 있다고 하면, 정답문의 어군의 배열과 동일하게 하기 위해서는, 해답문의 어군 C를 어군 B의 뒤로 이동시키는 하나의 단계가 필요하다. 이 필요 단계 수에 감점 포인트 5점을 곱하여, 어군의 배열 순서에 착안한 감점 포인트가 산출된다. 이와 같이 필요 단계 수가 하나의 단계일 경우에는, 감점 포인트가 5점이 된다.
- <119> 다음에, 품사 평가 수단(27)에서, 어군 평가 수단(25)에서의 평가 대상어를 제외한 각 단어에 대하여, 품사의 동일성에 착안한 채점이 행해진다(품사 평가 단계). 즉, 여기서, 실시예 1과 마찬가지로, 어군 평가 수단(25)에서 평가 대상이 되지 않았던 나머지 단어에 대하여, 사전 데이터 베이스(19)의 데이터를 기초로 하여, 품사의 판별이 행해지고, 해답문 및 정답문의 대비에 의해, 이들 사이에서 품사가 동일하게 되는 단어가 추출되고, 해당 단어가 품사 평가 수단(27)에서의 평가 대상이 된다. 전술한 예에서는, 도 14에 나타난 바와 같이, 사각형으로 에워싸인 단어 「in」 및 「for」, 및 단어 「minutes」 및 「hours」가, 각각 해답문과 정답문 사이에서 품사가 동일하게 되는 단어이며, 이들 단어가, 여기서의 평가 대상으로 된다.
- <120> 여기서, 이 품사 평가에 대한 감점 포인트는, 해답문과 정답문 사이에서 동일하게 된 품사의 종류에 따라 정해진다. 구체적으로, 동일하게 된 품사가 관사라면, 감점 포인트 3점, 전치사라면, 감점 포인트 7점, 그 외의 품사라면, 감점 포인트 10점이 각각 부여된다. 전술한 예에서는, 해답문 내의 단어 「in」과 정답문 내의 단어 「for」는, 품사가 전치사이므로, 감점 포인트가 7점이 된다. 한편, 해답문 내의 단어 「minutes」와 정답문 내의 단어 「hour」는, 전치사 이외의 그 외의 품사(명사)이므로, 감점 포인트가 10점이 되고, 품사 평가 수단(27)에서의 감점 포인트가 17점이 된다.
- <121> 그리고, 실시예 1과 같이, 평가 대상어에 대하여, 문두로부터의 위치가 해답문과 정답문 사이에서 서로 일치하는지의 여부도 판정하고, 그 일치·불일치에 의한 감점 포인트의 점수를 바꾸어도 된다.
- <122> 다음에, 최종 평가 수단(29)에서, 전술한 각 평가 수단(25, 27)에서 평가 대상어가 되지 않았던 나머지 단어에 대하여 채점이 행해지고(최종 평가 단계), 다음과 같이, 감점 포인트가 구해진다.
- <123> 즉, 아직 평가 대상이 되어 있지 않은 해답문 및 정답문의 각 단어에 대하여, 이들의 품사별로 감점 포인트가

산출되고, 이것을 총합계하여, 최종 평가 수단(29)에서의 감점 포인트가 구해진다. 구체적으로는, 본 실시예의 품사 평가 수단(27)과 마찬가지로, 관사이면 감점 포인트 3점, 전치사이면 감점 포인트 7점, 그 외의 품사이면 감점 포인트 10점이 된다. 여기서도, 품사 종류별의 판정에는, 전술한 사전 데이터 베이스(19)가 사용된다.

<124> 전술한 예에서는, 도 15에 나타낸 바와 같이, 해답문 내의 단어 「thirty」와, 정답문 내의 단어 「been」과 단어 「and」와 단어 「a」와 단어 「half」의 합계 5개의 단어에 대하여, 각각 품사에 의한 감점 포인트가 부여된다. 여기서, 단어 「a」는 관사이므로 감점 포인트는 3점, 그 이외의 4개의 단어는, 관사 및 전치사는 아니므로, 감점 포인트가 각 10점이 되고, 최종 평가 수단(29)에서의 감점 포인트는, 합계 43점이 된다.

<125> 그리고, 이상의 각 수단에서의 감점 포인트는, 특히 한정되지 않고, 각종 점수를 임의로 채용할 수 있다.

<126> 다음에, 이상의 각 수단에서 구해진 감점 포인트가 합산되고, 다음 식(2)에 의해, 정답문에 대한 해답문의 근사도가 구해진다.

<127> 여기서, 근사도를 N이라 하고, 감점 포인트의 합계를 P라 하고, 해답문의 단어 개수와 정답문의 단어 개수 사이에서 많은 쪽의 단어 개수를 Wmax라 하고, 설정되어 있는 최대 감점 포인트 Pmax(본 실시예에서는 10점)라 하면,

<128>
$$N = 1 - (P / (W_{max} \times P_{max})) \dots \text{식(2)}$$

<129> 그리고, 상기 식(2)에서, 근사도 N이 마이너스치가 되었을 경우에는, 근사도 N을 0로 한다.

<130> 전술한 예에서는, 도 13 ~ 도 15의 결과, 감점 포인트의 총합계는 62점이 되고, 해답문의 단어 개수는 11개, 정답문의 단어 개수는 14개이므로, Wmax는 14가 되고,

<131> 근사도 N은,

<132>
$$1 - (62 / (14 \times 10)) = 0.56$$
으로 구해진다.

<133> 이후는, 실시예 1과 마찬가지로, 이상의 채점 처리를 정답문 각각에 대하여 행하고, 그 중 가장 높은 근사도 N이 그 해답문의 점수가 되고, 해당 근사도 N에 대응하는 정답문이 해답문의 점수 대상이 된다.

<134> 따라서, 이와 같은 실시예 2에 의해서도, 실시예 1과 마찬가지로의 효과를 얻을 수 있다.

<135> 여기서, 실시예 2의 변형예로서 다음의 태양이 있다. 즉, 해답문 내의 단어의 스펠링 미스에 의해, 근사한 어군이 된 어군에 대하여, 스펠링 미스로 판단된 단어를 정답문 내의 정확한 단어로 치환한 다음, 전술한 품사 평가 수단(27)에 의한 품사 평가 단계를 행하도록 할 수도 있다. 도 13에 나타낸 전술한 예에 의하면, 스펠링 미스에 의해 근사한 어군으로 판단된 어군 B에 대하여, 해답문 내의 단어 「dlinking」을 해당하는 정답문 내의 단어 「drinking」으로 치환한 다음, 어군 평가 수단(25)에서 평가 대상으로 되지 않았던 나머지 단어에 대하여, 사전 데이터 베이스(19)의 데이터를 기초로 하여, 단어의 전후의 단어의 품사를 고려하여 품사의 판별이 행해진다.

<136> 이와 같은 변형예에 의하면, 품사의 판별을 행할 때의 정밀도를 높일 수 있다. 구체적으로, 복수 종류의 품사의 기능을 가지는 단어에 대하여 품사의 판별을 하는 경우에, 해당 단어의 전후의 단어의 품사를 고려하여 품사를 특정한다. 그런데, 스펠링 미스한 단어(도 13의 예에서는, 단어 「dlinking」) 뒤의 단어(도 13 내의 단어 「in」)의 품사의 판별을 행하고자 하는 경우에는, 스펠링 미스한 단어가 현존하지 않는 경우가 많기 때문에, 해당되는 스펠링 미스한 단어의 품사를 판별하지 못하고, 그 단어의 품사를 고려한 후의 단어의 품사 특정을 할 수 없다. 이 경우에는, 사용 빈도 등에 기초한 통계적인 관점으로부터, 하나의 품사로 한정할 수 밖에 없고, 이것으로는, 사용 태양에 맞는 정확한 품사의 판별을 행할 수 없다. 그 점은, 본 변형예에 의하면, 스펠링 미스한 단어를 정확한 단어로 치환함으로써, 스펠링 미스한 단어에 이어지는 단어의 품사를 특정할 때, 해당 단어가 복수 종류의 품사 기능을 가지는 경우라도, 해답문 작성자(이용자)의 사용 의도에 따른 품사의 종류를 보다 정확하게 특정할 수 있다.

<137> 그리고, 상기 각 실시예에서는, 컴퓨터 네트워크를 사용한 원격 교육형의 영작문 학습 시스템을 도시하여 설명하였으나, 본 발명의 문장 평가 장치는, 이에 한정되지 않고, 본 발명의 처리를 실행할 수 있는 프로그램이 인스톨된 컴퓨터라도 된다. 또한, 본 발명의 처리를 실행할 수 있는 프로그램을 각 단말기(11)에 인터넷 등에 의해 분배하거나, 해당 프로그램이 기억된 CD-ROM 등의 기록 매체를 이용자에게 반포함으로써, 이용자가 소유하는 단말기(11) 내에서, 전술한 모든 처리를 실행시킬 수도 있다.

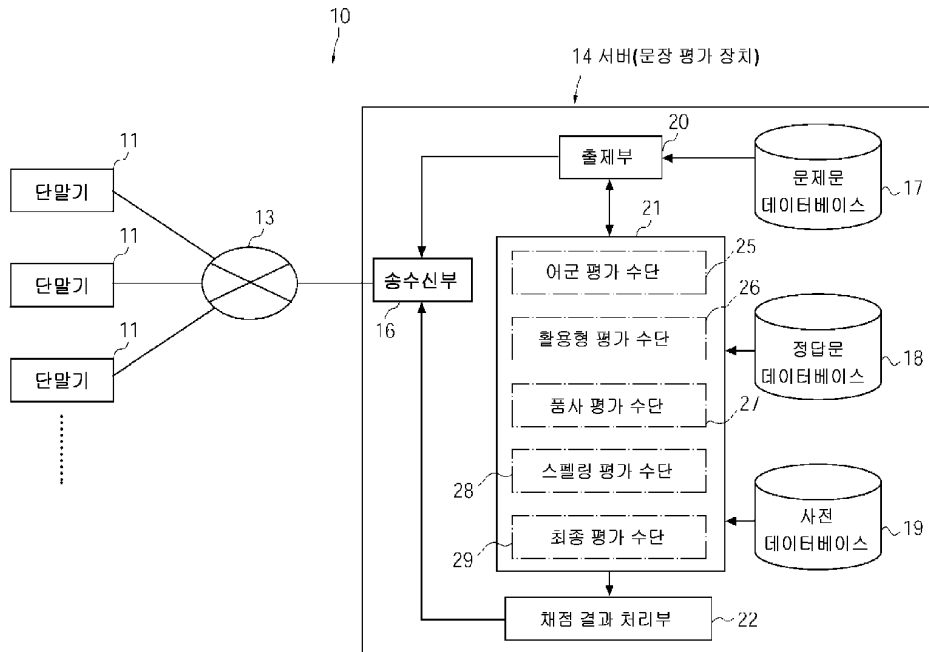
- <138> 또한, 상기 각 실시예에서는, 일본어의 문제문에 대하여 영어로 작성된 해답문의 평가를 행하고 있지만, 본 발명은 이에 한정되지 않고, 일본어 또는 다른 언어의 문제문에 대하여 다른 언어로 작성된 해답문의 평가를 행하도록 해도 된다.
- <139> 또한, 본 발명은, 번역자나 자동 번역기가 작성한 번역문에 대한 평가에 적용할 수도 있다. 자동 번역기의 번역 능력의 평가에는, 본 발명이 적용된 시스템을 자동 번역기와 연동시킴으로써, 일련의 시스템화도 가능하다.
- <140> 그 외에, 본 발명에서의 장치 각 부의 구성은 도시된 구성에 한정되는 것이 아니고, 실질적으로 마찬가지로의 작용을 가질 수 있는 한, 다양하게 변경될 수 있다.

도면의 간단한 설명

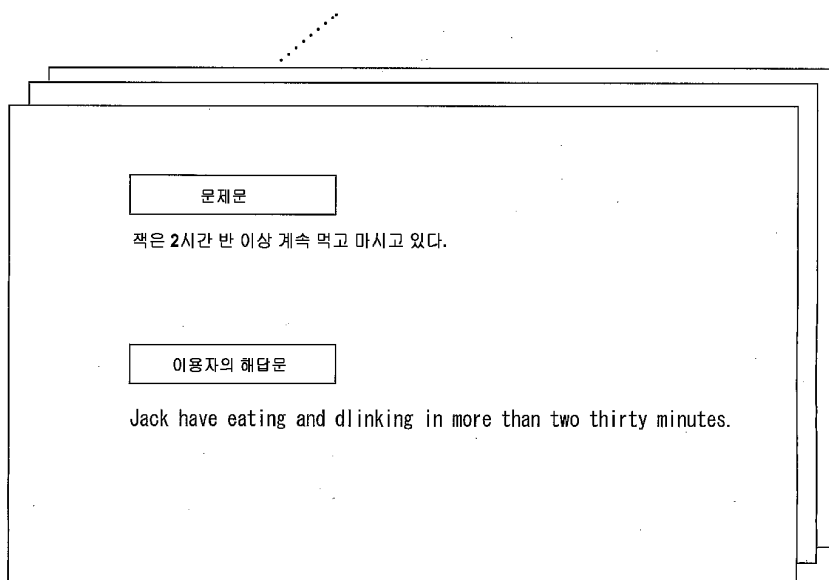
- <39> 도 1은 실시예 1에 따른 영작문 학습 시스템의 개략적인 시스템 구성도이다.
- <40> 도 2는 설명예에 사용하는 문제문 및 해답문을 적은 도면이다.
- <41> 도 3은 채점 프로세스를 나타낸 흐름도이다.
- <42> 도 4는 정답문 데이터 베이스의 기억 내용의 설명도이다.
- <43> 도 5는 어군 평가 단계에서의 채점을 설명하기 위한 도표이다.
- <44> 도 6은 활용형 평가 단계에서의 채점을 설명하기 위한 도표이다.
- <45> 도 7은 품사 평가 단계에서의 채점을 설명하기 위한 도표이다.
- <46> 도 8은 스펠링 평가 단계에서의 채점을 설명하기 위한 도표이다.
- <47> 도 9는 최종 평가 단계에서의 채점을 설명하기 위한 도표이다.
- <48> 도 10은 최종 평가 단계에서의 채점을 설명하기 위한 도표이다.
- <49> 도 11은 최종 평가 단계에서의 채점을 설명하기 위한 도표이다.
- <50> 도 12는 실시예 2에 따른 영작문 학습 시스템의 개략적인 시스템 구성도이다.
- <51> 도 13은 어군 평가 단계에서의 채점을 설명하기 위한 도표이다.
- <52> 도 14는 품사 평가 단계에서의 채점을 설명하기 위한 도표이다.
- <53> 도 15는 최종 평가 단계에서의 채점을 설명하기 위한 도표이다.

도면

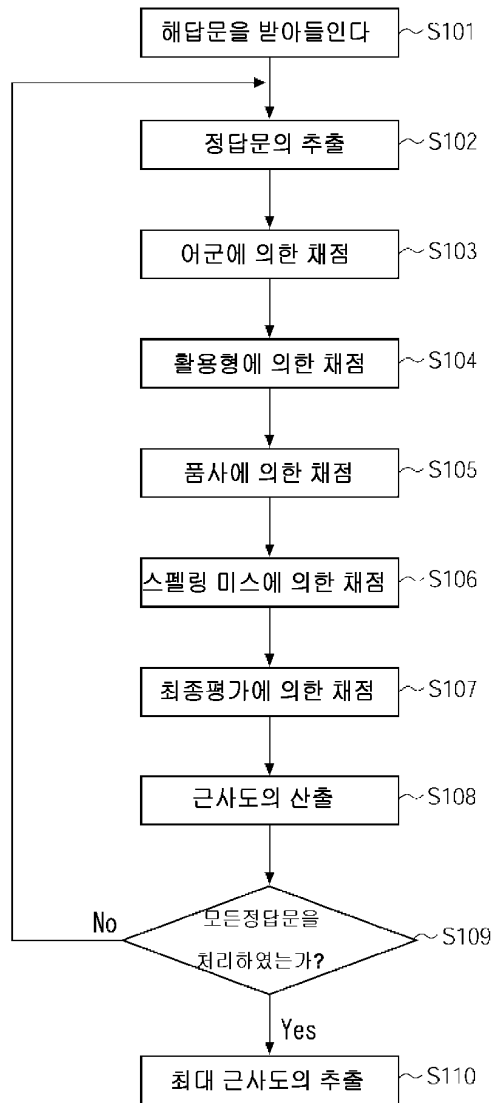
도면1



도면2



도면3



도면4

문제문 잭은 2시간 반 이상 계속 먹고 마시고 있다.

No.	
1	Jack has been eating and drinking for more than two and a half hours.
2	Jack has been eating and drinking since more than two and a half hours ago.
3	Jack has continued to eat and drink for more than two and a half hours.
4

도면5

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	dlinking	in	more	than	two	thirty	minutes.			
정답문 No.1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours.

평가 대상	위치 일치	감점포인트
Jack	○	0
eating and	×	1
more than two	×	1

⇒ 2점

도면6

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	dlinking	in	more	than	two	thirty	minutes			
정답문 No. 1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours

평가 대상		위치 일치	감점포인트	⇒ 1점
해답문	have	○	1	
정답문	has			

도면7

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	dlinking	in	more	than	two	thirty	minutes			
정답문 No. 1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours

평가 대상		위치 일치	감점포인트	⇒ 4점
해답문	in	×	2	
정답문	for			
해답문	minutes	×	2	
정답문	hours			

도면8

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	dlinking	in	more	than	two	thirty	minutes			
정답문 No. 1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours

평가 대상		위치 일치	감점포인트	⇒ 2점
해답문	dlinking	×	2	
정답문	drinking			

도면9

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	drinking	in	more	than	two	thirty	minutes			
정답문 No. 1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours.

평가 대상	감점포인트	⇒ 0점
없음	0	

도면10

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	drinking	in	more	than	two	thirty	minutes			
정답문 No. 1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours.

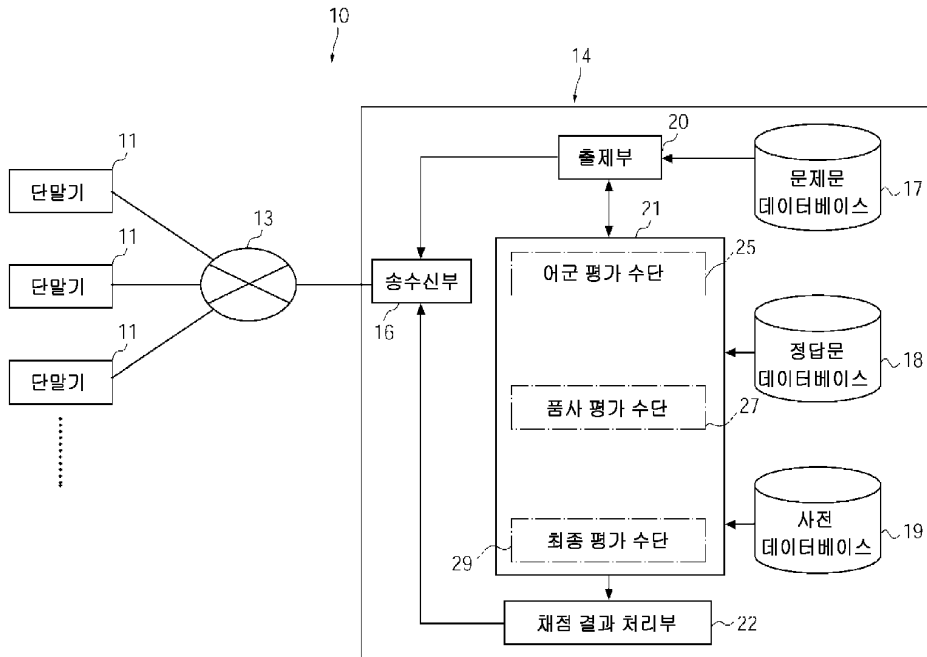
평가 대상	감점포인트	⇒ 2점
해답문	thirty	
정답문	been	

도면11

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	drinking	in	more	than	two	thirty	minutes			
정답문 No. 1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours.

평가 대상	감점포인트	⇒ 3점
and	1	
a	1	
half	1	

도면12



도면13

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	drinking	in	more	than	two	thirty	minutes.			
정답문 No. 1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours.

평가 대상	감점포인트
어군 A	1
어군 B	1
어군 C	0
배열 순서	0

⇒ 2점

도면14

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	drinking	in	more	than	two	thirty	minutes			
정답문 No. 1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours

평가 대상		감점포인트
해답문	in	7
정답문	for	
해답문	minutes	10
정답문	hours	

⇒ 17점

도면15

단어의 위치	1	2	3	4	5	6	7	8	9	10	11	12	13	14
해답문	Jack	have	eating	and	drinking	in	more	than	two	thirty	minutes			
정답문 No. 1	Jack	has	been	eating	and	drinking	for	more	than	two	and	a	half	hours

평가 대상	감점포인트
thirty	10
been	10
and	10
a	3
half	10

⇒ 43점