

지식의 깊이(DOK) 관점에 따른
중학교 문법 수행평가 문항 분석 연구
— ‘품사의 종류와 특성’ 성취기준을 중심으로

안이진 개포고등학교 교사(제1저자)

오태범 월촌중학교 교사(공동저자)

- * 이 논문은 제86회 국어교육학회 전국학술대회(2025.12.13.)에서 “지식의 깊이(DOK) 관점에 따른 중학교 문법 수행평가 문항 분석 연구-‘품사의 종류와 특성’ 성취기준을 중심으로-”라는 제목으로 발표한 내용을 수정·보완한 것이다. 당시 토론을 맡아 논의를 보완해주신 김은애 선생님께 감사의 뜻을 전한다.

- I. 서론
- II. 이론적 배경
- III. 연구 방법
- IV. 연구 결과
- V. 결론 및 제언

I. 서론

수행평가가 학교 현장에 도입된 지도 20년이 넘었다. 그동안 수행평가는 점차 교육과정 속에서 그 역할이 커져 왔으며, 특히 2015 개정 교육과정에서는 수행평가의 비중을 확대해야 함(교육부, 2015: 33)을 명시하였다. 이어 2022 개정 교육과정에서는 ‘확대’보다는 ‘내실화’에 방점을 두어, 수행평가의 질적 제고를 도모하고자 하였다(김형성, 2024: 44).

수행평가의 도입과 함께 국어과에서도 수행평가에 대한 다양한 연구가 진행되어 왔지만, 문법 영역의 경우 관련 논의가 상대적으로 부족하다(박형우, 2011: 148). 요컨대 지금까지의 문법 수행평가에 관한 연구는 주로 국어과 수행평가 전체 논의 속에서 문법 영역이 상대적으로 소홀히 다루어지고 있음을 지적하거나, 평가 도구 및 문항에 문법 교육의 목표를 반영할 필요가 있음을 제언하는 수준에 머물러 있다(박완주, 2022: 6). 수행평가의 내실화를 위해서는 성취기준을 반영한 타당한 수행평가 도구의 개발이 필수적이지만, 현재 학교 현장에서 활용되는 문법 수행평가 문항이 교육과정이 의도하는 목표와 얼마나 정합한지는 충분히 검토되지 못한 상황이다.

이러한 문제 의식을 가지고 본 연구는 Webb(1997)의 지식의 깊이(Depth of Knowledge, DOK)를 중심으로 문법 수행평가 문항이 성취기준 의도에 정합하게¹⁾ 구성되었는가를 검토하고자 한다. 본고에서 DOK에 주목하는 이유는 다음과 같다. 첫째, 문법 교육의 평가가 지향하고 있는 바가 문항에서 실제로 구현되고 있는지는, 해당 문항이 학습자에게 요구하는 사고 과정의 구조와 깊이를 분석함으로써 확인할 수 있기 때문이다. 둘째, 문법 수행평가는 학습자의 문법 지식 이해·탐구·적용 등 인지적 수행을 요구하므로 문항의 사고 수준에 대한 분석이 필수적이기 때문이다. 특히 문법 영역은 실제 국어 자료를 대상으로 언어 현상을 분석하고 설명하는 탐구 활동과 밀접하게 관련되므로, 문항이 학습자에게 어떠한 사고 과정을 요구하는지를 검토하는 것이 중요하다. 셋째, DOK는 과제가 요구하는 사고 과정의 구조를 수준별로 체계화하기 때문에, 문항 간 사고 수준의 질적 차이를 보다 선명하게 드러낼 수 있기 때문이다. 따라서 문법 수행평가에서 DOK는 학습자가 요구받는 사고 수준이 성취기준의 의도에 실제로 도달하고 있는지를 점검할 수 있는 분석 준거로 기능할 수 있다.

이에 본 연구는 문법 영역의 특성을 반영한 DOK 수준별 특징을 정리하고, 이를 중학교 국어 교과서 출판사 홈페이지에 공개된 수행평가 예시 문항에 적용하여 성취기준과의 정합성을 검토하고자 한다. 이를 통해 문법 수행평가 문항 설계에서 질적 개선이 필요한 지점을 구체적으로 드러내는 데 기여하고자 한다.

1) 선행 연구에서는 성취기준과 수행평가 문항의 DOK 수준이 동일하거나 더 높은 경우를 정합한 것으로 판단하는 기준을 제시해 왔다(천경록, 2009; 방길환, 2023). 다만 본 연구에서 사용하는 '정합성'은 판정적 의미라기보다는, 성취기준이 의도한 사고 수준이 문항에 실제로 구현되고 있는지를 살펴보는 질적 의미의 정합성을 가리킨다.

II. 이론적 배경

1. 문법 수행평가의 지향점

수행평가의 개념에 대해서는 연구자들마다 다소 차이를 보이지만, 공통적으로 강조되는 요소는 ‘과제의 수행’, ‘결과와 과정’, ‘직접적 관찰’, ‘전문적 판단’ 등이다(구민경, 2009: 8). 즉 수행평가에서는 학습자가 특정 과제를 수행하는 데 초점을 두며, 그 과정과 결과를 교사가 면밀히 관찰하여 전문적으로 평가하는 것을 중시한다.

문법 영역의 수행평가는 이러한 수행평가의 특성을 지니면서도 문법 교육의 목표가 반영되어야 한다. 이와 관련하여 고춘화(2024: 60)는 문법 교육의 목적을 학습자의 ‘문법 능력’을 향상시켜 미래 사회를 능동적으로 살아갈 수 있는 ‘언어적 주체’²⁾로 성장하는 데 있다고 보았다. 이때 ‘문법 능력’은 문법 교육의 목표를 뒷받침하는 핵심 개념으로, 그 구체적 성격에 대한 논의가 필요하다.

문법 능력은 전통적으로 언어학의 입장에서 문법적 정확성을 판단하는 능력(이도영, 1996: 124), 문장의 생성과 화용에 관여하는 능력(이춘근, 2000: 180) 등으로 규정되었다. 그러다가 이후 이관규(2008), 구본관(2010) 등에 의해 문법 지식이 문법 능력의 핵심 요소로 논의되었고, 최근에는 문법 교육의 목적에 비추어 문법 능력을 교육적 맥락에서 재정립해야 한다는 지적도 제기되고 있다(고춘화, 2024).

문법 능력을 교육적 맥락에서 논의하기 위해서는 이러한 논의가 반영된 교육과정의 진술을 함께 검토할 필요가 있다. 2022 개정 교육과정에서는

2) 여기서 ‘언어적 주체’란 언어에 관한 기본 지식과 흥미를 바탕으로 언어 현상을 분석·탐구하여 언어에 관한 인식과 태도를 주체적이고 능동적으로 형성해 나갈 수 있는 인간을 의미한다(신명선, 2007: 450).

핵심 아이디어를 중심으로 한 깊이 있는 학습을 강조한다. 문법 영역의 내용 체계 역시, 문법 교수·학습의 결과로 학습자의 마음속에 구성되기를 기대하는 앎으로 핵심 아이디어를 설정하고, 그러한 핵심 아이디어를 지향하여 문법 교수·학습이 이루어질 수 있도록 내용 체계 및 하위 범주가 설계 및 구성되었다(남가영, 2023: 371). 노은희 외(2022: 190-191)에서는 문법 영역의 핵심 아이디어가 다음과 같은 질문을 통해 도출되었음을 밝히고 있다.

〈표 1〉 2022 개정 교육과정 ‘문법’ 영역의 핵심 아이디어 도출의 원리(남가영, 2023: 374)

질문	핵심 아이디어
문법 영역에서 학습자가 경험하는 활동은 무엇인가?	문법은 국어의 형식과 내용을 이루는 틀로서 규칙과 원리로 구성·운영되며, 문법 탐구는 문법에 대해 사고하는 활동으로 국어에 대한 총체적 앎을 이끈다.
해당 활동을 통해 학습자들이 도달할 것이라 기대되는 국어에 대한 총체적 앎이란 무엇인가?	국어는 체계와 구조를 갖춘 의미 생성 자원이자, 사회적으로 구성된 관습적 규약이며, 공동체의 사고와 가치를 표상하는 문화적 산물이다.
해당 활동에서 다루어지는 ‘국어 자료’는 무엇인가?	국어 자료는 다양한 맥락에서 만들어지는 의사소통의 결과물로서, 국어 현상을 파악하고 국어 문제를 발견할 수 있는 문법 탐구의 대상이다.
해당 활동을 통해 학습자는 ‘어떠한 주체로 성장’하는가?	국어 사용자는 일상생활에서 국어 현상과 국어 문제를 탐구하고 성찰하면서 언어 주체로서의 정체성과 국어 의식을 형성한다.

요컨대, 문법 교육은 학습자가 언어 현상을 탐구하며 국어를 총체적으로 이해하도록 하는 것을 목표로 한다. 나아가 문법을 공동체적 가치가 반영된 문화적 산물로 인식하게 하고, 실제 국어 자료를 탐구하여 문제를 발견·해결하는 과정을 통해 주체적인 언어 사용자로 성장하도록 이끄는 것을 지향한다. 이러한 교육적 목표는 학습자가 갖추어야 할 문법 능력의 성격을 규정한다. 즉, 문법 능력은 문법 지식을 이해하고 언어 현상을 탐구하며 실제 국어 사용 맥락 속에서 이를 적절하게 사용할 수 있는 능력으로 정의된다(고춘화, 2024). 이러한 문법 능력은 문법 수행평가가 평가하고자 하는 핵심 능

력으로, 문법 수행평가는 이를 확인할 수 있는 방향으로 설계되어야 한다.

이러한 지향은 교육과정의 평가 원칙에도 반영되어 있으며, 실제로 2022 개정 교육과정에서는 영역별 평가의 주안점을 다음과 같이 제시하고 있다.

‘문법’ 영역에서는 문법 지식을 단순 암기하는 데 그치지 않고 국어의 구조와 문법의 작동 원리를 파악하고 이를 생활 속에 적용, 실천할 수 있도록 문법 지식의 이해와 탐구 및 적용 능력에 중점을 두어 평가한다. 태도와 같은 정의적 측면을 평가할 때는 자기 점검표나 성찰 보고서 등의 도구를 제공하여 지속적으로 누적 평가가 가능하도록 지원한다. 평가의 실제성을 확보할 수 있도록 다양한 국어 자료를 활용해 탐구 및 적용 과제를 설계하되, 과제 수행 결과뿐 아니라 국어 자료를 수집, 분석하고 언어 지식을 구성해 나가는 과정을 함께 평가할 수 있도록 평가를 설계한다.(교육부, 2022: 67, 밑줄은 연구자)

즉 문법 수행평가는 실제 국어 자료를 토대로 학습자의 문법 지식³⁾ 이해, 탐구, 적용 능력을 평가할 수 있어야 한다. 또한 결과뿐 아니라 과정에서 드러나는 학습자의 탐구와 사고를 함께 평가해야 한다.

2. 문법 영역의 ‘지식의 깊이’

지식의 깊이(Depth of Knowledge, DOK)는 Webb(1997)이 교육과정과 평가의 일치 정도를 판단하기 위하여 도입한 개념으로 평가가 성취기준에서 의도한 것과 동일한 수준의 지식이나 기능의 인지적 복잡성(Cognitive complexity)을 요구하고 있는가를 기준으로 문항을 분석한다(방길환, 2023:

3) 문법 능력이 인지적 영역에 제한되는 평가 상황에서는 ‘문법 수행’이 인지적 영역의 수행 수준에 따라 달리 작용한다(이경숙, 2021: 175). 이에 본 연구에서는 문법 수행평가의 인지적 측면에 초점을 두어, 정의적 영역은 논의에서 제외하였다.

46). 이는 Bloom의 교육목표분류체계를 보완한 것으로, Bloom의 교육목표 분류학이 학습자가 과제에 직면했을 때 나타나는 사고 과정의 유형을 분류한 것이라면, DOK는 과제를 완수하는 과정에서 요구되는 인지적 복잡성 수준에 주목한다(이은옥, 2017: 19-20). 요컨대, DOK는 Bloom의 분류체계가 같은 범주 안에서 인지적 깊이를 세분화하지 못하는 한계를 보완할 수 있다. 즉 Bloom의 체계하에서는 상대적으로 단순한 분류 활동과 고차원적 추론을 수반하는 분류 활동이 동일한 범주로 묶일 수 있다. 이는 문장이 실제로 요구하는 사고 수준의 차이를 충분히 드러내기 어렵다는 한계를 지닌다.

또한, Bloom의 분류체계는 과제의 서술어를 중심으로 분석한다는 점에서 방법론적 한계를 갖는다(정은선·김범진, 2023: 355). 과제의 서술어만을 분석 기준으로 삼을 경우 학습자에게 실제로 요구되는 인지 수준을 정밀하게 파악하기 어렵기 때문이다. 따라서 본 연구에서는 문법 수행평가 문항이 요구하는 사고의 깊이와 복잡성을 분석하기 위한 틀로서 DOK를 활용하고자 한다. 특히 문법 수행평가는 단순한 지식 확인을 넘어 탐구와 활용을 지향해야 하므로(박형우, 2011: 173), 학습자에게 요구되는 인지 수준을 정밀하게 파악할 수 있는 DOK 분석이 적합하다고 볼 수 있다.

DOK는 학습자가 과제를 수행하는 과정에서 요구되는 사고의 깊이를 네 수준으로 구분한다. 각 수준별 특징은 다음과 같다.

〈표 2〉 지식의 깊이(DOK)의 수준별 특징(Roach, Elliott & Webb, 2005: 223)

DOK 수준	주요 특징
1수준: 회상 (recall)	• 사실, 정의, 용어, 단순 절차와 같은 정보 기억 • 간단한 알고리즘 수행이나 공식 적용
2수준: 기능/개념 (skill/concept)	• 습관적 반응을 넘어서는 일정한 정신적 처리 포함 • 문제에 어떻게 접근할지에 대한 결정을 내리도록 요구 • 분류하기(classify), 조직하기(organize), 추정하기(estimate), 관찰하기(make observations), 자료 수집·제시하기(collect and display data), 자료 비교하기(compare data) 등

3수준: 전략적 사고 (strategic thinking)	- 추론(reasoning), 계획(planning), 근거 활용(using evidence) 등 앞선 두 수준보다 더 높은 수준의 사고를 요구 - 사고 과정 설명, 추측 제시, 정답이 하나 이상 존재하는 문제의 해결책 찾도록 요구
4수준: 확장된 사고 (extended thinking)	- 복잡한 추론(reasoning), 계획(planning), 발전(developing), 사고(thinking)를 요구하며, 대체로 오랜 시간 소요 - 과제가 요구하는 인지적 부담이 크고, 수행 또한 매우 복잡 - 교과 내 혹은 교과 간 연결 필요 - 실험 설계 및 수행, 발견과 관련 개념 사이의 연결, 아이디어를 결합·종합하여 새로운 개념을 창출, 문학 작품이나 실험 설계를 비판적으로 평가하기 등

이에 대해 방길환(2023)은 선행 연구를 바탕으로 DOK의 수준별 특징을 중학교 국어과 교육과정에 맞게 수정·보완하였다.

〈표 3〉 국어과 지식의 깊이(DOK)의 수준별 특징 및 준거(방길환, 2023: 53-54)

DOK 수준	주요 특징	준거
1수준: 회상 (recall)	- 간단한 완성형 문장을 쓴다. - 사실, 세부 사항, 용어, 원리를 정의한다(또는 기술한다). - 사실, 세부 사항, 용어, 원칙 등에 대한 짧은 답변 질문을 완성한다.	- 사고 기능: 회상,재인 - 글: 부분, 명시적 정보 - 사고 과정: 단순, 명료 - 사고 수준: 낮음
2수준: 기능/개념 (skill/ concept)	- 정보의 관계성 또는 사실과 의견을 구별한다. - 중심 아이디어를 뒷받침하기 위한 정보를 찾는다. - 문맥을 사용하여 단어와 단락의 의미를 확인한다. - 텍스트의 형식, 조직, 구조를 분석한다(또는 수정한다). - 문학적 요소, 용어, 사건들을 범주화한다(또는 비교한다).	- 사고 기능: 분류, 분석 - 글: 전체, 암시적 정보 - 사고 과정: 한 과정 이상, 비교적 명확 - 사고 수준: 보통
3수준: 전략적 사고 (strategic thinking)	- 도출된 결론을 정당화하거나 비판한다. - 의미, 아이디어의 진행 혹은 논리적 연결을 위한 최종안을 수정한다. - 단일 텍스트 내에서 작가의 영향력(문학적 장치, 관점, 대화)이나 사용법을 분석한다.	- 사고 기능: 추리, 추론 - 글: 텍스트 밖, 맥락 - 사고 과정: 두 가지 이상, 다소 복잡 - 사고 수준: 높음

4수준: 확장된 사고 (extended thinking)	• 새로운 문제를 연구하기 위해 많은 대안들 사이에서 방법을 선택하거나 고안한다. • 얻어진 결과나 사용된 전략의 일반화를 개발하고 새로운 문제 기반 상황에 적용한다.	- 사고 기능: 생성, 비평 - 글: 상호텍스트 - 사고 과정: 여러 과정, 여러 층위 - 사고 수준: 매우 높음
--	--	--

그런데 방길환(2023)의 논의는 주로 읽기 영역을 중심으로 한 것으로, 이를 문법 영역에 그대로 적용하기에는 한계가 있다. 이에 본 연구에서는 문법 평가의 행동 요소를 체계화한 논의⁴⁾를 참고하여, 문법 영역에 적합한 지식의 깊이 수준별 특징을 제시하고자 한다.

이경숙(2021)은 Merrill(1983)에서 학습의 결과를 인지적 영역의 수행 수준에 따라 ‘기억, 활용, 발견’으로 구분한 논의를 따랐다. Merrill(1983)은 요소 제시 이론(Component Display Theory: CDT)에서 학습의 결과를 내용 및 수행의 2차원으로 분류한다. 이때 내용은 ‘사실, 개념, 절차, 원리’로 구성되어 있으며, 수행은 ‘기억, 활용, 발견’으로 이루어져 있다(나동진·왕경수, 2000: 187). 그런데 Merrill(1983)의 논의는 지식과 인지적 능력이 교차하는 행렬표로 구성되어 있기 때문에 여기서 인지적 능력만을 차용하는 것은 문제가 있다(이경숙, 2022: 104). 이에 이경숙(2022)에서는 조수정(2014)의 ‘문법 행동 요소’ 논의에 근거하여 문법 능력의 ‘사용’에 해당하는 평가 구인을 설정하였다.

조수정(2014)은 ‘지식, 이해, 적용’에 해당하는 이원분류표의 행동 요소에 대해 예비 교사 및 현장 교사를 대상으로 설문조사를 실시하여 문법 영역 고유의 특성에 따라 다음과 같이 귀납적으로 재개념화하였다.

4) 이때 문법 수행의 인지적 깊이를 직접적으로 체계화한 연구를 검토하는 것이 가장 적절하겠으나, 현재까지는 이러한 연구가 드문 실정이다. 따라서 본 연구는 대안적 차원에서 문법 평가의 행동 요소를 체계화한 선행 연구를 검토하였다. 이들 연구는 Bloom의 교육목표분류학에 근거하고 있긴 하지만, 각 하위 요소가 구체적으로 무엇을 의미하는지 제시하고 있다는 점에서 문법 과제가 요구하는 구체적인 인지 활동을 보여 준다는 의미가 있다.

〈표 4〉 문법 영역의 행동 요소 재개념화(조수정, 2014: 28)

지식	단순히 문법 지식의 개념과 정의를 확인하고 아는 것
이해	문법 지식을 학습자가 자신의 것으로 받아들여 이를 자신의 언어로 재진술·재구성하거나 문법 지식을 응용하여 단어·문장·텍스트·표·그림 등의 자료를 분석할 수 있는 능력
적용	새로운 상황(문학 작품, 실생활, 타 교과)에 문법 지식을 활용하거나 반대로 그 상황 안에서 문법 지식을 파악할 수 있는 능력

조수정(2014)이 문법 영역의 행동 요소를 구체화한 데 이어, 권경일·김홍범(2018: 36)에서는 문법 평가 문항 유형 정립을 목적으로 전통적인 ‘지식·이해·적용·탐구’ 영역을 ‘개념 및 원리의 이해, 개념의 적용 및 명료화, 대안 탐색과 선택 및 일반화’라는 세 범주로 재구조화하였다.

〈표 5〉 문법 영역의 행동 영역 재분류화(권경일·김홍범, 2018: 37-49 참고하여 재구성)

개념 및 원리의 이해	- 기존의 행동 영역별 분류에서 ‘지식’과 ‘이해’ 영역에 해당되는 것을 통합한 것 - 교육과정에서 제시하고 있는 기본적인 개념과 원리에 대해 이해하고 기억하는 능력을 평가하는 것
개념의 적용 및 명료화	개념과 원리에 대한 이해를 바탕으로 실제 언어 자료에 대해 적용하고 명료화하는 능력을 평가하는 것
대안 탐색과 선택 및 일반화	문제 해결에 필요한 문법 정보를 바탕으로 다양한 대안을 탐색하고, 그 과정을 통하여 문제 해결에 필요한 필수 정보를 선택하여 일반화할 수 있는 능력을 평가하는 것

한편, 구분관(2010: 193)은 문법 교육 목표 능력을 ‘정확성, 적절성, 창의성’으로 설정하고, 이들이 평가의 결과를 판단하는 준거가 될 수 있음을 밝혔다. 최규홍(2009) 또한 학습자의 문법 현상 인식 단계를 정확성, 적절성, 창의성으로 구분하였는데, 정확성은 언어의 규칙성과 체계를 올바르게 파악하는 것을, 적절성은 문법 사용의 의도나 어휘 선택이 상황에 부합하는지를 판별하는 것을, 창의성은 새로운 문법 사용 현상을 만들어 내려는 것을 말한다. 이러한 준거는 문법 수행평가 문항에서 요구되는 사고 수준의 특징을 이해하는 데 참고적 틀로 활용될 수 있다. DOK의 2수준은 비교적 명확한 맥

락 안에서 규칙이나 절차를 적용하여 정답을 도출하는 활동을 포함한다는 점에서 정확성과 상응하며, DOK의 3수준은 정답이 하나 이상 존재할 수 있는 문제에서 학습자로 하여금 선택의 이유를 설명하도록 요구한다는 점에서 적절성과 관련된다. DOK의 4수준은 학습자가 새로운 개념을 창출하도록 요구한다는 점에서 창의성과 연결될 수 있다. 이상의 논의를 바탕으로 본 연구에서 문법 수행평가 문항을 분석하기 위한 지식의 깊이(DOK) 수준별 특징을 그 예와 함께 정리하면 다음과 같다.

〈표 6〉 문법 영역의 지식의 깊이(DOK)의 수준별 특징과 그 예

DOK 수준	주요 특징 ⁵⁾	예
1수준: 회상 (recall)	- 문법 지식의 개념 및 정의를 단순 기억·재생 - 문법 용어·사실을 외워 진술하는 수준	- 품사의 정의 진술하기 - 암기한 품사 예를 재생하기
2수준: 기능/개념 (skill/concept)	- 문법 지식을 활용하여 단어·문장·텍스트·표·그림 등의 자료를 정확하게 분석·분류 - 정확성에 기반해 규칙을 적용하고 판단하는 수준	- 자료에서 단어의 품사를 찾아 분류하기
3수준: 전략적 사고 (strategic thinking)	- 문법적 판단에 대해 근거를 제시하며 설명 - 맥락과 의도에 맞는 문법 지식의 전략적 사용 - 문법적 선택의 표현상 효과 설명 - 동일한 자료에 대해 다양한 해석 가능성 탐색 - 적절성에 기반해 문법적 판단을 내리고 정당화하는 수준	- 품사 분류의 근거 제시하기 - 특정 텍스트에서 많이 사용된 품사의 효과 설명하기
4수준: 확장된 사고 (extended thinking)	- 문법 지식을 종합하여 새로운 관점 구성 - 문법 지식을 새로운 맥락(문학 작품, 실생활, 타 교과 등)에 전이·적용 - 창의적 재구성과 종합적 판단을 수행하는 수준	- 품사 지식을 활용해 새로운 문장·텍스트 생성하기 - 품사 분류 기준 재구성하기

5) DOK 수준별 ‘주요 특징’ 제시된 ‘분석’, ‘설명’ 등의 표현은 동사 자체가 DOK 수준을 결정함을 의미하는 것은 아니다. 이는 문법 영역에서 자주 나타나는 사고 양상을 이해하기 위한 대표적 예시를 제시한 것으로, 수준 판정의 직접적인 기준으로 사용되지 않았다.

III. 연구 방법

1. 분석 대상 및 선정 기준

본 연구에서는 2022 개정 교육과정 중학교 성취기준인 ‘품사의 종류와 특성’에 해당하는 수행평가 문항을 분석 대상으로 하였다.

전영주·최은정(2019: 143)에 따르면 중학교 문법 수행평가 대상으로 가장 많이 언급된 문법 학습 요소가 ‘품사’이다. 이는 품사가 실제 국어생활에서의 어휘 사용과 밀접하게 연관되어 있어, 학습자가 자신의 언어생활 속 자료를 활용하는 수행 과제를 설계하기에 용이하기 때문으로 해석할 수 있다. 또한 품사는 문법 학습의 가장 기초적인 단원으로서 이후의 문법 개념 학습에 토대가 된다는 점에서 그 중요성이 크다. 특히 품사는 단순 개념 확인에서 실제 사용 맥락 분석에 이르기까지 다양한 인지 수준의 과제를 설계할 수 있으므로, DOK 분석을 통해 문항의 질적 수준을 점검하기에 적합하다.

한편, 현재 중학교 1학년은 자유학기제로 운영되고 있어 수행평가를 중심으로 평가가 이루어지고 있으며, 2022 개정 교육과정에 따른 모든 국어 교과서에서 품사 단원은 1학년에 배정되어 있다. 이러한 경향은 2015 개정 교육과정 시기에도 크게 다르지 않았기에, 품사를 소재로 한 문법 수행평가는 교육 현장에서 꾸준히 시행되어 왔으며, 최근까지도 평가 도구 개발과 관련된 논의가 이어지고 있다.

그런데 문법 수행평가의 설계는 성취기준에서부터 출발해야 하므로(전영주·최은정, 2019: 147), 품사와 관련된 성취기준을 검토할 필요가 있다.

〈표 7〉 2022 개정 교육과정, 2015 개정 교육과정 품사 관련 성취기준

2022 개정 교육과정	(9국04-03) 품사의 종류와 특성을 이해하고 국어 자료를 분석한다. 이 성취기준은 품사의 종류와 특성을 이해하고 이를 바탕으로 다양한 국어 자료를 분석할 수 있는 능력 을 기르기 위해 설정하였다. 품사 분류를 통해 우리말에 대한 이해의 폭을 넓힘은 물론 품사의 특성을 고려해 정확하고 적절하게 단어를 사용할 수 있도록 한다. 또한, 문장을 단어별로 나누어 문장 안에서의 쓰임을 분석하거나, 글이나 담화에 쓰인 단어의 정확성·적절성을 평가 하거나, 단어 부류의 빈도나 분포에 따라 글이나 담화의 특성을 간략히 기술하는 등, 품사 지식을 활용해 국어 자료의 특성을 간단히 분석해 봄으로써 품사가 국어 탐구를 위한 기초적이고 유용한 도구임을 이해하도록 한다. (밑줄은 연구자)
2015 개정 교육과정	(9국04-04) 품사의 종류를 알고 그 특성을 이해한다. (해설 없음)

2015 개정 교육과정과 비교했을 때, 2022 개정 교육과정에서 두드러지는 차이는 ‘국어 자료의 분석’이 성취기준에 명시되어 있는지와 성취기준 해설의 제공 여부이다.

2015 개정 교육과정의 성취기준에서는 품사의 종류와 특성에 대한 이해가 중심적으로 제시되어 있다. 한편 교수·학습 방법 및 유의 사항에서는 실제 국어 자료를 대상으로 일정한 기준에 따라 품사를 분류해 보는 탐구 활동이 제시되어 있다(교육부, 2015: 52). 그러나 성취기준의 진술 자체만을 기준으로 볼 때, 평가에서는 품사의 체계와 종류, 각 품사의 특성을 설명하거나 예시를 제시하는 활동이 주로 구성될 가능성이 크다. 이러한 활동은 개념적 지식에 기반한 이해 수준의 수행으로, 단순 회상 중심의 활동으로 구성될 가능성이 있다. 따라서 성취기준의 진술을 기준으로 보면 DOK 관점에서는 1수준에 가까운 평가가 설계될 가능성이 있다.

반면, 2022 개정 교육과정의 성취기준에는 ‘국어 자료의 분석’이 추가적으로 명시되어 있으며, 해설에서도 성취기준의 도달점을 단순히 품사 지식을 이해하는 것에서 나아가 ‘다양한 국어 자료를 분석하는 능력’으로 확장하여 제시하고 있다. 이때의 ‘분석’은 개별 단어의 품사 분류뿐 아니라, 정확성과 적절성의 평가, 더 나아가 글이나 담화의 특성 기술까지 포함한다. 따라서

2022 개정 성취기준은 문법 지식의 활용 과정(DOK 2수준)을 넘어, 근거를 들어 쓰임을 평가하거나 장르적 특성을 추론하는 과정(DOK 3수준)까지 요구한다. 즉, 2022 개정 성취기준은 DOK 2수준과 3수준을 지향한다고 볼 수 있다.

성취기준이 요구하는 사고 수준을 고려할 때, 수행평가 역시 이에 상응하는 수준의 사고 과정을 반영할 필요가 있다. 따라서 본 연구에서는 2022 개정 교육과정에 따른 품사 수행평가 문항을 수집하여, 개별 문항이 어떠한 깊이의 사고를 요구하는지를 검토하였다.

일반적으로 교과서 출판사에서는 각 단원에 따른 평가 예시 문항을 지도서나 홈페이지에 수록하여 현장의 교사들이 이를 활용할 수 있도록 한다. 이러한 평가 예시 문항은 평가지, 채점 기준표, 예시 답안을 갖추고 있어, 단순히 평가의 방향을 제시하는 것을 넘어 실제 현장에서 활용될 수 있을 정도의 구체성을 지니고 있다. 특히 박형우(2011: 166)에서 지적하듯, 현장의 교사들이 문법 수행평가를 위해 가장 필요로 하는 것이 ‘실질적으로 수행평가에 도움이 되는 평가 문항과 기준에 대한 자료’라는 점에서, 출판사가 공식적으로 제공하는 수행평가 자료는 교사들이 적극적으로 참조할 가능성이 높다. 또한 이러한 자료는 특정 교사의 수업 맥락에서 개별적으로 구성된 평가 문항과 달리, 교과서와 교육과정을 바탕으로 비교적 표준화된 형태로 제시된다는 점에서 문법 수행평가 문항의 전반적인 경향을 살펴보는 데에도 의미가 있다.

이에 본 연구에서는 2022 개정 교육과정에 따른 10종 교과서 출판사가 제공하는 수행평가 문항 가운데 품사 단원에 해당하는 문항을 분석 대상으로 삼았다.⁶⁾ 본 연구에서 연구의 대상으로 삼은 수행평가 문항 18개의 개요는 다음과 같다.

6) 출판사 홈페이지에 공개된 자료를 중심으로 수집하되, 지도서에 수록된 수행평가 자료를 함께 검토하였다.

〈표 8〉 분석 대상 수행평가 문항별 개요⁷⁾

출판사	수행평가 문항
비상 (박현숙 외) 1-1	(가) 명언이나 책의 구절에 쓰인 단어의 품사를 분석해 보고, 품사에 따라 문장의 의미가 어떻게 달라지는지 정리하기 (나) 영화 소개 글에 쓰인 단어의 품사를 분석해 보고, 아홉 개의 품사를 모두 활용하여 짧은 감상문 쓰기 (다) 모둠별로 우리 주변의 단어나 문장을 조사해 보고, 이를 품사별로 분류하고 그 결과를 발표하기
비상 (박영민 외) 1-2	(가) 자신의 주변 인물을 다양한 품사의 단어를 활용하여 표현하기
천재 (노미숙 외) 1-2	(가) 모둠원들과 좋아하는 노랫말을 정하여 노랫말에 나타난 단어들을 품사별로 분류하고, 그 결과를 발표하기 (나) 자신의 특징을 드러내는 단어를 품사별로 떠올려 보고, 떠올린 단어를 활용하여 자신을 소개하는 소재지를 만들기 (다) 모둠원들과 노랫말에 쓰인 품사를 분석해 보고, 분석한 내용을 바탕으로 노랫말의 특징을 정리하기
천재 (정호웅 외) 1-1	(가) 품사 노래 만들고 발표하기 (나) 품사를 사용해 작품 제목 정하고 작품 만들기
동아 (남궁민 외) 1-2	(가) 국어 자료에서 단어를 품사별로 구분하고, 품사와 관련지어 글이나 담화의 특성을 탐구하기
지학사 (서혁 외) 1-1	(가) 자신이 좋아하는 노랫말에 쓰인 단어들을 품사별로 분류하고, 단어를 적절하게 활용하여 문장 만들기
해냄에듀 (강양희 외) 1-2	(가) 품사의 종류와 특성을 반영하여 품사 노래를 만들기 (나) 품사 지식을 활용하여 제시된 국어 자료 분석하기
미래엔 (신유식 외) 1-1	(가) 품사를 설명하는 만화 만들기 (나) 노랫말에서 조사를 모두 찾아 쓰고, 조사의 종류 나누기
미래엔 (민병곤 외) 1-1	(가) 품사를 표현하는 타이포그래피 만들기 (나) 마음에 드는 동요의 가사를 찾고, 가사 속 단어의 품사를 분류해 보기 (다) 품사의 특성에 맞지 않게 잘못 쓰인 표현 찾아 고치기

7) 당초 분석 대상은 총 10종의 교과서였으나, 창비(이도영 외) 교과서는 출판사 홈페이지에 수행평가 예시 문항이 명시적으로 제공되어 있지 않아 최종 분석 대상에서 제외하였다.

분석 대상을 선정할 때는 다음과 같은 사항을 고려하였다.

첫째, 교과서 출판사에서 공식적으로 ‘수행평가’ 항목으로 분류한 평가 도구만 포함하였다. 따라서 학생 중심의 학습 활동이나 모둠 활동, 소단원· 대단원의 서술형 평가 등은 수행 활동의 형식을 취할 수는 있으나, 수행 과제를 통해 학습자의 사고 과정이나 수행 결과를 평가하는 수행평가의 특성에 부합하지 않는 경우가 많다고 판단하여 분석 대상에서 제외하였다.

둘째, 하나의 수행평가 안에 여러 개의 하위 문항이 포함된 경우, 상위 문항은 제외하고 하위 문항 각각을 독립적으로 분석하였다. 이는 각 하위 문항이 요구하는 인지적 수준이 다르기 때문에, 이를 독립적으로 분석해야 문항의 정합성을 보다 정확하게 파악할 수 있기 때문이다.⁸⁾

셋째, 하위 문항을 개별적으로 보았을 때 품사와 직접적인 관련이 없는 경우는 분석 대상에서 제외하였다.

문항
 다음 방법에 따라 모둠별로 품사 노래를 만들고 발표해 봅시다.

① 누구나 다 알 만한 노래를 골라 그 노래의 가락(멜로디)을 빌려 쓴다.
 ② 아홉 품사의 특성이 잘 드러나는 노랫말을 만든다.
 ③ 노래를 완성하면 노랫말을 컴퓨터의 화면에 띄우고 모둠원들과 함께 노래를 부른다.

1 가락(멜로디)을 빌려 쓸 노래를 골라 봅시다.
예시 199 동요 (동과 물)

2 아홉 품사의 특성이 잘 드러나는 노랫말을 만들어 봅시다.
예시 200

이름을 나타내는 명사	제언 뒤에 붙는 조사 관계언	노답 대답 감탄사는 독립언
명사를 대신하는 대명사	제언을 꾸며 주는 관형사	움직임을 나타내는 동사
수를 나타내는 수사	문장을 꾸며 주는 부사	상태나 성질만 행동사
이 것은 몸체라서 체언	관형사와 부사는 수식언	동사와 형용사는 용언
말라 활활라 활활라	말라 활활라 활활라	말라 활활라 활활라
말라 활활라 활활라	말라 활활라 활활라	말라 활활라 활활라

3 노랫말을 컴퓨터 화면에 띄운 뒤 다른 친구들 앞에서 모든 모둠원이 함께 노래를 불러 봅시다.

〈그림 1〉 천재(정호응 외) 수행평가 (가) 문항

8) 이때 하위 문항을 독립적으로 분석하되, 문항 간 관계 역시 검토하였다. 특히 문항 간의 배열과 연결이 학습자의 사고를 전략적 사고로 이끌도록 구조화되어 있는지를 보조적 분석 기준으로 활용하였다.

예컨대 <그림 1>의 4개 발문 중 분석의 대상이 되는 것은 ‘2. 아홉 품사의 특성이 잘 드러나는 노랫말을 만들어 봅시다.’만이다. 이는 상위 문항에 하위 문항이 포함된 경우 상위 문항은 분석 대상에서 제외하였기 때문이며, 1번과 3번 문항은 품사와 직접적으로 관련된 인지적인 수행을 요구한다고 보기 어려우므로 제외하였기 때문이다.

넷째, 일부 문항의 경우 발문은 단일하지만 하나의 문항 안에 여러 개의 조건을 제시하는 경우가 있었다. 이때 각 조건이 요구하는 인지적 수행이 다르다고 판단되는 경우, 이를 각각을 독립된 문항으로 간주하여 분석하였다.

1. 품사를 설명하는 만화를 만들어 보자.

| 조건 |

- 품사의 분류, 기준인 ‘형태, 기능, 의미’와 관련하여 설명하도록 할 것
- 각 품사에 해당하는 단어의 예를 2가지 이상 제시할 것

<그림 2> 미래엔(신유식 외) 수행평가 (가) 문항

예를 들어 <그림 2>는 단일한 발문으로 구성되어 있지만, 하위 조건 각각이 요구하는 인지적 수행의 깊이가 다르므로 2개의 문항으로 나누어 분석하였다.

2. 분석 방법 및 기준

수행평가 문항은 앞서 제시한 <표 6>의 DOK 수준을 기준으로 분석하였다.⁹⁾ 그 구체적인 양상은 다음과 같다.

9) 물론 문항이 요구하는 수행 수준과 학습자의 실제 수행 수준은 일치하지 않을 수 있다. 예컨대 품사의 예를 제시하는 활동의 경우 학습자가 수업에서 제시된 예를 재현하는 경우에는 회상 수준에 머무를 수 있으나, 새로운 단어를 떠올리거나 품사 판정이 어려운 사례에 대해 근거를 고려하여 예를 제시하는 경우에는 더 높은 수준의 사고가 나타날 가능성이

1수준의 ‘회상’에 해당하는 수행평가 문항은 문법 지식의 개념과 정의를 암기하여 진술하는 형태이다. 품사를 다룬 문법 수행평가에서는 주로 품사의 개념, 체계, 각 품사의 정의 및 특성을 서술하도록 요구하는 문항이 이에 해당한다.

2 1에서 선택한 품사에 대한 정보를 정리해 보자.

(1) 정의:

(2) 특성:

〈그림 3〉 미래엔(민병곤 외) 수행평가 (가) 문항

〈그림 3〉은 9개의 품사 중 1개의 품사를 골라 정의와 특성을 서술하도록 한 것으로, 문법 지식에 대한 단순 회상을 요구한다는 점에서 DOK 1수준에 해당한다.

2수준의 ‘기능/개념’에 해당하는 수행평가 문항은 문법 지식을 활용하여 다양한 국어 자료를 분석하도록 하는 문항이다. 이러한 문항들은 문법 지식을 단순히 진술하는 데 그치지 않고, 학습자가 이를 실제 자료에 정확하게 적용하는 과정을 요구한다. 예컨대 주변에서 찾아볼 수 있는 국어 자료에 쓰인 단어들의 품사를 분류하는 과제가 이에 해당한다.

도 있다. 그러나 DOK는 과제가 요구하는 인지적 복잡성에 주목하므로, 본 연구에서는 학습자의 실제 수행이 아니라 문항 자체가 요구하는 사고 과정을 중심으로 수준을 판정하였다. 다만 학습자의 실제 수행 과정에서 나타나는 사고의 깊이를 함께 살피는 일 역시 문법 수행평가 연구에서 중요한 과제가 될 수 있으며, 이는 향후 연구에서 보다 심층적으로 검토될 필요가 있다.

- 영화 소개 글에 쓰인 단어의 품사를 분석해 보고, 아홉 개의 품사를 모두 활용하여 짧은 감상문을 써 보자.

01 자신이 좋아하거나 보고 싶은 영화의 제목과 소개 글 일부를 쓰고, 어떤 품사를 가장 많이 사용하였는지 분석해 보자.

(1) 자신이 고른 영화의 제목과 소개 글에 쓰인 단어의 품사를 분석해 보자.

영화 제목	
제목에 쓰인 단어의 품사	

〈그림 4〉 비상(박현숙 외) 수행평가 (나) 문항

〈그림 4〉는 영화 제목과 소개 글에 쓰인 단어들의 품사를 분석하도록 하는 문항이다. 이는 실제 자료를 대상으로 하되, 품사 분류에 관한 지식을 적용하여 각 단어의 품사를 정확하게 밝혀야 한다는 점에서 DOK 2수준에 해당한다. 이때 학습자가 선택한 자료에 따라 분류의 난이도는 달라질 수 있다. 그러나 DOK 수준은 문항의 난이도와는 구별된다는 점에 유의하여야 하며, 본 문항은 자료의 난이도와 무관하게 DOK 2수준으로 판단할 수 있다.

3수준의 ‘전략적 사고’에 해당하는 문항은 학습자가 자신의 문법적 판단이나 수행에 대해 근거를 제시하거나, 특정한 맥락에서 문법 지식이 어떤 의도로 사용되었는지를 추론하도록 요구한다. 또한 문법 요소의 표현상의 효과를 설명하거나, 정답이 하나 이상 존재하는 문제에서 타당한 해석을 제시하도록 하는 등 보다 높은 수준의 사고를 포함한다. 품사 분류에 대한 근거를 제시하게 하거나 특정 품사의 사용 이유와 그 표현 효과를 추론하게 하는 문항이 DOK 3수준에 해당한다고 볼 수 있다.

(2) 영화의 제목과 소개 글에서 가장 많이 사용한 품사를 파악하고, 그 품사를 많이 사용하면서 얻는 효과가 무엇인지 써 보자.

가장 많이 사용한 품사	→	그 품사를 많이 사용하면서 얻는 효과

〈그림 5〉 비상(박현숙 외) 수행평가 (나) 문항

〈그림 5〉는 〈그림 4〉 바로 뒤에 이어지는 문항으로, 영화의 제목과 소개 글에서 가장 많이 사용된 품사를 밝히고, 그 품사를 많이 사용함으로써 얻어지는 표현상의 효과를 설명하도록 요구한다. 학습자는 해당 맥락에서 특정 품사가 자주 쓰인 이유를 추론하고, 그 과정에서 가설을 세우며 자신의 생각을 논리적으로 정당화해야 하므로 본 문항은 DOK 3수준에 해당한다고 볼 수 있다.

4수준의 ‘확장된 사고’에 해당하는 수행평가 문항은 주어진 상황을 넘어서 새로운 맥락에서 문법 지식을 활용하도록 요구한다. 이 수준의 문항은 국어의 다른 영역이나 타 교과와의 연계가 필요할 수도 있으며, 학습자가 문법 지식을 창의적으로 전이하고 재구성하는 과정을 포함한다. 예를 들어 특정한 품사를 활용해서 주어진 조건에 맞는 구(句)나 문장을 새로 구성하거나, 품사 지식을 고려하며 글을 쓰도록 하는 문항이 DOK 4수준에 해당한다고 볼 수 있다.

04 각자 찾은 문장에 다른 품사의 단어를 덧붙여 보고, 문장의 의미나 느낌이 어떻게 달라지는지 써 보자.



〈그림 6〉 비상(박현숙 외) 수행평가 (가) 문항

〈그림 6〉은 앞선 문항에서 학습자들이 수집한 자료(명언, 책의 구절)에서 다른 품사를 덧붙여 보고, 덧붙이기 이전과 이후의 문장의 의미와 표현 효과의 차이를 설명하도록 요구한다. 표면적으로는 〈그림 5〉와 같이 품사의 사용 효과를 물어본다는 점에서 DOK 3수준과 유사해 보이지만, 주어진 자료를 분석하는 것이 아니라 문장과 어울리는 단어를 직접 창안하여야 한다는 점에서 질적으로 구별된다. 즉 새로운 맥락에 문법 지식을 전이하고, 표현을 재구성하며 어떤 변화가 생겼는지 분석 및 정당화해야 하므로 본 문항은 DOK 4수준에 해당한다.

이러한 기준에 따라 문항의 DOK 수준을 판정하되, 실제 분석 과정에서 문항 해석과 판정 과정에서 고려해야 할 사항들을 함께 반영하였다.

3. 분석 시 유의 사항

본 연구에서는 수행평가 문항을 분석할 때 다음과 같은 점에 유의하였다.

첫째, DOK 수준은 문항의 난이도와 동일하지 않다는 점을 전제해야 한다. 난이도는 학습자의 배경지식, 제시 자료의 양과 복잡성 등 외적 요인에 따라 달라질 수 있지만, DOK 수준은 문항이 본질적으로 요구하는 인지적 사고 과정의 깊이에 의해 결정된다. 따라서 동일한 DOK 수준의 문항이라도 학습자에 따라 난이도가 다르게 인식될 수 있으며, 문항이 복잡해 보인다고 해서 반드시 높은 수준에 해당하는 것은 아니다.

둘째, 문항의 표면적 발문만으로는 DOK 수준을 판정하기는 어렵다. 따라서 본 연구에서는 발문에 드러난 표현이 아니라, 학습자가 과제를 수행하는 과정에서 요구되는 사고의 깊이를 분석하는 데 중점을 두었다.

〈표 9〉 유사한 발문이 사용되었지만 DOK 수준이 다른 경우

출판사 및 DOK 수준		수행평가 발문
천재 (노미숙 외) (나) 문항	2수준	◎ 자신의 특징을 드러내는 단어를 품사별로 떠올려 보고, 떠올린 단어를 활용하여 자신을 소개하는 소책자를 만들어 보자. ○ 자신의 특성을 잘 드러내는 체언 2가지를 떠올려 보고, 그 체언을 고른 <u>까닭을 서술하기</u>
비상 (박현숙 외) (다) 문항	3수준	◎ 모둠별로 우리 주변의 단어나 문장을 조사해 보자. 이를 품사별로 분류하고 그 결과를 발표해 보자. ○ 모둠별로 우리 주변의 단어나 문장에서 많이 사용된 품사를 찾고, 그 품사가 많이 사용되는 <u>이유 서술하기</u>

예를 들어 〈표 9〉의 두 문항은 발문에 모두 ‘이유를 서술하기’를 포함한다는 점에서 유사하다. 그러나 각 문항에서 실제로 요구하는 사고의 깊이는 차이를 보인다. 천재(노미숙 외) (나) 문항에서 학습자가 서술해야 하는 이유는 자신의 개인적 특성이나 성향과 관련된 것으로, 체언의 문법적 특성이나 쓰임을 설명하도록 요구하는 것은 아니다. 즉 이 문항에서 요구되는 ‘까닭 서술’은 체언의 문법적 특성이나 범주를 분석하기보다는 해당 체언이 학습자의 특성을 어떻게 드러내는지를 설명하는 데 초점이 있다. 물론 체언 가운데 명사, 대명사, 수사 등을 선택하는 과정에서 학습자가 품사의 의미적 특성을 부분적으로 고려할 가능성도 있다. 그러나 해당 문항에서는 이러한 품사 범주의 특성이나 사용 이유를 명시적으로 설명하도록 요구하지 않는다. 따라서 이 문항에서 확인되는 것은 체언에 해당하는 단어를 선택하는 수준에서의 지식의 정확성에 가까우며, 문법적 판단이나 근거 제시를 요구하는 수행으로 보기는 어렵다. 반면 비상(박현숙 외) (다) 문항에서 학습자가 서술해야 하는 이유는 국어 자료 전반에서 특정 품사의 사용 빈도가 달라지는 까닭과 관련된다. 이 과정에서는 단순한 적용을 넘어 가설을 세우고 근거를 제시하는 전략적 사고가 요구된다. 따라서 천재(노미숙 외) (나) 문항보다 더 높은 수준의 사고를 필요로 한다고 볼 수 있다.

셋째, 수행 결과물은 유사해 보이더라도 학습자가 과제를 해결하는 과

정에서 요구되는 인지적 수준은 다를 수 있다.

〈표 10〉 유사한 수행 결과를 요구하지만 DOK 수준이 다른 경우

출판사 및 DOK 수준		수행평가 발문
미래엔 (신유식 외) (가) 문항	1수준	◎ 품사를 설명하는 만화를 만들어 보자. ○ 품사의 분류 기준인 '형태, 기능, 의미'와 관련하여 설명하도록 할 것.
미래엔 (민병곤 외) (가) 문항	4수준	◎ 다음을 참고하여, 위에서 선택한 품사를 타이포그래피로 표현해 보자. ○ 타이포그래피(Typography)란 글자를 배열하고 디자인하여 꾸미는 기법이다. 타이포그래피의 방식에는 글자만 보고도 단어의 의미를 파악할 수 있도록 꾸미는 것도 포함된다.

〈표 10〉의 미래엔(신유식 외) (가) 문항과 미래엔(민병곤 외) (가) 문항은 모두 학습자가 그림을 그려 수행 결과를 제시하도록 요구한다는 점에서 유사해 보인다. 그러나 두 문항에서 요구하는 인지적 수준은 뚜렷한 차이가 있다. 미래엔(신유식 외) (가) 문항은 품사를 설명하는 만화를 그리되, 채점 기준이 품사의 체계(형태, 기능, 의미)의 포함 여부와 예시 제시 여부에 한정된다. 따라서 학습자는 단순히 암기한 품사 지식과 예시를 재현하면 되고, 이는 DOK 1수준에 해당한다. 만화로 표현한다는 과제가 학습자 개인의 적성에 따라 난이도의 차이를 보일 수 있으나, 이는 표현 방식의 어려움일 뿐 문법 지식 활용의 깊은 사고를 요구하는 것은 아니다.

반면 미래엔(민병곤 외) (가) 문항은 학습자가 선택한 품사를 타이포그래피로 구현하도록 요구한다. 타이포그래피는 단순한 장식이 아니라 글자의 시각적 배치를 통해 의미를 드러내는 방식이므로, 학습자는 해당 품사의 특성을 정확히 파악하는 것에 그치지 않고 이를 시각적 양식으로 재구성해야 한다. 즉 선택한 품사의 의미적·기능적 특성을 고려하여 이를 시각적 표현 방식으로 전환하는 수행이 요구된다. 이러한 과정은 문법 지식을 다른 표

현 양식으로 확장하여 활용하는 활동이라는 점에서 보다 복합적인 사고를 요구한다. 따라서 이 문항을 DOK 4수준, 즉 확장된 사고를 요구하는 수행으로 해석하였다. 이처럼 결과물의 형식이 모두 그림이라는 점에서는 두 문항이 유사해 보일 수 있으나, 실제 수행 과정에서 요구되는 사고의 깊이는 크게 다르다.

본 연구에서는 이러한 차이를 고려하여, 발문과 결과물 자체에만 주목하기보다 학습자가 과제를 해결하는 과정에서 거치게 되는 인지적 수준을 중심으로 문항을 분석하였다.

한편 이러한 분석은 연구자의 해석이 개입될 수 있는 분석이라는 점에서 판정의 일관성을 확보하는 것이 중요하다. 이에 문항이 요구하는 판단의 유형, 정답의 수렴 여부, 근거 제시 및 정당화 요구 여부 등을 판정 준거로 설정하였다. 두 연구자는 이러한 준거를 바탕으로 각 문항의 DOK 수준을 독립적으로 판정한 뒤 판정 결과를 비교·검토하였다. 이후 판정이 일치하지 않는 문항에 대해서는 반복적인 논의를 통해 수준을 합의하는 절차를 거쳤다. 또한 판정 기준과 함께 구체적인 문항 사례를 제시함으로써 분석 과정의 투명성을 확보하고자 하였다.

IV. 연구 결과

1. 분석의 개요

본 연구에서는 2022 개정 교육과정에 따른 10종 교과서의 출판사에서 제공하는 ‘품사의 종류와 특성’ 관련 수행평가 문항 42개를 DOK 수준에 근거하여 분석하였다. 각 수행평가에서 요구되는 최고 DOK 수준을 기준으로 전체 문항의 분포를 정리한 결과의 개요는 <표 11>과 같다.

〈표 11〉 수행평가 문항 분석 결과 개요

수준	1수준	2수준	3수준	4수준
문항 수	5(11.90%)	21(50.00%)	9(21.43%)	7(16.67%)

수행평가 문항의 최고 DOK 수준 분포를 살펴보면, 가장 높은 비율을 차지하는 것은 DOK 2 수준(21개, 50.00%)이었다. 반면, DOK 1수준과 4수준에 해당하는 문항은 상대적으로 적은 비중을 보였다.

DOK 1수준의 문항이 적은 것은 수행평가 특성과 관련이 있다. 개념을 단순히 암기하여 진술하는 능력은 지필평가 등 다른 평가 방식으로든 쉽게 확인할 수 있으므로, 수행평가에서 자연스럽게 낮은 비중을 차지하게 된다고 볼 수 있다.

DOK 4수준의 문항이 적은 이유는 성취기준 자체와 관련된다. 앞서 살펴본 바와 같이 ‘품사의 종류와 특성을 이해하고 국어 자료를 분석한다’는 성취기준은 DOK 3수준까지를 요구하는 것으로 이해할 수 있다. 따라서 성취기준의 도달 여부를 확인하기 위한 수행평가가 굳이 DOK 4수준까지 포함할 필요는 없으며, 이로 인해 문항 비율이 낮게 나타난 것으로 해석할 수 있다.

또한 본 연구에서는 상위 문항을 세부 하위 문항으로 나누어 DOK 수준을 분석하였으나, 실제 수행평가 과정에서 학습자들은 상위 문항 하나를 해결하기 위해 하위 문항을 순차적으로 풀이하게 된다. 따라서 하위 문항의 수와 각각의 DOK 수준, 그리고 제시 순서에 따라 학습자가 도달하게 되는 인지적 사고의 깊이는 달라질 수 있다. 아래 〈표 12〉에서는 총 18개의 수행평가 문항을 대상으로, 하위 문항별 DOK 분석 결과와 각 수행평가에서 확인된 최고 DOK 수준을 제시하였다.

〈표 12〉 수행평가 문항별 DOK 분석 결과

출판사	문항	하위 문항별 DOK 수준							최고 DOK
		i	ii	iii	iv	v	vi	vii	
비상(박현숙 외)	(가)	X	2	2	4				4수준
	(나)	2	3	4	X				4수준
	(다)	X	2	3	X	1			3수준
비상(박영민 외)	(가)	X	4						4수준
천재(노미숙 외)	(가)	2	3						3수준
	(나)	X	2	2	2	4	2		4수준
	(다)	X	X	X	2	2	3	X	3수준
천재(정호응 외)	(가)	X	1	X					1수준
	(나)	X	X	4	2	2			4수준
동아(남궁민 외)	(가)	X	2	2	3				3수준
지학사(서혁 외)	(가)	X	2	3	4				4수준
해냄에듀(강양희 외)	(가)	2							2수준
	(나)	3	3						3수준
미래엔(신유식 외)	(가)	1	2						2수준
	(나)	2	2	1					2수준
미래엔(민병곤 외)	(가)	1	4						4수준
	(나)	2							2수준
	(다)	3							3수준

(X는 '분석에서 제외된 하위 문항'을 의미함)

최고 DOK 수준을 기준으로 살펴보면, 18개 수행평가의 최고 수준 비율은 1수준 5.6%(1개), 2수준 22.2%(4개), 3수준 33.3%(6개), 4수준 38.9%(7개)로 나타났다. 이는 하위 문항을 기준으로 분석하였을 때와 비교하여 특히 4수준의 비율에서 큰 차이를 보인다. 이러한 결과는 수행평가 문항이 대체로 낮은 DOK 수준의 하위 문항을 먼저 제시하고, 이를 발판 삼아 학습자가 점

차 4 수준의 사고에 도달하도록 설계된 경우가 많기 때문에 해석할 수 있다. 실제로 비상(박영민 외) (가) 문항을 제외한 모든 4수준의 문항에서 하위 문항이 비계(scaffolding)로 작동하도록 의도된 구성을 확인할 수 있다.

문항별 DOK 수준의 비율은 표로 제시하여 양적으로 정리하였으나, 본 연구의 핵심은 각 문항의 특성과 맥락을 고려한 정성적 해석에 있다. 이에 따라 DOK 수준을 기준으로 한 수행평가 문항의 정성적 분석 결과를 다음에 제시하였다.

2. 분석의 결과 및 논의

전술한 바와 같이, 품사의 종류와 특성 성취기준은 DOK 3수준에 해당하는 깊이 있는 사고를 요구한다. <표 12>의 결과를 보면, 상위 문항을 하나의 수행평가 도구로 간주하고 최고 DOK 수준을 기준으로 비율을 산출했을 때 DOK 3수준 이상에 해당하는 수행평가의 수는 적지 않다. 그러나 개별 문항 단위로 분석해 보면, 일부 문항은 깊이 있는 문법적 사고를 유도하지 못하거나, 평가 대상인 ‘품사’와 직접적인 관련성이 낮아 문법 지식의 이해, 적용, 사용 능력을 적절히 드러내지 못하는 경우도 있었다.

이처럼 수행평가에서 요구되는 DOK 수준이 상대적으로 낮게 나타난 사례로는 앞서 <그림 1>에서 살펴본 천재(정호응 외) (가) 문항을 꼽을 수 있다.

● 평가 기준		
평가 요소	배점	채점 기준
품사의 종류와 특성 이해하기	상	아홉 품사의 특성을 잘 드러내는 노랫말을 만들었다.
	중	여섯 개 이상 여덟 개 이하 품사의 특성을 잘 드러내는 노랫말을 만들었다.
	하	다섯 개 이하 품사의 특성을 잘 드러내는 노랫말을 만들었다.
모둠원과의 협력도	상	노랫말과 가락(멜로디)을 잘 맞추었으며 모든 모둠원이 협력하여 적극적으로 발표하였다.
	중	노랫말과 가락(멜로디)을 잘 맞추었으며 모둠원들이 협력하여 발표하였다.
	하	노랫말과 가락(멜로디)을 잘 맞추지 못하였으나 모둠원들이 함께 발표하였다.

※ 노래를 만들고 발표하는 과정에 참여한 정도를 점검하는 질문에 대한 자기 평가와 상호 평가 결과를 반영한다.

<그림 7> 천재(정호응 외) 수행평가 (가) 문항에 첨부된 평가 기준표

천재(정호응 외) (가) 문항은 가락을 빌려 쓸 노래를 고르고, 아홉 품사의 특성이 잘 드러나는 노랫말을 만들고, 이를 발표하며 자기 평가를 하는 과제이다. 이는 모둠 협력 활동을 포함하고, 지필평가와 구별되는 형식을 갖추어 겉으로 보기에는 수행평가의 외형적 틀을 충실히 따른 것처럼 보인다. 그러나 실제로 학습자들에게 요구되는 인지적 깊이는 DOK 1수준에 머무른다. 학습자가 노랫말을 구성하는 과정에서도 품사의 쓰임을 맥락 속에서 분석하거나 판단할 필요는 없으며, 이미 학습한 품사의 정의나 특징을 재현하는 수준의 사고가 주로 요구되기 때문이다. 이러한 점에서 이 문항은 문법 지식의 단순 회상에 가까운 인지 과정을 요구한다고 볼 수 있다.

더욱이 <그림 7>의 평가 기준표에는 성취기준과 무관한 ‘모둠원과의 협력도’, ‘노랫말과 가락’ 등이 포함되어 있어, 문법 교육의 본래 목표와의 괴리가 드러난다.¹⁰⁾ 이는 전영주 외(2019: 148)에서 현장에서의 문법 수행평가의 개선 방안으로 지적했던, ‘평가 요소가 문법 교육의 내용과 상당히 거리가 먼 것들이 많은’ 상황이 여전히 개선되지 못한 것으로 해석된다. 이러한 한계는 ‘품사 노래 만들기’라는 유사한 형식을 취하더라도 <그림 8>의 해냄에듀(강양희 외) (가) 문항과 비교했을 때 더 분명히 드러난다.

10) 노랫말 만들기나 모둠 활동 역시 의사소통 역량, 공동체·대인 관계 역량, 문화 향유 역량 등을 신장하는 데 교육적 의미가 있다. 다만 본 연구에서 주목한 지점은 이러한 활동의 가치 자체가 아니라, 해당 요소들이 문법 수행평가의 평가 기준으로 제시될 때의 문제이다. 요컨대 문법 수행평가에서는 성취기준이 요구하는 문법적 사고의 구현 여부가 핵심이므로, 다른 역량을 기르는 활동 요소와 문법적 사고를 평가하는 기준은 평가 차원에서 구분될 필요가 있다.

1 품사의 종류와 특성을 반영하여 품사 노래를 만들어 보자.

조건

- 품사의 종류(명사, 대명사, 수사, 동사, 형용사, 관형사, 부사, 조사, 감탄사) 중 5가지 이상을 사용할 것.
- 품사의 특성을 설명하는 내용과 품사 예시를 포함할 것.

평가하기 1

평가 방식 및 평가 목표

[개별 평가] 품사의 종류와 특성을 설명할 수 있고 국어 생활에 적용할 수 있다.

평가 기준

평가 영역	평가 내용	배점
품사의 종류와 특성 설명하기	품사의 종류를 다양하게 사용했는가?(5가지 이상)	30
	품사의 특성을 설명하였는가?	35
품사 예시 활용하기	품사에 해당하는 예시를 활용하여 설명하였는가?	35
합계		100

평가상의 유의 사항

- 품사의 종류에 맞는 특성을 정확하게 설명하였는지 평가한다.
- 품사에 해당하는 알맞은 예시를 활용하였는지 평가한다.

〈그림 8〉 해냄에듀(강양희 외) 수행평가 (가) 문항과 평가 기준표

해냄에듀(강양희 외) (가) 문항은 상위 발문에서 ‘품사의 종류와 특성’을 고려하도록 하고, 하위 〈조건〉에서 ‘설명과 예시’를 함께 제시하도록 요구한다. 따라서 단순히 품사 지식을 암기하여 진술하는 수준을 넘어, 학습자가 조건에 따라 여러 품사를 선택하고 각각의 특성을 설명하면서 적절한 예시까지 연결해야 한다는 점에서 DOK 2수준에 해당한다고 볼 수 있다. 또한 노래 부르기의 형식을 활용하고 있으나, 실제 평가 기준표에서는 품사 사용의 다양성, 특성 설명, 예시 활용 여부 등 품사와 관련된 채점 요소만을 제시하고 있다는 점에서 평가의 목적에도 부합한다고 볼 수 있다. 이러한 점에서 학습자는 단순한 회상을 넘어 품사 지식을 실제 언어 자료에 적용하고 설명해야 한다는 점에서 일정한 개념적 이해와 적용을 요구받는다. 이러한 사례는 활동 중심의 수행 과제를 활용하면서도 문법 지식의 이해와 적용을 실제

수행 과정에서 드러낼 수 있도록 설계된 문항의 가능성을 보여 준다.

〈표 12〉에서 ‘X’로 표시된, 성취기준의 달성이나 문법적 사고력과 직접적인 관련성이 부족한 하위 문항 중 하나로 다음 문항을 들 수 있다.

품사의 쓰임을 바탕으로 노랫말 분석하기

● 모둠원들과 노랫말에 쓰인 품사를 분석해 보고, 분석한 내용을 바탕으로 노랫말의 특징을 정리해 보자.

1 우리 반 친구들에게 어떤 노래를 추천하고 싶은지 모둠원들과 이야기해 보자.

모듬원	노래 제목, 가수	추천하는 까닭

2 모듬원들이 추천한 노래 가운데 한 곡을 골라 노랫말을 적어 보자.

♪ 노래 제목(가수): _____

♪ 노랫말 _____

3 모듬에서 고른 노래의 특징을 정리해 보자.

장르	
주제	
멜로디	
기타	

〈그림 9〉 천재(노미숙 외) 수행평가 (다) 문항

천재(노미숙 외) (다) 문항은 노래 가사를 국어 자료로 삼아, 가사에 쓰인 단어들의 품사를 분석하고 이러한 품사의 쓰임이 노래의 분위기에 미치는 영향을 탐구하도록 하는 문항이다. 이 과제는 품사의 분류를 넘어, 품사의 쓰임이 텍스트 전반에 어떤 영향을 주는지를 파악하도록 한다는 점에서 최고 DOK는 3수준으로 분석되었다. 그러나 전반부의 하위 문항들은 노래 선정 과정에 치중되어 있으며, 학습한 품사 지식을 활용하는 과제라고 보기 어렵다. 특히 〈그림 9〉의 3번 문항은 ‘노래의 특징’을 미리 정리하도록 요구하는데, 이는 가사 분석을 넘어 ‘장르, 멜로디’까지 포함한다는 점에서 품사 단

원의 평가 문항으로 적절한지 의문이 제기된다. 나아가 이 문항으로 인해 학습자들은 최종 과제인 ‘품사의 쓰임이 노래의 주제, 분위기에 미치는 영향’을 탐구할 때 스스로 가설을 세우고 검증하기보다는, 이미 정리된 주제와 분위기를 전제로 답을 구성할 가능성이 있다. 다시 말해 하위 문항의 구성이 오히려 상위 문항이 의도한 DOK 수준에 도달하는 것을 방해할 수 있다. 즉 3번 문항이 비계로 기능하기 위해서는 이후 문항에서 그 내용을 재검토하거나 수정·정당화하도록 요구하는 구조가 함께 제시될 필요가 있으나, 해당 수행평가에서는 이러한 문항 간 연결이 충분히 명시적으로 드러나지 않는다는 점에서 한계를 지닌다. 이러한 점에서 해당 수행평가는 상위 문항에서 요구되는 전략적 사고 수준의 탐구 활동을 충분히 뒷받침하지 못하며, 하위 문항의 구성 방식이 오히려 학습자의 문법적 사고 과정을 제한할 가능성이 있다.

따라서 수행평가 문항을 설계할 때에는 단순히 활동 형식을 나열하기보다는, 문항이 의도하는 사고의 깊이에 실제로 도달할 수 있도록 하위 문항이 유기적으로 배치되어 있는지를 면밀히 고려할 필요가 있다. 학습자 활동이나 모둠 활동을 포함하더라도 문법적 사고를 충분히 요구하지 않거나, 불필요한 하위 문항으로 인해 능동적 탐구가 제약되는 일이 없도록 해야 한다. 이러한 점에서 성취기준과 문항의 DOK를 함께 분석하는 것은 수행평가 문항이 성취기준이 요구하는 문법적 사고 과정과 얼마나 정합하게 구성되어 있는지를 점검하는 데 유용하다. 또한 품사 단원을 중심으로 한 본 연구의 분석 결과는 다른 문법 단원의 수행평가 문항을 설계할 때에도 성취기준이 의도한 사고 과정이 실제 평가 과제에서 어떻게 구현되는지를 검토할 필요가 있음을 시사한다.

V. 결론 및 제언

본 연구는 2022 개정 교육과정 중학교 국어 ‘품사의 종류와 특성’ 성취 기준을 중심으로, 교과서 출판사에서 제공하는 문법 수행평가 문항이 학습자에게 요구하는 사고의 깊이를 지식의 깊이(DOK) 관점에서 분석하였다. 분석 결과 수행평가 문항의 다수는 DOK 2수준에 분포하고 있었으며, 형식적으로는 수행 활동의 외형을 갖추고 있으나 실제로는 단순 회상이나 제한된 적용 수준에 머무르는 문항도 적지 않음을 확인할 수 있었다. 일부 문항에서는 하위 문항을 비계로 활용하여 전략적 사고나 확장된 사고에 도달하도록 설계된 사례도 확인되었다. 이러한 분석 결과는 수행평가의 활동 형식만으로는 문항이 요구하는 인지적 사고 수준을 충분히 설명하기 어렵다는 점을 보여 준다.

본 연구의 의의는 문법 수행평가 문항을 활동의 유형이나 결과물의 형식 중심이 아니라, 학습자가 과제를 수행하는 과정에서 요구되는 인지적 복잡성의 구조를 중심으로 분석하였다는 데 있다. 특히 기존의 DOK 논의를 읽기 영역 중심에서 확장하여 문법 영역에 적용하고, 문법 지식의 이해·적용·사용이라는 특성을 고려하여 수준별 특징을 구체화하였다는 점에서 문법 수행평가 문항 분석을 위한 하나의 준거를 제시하였다는 의미를 지닌다.

다만 본 연구는 교과서 출판사에서 제공하는 수행평가 예시 문항을 중심으로 분석하였다는 점에서 한계를 지닌다. 실제 학교 현장에서 교사가 재구성하여 활용하는 수행평가 문항이나, 학습자의 실제 수행 과정에서 나타나는 사고의 깊이까지는 충분히 포착하지 못하였다. 또한 본 연구는 ‘품사’ 단원을 중심으로 수행평가 문항을 분석하였다는 점에서, 그 결과를 문법 영역 전체에 그대로 일반화하기에는 일정한 한계가 있을 수 있다. 품사는 실제 국어 자료에서 단어를 분류하거나 쓰임을 분석하는 과제를 비교적 쉽게 설계할 수 있는 문법 요소인 반면, 다른 문법 요소의 경우 요구되는 인지적 수

행의 양상이 서로 다를 수 있기 때문이다. 따라서 향후 연구에서는 실제 수업 장면의 수행평가 문항과 학습자의 수행 결과를 포함하여 분석하는 한편, 다양한 문법 학습 요소를 대상으로 수행평가 문항의 사고 수준을 비교·분석함으로써 DOK 기반 문법 수행평가 분석의 타당성을 보다 확장할 필요가 있을 것이다. 아울러 판정자 간 일치도를 수치화하는 양적 검증 절차를 병행함으로써 분석 결과의 신뢰도를 보다 체계적으로 확보할 필요가 있다.

아울러 문법 수행평가 문항을 설계할 때에는 수행 활동의 외형적 다양성에만 주목하기보다, 해당 문항이 성취기준이 요구하는 사고 수준에 실제로 도달할 수 있도록 하위 문항의 배열과 평가 기준을 정교하게 구성할 필요가 있다. 이를 위해 문법 수행평가 문항 개발 시 인지적 수준을 고려한 설계 원리를 마련하고, 교사 연수나 평가 자료 개발 과정에서도 이러한 관점이 적극적으로 반영될 필요가 있을 것이다.

* 본 논문은 2026.01.31. 투고되었으며, 2026.02.08. 심사가 시작되어 2026.03.07. 심사가 종료되었음.

참고문헌

- 고춘화(2024), 「국어 문법 능력의 재개념화와 이론적 구성」, 『국어교육연구』 85, 59-82.
- 교육부(2015), 『국어과 교육과정(제2015-74호)』 세종: 교육부.
- 교육부(2022), 『국어과 교육과정(제2022-34호)』 세종: 교육부.
- 구민경(2009), 「중학교 국어과 수행평가의 실태와 개선방안 연구」, 경남대학교 석사학위논문.
- 구본관(2010), 「문법 능력과 문법 평가 문항 개발의 방향」, 『국어교육학연구』 37, 185-218.
- 권경일·김홍범(2018), 「문법평가 문항의 유형 정립을 위한 연구」, 『한말연구』 49, 35-54.
- 김형성(2024), 「예비 국어 교사의 수행평가 설계 양상 분석 - 2022 개정 교육과정 <독서와 작문> 평가계획서 분석을 중심으로 -」, 『청람어문교육』 99, 43-85.
- 나동진·왕경수(2000), 「인지적 교육목표의 분류 - 하나의 대안」, 『교육과정연구』 18(1), 181-200.
- 남가영(2023), 「2022 개정 국어과 교육과정 ‘문법’ 영역 개발의 논리와 교육적 대응」, 『청람어문교육』 92, 369-407.
- 노은희 외(2022), 『2022 개정 국어과 교육과정 시안(최종안) 개발 연구』(연구보고 CRC 2022-14), 한국교육과정평가원.
- 박완주(2022), 「중학교 국어 문법 수행평가의 루브릭 개발 연구 - 품사 단원을 중심으로 -」, 전남대학교 석사학위논문.
- 박형우(2011), 「‘문법’ 영역 수행평가의 문제점과 개선 방안」, 『문법교육』 15, 147-177.
- 방길환(2023), 「지식의 깊이(DOK)에 따른 성취기준과 수행평가 문항의 정합성 분석」, 『교과와 교과서연구』 3(2), 43-72.
- 신명선(2007), 「문법교육에서 추구하는 교육적 인간상에 관한 연구」, 『국어교육학연구』 28, 423-458.
- 이경숙(2021), 「문법 수행의 재개념화와 평가 내용 선정을 위한 교육적 방향 탐색」, 『한국어문교육』 37, 163-189.
- 이경숙(2022), 「학습자 문법 능력의 평가 구인과 발달 양상에 대한 연구」, 고려대학교 박사학위논문.
- 이관규(2008), 『학교 문법 교육론』, 서울: 고려대학교 민족문화연구원.
- 이도영(1996), 「문법 교육의 목표 정립을 위한 제언」, 『국어교육』 91, 109-128.
- 이은옥(2017), 「Cognitive Rigor Matrix에 따른 교육과정 성취기준과 서술형 평가 문항의 인지 수준 정합성 분석 -서울시교육청 초등 과학과 서술형 평가 문항을 중심으로-」, 경인교육대학교 석사학위논문.
- 이춘근(2000), 「문법교육의 목적」, 『우리말연구』 10, 169-191.
- 전영주·최은정(2019), 「문법 수행평가의 실행 양상과 과제: 중학교 국어과 평가 계획을 중심으로」, 『교육과정평가연구』 22(4), 127-156.
- 정은선·김범진(2023), 「‘지식의 깊이’ 관점에서의 고등학교 『문학』 교과서 고전시가 학습활동

- 분석 연구」, 『국어교육연구』 52, 351-397.
- 조수정(2014), 「이원목적분류표의 행동 요소에 근거한 교과서 학습 활동 유형 연구 - 중·고등학교 국어 교과서 문법 단원을 중심으로 -」, 서강대학교 석사학위논문.
- 천경록(2009), 「국어과 평가의 정합성 분석의 기초」, 『국어교육』 129, 113-141.
- 최규홍(2009), 「문법 현상 인식 중심의 초등학교 문법 교육 연구」, 한국교원대학교 박사학위논문.
- Roach, A. T., Elliott, S. N., & Webb, N. L. (2005), "Alignment of an Alternate Assessment with State Academic Standards: Evidence for the Content Validity of the Wisconsin Alternate Assessment", *The Journal of Special Education* 38(4), 218-231.
- Merrill, M. D.(1983), Component Display Theory. In C. M. Reigeluth(Ed), *Instructional Design Theories and Models: An Overview of their Current States*, Lawrence Erlbaum.
- Webb, N. L.(1997), *Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education*, Research Monograph No. 6.

* 수행평가 문항 참고 인터넷 사이트(검색 일자 2026.01.30.)

두클래스 <https://mh.douclass.com/>

비바샘 <https://v.vivasam.com/main.do>

엠티치 <https://m.m-teacher.co.kr/pages/mid/Main.mrn>

티솔루션 <https://tsol.jihak.co.kr/>

해냄에듀 <https://www.hnedu.co.kr/>

T셀파 <https://mh.tsherpa.co.kr/>

지식의 깊이(DOK) 관점에 따른 중학교 문법 수행평가 문항 분석 연구 — ‘품사의 종류와 특성’ 성취기준을 중심으로

안이진 · 오태범

본 연구는 2022 개정 교육과정 중학교 국어 ‘품사의 종류와 특성’ 성취기준을 중심으로, 교과서 출판사에서 제공하는 문법 수행평가 문항이 요구하는 사고의 깊이를 지식의 깊이(DOK) 관점에서 분석하는 데 목적이 있다. 이를 위해 출판사 홈페이지에 공개된 수행평가 예시 문항을 수집하여 검토하였다. 분석 결과 수행평가 문항의 다수는 DOK 2수준에 분포하였으며, 수행 활동의 외형과 실제 요구되는 인지적 사고 수준이 일치하지 않는 경우도 확인되었다. 반면 일부 문항은 하위 문항을 비계로 활용하여 전략적·확장된 사고에 도달하도록 설계되어 있었다. 본 연구의 의의는 문법 수행평가 문항을 인지적 복잡성의 구조를 중심으로 분석하고 DOK 논의를 문법 영역에 적용하여 수준별 특징을 구체화하였다는 데 있다. 다만 출판사 제공 예시 문항 중심의 분석이라는 한계를 지니며, 향후 실제 수업 자료와 학습자 수행 결과를 포함한 확장 연구가 요구된다.

핵심어 문법 수행평가, 지식의 깊이(DOK), 인지적 복잡성, 문항 분석

Middle School Grammar Performance Assessment Items: From the Depth of Knowledge (DOK) Perspective

— Focusing on the Achievement Standard
“Types and Characteristics of Parts of Speech”

Ahn Yijin · Oh Taebeom

This study analyzes the depth of thinking required by grammar performance assessment items provided by textbook publishers. It focuses on the middle school Korean language achievement standard “Types and Characteristics of Parts of Speech” in the 2022 Revised National Curriculum, from the perspective of Depth of Knowledge (DOK). Sample items publicly available on publisher websites were collected and analyzed.

The results show that most items were distributed at DOK Level 2 and that the external format of activities did not always correspond to the actual cognitive level required. However, some items were structured with sub-items functioning as scaffolds that guided learners toward strategic or extended thinking.

This study contributes to the literature by examining grammar performance assessment in terms of cognitive complexity and extends DOK discussions to the grammar domain. However, because the analysis was limited to publisher-provided samples, future research should include classroom assessment materials and learner performance data to further validate DOK-based analyses.

KEYWORDS Grammar performance assessment, Depth of Knowledge (DOK), cognitive complexity, item analysis