

외국어 표기법 연구안

훈민바름 선언본

뜻과 미래 · IPA로 세계의 소리를 쓰는 살아 있는 한글
영어확장판 제1본 · 역사·문화·사회·정책 선언서

훈민바름 영어확장판

제1본 선언본

2026년 10월 12일

연구·검토용 비구속 문서

우리 글, 바른 소리

Park Il Hwan

2026년 10월 12일 개정판 · 연구·검토용

연결 발음과 GA t 변이 개정 부록 · 37-38쪽

훈민바름 영어확장판 4본의 구성

훈민바름 영어확장판은 하나의 체계를 네 역할로 나눈다. 선언본은 뜻과 공공적 필요를 밝히고, 해례본은 손으로 쓸 수 있는 정보와 발음 원리를 풀며, 전산본은 정보의 디지털 보존·처리 규격을 세우고, 구현본은 웹 참조구현과 시험으로 실제 작동을 검증한다.

문서	핵심 내용	답하는 질문
제1본 선언본	살아 있는 확장 한글 문자체계의 역사·문화·사회·정책·교육·접근성·AI·표준화와 세계화 비전	왜 필요한가, 대한민국과 세계에 무엇을 열 수 있는가
제2본 해례본	영어 RP·GA 정보의 음가-자모 대응, 손글씨 블록 문법, 현대 해례식 조음 안내	어떤 소리를 어떻게 만들고 적고 다시 읽는가
제3본 전산본	정보의 Unicode·구조 데이터·글꼴·셰이핑·검색·복사·역 변환 규격	디지털 환경에서 훈민바름 정보를 어떻게 변환하고 표시하는가
제4본 구현본	웹 참조구현, 사용자 흐름, 예제, 시험, 감사, 공개 배포와 오류 제보 절차	변환규칙이 실제로 웹앱을 통해 어떻게 작동하고 검증되는가

이 문서의 지위와 명칭

‘훈민바름’은 옛한글·확장 한글 자모를 활용한 「외국어 표기법」의 가능성을 연구·개발하는 동안 사용하는 내부 코드명이다. 공적 규범의 명칭을 선점하지 않으며, 향후 공공 연구·검토와 제도화 과정에서는 별도의 공식 명칭이 정해질 수 있다.

현재 훈민바름은 정밀 외국어 표기체계의 구현 가능성을 시험하기 위한 하나의 가표본(working specimen) 연구이자 참조 실험(reference experiment)이다. 제시된 자모 배당과 표기규칙은 그대로 채택하도록 요구하는 규범이 아니며, 공개 검토와 시험을 통해 수정·보완하거나 더 나은 대안으로 대체할 수 있다.

첫 참조모듈은 영어 RP·GA 영어확장판이다. 영어는 전체 구조를 검증하기 위한 첫 시험 언어이며, 다른 언어는 각 언어의 원문·source IPA·음운체계와 실제 화자 검토를 갖춘 독립 모듈로 연구한다. 현재 영어확장판의 세부 규칙 역시 잠정 고정·시범·미결 상태를 포함하며 최종 답안을 선점하지 않는다.

현재 개발·구현 상태

제1-4본은 연구·검토용 판본이다. 2026년 10월 9일 교정판은 문언·예시·상호 참조와 구현 상태를 대조한 판본이며, Broad 웹앱은 공개 배포됐다. 미결 음가·최종 조자법·Narrow 구현과 교육 효과 검증은 후속 과제다.

<https://github.com/ivanpark/hmbr>

차례

권두 선언 · 세계의 소리를 받아들이고, 세계가 쓰는 글자로	P4
머리말 · 살아 있는 글자체계로 세계의 소리를 쓰는 일	P7
제1부 역사적 정당성	P10
1. 훈민바름: 개발 코드명과 영어 확장판의 현재 범위	
2. IPA: 세계 언어와 한글을 잇는 국제 음성 표기체계	
3. 구한말의 외국어 표기 실험: 『국문정리』와 『아학편』	
4. 구한말부터 1986년까지: 확장과 제한의 교차	
5. 1948년 국가안이 남긴 방법론과 과제	
6. 왜 지금 다시 가능한가	
제2부 문화적 계승성	P16
7. 훈민정음의 생성 원리를 오늘에 잇는다는 것	
8. 옛한글의 복원이 아니라 살아 있는 조자법의 계승	
9. 손으로 완성되는 정본과 문자 주권	
10. 한글의 세계화·국제화와 영어화의 차이	
제3부 사회적 필요성	P19
11. 외국어 기반 신개념 수용 지연과 판교체	
12. 교육·사전·고유명사·학술·접근성·언어기술의 공백	
13. 이름과 지식이 서로를 찾게 하는 데이터 층	
제4부 공공정책적 의미	P23
14. 외래어 표기와 정밀 외국어 표기의 두 층	
15. 규제 완화가 아니라 적용 범위의 합리화	
16. 공개·비독점·기술중립의 원칙	
17. 반대 논거를 설계 제약으로 받아들이기	
제5부 교육과 접근성의 가치	P25
18. 학습자가 소리의 차이를 먼저 볼 수 있게 하는 표기	
19. IPA를 대체하지 않고 IPA로 가는 다리 놓기	
20. 접근성은 부가 기능이 아니라 규격의 조건	
21. 무엇을 시험해야 하는가	
제6부 AI와 표준화의 미래	P27
22. 공개 발음 데이터라는 공공 인프라	
23. TTS·G2P·ASR·LLM에 생기는 변화	
24. IPA와 국제표준 호환 체계를 통한 세계 언어 확장	
25. 응용 프로파일에서 국가표준까지	
제7부 한계와 과제	P30
26. 과장하지 않기: 부호화와 상호운용성은 다르다	
27. 복잡성·독점·영어중심성의 위험	
28. 향후 10년을 위한 단계적 로드맵	
맺음말 · 전환이 그지없는 글자의 공공적 미래	P32
부록 A. 공공성 원칙 12조	P33
부록 B. 핵심 연표	P34
참고문헌 및 작성 근거	P35

권두 선언 · 세계의 소리를 받아들이고, 세계가 쓰는 글자로

나랏말싸미 中國에 달아

文字와로 서르 스몏디 아니홀씨¹

훈민정음은 말과 문자가 서로 맞지 않는다는 문제에서 출발하였다. 우리말은 당시 문자생활의 중심이던 중국의 문자와 달랐고, 그 때문에 우리말의 소리와 뜻을 온전히 적고 펴는 데 어려움이 있었다. 세종은 우리말을 빌려 쓰던 문자에 계속 맞추는 대신, 우리말의 소리를 우리 글로 적을 수 있는 새로운 문자체계를 열었다.

오늘날 그 문제는 사라진 것이 아니라 범위가 넓어졌다.

조선시대의 문제가 우리말과 중국 문자 사이의 불일치였다면, 세계화된 오늘날에는 세계 각국에서 들어오는 말소리와 이를 받아들이는 문자표기 사이의 불일치가 새로운 과제가 되었다. 이제 한국어는 영어를 비롯한 수많은 외국어의 이름·용어·개념을 교육·산업·과학·기술·문화 속에서 끊임없이 받아들인다. 세종 시대의 문자 문제가 주로 中國을 향하고 있었다면, 오늘의 문자 문제는 世界 전체를 향한다.

그러나 세계의 소리를 받아들이는 두 통로에는 각각 한계가 있다. 영어의 로마자 철자는 자소와 소리값의 대응이 일정하지 않아 한국어 사용자가 처음 보는 낱말의 실제 발음을 직관적으로 알기 어렵다. 반대로 현행 외래어 표기는 한국어 문자생활의 편의와 통일을 위해 외국어의 여러 음가 차이를 현용 한글 안에서 합리적으로 단순화한다. 일반 외래어 표기로서는 필요한 기능이지만, 원어의 실제 소리를 구별하고 배우고 기록해야 하는 경우에는 중요한 음가 정보가 사라질 수 있다.

세계의 이름과 소리와 개념이 일상적으로 국경을 넘는 오늘날에도 이처럼 말과 문자가 ‘서르 스몏디 아니하는’ 지점은 여전히 남아 있다.

따라서 현행 일반 외래어 표기의 편의와 통일성은 그대로 지키면서, 교육·사전·학술·고유명사·접근성·언어기술처럼 원어 음가의 보존이 필요한 영역에는 별도의 정밀 외국어 표기층을 연구할 필요가 있다. 그 한 방법은 세계의 소리를 현용 한글 24자모의 범위에 계속 맞추는 대신, 훈민정음이 본래 지닌 조자와 확장의 원리를 활용하여 한글이 받아들일 수 있는 소리의 범위를 넓히는 것이다.

여기서 다시 물어야 한다. 한글을 보존한다는 것은 무엇인가.

1950년대 이후 외래어 표기 규범이 현용 자모 중심으로 정착해 온 것은 당시의 문자생활과 인쇄·교육 환경에서 충분한 이유가 있었던 선택이다. 그러나 그 시대의 기술적·사회적 조건에서

¹ 『훈민정음』 언해본은 ‘中國’에 대하여 “皇帝 거신 나라히니 우리나라 常談에 江南이라 호느니라”라고 협주한다. 본문의 서문 인용은 가독성을 위해 언해본 원문의 방점과 동국정운식 한자음 표기를 덜어낸 형태다.

정해진 운용 범위를 오늘날까지 그대로 고정하는 것이 곧 훈민정음을 보존하는 일이라고 볼 수는 없다.

보존은 동결이 아니다.

지켜야 할 것은 어느 한 시대에 사용하던 자모의 목록만이 아니다. 소리를 관찰하고 그 차이를 문자로 드러내며, 필요한 경우 기존 자모를 가획·합용·합성·병서·연서하여 새로운 소리를 적게 했던 훈민정음의 철학과 조자 원리도 함께 계승되어야 한다. 옛한글 자모 또한 박물관 속에 보존해야 할 과거의 흔적만이 아니라, 그러한 생성 원리가 남긴 문자 자원이다.

현용 24자모만을 한글의 영구적인 경계로 삼는다면 한글은 한 시기의 모습으로 보존될 수는 있어도, 새로운 소리와 새로운 언어환경에 대응하며 성장하는 살아 있는 문자가 되기는 어렵다. 반대로 옛한글 자모와 훈민정음의 조자 원리를 오늘의 음성학·IPA·Unicode 환경에서 다시 연구하고 검증한다면, 우리는 옛한글을 단순히 복원하는 것이 아니라 한글을 살아 있는 문자로 계승하는 일을 시작할 수 있다.

과거에 만들어진 글자를 보존하는 데서 한 걸음 더 나아가, 글자를 만들어 낸 철학과 원리를 다시 살아 움직이게 하자.

옛한글·확장 한글 자모를 국제음성기호 IPA와 연결하면 한국어에 없는 자음과 모음, 장단과 강세, 지역적 발음 차이 등을 필요에 따라 구별하여 기록하는 한글 기반 정밀 표기체계를 시험할 수 있다. IPA는 세계 언어의 소리를 비교할 수 있는 국제적으로 통용되는 음성 공통층을 제공하고, 한글은 그 음가를 사람이 쓰고 읽을 수 있는 초성·중성·종성의 문자구조로 조직할 수 있다.

그러나 이 가능성은 한국인이 외국어를 쉽게 읽게 하는 데서 끝나지 않는다.

세계의 다양한 말소리를 한글로 일관되게 기록할 수 있다면, 확장된 한글은 한국어만을 기록하는 문자에서 한 걸음 더 나아갈 수 있다. 다른 언어권의 사람들도 자신의 언어와 다른 언어의 소리를 기록하고 배우는 데 필요에 따라 선택해 사용할 수 있는 문자, 곧 세계의 소리를 담을 수 있기 때문에 세계가 사용할 이유를 가진 문자로 발전할 가능성이 생긴다.

이것이 이 문서가 말하는 한글의 세계화다. 한글을 해외에 알리고 보급하는 문화홍보에만 머무르는 것이 아니라, 세계의 언어를 실제로 담을 수 있는 기능을 넓혀 세계가 한글을 사용할 이유를 만드는 것이다. 세계의 소리를 한글 안으로 받아들이는 일과, 그렇게 확장된 한글을 다시 세계에 열어 놓는 일은 서로 반대 방향이 아니라 하나의 순환이다.

세계에서 한글로, 외국어의 소리·이름·개념을 원문과 IPA를 보존하면서 더 정확하게 받아들이는 일이다. 그리고 한글에서 세계로, 그 과정에서 확장된 문자 능력을 특정 언어나 국가에 가두지 않고, 세계의 여러 언어에 적용해 보며 다양한 사용자가 검토하고 활용할 수 있도록 개방한다.

훈민바름은 이 비전을 이미 완성했다고 선언하는 체계가 아니다. 훈민정음의 조자 원리와 옛 한글·확장 한글 자모를 IPA와 연결했을 때 이러한 양방향 문자 확장이 실제로 가능한지를 시험하기 위해 만든 하나의 가표본(working specimen) 연구이자 참조 실험(reference experiment)이다. 현재 영어 RP·GA 확장판의 자모 배당과 세부 규칙 역시 최종 답안을 선점하지 않으며, 공개 검증과 전문가·관계기관·사용자의 참여를 통해 수정·보완되거나 더 나은 대안으로 대체될 수 있다.

이 제안이 요구하는 것은 훈민바름이라는 하나의 표기법을 국가와 국민에게 강제하는 일이 아니다. 현행 일반 외래어 표기는 유지하되, 정밀한 외국어 음가 기록이 필요한 영역에서 옛 한글·확장 한글의 가능성을 다시 연구하고 실험할 제도적 공간을 여는 것이다.

中國에서 世界로

훈민정음이 우리말과 빌려 쓰던 문자 사이의 불일치를 풀어 우리말을 우리 글로 쓰게 했다면, 오늘의 확장 한글은 세계에서 들어오는 소리를 우리 글로 더 정확히 받아들이고, 그 과정에서 넓어진 한글의 표현능력을 다시 세계에 열어 놓는 문자 확장의 가능성을 시험한다.

한글을 보존한다는 것은 한 시대의 자모 목록을 동결하는 일이 아니라, 새로운 소리를 만나면 다시 확장할 수 있었던 훈민정음의 정신을 살아 있게 하는 일이다.

한글의 세계화란 한글을 세계에 알리는 데서 끝나는 것이 아니라, 전 세계의 사용자가 필요에 따라 소리를 기록하는 문자로 한글을 선택할 수 있는 방안을 마련하여, 세계가 한글을 사용할 이유와 가능성을 만드는 것이다.

머리말 · 살아 있는 글자체계로 세계의 소리를 쓰는 일

앞서 권두 선언이 제기한 문제는 결국 한 가지 질문으로 모인다. 한글을 보존한다는 것은 한 시대의 자모 목록을 그대로 지키는 일인가, 아니면 새로운 소리와 언어환경을 만날 때 다시 확장할 수 있었던 훈민정음의 철학과 조자 원리까지 살아 있게 하는 일인가.

이 안은 후자의 가능성을 검토한다. 현행 일반 외래어 표기의 편의와 통일성은 그대로 존중하면서, 원어의 실제 음가를 구별하고 기록해야 하는 영역에서는 옛한글·확장 한글 자모와 훈민정음의 조자 원리를 다시 활용한 별도의 정밀 외국어 표기체계를 마련할 수 있는지를 역사·문화·사회·정책·교육·접근성·언어기술의 관점에서 살핀다.

여기서 훈민바름은 이미 완성된 규범이나 유일한 해답을 뜻하지 않는다. 현재의 훈민바름 영어확장판은 이러한 확장 한글 외국어 표기체계가 실제로 구성되고, 사람이 쓰고 읽으며, 디지털 환경에서 처리될 수 있는지를 보여 주기 위한 하나의 가표본(working specimen) 연구이자 참조 실험(reference experiment)이다. 일부 규칙은 현재 단계에서 충분히 검토되어 잠정적으로 고정될 수 있고, 일부는 시험 중이며, 또 일부는 향후 전문가·관계기관·사용자의 공동 검토를 위해 열어 둘 수 있다. 중요한 것은 모든 세부규칙을 한 사람이 미리 완결하는 것이 아니라, 다양한 대안을 실제로 구현하고 비교하고 검증할 수 있는 공개된 문자적 공간을 만드는 데 있다.

현재 구체적으로 문서화하고 구현하는 첫 참조모듈은 영어의 Received Pronunciation(RP, 영국식)과 General American(GA, 미국식)이다. 영어는 국제적 사용범위가 넓고 한국어 사용자에게 교육·산업적 수요가 크며, IPA(International Phonetic Alphabet, 국제음성기호)·발음사전·음성자료가 풍부하여 확장 한글의 가능성을 시험하기에 적합한 출발점이다.

현재 영어확장판은 RP와 GA를 기준으로 영어의 음절구조, 자음과 모음의 음가, 장단, 강세와 주요 발음 변이를 서로 구별하여 한글 기반 정본으로 기록하는 구조를 시험한다. 이는 단순히 영어 단어를 비슷하게 한글로 옮겨 적는 작업이 아니라, 원어 철자와 source IPA, 음소와 음절의 구조, 한글 자모값과 블록 관계를 서로 연결하고 다시 역검증할 수 있는 표기체계를 만드는 실험이다. 구체적인 음소-자모 대응과 음절·블록 구성, Broad와 Narrow 표기의 세부 원리는 제2본 해례본에서 다룬다.

그러나 RP와 GA는 영어의 유일한 표본이 아니라 현재 구현하는 첫 두 참조 방언(dialect)이다. 영어에는 지역·사회·세대에 따라 다양한 발음체계가 존재하며, 향후 Australian English, Canadian English를 비롯한 다른 영어 방언과 발음 변이도 별도의 프로파일 또는 모듈로 추가할 수 있다. 새로운 모듈은 기존 RP·GA 값을 기계적으로 복사하는 것이 아니라 해당 방언의 source IPA, 음소체계와 실제 사용을 근거로 독립적으로 구성되어야 한다.

영어 자체도 훈민바름의 목적지가 아니다. IPA가 범언어적으로 기술할 수 있는 음성 정보를 공통 매개층으로 삼으면, 다른 언어 역시 각자의 원문 문자, source IPA, 음소체계, 음절구조, 장단, 강세, 성조, 지역변이와 실제 화자 검토를 갖춘 독립 언어모듈로 연구할 수 있다. 확장성의 의미는 영어를 위해 만든 표를 세계 언어에 강제하는 데 있지 않다. 각 언어의 차이를 보존하면서 그 그것을 한글의 문자구조와 연결할 수 있는 공통 방법을 마련하는 데 있다.

따라서 현재의 영어 RP·GA 확장판은 ‘한글로 영어를 적는 완성 규격’이라기보다, 훈민정음의 조자 원리와 옛한글·확장 한글 자모가 새로운 언어와 방언을 계속 받아들일 수 있는지를 검증하는 첫 가표본이다. 영어에서 시험한 원리가 다른 영어 방언과 다른 언어에서도 검증되고, 필요한 경우 새 자모값과 조자규칙이 공개적으로 추가·수정될 수 있다면 한글은 고정된 자모 목록이 아니라 새로운 언어환경에 대응하며 성장하는 살아 있는 문자체계로 기능할 수 있다.

훈민바름 영어확장판은 이 실험을 네 문서로 나누어서 보여준다. 제1본 선언본은 왜 이러한 표기체계가 필요한지와 그 역사·문화·사회·정책적 의미를 밝힌다. 제2본 해례본은 사람이 손으로 쓸 수 있는 영어 정본과 음소·조음·표기 원리를 풀고, 제3본 전산본은 그 정본을 디지털 환경에서 Unicode와 구조 데이터로 저장·교환·검색·역변환하는 전산 규격을 다룬다. 제4본 구현본은 웹 참조구현과 시험을 통해 실제 작동 가능성과 한계를 검증한다.

뜻·원리·전산·구현을 나눈 이유는 하나의 실험적 매핑을 곧바로 완성된 규범으로 오인하지 않게 하기 위해서다. 어느 부분이 문자철학과 정책 제안이고, 어느 부분이 현재의 음성학적 판단이며, 어느 부분이 전산 구현상의 선택인지를 분리해야 각각을 독립적으로 검토하고 수정할 수 있다.

이 책은 세 가지 질문을 중심으로 전개된다.

첫째, 무엇을 보존할 것인가.

현용 24자모의 한글만을 한글의 절대적이고 영구적인 체계로 삼을 것인가, 아니면 새로운 소리를 관찰하고 필요한 글자를 조합해 내던 훈민정음의 철학과 조자 원리까지 계승할 것인가.

둘째, 무엇을 받아들일 것인가.

세계에서 들어오는 소리·이름·개념을 원어 철자와 현용 한글 사이의 간극에 그대로 둘 것인가, 아니면 IPA와 확장 한글을 매개로 그 음가를 더 정확하게 읽고 기록할 수 있는 별도의 통로를 만들 것인가.

셋째, 어디까지 열 것인가.

그 통로를 한국인의 외국어 학습을 위한 보조표기에만 머물게 할 것인가, 아니면 다른 영어 방언과 세계의 여러 언어로 계속 확장하여 다른 언어권의 사용자도 자신의 언어와 다른 언어의 소리를 기록하고 배우는 데 선택하여 사용할 수 있는 열린 한글 문자체계로 발전시킬 것인가.

물론 이 세 질문의 답은 정해져 있지 않다. 훈민바름은 답을 명령하기보다 하나의 가능한 답을 실제로 작동시켜 보여 주고, 그 결과와 미결사항을 함께 공개하고 사회적 담론으로 열어, 더 나은 답을 공동으로 찾아가기 위한 출발점이다.

字雖簡要 轉換無窮 — 글자는 비록 간단하고 요약되나 전환이 그지없다.

『세종실록』 세종 25년 12월 계해조; 홍운표(2012: 2) 재인용

이 연구가 시험하는 것

훈민바름은 옛한글·확장 한글과 IPA를 연결하여 세계의 소리를 한글로 더 정확히 받아들이고, 영어 RP·GA에서 시작해 다른 영어 방언과 세계 언어로 확장하며, 그 과정에서 되살아난 훈민정음의 조자 원리를 다시 세계에 열어 놓을 수 있는지를 시험하는 하나의 가표본 연구이다.

제1부

역사적 정당성

IPA라는 국제 음성 공통층과, 확장과 제한이 교차해 온
외래어 표기 규범의 역사 속에서 훈민바름의 자리를 찾는다

1. 훈민바름: 개발 코드명과 영어 확장판의 현재 범위

전술한 바와 같이 훈민바름은 「외국어 표기법」을 연구·개발하는 동안 사용하는 내부 코드명이다. 이 명칭은 ‘우리 글로 바르게 소리를 적는다’는 개발 철학을 압축하지만, 공적 규범의 최종 공식 명칭을 선점하지 않는다. 본 문서는 이 지위를 밝힌 뒤 가독성과 연속성을 위해 훈민바름이라는 명칭을 계속 사용한다. 공적 제도화 과정에서는 별도 공식 명칭이나 대중적 이름도 별도로 검토될 수 있다.

현재 기술적으로 구체화하여 문서화·시험하고 있는 첫 훈민바름 참조모듈은 영어 RP·GA 확장판이다. 영어 자음·모음·복모음, 음절과 자음군, 장단 및 강세와 주요 Narrow 변이를 source IPA에서 훈민바름 정본으로 변환하고 다시 IPA로 역변환하는 구조를 다룬다. 그러므로 현재 문서가 곧 세계 모든 언어의 완성 매핑이라는 뜻은 아니다. 영어판은 IPA 음소·음성 분석, 음소 ID, 자모값 ID, 언어모듈, Broad/Narrow, 손글씨 정본, 전산 규격·표시층과 공개 시험이 실제로 작동하는지를 검증하는 첫 참조모듈이다.

4본은 이 범위와 확장 경로를 분명히 나눈다. 제1본 선언본은 뜻과 미래를 밝히고, 제2본 해례본은 영어 정본과 음소 해례를 규정하며, 제3본 전산본은 디지털 저장·교환·처리·표시 규격을 맡고, 제4본 구현본은 웹 참조구현과 공개 시험을 맡는다. 다음 언어판은 영어 매핑을 복사하지 않고 해당 언어의 원문과 IPA 기술, 음운분석, 화자 공동검토를 거쳐 독립 모듈로 추가한다. IPA는 언어 사이의 국제 음성 공통층이며, 원문 문자·로마자 전사·언어코드·Unicode와 함께 보존되어야 한다.

현재 기술 규칙은 영어 RP·GA 확장판이다. 세계 확장은 영어표의 복제가 아니라 ‘원문 문자 → source IPA → normalized IPA → 언어별 음운분석 → 훈민바름 정본 → 디지털 표시·데이터’의 공통 절차를 각 언어에 적용하는 일이다.

2. IPA: 세계 언어와 한글을 잇는 국제 음성 표기체계

국제음성기호(IPA)는 한 언어의 철자법이나 로마자 읽기 규칙이 아니다. 인간이 발음하는 소리를 조음과 청각의 공통 기준으로 분석하고 기록하기 위한 국제 표기체계다. International Phonetic Association은 음성학의 과학적 연구와 실제 응용을 위해 이를 발전시켜 왔으며, Unicode Standard는 IPA가 모든 언어의 음성을 나타낼 수 있도록 라틴 문자를 확장한 체계라고 설명한다(International Phonetic Association 1999·2018; The Unicode Consortium 2025, ch. 7).

IPA의 자음표는 공기가 막히거나 좁아지는 위치와 방법, 성대의 울림을 기준으로 소리를 배열한다. 모음표는 혀의 높이와 앞뒤 위치, 입술의 둥글림을 좌표로 삼는다. 여기에 장단, 강세, 성조, 비음화·기식·권설화 같은 보조기호를 더해 Broad 음소 전사부터 Narrow 음성 전사까지 기록할 수 있다. 이 체계 덕분에 라틴·키릴·아랍·한자·데바나가리 등 서로 다른 문자를 사용하는 언어도 발음정보를 비교 가능한 한 층에 놓을 수 있다.

그러나 IPA가 존재한다고 해서 어느 언어든 자동으로 한글로 변환되는 것은 아니다. 같은 IPA 기호도 언어 안에서 차지하는 음운적 지위와 결합 방식이 다를 수 있고, 언어마다 음소 목록·음절구조·이음 규칙·강세·성조·지역변이가 다르다. 그러므로 훈민바름의 세계 확장은 IPA 기호표를 기계적으로 치환하는 일이 아니라, 각 언어의 source IPA와 음운분석을 보존한 독립 모듈을 공개 연구와 화자 검토로 구축하는 일이다.

훈민정음은 초성·중성·종성의 자리와 가획·합용·합성·병서·연서의 조자 원리를 통해 소리의 관계를 한글 블록 안에 조직한다. 현용하지 않는 옛한글과 확장 자모는 이 생성 능력의 역사적 자원이다. 훈민바름은 IPA가 제공하는 국제 음성 좌표를 이 자모 자원과 연결하여, 원문을 지우지 않고도 그 발음구조를 한글로 손수 적고 다시 IPA로 역검증할 수 있는 정본을 세우려 한다.

연결 층	맡는 일	보존하는 정보	한글 세계화의 효과
원어층	언어가 실제로 쓰는 문자와 명칭을 보존	원문 문자·표제어·의미·출처	원어를 한글로 대체하지 않고 연결
IPA층	서로 다른 언어의 소리를 국제 공통 좌표로 기술	source IPA·Broad/Narrow·강세·성조	언어 간 비교와 국제적 역검증
언어모듈	각 언어의 음소·음절·변이 규칙을 해석	음소 ID·환경규칙·지역변이	영어중심성을 피하고 언어별 정확성 확보
훈민바름 정본	IPA 분석을 확장 한글 자모와 블록으로 기록	자모값·위치형·블록·역변환 규칙	세계의 소리를 한글 문자생활 안에 높은 통합성으로 수용
디지털 공공데이터	정본을 저장·검색·교환·음성기술에 활용	Unicode·구조 데이터·음성·라이선스·버전	교육·사전·접근성·AI·표준의 공공 인프라

이 구조에서 한글의 세계화는 외국 문자를 몰아내거나 모든 언어를 한국어식으로 바꾸는 일이 아니다. 원문 문자와 IPA를 보존하면서 세계 언어의 소리정보를 한글의 초·중·종성 구조 안에 높은 일관성으로 통합하고, 사람이 쓰고 기계가 처리하며 국제적으로 다시 검증할 수 있게 하는 일이다. 여기서 말하는 높은 통합성은 단순한 발음 유사성이 아니라 원문·음가·언어모듈·한글 정보·디지털 데이터가 한 레코드 안에서 연결되는 구조적 통합을 뜻한다.

훈민정음이 백성이 자기 말을 빌린 문자에 맞추지 않고 자기 글로 펴게 한 문자 해방이었다면, 훈민바름은 세계의 말소리를 현용 자모의 범위에서 합류시키지 않고 필요한 음가 차이를 한글로 구별해 기록할 수 있게 하는 문자 확장이다. 그 역사적 평행은 과거를 그대로 반복한다는 뜻이 아니라, ‘말과 문자가 맞지 않아 생기는 고통을 줄인다’는 창제 정신을 오늘의 국제 문자생활에 적용한다는 뜻이다.

규제합리화의 핵심

현행 일반 외래어 표기에서는 현용 24자모 원칙을 유지한다. 다만 교육·사전·학술·접근성·언어기술과 신규 개념의 초기 수용을 위한 정밀 외국어 표기층에서는 옛한글·확장 한글 자모를 허용한다. 이 선택적 제한 해제가 IPA와 한글을 연결하는 제도적 통로이며, 한글 세계화의 출발점이다.

3. 구한말의 외국어 표기 실험: 『국문정리』와 『아학편』

한글로 외국어의 소리를 기록하려는 시도는 1930년대의 어문 규범에서 갑자기 시작된 것이 아니다. 한자음을 바로잡아 적으려 했던 『동국정운』까지 넓게 거슬러 올라갈 수 있지만, 오늘날의 외국어 표기 문제와 직접 맞는 시도는 개항 이후 세계의 언어와 문물을 본격적으로 접하게 된 구한말에 더욱 뚜렷하게 나타났다. 1897년 이봉운의 『국문정리』와 1908년 지석영이 주석·증보한 『아학편』은 이러한 근대 외국어 표기의 대표적인 선례다.

1897년 간행된 이봉운의 『국문정리』는 국문의 역사와 자모·문법을 논하면서 일본어의 탁음을 구별하여 적는 방안을 제안하였다. 이는 외국어의 음운 대립이 한국어의 통상적인 자모 운용만으로 충분히 드러나지 않을 때, 한글 표기 안에 별도의 구별 장치를 마련하려 한 초기 실험이었다. 오늘날의 기준으로 완성된 정밀 전사체계는 아니었으나, 외국어의 서로 다른 소리를 한글에서 어떻게 보존할 것인가라는 문제가 이미 대한제국기에 제기되고 있었음을 보여 준다.

1908년의 『아학편』은 이러한 문제의식을 실용적인 외국어 학습의 장으로 옮겼다. 지식영은 정약용의 원저에 중국어·일본어·영어의 발음 정보를 더하여, 한자를 중심으로 한국어·중국어·일본어·영어를 한 권에서 함께 익힐 수 있도록 구성하였다. 영어 부분에는 로마자 철자뿐 아니라 그 발음을 한글로 병기하여, 학습자가 영어 철자를 모르는 상태에서도 한글을 통해 소리에 접근할 수 있게 하였다.

특히 『아학편』은 당시의 현용 자모에만 머물지 않고, ㅞ를 비롯한 합용병서와 옛자모를 활용하여 영어의 발음을 보다 세밀하게 나타내려 했다. 예컨대 영어 /v/를 현용 ㅂ과 구별되는 자형으로 적은 사례는, 옛자모가 외국어 발음교육을 위한 실용적 문자 자원으로 다시 사용되었음을 보여 준다. 다만 개별 자모의 정확한 음가 대응과 표기 체계 전체는 원문 판본과 후속 연구를 통해 별도로 검증할 필요가 있다.

대표적인 예를 간단히 정리하면 다음과 같다.

사례	『아학편』의 표기 특징	역사적 의미
영어 /v/	ㅞ 계열의 합용병서·옛자모 활용	현용 ㅂ과 구별되는 /v/를 한글로 나타내려 한 시도
영어 발음 전반	로마자 철자와 한글 발음을 함께 제시	한글을 외국어 발음교육의 매개 문자로 활용한 선례

이러한 표기가 오늘날의 IPA 분석이나 훈민바름 정본과 동일한 것은 아니다. 당시의 표기에는 한국어 음절구조와 영어 발음교육의 관습이 함께 반영되었다. 오늘날처럼 출처 IPA·Unicode 부호·언어별 등록부·역변환 이력을 연결한 디지털 체계도 아니었다. 그럼에도 옛자모를 외국어의 실제 음가를 기록하는 실용적 자원으로 활용했다는 의의가 있다.

『아학편』의 의의는 개별 표기가 오늘날의 음성학적 기준에서 얼마나 정확했는가에만 있지 않다. 더 중요한 점은 새로운 언어와 지식이 밀려들던 시대에 외국어의 철자만 제시하는 데 그치지 않고, 그 소리를 한글로 읽고 배울 수 있게 하였으며, 필요한 경우 훈민정음의 확장 자원을 적극적으로 사용했다는 데 있다. 외국어의 소리를 한글 문자생활 안으로 받아들이려는 요구와 확장 한글의 실용 가능성이 근대 교육 현장에서 이미 함께 나타났던 것이다.

『국문정리』와 『아학편』을 훈민바름의 직접적인 규격상 선조라고 단정할 수는 없다. 그러나 두 문헌은 중요한 역사적 연속성을 보여 준다. 세계와의 접촉이 넓어질 때마다 현용 한글만으로는 충분히 구별하기 어려운 외국어 소리를 어떻게 기록할 것인가라는 문제가 되풀이되었고, 그때마다 한글의 자모와 조자 가능성을 활용한 해결책이 모색되었다.

구한말에서 이어받는 것

구한말의 외국어 표기 실험이 합용병서와 옛자모를 활용하여 외국어의 소리를 한글로 배우고 읽게 하려는 실용적 시도였다면, 훈민바름은 그 가능성을 IPA·언어별 음운모듈·확장 한글·공개 데이터·역변환 규칙으로 체계화하려는 현대적 계승이다.

4. 구한말부터 1986년까지: 확장과 제한의 교차

연도	규범·자료	확장 자모의 위치	역사적 의미
1897	이봉운 『국문정리』	일본어 탁음 구별 표기 제안	외국어 음운 대립을 한글 안에서 구별하려 한 근대 초기 실험
1908	지석영 주석·증보 『아학편』	ㅕ 등 합용병서·옛자모를 외국어 발음 표기에 활용	확장 한글을 실제 다언어 교육에 적용한 선례
1933	「한글 맞춤법 통일안」	현용 자모 중심	표기 통일 우선
1937	이종극 『모던조선외래어사전』	ㅍ, ㅑ 등 사용	민간 사전의 실사용 선례
1941	「외래어 표기법 통일안」	현용 자모 원칙	조선어학회안
1948/1952	문교부 「들온말 적는 법」	ㅏ, ㅑ, ㅓ, ㅕ 계열 확장	IPA 매개 국가안
1958	「로마자자의 한글화 표기법」	확장 자모 폐기	현용 자모로 회귀
1986	문교부 고시 제85-11호	현용 24자모 명문화	현행 외래어 표기 체계
2026	외국어 표기법 시안·영어확장판 4본	두 층·영어 RP·GA·공개 데이터·표준화	정책·기술·세계화 통합 제안

이 연표가 보여 주는 것은 단순한 진보나 퇴보가 아니다. 외국어 음가를 더 세밀하게 적고자 하는 필요와, 일반 사용자가 쉽게 읽고 입력할 수 있어야 한다는 필요가 서로 다른 시기마다 다른 균형점을 찾은 역사다. 훈민바름은 현용 자모 중심 규범을 실패로 규정하지 않는다. 일반 외래어 표기의 통일성과 편의는 여전히 공공적 가치다. 다만 그 규범의 목적을 정밀 전사까지 확장해 해석할 때 생기는 공백을 별도 층으로 채우려 한다.

1937년 『모던조선외래어사전』에서 /f/와 /v/를 구별하기 위해 ㅍ, ㅑ를 사용한 사례는 특히 중요하다. 현용 자모 제한 아래에서도 사전 편찬 현장에는 구별해야 할 소리가 있었고, 실무자가 한글의 기존 조자 자원을 이용해 해결을 시도했다는 뜻이기 때문이다(박창원·김수현 2004: 61).

5. 1948년 국가안이 남긴 방법론과 과제

1948년 문교부가 제정하고 1952년에 간행한 「들온말 적는 법」의 가장 중요한 유산은 개별 자모 목록보다 방법론이다. 이 안은 한국인이 외국어를 듣고 주관적으로 옮기는 방식보다, 외국어의 음가를 국제음성기호로 기술한 뒤 이를 한글로 변환하는 간접 방식을 택했다. 오늘의 훈민바름이 IPA를 공통 매개층으로 삼는 구조와 같은 방향이다(박창원·김수현 2004: 63).

그러나 1948년 안은 정착하지 못했다. 선행 연구가 지적한 원인은 체계의 전문성과 복잡성, 판독과 입력의 부담이었다. 당시의 활자·인쇄·타자 환경도 확장 자모의 보급에 불리했을 가능성이 크다(박창원·김수현 2004: 64). 이 실패는 정밀 표기의 필요 자체가 틀렸다는 증거라기보다, 일반 외래어 표기와 전문적 정밀 표기를 하나의 규범에 동시에 맡기고 당시 매체가 감당하기 어려운 체계를 대중 표기로 요구한 설계의 한계를 보여 준다.

1948년 안에서 배우는 두 가지

계승할 것은 IPA 매개와 확장 자모의 방법론이다. 반복하지 말아야 할 것은 일반 표기와 정밀 표기의 목적을 구분하지 않고 모든 사용자에게 같은 복잡성을 요구하는 방식이다.

6. 왜 지금 다시 가능한가

오늘날 옛한글 자모는 Unicode/ISO 10646에 등록되어 있고, 확장 자모를 포함한 글꼴과 입력 기술이 존재한다. 이것은 1948년과 비교해 근본적인 조건 변화다. 한편 부호가 등록되어 있다는 사실만으로 모든 브라우저·문서도구·모바일에서 입력·검색·정렬·셰이핑 결과가 자동으로 같아지는 것은 아니다. 훈민바름의 주장은 “이미 완벽히 된다”가 아니라 “이제 금지와 추측이 아니라 공개 시험으로 판단할 수 있다”는 것이다.

제2본과 제3본의 분리는 이 역사적 교훈을 반영한다. 정본은 사람이 손으로 적을 수 있어야 하며 특정 장치나 코드에 종속되지 않는다. 제3본 전산본은 디지털 환경에서 Unicode와 구조 데이터, 글꼴과 셰이핑을 이용해 정본을 저장·교환·처리하고 표시한다. 기술이 바뀌어도 정본의 원리가 무너지지 않고, 정본의 원리가 기기의 한계를 무시하지도 않는 이중 구조다.

제2부

문화적 계승성

옛글자를 장식으로 되살리는 것이 아니라 훈민정음의 생성 원리를 현재의 필요에 적용한다

7. 훈민정음의 생성 원리를 오늘에 잇는다는 것

훈민정음의 서문이 우리말과 중국 문자의 불일치 때문에 백성이 뜻을 펴기 어려웠던 현실을 출발점으로 삼았다면, 훈민바름은 세계 각국의 말소리와 현용 한글만으로 제한된 외국어 표기의 불일치를 현대의 출발점으로 삼는다. 전자는 우리말을 중국 문자에 종속시키지 않고 자기 글로 쓰게 했고, 후자는 세계 언어의 원문과 IPA를 보존하면서 그 소리를 한글로 정확히 읽고 씀으로써 한글 세계화의 길을 열고자 한다.

훈민바름의 문화적 정당성은 “옛글자도 한글이니 무엇에든 다시 쓰자”는 소유권 주장에 있지 않다. 핵심은 훈민정음이 기본자만 제시한 닫힌 목록이 아니라, 가획·합용·합성·병서·연서의 방법으로 새로운 대립과 결합을 표현하도록 설계된 생성 체계였다는 데 있다(홍운표 2012: 22-24). 살아 있는 문자체계란 자형을 무한히 늘린다는 뜻이 아니라, 새로 분석된 소리를 기존의 원리와 공개된 대응 규칙으로 받아들일 수 있다는 뜻이다.

병서형 ㅁ.ㅂ과 연서형 ㅁ.ㅂ은 이 원리를 서로 다른 방식으로 보여 준다. 훈민바름은 기존 한글 자모와의 시각적 일관성을 위해 병서를 기본형으로 삼고 연서를 역사적·대안형으로 보존하지만, 어느 쪽도 체계 밖에서 새로 고안한 사유(私有) 문자가 아니다. 이미 한글의 역사와 Unicode 문자집합 안에 존재하는 조자 자원을 오늘의 음가 구별에 다시 배치하는 것이다.

훈민정음의 기본글자 이외의 글자는 가획과 합용과 합성, 그리고 연서에 의해 만들어졌다.

홍운표(2012: 22)의 조자 원리 설명 요지

8. 옛한글의 복원이 아니라 살아 있는 조자(造字)법의 계승

훈민바름은 중세국어의 음가를 그대로 복원하는 역사 표기법이 아니다. 같은 자형을 쓰더라도 외국어 표기를 위한 현대의 자모값은 중세국어에서의 값과 다를 수 있다. 그러므로 훈민바름은 과거의 소리를 부활시키는 일이 아니라, 훈민정음의 살아 있는 조자법을 IPA와 연결해 현재와 미래의 세계 언어를 기록하는 일이다.

이 구분은 문화적 존중을 위해서도 중요하다. 옛한글을 권위의 장식이나 민족주의적 표상으로 소비하면 역사적 문헌과 현대 규약의 차이가 흐려진다. 반대로 옛 자모의 형태와 조합 원리를 정확히 기록하고, 현대 자모값의 선택 근거와 역변환 규칙을 공개하면 문화유산은 고정된 유물이 아니라 검증 가능한 창작 원리가 된다.

9. 손으로 완성되는 정보와 문자 주권

훈민바름 정보는 컴퓨터가 있어야만 성립하는 부호열이 아니다. 사람이 IPA와 대응표를 보고 종이에 직접 적었을 때에도 음소와 위치, 순서와 블록 관계가 드러나야 한다. 이 원칙은 기술적 편의보다 문화적 의미가 더 크다. 문자가 특정 회사의 글꼴, 특정 운영체제의 셰이핑, 특정 앱의 내부 코드가 있어야만 존재한다면 그것은 공공 문자의 기반이 되기 어렵다.

손글씨 정보는 디지털 기술을 거부하는 원시주의가 아니다. 오히려 디지털 규격이 무엇을 보존해야 하는지 기준을 제공한다. Unicode 평문이 한 구조를 완벽하게 담지 못하면 구조형 데이터가 보완하고, 글꼴이 결합에 실패하면 fallback이 나타나야 한다. 그러나 어느 경우에도 기기의 한계 때문에 손으로 적을 수 있는 정보 원리를 바꾸어서는 안 된다.

문자 주권의 의미

문자 주권은 모든 것을 국산 기술로 고립시키는 일이 아니다. 사람이 이해할 수 있는 정보, 공개된 변환 규칙, 국제 표준 부호, 여러 구현이 재현할 수 있는 시험을 함께 갖추어 어느 한 장치나 기업도 표기 자체를 독점하지 못하게 하는 일이다.

10. 한글의 세계화·국제화와 영어화의 차이

현재 훈민바름 문서가 영어 RP·GA를 첫 적용 언어로 삼는다는 사실은 한글을 영어에 종속시키려는 뜻이 아니다. 영어는 국제적으로 널리 쓰이고 한국어 학습자에게 수요가 크며 발음사전과 연구자료가 풍부해 첫 검증 모듈을 만들기 좋은 언어다. 영어 모듈의 현재 구현은 목적지가 아니라, 살아 있는 문자체계가 한 언어의 음소 목록과 변이를 어떻게 받아들이고 시험하고 개정하는지를 보여 주는 첫 사례다.

한글의 세계화는 세계의 이름과 소리를 모두 영어식으로 바꾸는 일이 아니다. IPA를 국제 음성 공통층으로 삼고, 각 언어의 원문·source IPA·음운모듈을 보존한 채 확장 한글 정보와 연결하는 일이다. 한국어 관용명, 국제통용명, 원어명, 영어권 발음과 원어 발음을 서로 다른 필드로 보존하고, 원문 문자·로마자 전사·언어코드·Unicode와 함께 묶는다. 이때 한글은 세계의 차이를 지우는 문자가 아니라, 그 차이를 한글의 블록 구조 안에 높은 통합성과 가역성으로 기록하면서 서로 연결하는 국제 문자 인터페이스가 된다.

대한민국의 국제화는 외국어를 많이 섞어 쓰는 데서 끝나지 않는다. 세계에서 태어난 소리·이름·개념을 원문과 IPA에 기대어 한국어 안에 빠르고 정확하게 받아들이고, 한국어로 생산한 지식도 국제적으로 호환되는 데이터로 내보낼 수 있어야 한다. 현용 자모 제한을 정밀 표기층에서 선택적으로 풀고 확장 한글을 공공 규격으로 검증하면, 한글은 국내 문화유산의 상징을 넘어 세계 언어를 연결하는 교육·지식·언어기술의 기반으로 국제무대에 설 수 있다.

이 확장은 한국어 사용자의 외국어 수용에서 끝나지 않는다. 언어별 모듈과 원문·IPA를 함께 보존하는 구조가 충분히 검증된다면, 다른 언어권의 사용자 역시 자신의 언어나 다른 언어의 소리를 기록하고 학습하는 데 확장 한글을 선택적으로 활용할 수 있다. 세계의 소리를 한글로 받아들이는 능력은 곧 한글을 세계에 다시 열어 놓는 기반이 된다.

세계의 소리를 받아들이고, 다시 세계로

언어별 원문과 IPA를 보존하는 구조가 충분히 검증된다면, 확장 한글은 한국어 사용자의 외국어 표기를 넘어 다른 언어권의 사용자도 자신의 언어나 다른 언어의 소리를 기록하고 학습하는 데 선택할 수 있는 문자 도구로 시험될 수 있다.

한글의 세계화는 한글을 세계에 알리는 데서 끝나는 것이 아니라, 세계가 한글을 사용할 이유와 가능성을 만드는 데까지 나아간다.

제3부

사회적 필요성

현행 규범을 비난하지 않고, 그 규범이 원래 말지 않은 기록 과제를 분명히 한다

11. 외국어 기반 신개념 수용 지연과 판교체

현행 외래어 표기법은 국어생활에 들어온 외래어를 현대 한국어의 자모와 음절구조 안에서 간편하고 통일되게 적는 데 강점이 있다. “픽업”, “패션”, “리버” 같은 표기는 한국어 단어로 읽고 쓰기에 안정적이다. 이 기능을 원어 음가의 정밀 보존 여부만으로 평가하면 규범의 목적을 잘못 잡게 된다.

문제는 그 표기를 외국어 발음 안내, 사전의 원음 병기, 학술 전사, 접근성 자료, 음성 AI의 학습 레이블까지 담당하는 유일한 한글 표기로 사용할 때 생긴다. /l/과 /ɪ/, /f/와 /p/, /b/와 /v/, /θ/와 /s/ 같은 대립이 합류하고, 사용자는 표기만 보고 구별해야 할 소리가 있다는 사실조차 알기 어렵다. 일반 표기가 실패한 것이 아니라 별도의 정밀 기록층이 부재한 것이다.

또 하나의 문제는 속도다. 외국에서 새로 태어난 개념이 한국어 어휘로 정착하려면 번역어와 음역형이 경쟁하고, 관용이 형성되고, 사전과 규범이 이를 뒤따르는 시간이 필요하다. 안정된 개념에는 이 숙성 과정이 유익하지만, 제품·서비스·개발방법론이 짧은 주기로 바뀌는 IT 산업에서는 한국어 표현이 자리 잡기 전에 다음 개념이 도착한다. 그 사이 원어의 음가와 의미가 손실되거나, 아예 로마자 원어를 문장 속에 삽입하는 방식이 일상화된다.

이른바 ‘판교체’ 또는 ‘판교어’는 이러한 지연을 드러내는 상징적 사례다. 판교 IT·스타트업 현장에서는 영어 업계용어를 한국어 문법에 바로 끼워 넣는 한영혼용체가 빠르게 퍼졌고, 내부에서는 속도와 개념 동일성을 지키는 수단이 되지만 외부 사용자에게는 의미와 발음을 가리는 장벽이 되기도 한다(이서현·김소현 2025). 이를 단순한 외국어 남용으로만 꾸짖으면 산업현장이 왜 원어를 놓지 못하는지 설명하지 못한다.

정밀 외국어 표기층은 이 딜레마에 ‘초기 수용 통로’를 제공할 수 있다. 새 개념이 등장하면 원철자·source IPA·훈민바름·간단한 정의·출처를 함께 기록하여 최소한의 왜곡으로 한국어 문장과 데이터 안에 즉시 들여오고, 이후 사회적 사용을 거쳐 번역어 또는 일반 외래어가 정착하면 그 관계를 연결해 보존한다. 이렇게 하면 한국어는 세계의 변화를 늦게 번역해 따라가는 언어가 아니라, 새 개념을 빠르게 수용하고 한국에서 태어난 개념을 다시 세계에 내보내는 선도 언어가 될 수 있다.

목적과 시간의 분리

일반 표기는 정착된 외래어의 통일성과 한국어 적합성을 맡는다. 정밀 외국어 표기는 새 개념의 초기 수용, 음가 식별과 역변환을 맡는다. 빠른 임시 수용과 장기적 어휘 정착을 한 규범에 억지로 합치지 않는 것이 속도와 공공성을 함께 지키는 방법이다.

12. 교육·사전·고유명사·학술·접근성·언어기술의 공백

영역	현재의 공백	훈민바름이 제공할 수 있는 층
외국어 교육	표기에서 음소 대립이 합류하여 학습자가 차이의 존재를 보기 어렵다.	IPA와 함께 병기되는 시각적 발음 지도
사전 편찬	사전마다 원어 발음의 한글 병기 관행이 달라진다.	공개된 음소-자모 등록부와 출처 메타데이터
고유명사	한국어 관용명·국제통용명·원어명이 서로 연결되지 않는다.	명칭층과 발음층을 분리한 상호 검색 레코드
학술 전사	방언·언어기록 연구자가 사적 표기체계를 만든다.	언어모듈별 공개 전사 프로파일
접근성	점자·화면낭독·음성자료에서 원어 발음 정보가 소실된다.	IPA·대체텍스트·선형 fallback을 함께 보존
언어기술	TTS·G2P·ASR가 외국어 발음을 내부 규칙으로 각각 처리한다.	공개 발음 스키마와 시험 가능한 레이블
신규 개념·산업 용어	번역과 관용이 정착되기 전 로마자 원어·판교체에 의존하여 속도와 공공성 중 하나를 잃는다.	원철자·IPA·훈민바름·정의·출처를 연결한 초기 수용 레코드

공백의 공통 구조

여섯 영역과 신규 개념 수용의 공백은 서로 달라 보이지만 공통 원인이 있다. 원철자와 실제 발음, 한국어 표기와 의미, 출처와 변이를 한 레코드 안에서 연결하는 공통층이 없기 때문에 각 기관과 제품이 같은 정보를 반복해서 만들고 서로 다른 방식으로 잃는다.

외국어 교육

한글을 사용하는 외국어 학습자는 현행 표기만으로 /l/과 /ɹ/, /f/와 /p/, /v/와 /b/, /θ/와 /s/ 같은 대립의 존재를 보기 어렵다. 영어 정본의 해례식 조음 안내, RP·GA 분리, 음성자료와 최소대립 연습을 함께 제공하면 한글 기반의 친숙함을 유지하면서도 IPA로 이동하는 발판을 만들 수 있다. 효과는 발음 인상평이 아니라 구별 정확도·학습시간·IPA 전이로 시험해야 한다.

사전 편찬

사전은 표제어 하나에 원철자, 언어·지역, source IPA, 훈민바름, 일반 외래어 표기, 음성자료와 출처를 함께 담을 수 있다. 이렇게 하면 사전마다 임의로 달라지던 한글 발음 병기를 줄이고, 발음이 바뀌거나 새 변이가 추가될 때 전체 표제어를 다시 쓰지 않고 해당 모듈과 출처만 갱신할 수 있다.

고유명사

사람·도시·국가·작품·역사인물의 이름은 한국어 관용명, 국제통용명, 원어명과 발음이 서로 다른 층을 가진다. 어느 하나를 ‘진짜 이름’으로 강제하기보다 하나의 개체 ID 아래 각각의 이름·언어·발음을 연결하면 지도, 항공, 번역, 학술, 검색과 교육에서 동일한 대상을 더 정확히 찾을 수 있다.

학술 전사

방언·소수언어·현장 언어기록에서는 연구자마다 사적 한글 전사를 만들기 쉽다. IPA 원문을 보존하고 언어모듈·자모값·역변환 규칙을 공개하면 한글 기반 전사를 쓰더라도 국제 연구자가 원자료를 다시 확인할 수 있다. 훈민바름은 IPA를 대체하는 폐쇄적 부호가 아니라 IPA와 지역 연구 사이의 가독성 층이 된다.

접근성

정밀한 자형이 보이더라도 화면낭독·점자·저시력 환경에서 정보가 사라지면 공공 표기가 될 수 없다. source IPA, 자모명, 선형 fallback, 대체텍스트와 구조 데이터를 함께 보존하고 강제·장단·성조를 크거나 색 하나에만 맡기지 않아야 한다. 접근성 사용자는 완성 뒤의 검수자가 아니라 규격 설계 단계의 공동 참여자여야 한다.

언어기술과 신규 개념 수용

음성합성(Text-to-Speech, TTS), 자소-음소 변환(Grapheme-to-Phoneme, G2P), 자동음성인식(Automatic Speech Recognition, ASR), 대규모 언어모델(Large Language Model, LLM) 등의 언어기술은 외국어·고유명사·코드혼용을 시스템별 규칙과 예외사전에 의존하여 처리해 왔다. 공개된 원철자-IPA-훈민바름-의미 레코드는 시스템 간 오류를 비교할 공통 기준이 된다. 특히 새 IT 개념은 정식 번역어가 생기기 전에도 이 레코드로 한국어 검색·교육·음성처리에 들어올 수 있어, 판교체의 속도는 살리고 배타성과 정보 손실은 줄일 수 있다.

훈민바름 하나가 각 영역의 전문성을 대신한다는 뜻은 아니다. 사전에는 편찬 원칙과 출처 검증이, 교육에는 교수법과 학습 실험이, 접근성에는 장애 당사자와의 공동 설계가 필요하다. 훈민바름이 제공하는 것은 각 영역이 반복해서 새로 만들어야 했던 “음가와 개념을 한글로 식별해 기록하고 국제 표준 정보와 연결하는 공통층”이다.

13. 이름과 지식이 서로를 찾게 하는 데이터 층

고유명사와 신규 개념의 문제는 발음만의 문제가 아니다. 한국어 관용명, 국제적으로 통용되는 이름, 원어명과 분야별 개념어는 서로 다른 역사와 기능을 가진다. “아테네”를 버리고 Athens로 바꾸거나, 새 IT 용어를 무조건 번역 또는 로마자로만 고정하는 방식은 해결책이 아니다. 필요한 것은 어느 이름과 표현으로 검색하더라도 같은 개체·개념에 도달하고, 각 이름의 언어·발음·정의·사용시기를 별도로 확인할 수 있는 데이터 구조다.

훈민바름은 이 구조에서 발음층과 초기 수용층을 맡는다. 한 레코드가 한국어 관용명, 국제통용명, 원어명, 영어권 IPA와 원어 IPA, 해당 언어모듈의 훈민바름, 분야·정의·출처·최초 사용시기를 함께 가질 수 있다. 이는 지도·항공·학술문헌·번역·검색 및 검색 증강 생성(Retrieval-Augmented Generation, RAG) 시스템에서 같은 대상을 서로 다른 이름이 가리키는 문제를 줄이고, 새 개념이 번역어 정착을 기다리지 않고 한국어 지식망에 들어오게 한다. 국제 정합성은 한 이름으로 통일하는 것이 아니라 여러 이름과 개념 사이의 관계를 표준화하는 데서 생긴다.

제4부

공공정책적 의미

새 규제를 만드는 것이 아니라 목적이 다른 표기 영역에 선택권과 검증의 절차를 부여한다

14. 외래어 표기와 정밀 외국어 표기의 두 층

층	목적	문자 자원	주요 적용 영역
일반 외래어 표기	한국어 문자생활의 편의와 통일	현용 24자모와 현행 받침 원칙	일반 문서·언론·행정·생활어
정밀 외국어 표기	원어 음가의 식별·학습·기록·데이터화	옛한글·확장 자모와 언어모듈	교육·사전·학술·고유명사·접근성·언어기술·신규 개념의 초기 수용

두 층은 대체 관계가 아니다. 정밀 표기가 일반 표기를 밀어내지 않고, 일반 문서에서 옛한글을 의무적으로 쓰게 하지 않는다. 같은 대상에 “일반 표기 + 정밀 표기”를 필요에 따라 병기할 수 있다. 이 원칙은 1948년 안과 달리 전문적 정밀 표기의 복잡성을 모든 국민에게 전가하지 않는다.

15. 규제 완화가 아니라 적용 범위의 합리화

공공정책에서 가장 중요한 질문은 “현행 규칙을 폐지할 것인가”가 아니라 “그 규칙이 어떤 목적까지 지배해야 하는가”다. 현행 자모 제한은 일반 외래어 표기의 통일을 위해 타당하다. 그러나 외국어 교육·사전·학술·접근성·언어기술에도 같은 제한을 사실상 적용하면, 원음 정보를 기록할 수단이 규범 안에 존재하지 않게 된다.

따라서 제안의 실질은 자모 제한을 전면 철폐하는 것이 아니다. 일반 외래어 표기에는 현용 24자모 원칙을 유지하되, 원어 음가를 식별해야 하는 별도 정밀 외국어 표기편 또는 부속 프로파일에는 옛한글·확장 한글 자모를 허용하는 것이다. 곧 ‘현용 한글을 없애는 개정’이 아니라 ‘현용 자모만을 쓸 수 있다는 제한의 적용 범위를 합리화하는 개정’이다. 이 제도적 공간이 열려야 IPA로 기술된 세계 언어를 언어별 훈민바름 모듈과 연결할 수 있다.

16. 공개·비독점·기술중립의 원칙

훈민바름은 「외국어 표기법」 연구·개발 과정의 내부 코드명이자, 현재의 가표본 연구와 참조 실험을 가리킨다. 공공 규격의 공식 명칭은 별도의 합의로 정할 수 있으며, 특정 브랜드·글꼴·앱을 독점적으로 채택해서는 안 된다. 음소-자모 매핑, 언어모듈, 데이터 스키마, 테스트셋, 변환엔진과 API는 공개 검증이 가능해야 하고, 다른 연구자와 기업이 같은 규칙을 독립적으로 구현할 수 있어야 한다.

기술중립은 아무 기술도 정하지 않는다는 뜻이 아니다. 저장·교환의 기준, 역변환 규칙, 접근성 fallback, 정규화와 그래픽 경계, 시험 항목은 명확히 정하되, 특정 운영체제나 특정 글꼴의 결과만을 규범으로 만들지 않는다는 뜻이다. 제2본의 손글씨 정본과 제3본의 전산 구현·표시층 분리는 이 원칙을 문서 구조로 구현한다.

비독점 선언

목적은 내부 코드명 “훈민바름”을 국가 브랜드로 고정하는 데 있지 않다. 한글로 외국어 음가와 새 개념을 정밀하게 기록할 수 있는 공개 규격과 공개 데이터가 존재할 제도적·기술적 자리를 만드는 데 있다.

17. 반대 논거를 설계 제약으로 받아들이기

심재기(2008)는 옛자모를 이용한 정밀 표기가 현대 한국어의 외래어 표기가 아니라 별도의 외국어 표기 또는 국제음성부호화 작업이라고 지적했다. 훈민바름은 이 반론을 피하지 않는다. 오히려 그 구분을 체계의 이름과 적용 범위로 받아들인다.

반대론을 수용한다는 것은 가능성을 포기한다는 뜻이 아니다. 일반 외래어 표기에는 사용하지 않고, 외국어 학습과 정밀 기록에 선택적으로 사용하며, 국어 문자생활에 강제하지 않는다는 세 제약을 명문화하는 것이다. 이러한 설계는 정책적 신뢰를 높인다. 훈민바름의 정당성은 “한 글은 모든 소리를 완벽히 적을 수 있다”는 과장보다, 어디에 쓰고 어디에는 쓰지 않을지를 분명히 하는 절제에서 나온다.

제5부

교육과 접근성의 가치

학습자가 음가의 차이를 보고, 어떤 환경에서도 정보가 사라지지 않게 한다

18. 학습자가 소리의 차이를 먼저 볼 수 있게 하는 표기

외국어 발음 교육에서 가장 먼저 필요한 것은 완벽한 발음이 아니라 “구별해야 할 소리가 둘 이상 존재한다”는 인식이다. liver와 river, fan과 pan, vest와 best가 같은 한글 표기로 수렴하면, 학습자는 교사의 설명이나 로마자 철자에 의존해 차이를 다시 복원해야 한다. 훈민바름은 그 차이를 한글 기반 자형으로 눈앞에 남겨 학습의 출발점을 바꾼다.

이것은 한국어식으로 외국어를 읽게 하려는 표기가 아니다. 자모값은 원음과 자연적으로 동일한 소리가 아니라 IPA를 매개로 한 변환 규약이다. 학습자는 훈민바름을 보고 음소의 구별과 위치를 확인한 뒤, 반드시 음성자료와 조음 설명을 함께 사용해야 한다. 시각 표기는 소리를 대체하지 않고 소리를 찾아가는 지도다.

19. IPA를 대체하지 않고 IPA로 가는 다리 놓기

IPA는 인간 언어의 음성을 범언어적으로 기술하기 위해 발전한 국제 표기체계이며, 훈민바름의 근거층이자 세계 확장의 관문이다. 훈민바름이 IPA를 없애거나 한국어 사용자만의 폐쇄적 기호로 대체하면 국제적 검증 가능성과 다언어 확장성을 함께 잃는다. 모든 정본 레코드는 원문 문자와 source IPA를 보존하고, 훈민바름에서 IPA로 되돌아가는 규칙을 가져야 한다. 이 가역성이 있어야 새로운 언어의 소리를 기존 모듈과 혼동하지 않고 추가하며, 국제 연구와 사전·교육·음성기술 자료에 다시 연결할 수 있다.

교육적 역할은 IPA를 배우기 전후의 간극을 줄이는 데 있다. 초보자는 익숙한 한글 블록과 새 자모의 대비를 통해 음소 구조를 볼 수 있고, 중급 이상 학습자는 IPA와 훈민바름을 나란히 보며 국제 표기로 이동할 수 있다. 장기적으로 훈민바름이 IPA의 대체물이 아니라 세계 언어의 소리를 IPA와 함께 이해하게 하는 시각적 동반자가 되는지를 실험으로 확인해야 한다.

20. 접근성은 부가 기능이 아니라 규격의 조건

정밀 표기를 도입하면서 특정 글꼴이 없는 사용자, 화면낭독기를 쓰는 사용자, 저시력 사용자, 점자 출력을 이용하는 사용자를 배제한다면 적용 범위에 접근성을 넣은 취지와 모순된다. 따라서 정상 블록 표시가 실패해도 source IPA, 자모명, 선형 fallback, 대체텍스트로 같은 정보를 전달할 수 있어야 한다.

강제·장단·성조도 색상이나 글자 크기 하나에만 의존해서는 안 된다. 구조 데이터와 수치·기호·텍스트 설명을 함께 제공하고, 사용자가 시각 표시 방식을 선택할 수 있어야 한다. 접근성은 완성된 제품에 나중에 붙이는 기능이 아니라 어떤 정보를 정본에 보존해야 하는지를 결정하는 설계 원칙이다.

21. 무엇을 시험해야 하는가

시험 축	핵심 질문
음성학적 적합성	IPA와 훈민바름 사이의 정보 손실, 역변환 정확도
교육 효과	음소 구별 정확도, 발음 재현, 학습시간, IPA 전이 효과
인지 부담	새 자모 판독, 표기 길이, 초보·고령 학습자의 부담
접근성	화면낭독·점자·저시력·키보드 입력 환경의 이용 가능성
상호운용성	OS·브라우저·문서·모바일의 입력·검색·복사·정렬·렌더링
기술중립성	서로 다른 구현이 같은 테스트셋에서 같은 정본을 생성하는지

좋은 결과만 공개하는 시험은 공공 표준의 근거가 되기 어렵다. 실패한 글꼴, 혼동되는 자모, 학습자의 낮은 판독률도 버전과 함께 공개해야 한다. 사용자 피드백은 중요하지만 즉시 매핑을 바꾸는 투표가 아니라, 변경 제안·근거자료·호환성 영향·공개검토·승인의 절차로 들어가야 한다.

현재의 미결 항목은 체계의 실패가 아니라 공개 연구의 의제다. 예컨대 일부 반모음 결합표기처럼 복수의 합리적 대안이 존재하는 경우에는 성급히 하나의 규칙으로 고정하지 않고, 잠정 고정·시범·미결의 상태를 명시한 채 음성학·문자학·교육·정보기술 분야의 검토와 사용자 시험을 거칠 수 있다. 미결사항을 공개하는 것 역시 이 가표본 연구의 검증 절차다.

제6부

AI와 표준화의 미래

발음 표기를 앱 기능이 아니라 공개 언어자원과 상호운용 규격으로 확장한다

22. 공개 발음 데이터라는 공공 인프라

훈민바름의 디지털 미래는 화면에 옛한글을 잘 그리는 데서 끝나지 않는다. 더 중요한 산출물은 한 낱말의 원철자, 언어와 지역, source IPA, normalized IPA, Broad/Narrow, 자모값 ID, 음절과 자음군, 강세와 운율, 음성자료, 출처와 라이선스, 버전을 함께 담은 공개 레코드다. 이 레코드가 언어별로 축적될수록 훈민바름은 고정된 영어 표기표가 아니라 세계의 소리를 계속 받아들이는 살아 있는 언어자원이 된다.

데이터 층	최소 항목
식별	stable_id, 표제어, 별칭, 개념 ID, 분야·정의, 언어코드, 지역변이
발음	source IPA, normalized IPA, 음소토큰, Broad/Narrow, 강세·운율
훈민바름	자모값 ID, 손글씨 정본 구조, Unicode 직렬화, 표시 별칭(alias)
증거	음성자료, 사전·화자·문헌 출처, 갱신일
거버넌스	라이선스, 규격버전, 검토상태, 변경이력

이 구조가 공공 인프라인 이유는 여러 기관이 반복 구축하던 발음사전과 명칭사전을 연결할 수 있기 때문이다. 학교·사전·방송·접근성 서비스·AI 기업은 같은 IPA와 자모값 ID, 출처와 언어 모듈을 참조하되 각자의 표시 프로파일과 제품 경험을 만들 수 있다. 한 언어의 새 음소나 변이가 검증되면 등록부와 모듈을 확장하고, 기존 구현은 같은 변경 이력과 역변환 규칙을 따라갈 수 있다.

공개 레코드에는 음소토큰과 자모 요소의 대응 관계도 함께 기록한다. 한 음소를 여러 자모로 적거나 여러 음소를 하나의 복합값으로 적는 경우를 구별하여 제3본의 구조화와 역변환에 연결한다.

23. TTS·G2P·ASR·LLM에 생기는 변화

한국어 TTS와 G2P를 비롯한 언어기술은 외래어·외국어 발음의 공통 기준이 부족하여 시스템별 예외 규칙과 사전에 의존해 왔다(임현열 2020; 문성민·김수한·고연숙 2022). 훈민바름 레코드는 원철자에서 IPA, 음소 토큰, 한글 기반 표기로 이어지는 중간층을 공개하여 오류를 비교하고 교정할 기준을 제공할 수 있다.

- TTS: 고유명사와 외국어 구절의 발음변이·강세·지역정보를 명시해 잘못된 한국어식 독음을 줄인다.
- G2P: 로마자 또는 혼용문자에서 한국어 발음으로 가는 과정에 source IPA와 언어모듈을 분리한다.
- ASR: 인식 후보의 원철자·IPA·훈민바름을 연결해 외국인 이름과 전문용어의 사전 확장을 돕는다.
- LLM·RAG: 한국어 관용명·국제통용명·원어명·발음 필드를 하나의 개체 ID로 연결해 검색과 답변의 일관성을 높인다.
- 교육 AI: 사용자가 선택한 Narrow 기능, 조음 설명, 음성자료와 시각 표기를 결합해 맞춤형 피드백을 만든다.

다만 확장 자모를 모든 LLM의 본문 토큰으로 강제할 필요는 없다. 우선은 표준 메타데이터와 별칭층, 검색 인덱스, 도구 호출 결과로 제공하고, 필요한 모델과 서비스가 토큰나이저 확장과 학습을 선택할 수 있다. 기술의 채택 방식까지 하나로 강제하지 않는 것이 기술중립 원칙에 맞다.

24. IPA와 국제 표준 호환 체계를 통한 세계 언어 확장

현재 훈민바름의 기술 규칙은 영어 RP·GA 확장판이다. 그러나 세계 확장의 기반은 영어의 국제적 지위가 아니라 IPA의 범언어적 기술 능력이다. 각 언어는 고유한 음소 목록, 이음 규칙, 음절구조, 강세·성조와 지역변이를 가지므로, 다음 언어판은 원문·source IPA·음운분석·화자 공동검토를 갖춘 독립 모듈로 추가되어야 한다. 확장성은 영어표를 복사하는 능력이 아니라 IPA라는 공통 좌표 위에서 새로운 언어의 소리 체계를 존중하는 한글 정본을 계속 세우는 능력이다.

다언어·다문자 확장에서 IPA는 발음의 핵심 국제 공통층이지만 유일한 정보층은 아니다. 원문 문자, 언어·지역 식별, 국제 로마자·전사 체계, Unicode 부호, 의미·개념 ID와 출처를 함께 연결해야 한다. 훈민바름은 이 국제 체계들을 대체하지 않고 한글 사용자가 같은 정보에 접근하고 다시 국제적으로 역검증할 수 있게 하는 인터페이스가 되어야 한다. 아직 모든 IPA 음가의 훈민바름 값이 확정된 것은 아니며, 바로 그렇기 때문에 공개 등록부·근거·시험·버전 관리가 살아 있는 문자체계의 핵심 제도가 된다.

25. 응용 프로파일에서 국가표준까지

단계	산출물과 판단	주체
1. 공개 응용 프로파일	처리 규칙, 등록부, 참조구현, 테스트셋을 공개	제안자·공개 참여자
2. 기술보고서	시험 결과와 구현 사례를 TR 형식으로 정리	연구 참여자·표준화 기구
3. 단체표준	합의된 부분을 실제 제품과 기관에 적용	TTA 등
4. 국가표준 검토	운용 실적을 근거로 KS X 1026-1 보완 검토	과기정통부·국가표준 체계

지금 필요한 것은 마지막 단계의 즉시 개정이 아니라 첫 단계의 공개 실험이다. 정본과 전산·표시 규격, 언어모듈과 테스트셋을 공개하고 실제 사용 사례를 쌓은 뒤, 합의된 최소 부분만 상위 표준으로 올리는 것이 순서다. 각 단계는 되돌릴 수 있어야 하고 이전 버전의 데이터 변환 규칙을 함께 제공해야 한다.

제7부

한계와 과제

가능성을 크게 말되, “세계의 모든 소리를 이미 적는다”는 완성 선언과
“세계의 소리로 계속 확장할 수 있다”는 구조적 가능성을 구분한다

26. 과장하지 않기: 부호화와 상호운용성은 다르다

옛한글 자모가 Unicode에 등록되어 있다는 사실은 출발 조건이지 완성의 증거가 아니다. 글꼴에 글리프가 있어도 ljmo·vjmo·tjmo 셰이핑이 다를 수 있고, 완성형과 조합형의 정규화·검색·커서 이동이 환경마다 어긋날 수 있다(안대혁·박영배 2007; 양왕성 2018). “표준 부호가 있다”와 “모든 사용자가 같은 결과를 얻는다”를 구분해야 한다.

마찬가지로 수십만 회의 엔진 테스트가 통과해도 그것은 정해진 규칙을 코드가 일관되게 실행했다는 뜻이지, 그 자모 선택이 학습자에게 최선이라는 증거는 아니다. 기계적 적합성, 음성학적 타당성, 학습효과, 접근성, 사회적 수용성은 서로 다른 시험이다.

27. 복잡성·독점·영어중심성의 위험

위험	어떻게 나타나는가	안전장치
복잡성	새 자모와 여러 모드는 일반 사용자의 부담을 늘릴 수 있다.	선택적 사용, 단계별 교육, 단순 표시 프로파일
규범 충돌	일반 표기를 대체하는 것으로 오해받을 수 있다.	두 층의 적용 범위와 병기 원칙 명문화
기술 종속	특정 글꼴·앱이 사실상 표준을 독점할 수 있다.	손글씨 정보, 공개 등록부, 다중 참조구현
영어중심성	영어 시범값이 다른 언어에 강제될 수 있다.	언어모듈 독립, source IPA 보존
민족주의적 과장	한글이 모든 소리를 “완벽히 낸다”는 주장으로 흐를 수 있다.	변환 규약임을 명시하고 음성자료와 함께 사용
데이터 위험	출처·라이선스 없는 발음자료가 공공 데이터로 유통될 수 있다.	provenance·license·review 필드 의무화

28. 향후 10년을 위한 단계적 로드맵

기간	주요 과제	도달점
0-1년	영어 RP·GA 핵심 정보 규칙과 전산·표시 규격의 가표본 정리·공개, Australian English 등 영어 방언·발음 프로파일 추가, 미결 항목의 시험	재현 가능한 참조구현
1-2년	교사·학습자 파일럿, IPA 전이·인지부담·발음 구별 실험, 접근성 공동검토	교육·접근성 근거
2-4년	사전·고유명사·TTS·G2P 기관 협력, 공개 명칭·발음 레코드 확장	공공 데이터 생태계
4-6년	프랑스어·그리스어·중국어 등 독립 언어모듈, 원문 문자·국제 전사체계 연계, 기술보고서·단체표준 검토	다언어·다문자·표준화 확장
6-10년	운용 실적에 기반한 KS 부속 프로파일 또는 개정 검토, 국제 연구 협력	지속 가능한 표준 거버넌스

이 로드맵의 성공 기준은 옛한글을 많이 보급하는 것이 아니다. 일반 표기를 흔들지 않으면서도 필요한 영역에서 음가 정보가 덜 소실되고, 학습자가 더 잘 구별하며, 접근성 사용자가 배제되지 않고, 서로 다른 구현이 같은 정보를 재현하고, 공개 데이터가 실제 언어기술의 오류를 줄이는가가 기준이다.

맏음말 · 전환이 그지없는 글자의 공공적 미래

훈민바름의 역사적 정당성은 과거에 비슷한 자모가 있었다는 사실만으로 완성되지 않는다. 훈민정음이 우리말과 중국 문자 사이의 불일치로 인한 문자생활의 어려움에서 출발했듯, 훈민바름은 세계 언어의 소리와 현용 자모로 제한된 외국어 표기 사이의 불일치에서 출발한다. 1897년 『국문정리』와 1908년 『아학편』의 외국어 표기 실험, 1937년의 확장 자모 실사용, 1948년 국가안의 IPA 매개 방법론, 훈민정음의 조자 원리와 Unicode 시대의 조건을 계승하되, 복잡성·강제성·매체 한계를 오늘의 설계 제약으로 받아들일 때 비로소 현재의 정당성이 생긴다.

문화적 계승성은 옛글자를 박물관에서 꺼내는 데 있지 않다. 사람이 손으로 적을 수 있는 정본을 세우고, IPA의 국제 음성 좌표와 한글의 가획·합용·합성·병서·연서 원리를 연결하며, 디지털 장치가 그 정본과 source IPA를 함께 보존하도록 요구하는 데 있다. 훈민정음이 백성이 자기 말을 자기 글로 쓰게 했다면, 훈민바름은 한글이 세계의 말을 그 소리 정보와 함께 받아 적게 하는 현대적 문자 확장이다. 동시에 그렇게 넓어진 한글의 표현 능력을 다른 언어권 사용자에게도 선택 가능한 문자 자원으로 다시 열어 놓는 실험이다.

사회적 필요성과 정책적 의미는 더 절제된 곳에 있다. 현행 일반 표기를 폐기하지 않고, 그것이 원래 말지 않은 교육·사전·학술·접근성·언어기술과 신규 개념의 초기 수용 영역에 선택적 정밀 기록층을 마련한다. 특히 판교체가 보여 주듯 빠르게 변하는 산업은 원어의 속도를 포기할 수 없지만, 로마자 혼용만으로는 국민 다수에게 의미와 발음의 장벽을 만든다. 훈민바름은 새 개념을 원철자·IPA·한글 기반 표기·정의와 함께 즉시 수용하고, 훗날 정착한 번역어·외래어와 연결하는 중간층을 제공한다.

AI 시대의 한글 세계화와 한국어 국제화는 모든 언어를 한글로 바꾸는 사업이 아니다. 세계의 원문·이름·소리·개념을 source IPA와 함께 보존하고, 언어별 모듈을 거쳐 확장 한글 정본으로 연결하는 공공적 능력이다. 이 통로가 갖춰지면 한글은 세계 언어의 소리를 쓰고 읽고 검색하고 계산하는 국제 인터페이스가 되고, 한국어는 새 지식과 기술을 덜 손실되게 수용해 다시 세계로 내보내는 선도 언어가 될 수 있다.

최종 명제

훈민정음이 우리말과 빌려 쓰던 문자의 불일치를 풀어 백성이 자기 말을 자기 글로 쓰게 했다면, 훈민바름은 세계 언어의 소리와 현용 자모로 제한된 외국어 표기의 불일치를 줄이고, 세계 언어의 음가 차이를 필요에 따라 한글로 더 정밀하게 구별해 기록할 수 있는 가능성을 시험한다. 일반 외래어 표기는 유지하되 정밀 외국어 표기층에서 옛한글·확장 한글을 허용하고 IPA를 매개층으로 삼는 것이, 한글을 살아 있는 세계 문자체계로 확장하는 규제합리화의 핵심이다.

부록 A. 훈민바름 공공성 원칙 12조

1. 훈민바름은 새 소리·개념·언어환경에 따라 공개 검증을 통해 진화하는 살아 있는 문자체계이며, 그 정보는 손으로도 완성되어야 한다.
2. 현행 일반 외래어 표기에는 현용 24자모 원칙을 유지하고, 정밀 외국어 표기층에는 옛한글·확장 한글 자모의 선택적 사용을 허용한다.
3. 원문 문자와 source IPA, 언어·지역·출처를 함께 보존하고 IPA를 세계 언어와 한글을 잇는 국제 음성 공통층으로 삼는다.
4. 자모 대응은 자연적 동일성의 주장이 아니라 공개된 변환 규약으로 다룬다.
5. 새 자형보다 이미 등록된 한글 자모와 조자법을 우선하되, 실제 적합성은 시험한다.
6. 현재 기술 문서는 영어 RP·GA 확장판을 첫 참조범위로 삼으며, 이 범위를 넘어선 언어·방언의 완성 매핑을 주장하지 않는다.
7. 정상 표시가 실패해도 IPA·선형 표시·대체텍스트로 정보가 사라지지 않게 한다.
8. 교육 효과·인지 부담·접근성·상호운용성의 부정적 결과도 버전과 함께 공개한다.
9. 등록부·스키마·테스트셋·참조구현은 공개 검증과 독립 재구현이 가능해야 한다.
10. 내부 코드명·공식 명칭·대중 브랜드를 구분하고, 특정 명칭·글꼴·소프트웨어·기업의 독점적 채택을 요구하지 않는다.
11. 변경은 근거·호환성 영향·공개 검토·버전 승인 절차를 거치고 이전 데이터의 이행 규칙을 제공한다.
12. 다른 영어 방언은 독립 프로파일로, 다른 언어는 각자의 IPA 분석과 화자 검토를 갖춘 독립 언어모듈로 관리한다.

부록 B. 핵심 연표

연도	사건	의미
1443	세종실록의 “字雖簡要 轉換無窮”	한글의 전환·확장 가능성을 상징
1897	『국문정리』의 일본어 탁음 구별 표기 제안	외국어 음운 대립을 한글로 구별하려 한 초기 실험
1908	지석영 주석·증보 『아학편』의 다언어 병기와 ^ㅅ 등 옛 자모 활용	외국어 발음을 확장 한글로 기록하고 교육에 사용한 근대적 선례
1937	『모던조선외래어사전』의 ^ㅅ , ^ㅅ 실사용	민간 편찬 현장의 음가 구별 필요
1948/1952	「들온말 적는 법」	IPA 매개 + 확장 자모의 국가안
1958	「로마자자의 한글화 표기법」	확장 자모 폐기, 현용 자모 회귀
1986	현행 「외래어 표기법」	24자모·일반 표기의 통일
1990s – 2000s	Unicode·옛한글 처리 연구	부호화와 정규화·셰이핑의 기반 및 과제
2026	외국어 표기법 시안·규제합리화 패키지	두 층, 영어 RP·GA, 공개 데이터, 단계적 표준화
2026	훈민바름 영어확장판 4본: 선언본·해례본·전산본·구현본	뜻·원리·전산·구현의 역할 분업

참고문헌 및 작성 근거

제1본은 2026년 규제합리화 제안을 위해 내부 작성한 「외국어 표기법 시안」과 관련 연구자료를 바탕으로, 훈민바름 영어확장판 제2본 해례본·제3본 전산본·제4본 구현본의 기술 구조와 상호 정합성을 검토하여 다시 구성한 선언본이다.

문서의 역사적·언어학적·정책적 논의는 아래에 열거한 문헌과 공공 규범·표준자료를 참고하였다. 특히 훈민정음의 조자 원리와 한글 확장 가능성, 근현대 외국어 표기 규범의 변천, IPA와 Unicode, 외국어 교육·사전·음성기술 및 문자처리에 관한 주장은 해당 문헌의 연구 결과와 공공자료를 근거로 삼는다.

한편 훈민바름의 구체적인 음소-자모 대응, 정본 구조, 언어모듈, 역변환 규칙, 전산 스키마와 구현방식은 기존 연구에서 그대로 가져온 표준이 아니라 본 연구에서 제안하고 시험하는 가표본 연구의 설계안이다. 따라서 역사적 사실이나 선행연구의 견해와 훈민바름 자체의 제안·추론·시험규칙을 구분하여 읽어야 한다.

현재 기술적으로 구체화한 첫 참조범위는 영어 RP·GA이며, 일부 세부 규칙은 잠정 고정·시범·미결 상태를 포함한다. 이후 다른 영어 방언과 다른 언어로의 확장은 각 언어와 방언의 source IPA, 음운체계, 실제 화자 자료와 공개 검토를 바탕으로 별도의 프로파일 또는 언어모듈에서 이루어져야 한다.

아래 참고문헌은 이 제1본에서 직접 인용하거나 주요 논거의 근거로 활용한 자료를 중심으로 정리하였다. 훈민바름의 세부 기술규칙과 구현 결과에 관한 근거·상태·시험내용은 제2본·제3본·제4본에서 별도로 제시한다.

- International Phonetic Association(1999), Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet, Cambridge University Press.
- International Phonetic Association(2018), International Phonetic Alphabet Chart and Interactive IPA Chart.
- The Unicode Consortium(2025), The Unicode Standard, Version 17.0, Chapter 7: Europe-I.
- 박창원·김수현(2004), 「외래어 표기 양상의 변천」, 『새국어생활』 14(2), 59-102.
- 심재기(2008), 「외래어 표기법의 문제점과 그 해결책」, 『새국어생활』 18(4), 105-124.
- 홍윤표(2012), 「훈민정음에 대한 몇 가지 주장」, 『2012년 훈민정음학회 국내학술대회 발표논문집』, 1-29.
- 김상태(2012), 「훈민정음의 자소 결합 방식과 구조 연구」, 같은 발표논문집, 33-59.
- 이현복(2011), 「국제한글음성문자의 일본어, 중국어, 태국어 표기」, 『東亞文化』 49, 91-117.
- 김미애·김민(2023), 「IPA 국제음성기호 표기를 위한 한글 형태 연구 — 한글 신자소 디자인을 중심으로」, 『기초조형학연구』 24(4), 19-32.
- 김아영(2023), 「외래어 표기법과 표음주의」, 『언어사실과 관점』 59, 309-331.
- 김아영(2024), 「외래어 오탐기에 대한 음운론적 고찰」, 『청람어문교육』 100, 187-217.
- 이정복(2024), 「어문 규범에서 보이는 ‘외래어’와 ‘외국어’의 혼란과 대책」, 『한말연구』 65, 1-20.
- 이은경(2025), 「〈외래어 표기법〉에 대한 고찰: 〈한글 맞춤법〉과의 비교를 중심으로」, 『지식과 교양』 19, 301-340.

- 공나형(2022), 「외래어 표기법에서 인명의 관용 표기에 대한 비판적 검토와 대안 모색」, 『언어와 문화』 18(1), 1-29.
- 한명숙(2016), 「대학생들의 외래어 표기와 발음의 상관성 연구」, 『한말연구』 41, 205-240.
- 함계임(2023), 「외국인 관점에서 본 외래어 표기법의 현황과 과제」, 『언어와 언어학』 102, 121-142.
- 임현열(2020), 「한국어 TTS 외래어 발음 처리의 문제점과 개선 방안」, 『어문론집』 83, 35-62.
- 문성민·김수한·고연숙(2022), 「한국어 발음 변환기(G2P)의 현황과 성능 향상에 대한 언어학적 제안」, 『언어와 정보』 26(2), 27-46.
- 안대혁·박영배(2007), 「유니코드의 한글 인코딩 표준안」, 『정보과학회논문지: 소프트웨어 및 응용』 34(12), 1083-1092.
- 양왕성(2018), 「옛한글 코드의 변천사」, 『정신문화연구』 41(4), 213-238.
- 전광진(2019), 「타이완 세덕민족 한글 서사체계 활용의 어휘 기반 연구」, 『中語中文學』 77, 209-269.
- de Jong, Kenneth & Mi-Hui Cho (2012), “Loanword phonology and perceptual mapping: Comparing two corpora of Korean contact with English”, *Language* 88(2), 341-368.
- An, Jinsan, Kilim Nam & Hae-Yun Jung (2025), “Foreign Words, Loanwords, Korean Words: A Corpus-Based Reassessment of Lexical Resources for Korean Lexicography”, *Lexicography* 12(2), 154-180.
- Cho, Won Ik, Seok Min Kim & Nam Soo Kim (2020), “Towards an Efficient Code-Mixed Grapheme-to-Phoneme Conversion in an Agglutinative Language: A Case Study on To-Korean Transliteration”, *CALCS* 2020, 65-70.
- 이봉운(1897), 『국문정리』.
- 지석영 주석·증보(1908), 『아학편』.
- 이종극(1937), 『모던조선외래어사전』, 한성도서주식회사.
- 문교부(1948 제정, 1952 간행), 「들온말 적는 법」.
- 문교부(1958), 「로마자자의 한글화 표기법」.
- 문교부 고시 제85-11호(1986), 「외래어 표기법」.
- KS X 1026-1, 『정보교환용 한글 처리 지침』.
- The Unicode Consortium, Unicode Standard 및 UAX #15·#29: 정규화·그래픽·문자처리 관련 규정.
- Microsoft, “Developing OpenType Fonts for Korean Hangul Script”.
- 이서현·김소현(2025), 「판교 스타트업부터 정치권까지... 외국어 남용에 유권자 ‘아리송’」, 『경기일보』, 2025년 5월 31일.

연결 발음 학습과 선택권

2026년 10월 12일 개정 부록. 이 부록은 선언본의 교육적 적용 범위와 최신 참조구현 상태를 갱신한다. 본문 2쪽의 개발 현황과 28쪽의 음성 기술 전망은 그 작성 시점의 기록이며, 이번 부록의 현재 상태와 구분해서 읽는다.

일상 표기와 발음 연습의 관계

단어의 사전형만 익혀서는 회화에서 실제로 이어지는 소리를 이해하기 어렵다. 훈민바름 웹앱은 원형 IPA와 훈민바름을 보존한 채, 사용자가 연결 발음을 선택하고 비교하는 연습층을 제공한다. 연결 가능한 곳을 모두 강제로 연결하지 않는다. 일반 변환은 원형을 보여 주고 별도의 연음화 버튼이 연결 후보를 연다. 각 후보의 기본값은 OFF이며 추천 표시를 먼저 제시한다.

학습자는 각 연결을 켜거나 끌 수 있고, 말하기 속도·단어 강조·숨을 조절할 수 있다. 추천대로 적용하면 해당 설정의 후보가 켜진다. 모두 해제하면 원형으로 돌아간다. 추천은 교사의 판단을 돕는 규칙 기반 제안이며 문맥상 유일한 정답이나 원어민 발화의 복제는 아니다.

GA의 t 변이

이번 갱신은 일반 미국 영어를 중심으로 t 탈락, 성문음화, 미파열을 구분한다. button의 [ʔn], winter의 t가 들리지 않는 실현, 자음군 끝의 t/d 탈락을 선택적으로 비교한다. 소리가 약하게 들린다는 이유만으로 모두 같은 현상이라고 설명하지 않는다.

성문음화는 성문에서의 조절이고, 미파열은 혀의 폐쇄를 유지하며 뚜렷한 파열을 내지 않는 것이다. 탈락은 별도의 t 실현을 생략하는 연습형이다. 실제 발화에는 성문 강화와 구강 폐쇄가 함께 나타날 수도 있다. 해례 본 개정 부록에서 조건과 조음법을 설명한다.

표기의 책임 범위

[ʔ] 등 아직 배당하지 않은 소리에 임의의 옛한글 자형을 만들지 않는다. 해당 표면 IPA와 설명을 보여 주고 훈민바름은 원형 참조임을 표시한다. 일반 Broad 조자와 정본 NFD를 바꾸지 않으며, 표면 발음의 선택을 정본의 확정으로 해석하지 않는다.

교실 적용과 음성 재생의 다음 단계

관찰하고 선택하고 비교하기

교사는 먼저 원형과 선택형을 소리 내어 비교하고, 학생이 어떤 기관의 움직임을 바꾸었는지 말하게 할 수 있다. t 탈락에서는 별도의 혀 폐쇄·파열을 더하지 않는지, 미파열에서는 폐쇄를 유지하는지, 성문음화에서는 목청에서의 가벼운 끊김을 만드는지를 구분한다. 글자 모양의 암기만으로 발음 학습을 평가하지 않는다.

추천이 꺼져 있어도 가능한 발음이 틀린 것은 아니다. 말의 강조와 속도, 경계, 화자·지역 차이가 결과를 바꾼다. 정확한 대조를 위해 같은 문장에서 한 선택씩 바꾸고 원형도 함께 보존한다. 이 수업 구성은 활용 제안이며 학습 효과가 실험으로 입증되었다는 주장은 아니다.

현재 구현과 아직 남은 일

현재 제공	다음 연구와 구현
연결별 선택과 추천, 강조·쉽 설정	문장 의미와 운율을 고려한 추천
GA t 변이의 표면 IPA와 조 음 안내	자동 flapping 및 미결 자형 판정
원형과 선택형을 함께 담는 JSON	선택된 발음과 일치하는 TTS 재생

이 판에는 음성 재생을 넣지 않았다. 다음 단계의 TTS는 영어 철자를 읽는 기능과 선택된 표면 발음을 재현하는 기능을 구분해야 한다. 일반 TTS가 button을 자연스럽게 읽어도 사용자가 선택한 t 탈락·성문음화·미파열을 정확히 반영했다고 단정할 수 없다.

음성 재생을 도입할 때에는 원형과 선택형의 비교, 구간 반복, 속도 조절, 중지, 발음 일치 여부를 함께 검증한다. 음소 제어를 지원하지 않는 음성은 선택형의 확정된 시범으로 표시하지 않는다. 합성음의 적절성과 교육 효과는 별도의 화자·교사 검토가 필요하다.

다른 본과의 관계

제2본은 음성학적 구별과 조음 연습을, 제3본은 원형 토큰과 표면 주석의 보존을, 제4본은 화면 조작과 구현 한계를 갱신한다. 기존 본문의 역사·제도·표기 연구 결정을 이 기능 추가로 변경하지 않는다.

외국어 표기법 연구안

훈민바름 해례본

영어 RP·GA 정본과 음소 해례

손으로 적고 입으로 되살리는 영어확장판 비디지털 해례

훈민바름 영어확장판

제2본 해례본

2026년 10월 12일

연구·검토용 비구속 문서

우리 글, 바른 소리

Park Il Hwan

2026년 10월 12일 개정판 · 연구·검토용

연결 발음과 GA t 변이 개정 부록 · 60-62쪽

훈민바름 영어확장판 4본의 구성

훈민바름 영어확장판은 하나의 체계를 네 역할로 나눈다. 선언본은 뜻과 공공적 필요를 밝히고, 해례본은 손으로 기록할 수 있는 정본의 구조와 발음 원리를 풀며, 전산본은 정본의 디지털 보존·처리 규격을 세우고, 구현본은 웹 참조구현과 시험으로 실제 작동을 검증한다. 이 제2본은 사람이 디지털 기기 없이 적용할 수 있는 정본의 음가-자모 대응, 음절블록의 내부 순서와 손글씨 시범 조자법, RP·GA 음소와 조음법을 다룬다.

문서	핵심 내용	답하는 질문
제1본 선언본	살아 있는 확장 한글 문자체계의 역사·문화·사회·정책·교육·접근성·AI·표준화와 세계화 비전	왜 필요한가, 대한민국과 세계에 무엇을 열 수 있는가
제2본 해례본	영어 RP·GA 정본의 음가-자모 대응, 손글씨 블록 문법, 현대 해례식 조음 안내	어떤 소리를 어떻게 만들고 적고 다시 읽는가
제3본 전산본	정본의 Unicode·구조 데이터·글꼴·셰이핑·검색·복사·역변환 규격	디지털 환경에서 훈민바름 정본을 어떻게 변환하고 표시하는가
제4본 구현본	웹 참조구현, 사용자 흐름, 예제, 시험, 감사, 공개 배포와 오류 제보 절차	변환규칙이 실제로 웹앱을 통해 어떻게 작동하고 검증되는가

이 문서의 지위와 명칭

‘훈민바름’은 옛한글·확장 한글 자모를 활용한 「외국어 표기법」의 가능성을 연구·개발하는 동안 사용하는 내부 코드명이며, 공적 규범의 명칭을 선점하지 않는다.

현재 훈민바름은 정밀 외국어 표기체계의 구현 가능성을 시험하기 위한 하나의 가표본(working specimen) 연구이자 참조 실험(reference experiment)이다. 자모 배당과 표기규칙은 공개 검토와 시험을 통해 수정·보완하는 연구안이다.

제2본의 세부 규칙은 잠정 고정·시범·미결 세 단계로 공개한다. 미결사항은 감추어야 할 결함이 아니라 공개 검토와 시험을 통해 함께 풀어 갈 공동 연구 의제다.

RP·GA는 현재 선택한 첫 참조범위일 뿐 훈민바름의 최종 목적이 아니다. 다른 영어 방언은 독립 프로파일로, 다른 언어는 각 언어의 원문·source IPA·음운체계와 실제 화자 검토를 갖춘 독립 언어모듈로 연구·검토한다.

현재 개발·구현 상태

제1-4본은 연구·검토용 판본이다. 2026년 10월 9일 교정판은 문언·예시·상호 참조를 대조했다. Broad 웹앱은 공개 배포됐으며, 미결 음가·최종 조자법과 교육 효과 검증은 후속 과제다.

최신판 및 변경 이력

<https://github.com/ivanpark/hmbr>

교정판을 읽는 기준

2026년 10월 9일 교정판 · 제2본 해례본

이 판본은 기존 59쪽의 구성과 쪽수를 유지한다. 2026-09-17 저자 결정으로 뒤 모음 요소에 읽기 용 무음 초성 ㅇ을 붙인다. 부우ㅅㅌ의 우도 같은 음절핵에 속한다.

표시와 저장을 구별한다

본문의 단어 정본은 읽기용 표시다. 부우ㅅㅌ처럼 뒤 모음에 무음 ㅇ을 붙여도 영어 음소·음절은 늘지 않는다. 표의 날자값·핵 요소와 저장용 V 배열에는 표시용 ㅇ을 넣지 않는다. 저장 부호열은 제3본의 등록부와 직렬화를 따른다.

종성 /s/는 앞 모음의 한국어 받침으로 숨기지 않고 같은 음절블록 안에 음소 순서대로 드러낸다. /l/은 초성군에서 한 번만 기록한다. 장모음은 같은 음절핵의 두 요소이며, 두 영어 음절을 뜻하지 않는다.

예시의 범위

RP·GA는 이 연구가 선택한 참조모들이다. 모든 영국·미국 화자를 대표하지 않는다. Broad /r/ 표기는 이 책에서 영어 접근음 /ɹ/의 관례적 음소 표기이며 Narrow [r]의 설전음을 뜻하지 않는다. Narrow 예시는 특정 표면 실현의 설명이다.

GA /ɜ̃/의 기저 모음, 반모음의 최종 결합표기와 복합 블록의 최종 압축 자형은 미결이다. 제3본에서 종성 부호를 확정하지 않은 자모도 손글씨 설명에 나타날 수 있으며, 그것이 완성 평문을 내보낼 수 있다는 뜻은 아니다.

대조 자료

음가·시범 상태: 이 책 §§3-16 및 부록 A·B. 자리별 부호와 기준형: 제3본 §§4-9. 현재 프로그램의 지원 범위: 제4본 §§5-8-11.

어휘집합의 방법론: J. C. Wells, *Accents of English* (1982), Cambridge University Press. 이 책 §7.1은 그 집합명을 사용한 선택 예시 대조이며 전체 집합 목록이 아니다.

이 문서의 읽는 순서

이 해례본은 개별 규칙부터 나열하지 않고, 전체 변환 지도를 먼저 보여 준 뒤 각론으로 들어간다. 먼저 영어 원어가 RP·GA source IPA를 거쳐 훈민바름 정본이 되는 흐름을 보고, 자음·모음·복모음 대조표와 충분한 변환 예시를 살핀다. 그 뒤 자모값·블록 문법·위치형·Broad/Narrow·운율·연결발화를 규정하고, 제2부에서 각 음소의 실제 조음법을 해례식으로 풀이한다.

제1부는 “무엇을 어떻게 적는가”를 먼저 파악하고, 제2부에서 “그 소리를 입에서 어떻게 만드는가”를 확인한다.

순서	내용
제1부	규정과 예시
§ 2	변환 흐름 한 장 — 영어에서 정본까지
§ 3 – § 5	IPA – 훈민바름 대조표 세 장 (자음·모음·복모음)
§ 6	변환 갤러리 — 눈으로 보는 예
§ 7	RP와 GA는 어떻게 갈리는가 — 표준 어휘집합 참조어 대조
§ 8 – § 16	각론 (자모값·블록문법·ㄹ·/s·/t+s·/Broad/Narrow·강세·연결발화)
제2부	영어 음소 해례

차례

훈민바름 영어확장판 4본의 구성	P2
이 문서의 읽는 순서	P4
1. 4본 안에서 해례본과 정본의 지위	P6
2. 훈민바름 변환은 어떻게 이루어지는가	P7
3. IPA - 훈민바름 영어자음 대조표	P11
4. IPA - 훈민바름 영어모음 대조표	P13
5. IPA - 훈민바름 영어복모음 대조표	P14
6. 변환 갤러리 — 눈으로 보는 예	P15
7. 영국식 RP와 미국식 GA는 어떻게 다른가	P18
8. 기본 원칙과 자모값·위치형	P25
9. 블록 문법	P26
10. ㄹ 계열의 다중 대응	P29
11. /s/ 의 위치형	P29
12. /t+s/ 복합값 ㅅ	P30
13. Broad 와 Narrow	P31
14. 강세·장단·성조	P32
15. 어휘형과 연결발화형	P34
16. 잠정 고정·시범·미결 목록	P35
제2부 · 영어 음소 해례	P38
17. 발음기관과 해례 용어	P38
18. 자음 총론	P39
19. 파열음	P39
20. 파찰음과 복합 자음값	P40
21. 마찰음	P41
22. 비음	P43
23. 접근음과 ㄹ 계열	P44
24. 모음 총론	P46
25. 공통 핵심 모음	P46
26. RP의 모음 체계	P48
27. GA의 모음 체계	P49
28. 복모음 총론: 한 음절 안의 이동	P50
29. 공통 복모음 해례	P51
30. RP 복모음 해례	P52
31. GA 복모음과 vowel+/ɹ/	P53
32. RP와 GA의 조음 차이 요약	P54
33. 발음 점검과 대립 연습	P55
34. 땀음말 — 훈민바름 정본의 강점	P56
부록 A. 손글씨 작성 규칙	P58
부록 B. 현재 가표본 판정 요약	P59

1. 4본 안에서 해례본과 정본의 지위

훈민바름 정본은 사람이 종이에 직접 손으로 적을 수 있는 확장 한글기반 외국어 표기체계이다. Unicode 부호열이나 글꼴은 정본을 저장하고 나르는 수단이지 정본 자체가 아니다. 이 순서가 뒤집히면 기술이 바뀔 때마다 표기 원리가 흔들린다.

따라서 제2본의 정본 규칙은 디지털 기기 없이 적용할 수 있다. 현재 판정이 끝난 범위에서는 손과 종이와 IPA만 있으면 음가-자모 대응과 「한 영어 음절 = 한 훈민바름 음절블록」 원칙을 따라 적을 수 있다. 복합 블록은 현 단계에서 [] 안에 초성군·음절핵·종성군을 음소 순서대로 이어 적는 손글씨 시범 조자법을 쓴다. 이를 하나의 압축된 자형으로 배치하는 최종 조자형은 미결로 공개한다.

층위	담당	권위
손글씨 정본	제2본 해례본	원본
구조 데이터·Unicode 직렬화	제3본 전산본	정본의 계산 가능한 보존
화면 표시·대체 표시	제3본 전산본	여러 출력 가운데 사람이 보는 결과
참조구현·시험	제4본 구현본	검증

2. 훈민바름 변환은 어떻게 이루어지는가

이 책이 쓰는 표기 기호

// Broad IPA — 음소 대립만 적는다. 정보 변환의 입력이다.

[] Narrow IPA — 이음·조음 세부를 적는다. 훈민바름 블록에는 쓰지 않는다.

[] 훈민바름 음절블록 경계 — 영어 한 음절이 한 블록임을 보인다.

source IPA — 출처 사전의 IPA를 그대로 보존한 것.

normalized IPA — 해당 언어모듈의 규칙에 따라 훈민바름 매핑 직전에 정규화한 IPA.

2.1 변환 흐름

훈민바름 영어 확장판에서 영어 단어 하나가 훈민바름 정본이 되기까지는 여덟 단계의 변환 과정을 거치게 된다.

영어 원어 → 언어모듈 선택 (RP 또는 GA) → source IPA → normalized IPA
→ 음소 분해 → IPA-훈민바름 자모값 대응 → 영어 음절별 블록 구성 → 훈민바름 정본

pick 을 예로 들면, 정규화 (normalisation)가 필요 없는 단어에서는 source IPA와 normalized IPA가 같다.

단계	결과
영어 원어	pick
언어모듈※	영국식 RP(Received Pronunciation), 미국식 GA(General American)
source IPA	/pɪk/
normalized IPA	/pɪk/
음소 분해	/p/ + /ɪ/ + /k/
자모값 대응	ㅍ + ㅣ + ㅋ
블록 구성	초성 ㅍ · 중성 ㅣ · 종성 ㅋ 한 블록
정본	픽
※ Australian, Canadian, South African, Irish, Scottish 등의 다른 영어 방언은 향후 독립 언어모듈로 확장할 수 있다.	

pick의 /ɪ/에 ㅣ를 배당한 음성학적 근거는 § 4와 제2부에서 자세히 설명한다.

2.2 영어 음절과 훈민바름 음절블록

한국어와 영어는 모두 음절을 이루어 발음하지만, 한 음절 안에 들어갈 수 있는 소리의 배열은 서로 다르다. 현대 한국어는 영어에 비해 한 음절 안에서 허용되는 자음군의 범위가 좁고, 현대 한글 표기에서는 한 음절을 기본적으로 초성·중성·종성의 틀 안에 묶어 적는다. 반면 영어에서는 하나의 음절 안에 둘 이상의 자음이 이어지는 자음군이 흔하고, 음절핵도 복모음처럼 음절이 이동하거나 장모음처럼 길게 지속될 수 있다. 훈민바름은 이러한 복합 음절핵을 한 음절블록 내부의 중성 요소로 나타낸다.

예를 들어 pick /pɪk/은 단순한 초성 - 음절핵 - 종성 구조이지만, friend /frɛnd/에는 /fr-/와 /-nd/라는 자음군이 있고, boost /buːst/에는 장모음 /u:/와 /-st/ 종성군이 한 음절 안에 함께 있다. 이 소리들을 현대 한국어의 음절구조에 그대로 맞추려고 하면, 원래 영어에는 없는 모음을 넣거나 하나의 영어 음절을 여러 표기블록으로 나누게 되기 쉽다.

훈민바름은 이러한 차이를 줄이기 위해 영어의 음절 경계를 먼저 보존한다. 즉 영어 한 음절을 훈민바름 한 음절블록에 대응시키고, 그 안에서 초성군·음절핵·종성군의 음소 순서를 그대로 보존한다. 이는 영어의 음절구조를 한국어식으로 바꾸는 것이 아니라, 한글의 음절블록 원리를 유지하면서 블록 내부의 조자 범위를 영어의 음절구조에 맞게 확장하는 방식이다.

따라서 훈민바름에서 중요한 것은 자모의 수가 아니라 음절의 수이다. 한 블록 안에 여러 자모가 들어가더라도 그것들이 하나의 영어 음절을 이루는 한 하나의 훈민바름 음절블록으로 적는다.

핵심 원칙 · 한 영어 음절 = 한 훈민바름 음절블록

음절블록 안에는 초성군·음절핵·종성군이 음소 순서대로 들어간다. 초성군과 종성군에는 자모가 둘 이상 올 수 있다.

영어의 자음군은 별도 블록으로 떼어 내지 않고 같은 블록의 초성군 또는 종성군으로 받는다.

복모음과 장모음의 음절핵에는 모음 요소가 둘 이상 온다. 이 역시 같은 블록의 중성 자리에 순서대로 놓는다.

이 원칙은 선택된 언어모듈의 음절분석을 기준으로 영어 음절 수와 훈민바름 음절블록 수를 일치시킨다.

이 구별이 필요한 이유는 곧바로 드러난다.

단어	source IPA	영어 음절 수	훈민바름 음절블록 수	내부 구성	정본
pick	/pɪk/	1	1	초성 ㅍ · 핵 ㅣ · 종성 ㄱ	픽
bank	/bæŋk/	1	1	초성 ㅂ · 핵 ㅏ · 종성군 ㅇㄱ	뱅크
friend	/frɛnd/	1	1	초성군 ㅍㄹ ㅐ · 핵 ㅓ · 종성군 ㄴㄷ	프렌드
boost	/buːst/	1	1	초성 ㅂ · 핵 ㅜ · 종성군 ㅍㅌ	부우스트

bank·friend·boost 는 영어에서 모두 한 음절이므로 훈민바름에서도 각각 한 음절블록이다. 저장 자모가 넷·다섯·다섯으로 늘어난 것은 음절이 늘어난 것이 아니라 초성군과 종성군의 음소 순서를 그대로 보존했기 때문이다.

2.3 복모음과 장모음

복모음은 한 음절 안에서 한 모음 위치에서 다른 모음 위치로 음질이 이동하는 하나의 음절핵이다. 훈민바름은 구성요소의 자모값을 보존하기 위해 같은 음절블록의 중성 자리에 두 모음 요소를 순서대로 놓는다. 장모음은 하나의 모음이 길게 지속되는 음절핵이며, 현재 가표본에서는 같은 음절블록 안에서 해당 모음값을 반복하여 그 길이를 나타낸다.

2026-09-17 저자 결정으로 복모음·장모음의 뒤 모음 요소에는 읽기용 무음 초성 ㅇ을 붙인다. 초성 없이 놓인 모음의 시각적 혼동을 줄이기 위한 조자 원칙이다. 따라서 boost는

〔부우ㅍㅌ〕으로 적는다. 우는 /u:/의 뒤 모음 요소이며 새 음절이 아니다. 옛한글 모음도 무음 ㅇ과 함께 조합하되, 자모값과 요소 순서를 보존한다.

표시용 ㅇ은 새 자음 음소가 아니며 같은 음절블록 안의 읽기용 자리다. 저장 부호열에는 이를 더하지 않는다.

이 책의 복모음·장모음 정본 예시는 별도 표시가 없는 한 이 시범 표기를 따른다.

복모음은 두 모음값의 순서를 보존하고, 장모음은 같은 모음값의 반복과 길이 속성으로 길이를 보존한다. 두 경우 모두 전체 음절핵은 하나의 영어 음절에 속한다.

자음군은 하나의 영어 음절 안에서 연속되는 자음들의 결합이다. 자음군은 별도의 블록으로 떼어 내지 않고, 모음 앞의 것은 초성군에, 모음 뒤의 것은 종성군에 음소 순서대로 넣는다. 따라서 자음군 사이에 임의의 모음을 넣거나 음절을 나누지 않는다.

예를 들어 bank의 〔뱅크〕, friend의 〔프렌드〕, boost의 〔부우ㅍㅌ〕은 모두 한 음절이고 한 음절블록이다.

단어	source IPA	음절 수	훈민바름 블록 구조	설명
boy	/bɔɪ/	1	[보 _이]	음절핵의 두 모음 요소
boost	/buːst/	1	[부 _우 스 _ㅌ]	음절핵 두 모음 요소 + 종성군 두 자모
low (RP)	/ləʊ/	1	[로 _우]	두 모음값을 각각 보존한 음절핵

복합 음절핵의 현재 처리

잠정 고정 — 복모음의 구성요소는 1:1로 보존하고, 복모음·장모음 전체는 하나의 영어 음절핵으로 본다.

저자 결정 — 뒤 모음 요소에 읽기용 무음 초성 ㅇ을 붙인다. 같은 음절핵의 요소이며 새 음소나 음절을 더하지 않는다.

저장과 표시 — 핵의 자모값·V 배열은 그대로 보존한다. 표시용 ㅇ은 저장 부호열과 구별한다.

위의 [보_이] , [부_우스_ㅌ] , [로_우] 은 모두 이 시범 표기를 따른다.

3. IPA—훈민바름 영어자음 대조표

아래 표는 영어 RP·GA 참조모듈의 자음 음소와 주요 Narrow 이음을 훈민바름 정본 자모에 대응한 총괄표다. Unicode 코드포인트와 입력·셰이핑 규격은 제3본 전산본에서 별도로 규정한 다.

IPA	IPA No.	조음 분류	정본	위치·상태	분류	단어 예시	정본 예
/p/	101	무성 양순 파열음	ㅍ	초성·종성	Broad	pun /pʌn/	펀
[p]	101	무성 양순 무기 파열음	ㅍ	/s/ 뒤	Narrow	spin [spɪn]	ㅍ스핀
/b/	102	유성 양순 파열음	ㅂ	초성·종성	Broad	boost /bu:st/	부우스테
/t/	103	무성 치경 파열음	ㅌ	초성·종성	Broad	task /tɑ:sk/	타아스크
[t]	103	무성 치경 무기 파열음	ㅌ	/s/ 뒤	Narrow	stool [stu:t]	ㅌ스투울
[r]	124	치경 탄설음	ㄹ	GA 모음 사이	Narrow	water ['wɔ:tə]	워터
/d/	104	유성 치경 파열음	ㄷ	초성·종성	Broad	day /deɪ/	데이
/k/	109	무성 연구개 파열음	ㅋ	초성·종성	Broad	kill /kɪl/	킬
[k]	109	무성 연구개 무기 파열음	ㅋ	/s/ 뒤	Narrow	skill [skɪl]	ㅌ스킬
/g/	110	유성 연구개 파열음	ㄱ	초성·종성	Broad	good /gʊd/	굳
/tʃ/	103/134	무성 후치경 파찰음	ㄷ	한 음소 /t+s/ 와 구별	Broad	catch /kætʃ/	캐ㄷ
/dʒ/	104/135	유성 후치경 파찰음	ㄷ	한 음소	Broad	judge /dʒʌdʒ/	ㄷㄷ
/t+s/	103+132	파열음 + 마찰음	ㅌ	두 음소 복합값 종성군 한 자리	Broad	cats /kæts/	캐ㅌ
/f/	128	무성 순치 마찰음	ㅍ	초성·종성	Broad	fish /fɪʃ/	페ㅣ
/v/	129	유성 순치 마찰음	ㅍ	초성·종성	Broad	very /'veri/	페ㅣ
/θ/	130	무성 치 마찰음	ㅌ	초성 · 종성	Broad	three /θri:/	ㄷ 페ㅣ
/ð/	131	유성 치 마찰음	ㅌ	초성 · 종성	Broad	this /ðɪs/	페ㅣ
/s/	132	무성 치경 마찰음	ㅌ	위치형 · 종성	Broad	sun /sʌn/ pass (GA) /pæs/	스핀 패ㅌ
/z/	133	유성 치경 마찰음	ㅌ	초성·종성	Broad	noise /nɔɪz/	노ㅣ
/ʃ/	134	무성 후치경 마찰음	ㅌ	초성·종성	Broad	fish /fɪʃ/	페ㅣ

IPA	IPA No.	조음 분류	정보	위치·상태	분류	단어 예시	정보 예
/ɜ:/	135	유성 후치경 마찰음	ㅅ	초성·종성	Broad	measure /'meɜ:/	메 ㅅ
/h/	146	무성 성문 마찰음	ㅎ	초성	Broad	home (GA) /houm/	호 ㅅ
/m/	114	양순 비음	ㅁ	초성·종성	Broad	milk /mɪlk/	밀 ㄴ
/n/	116	치경 비음	ㄴ	초성·종성	Broad	now /naʊ/	나 ㅅ
/ŋ/	119	연구개 비음	ㅇ	종성 전용	Broad	sing /sɪŋ/	씁
/l/	155	치경 설측 접근음	ㄹ	초성·종성 C+/l/은 초성군	Broad	play /pleɪ/	플레이
[ɫ]	209	연구개화 설측음 (dark L)	ㄹ	어말·자음 앞	Narrow	milk [mɪɫ]	밀 ㄴ
/ɹ/	151	치경 접근음	ㄹ	초성 · 종성 ㄹ	Broad	very /'veri/ car (GA) /kɑɹ/	예 ㄹ 칼
/w/	170	양순 연구개 접근음	ㅇ	초성	Broad	wood /wʊd/	웁
/j/	153	경구개 접근음	ㅇ	초성	Broad	yes /jes/	예 ㅅ
[ʔ]	113	성문 파열음	미정	미결	Narrow	button ['bʌʔŋ]	—

※ source IPA 또는 Narrow IPA에 /ɹ/가 나타나는 경우, 훈민바름 자모 매핑 전 normalized IPA 단계에서 /əɹ/로 분해한다. 표의 source/Narrow 표기는 출처형 또는 표면형을 보존한 것이다.

이 표를 읽는 세 가지 요령

1. Broad의 /p·t·k/

Broad에서는 /p·t·k/의 기본 자모값을 ㅍ·ㅌ·ㄱ으로 적는다. 영어의 /p·t·k/는 특히 강세 음절의 첫머리에서 숨을 동반해 실현되는 경우가 많으며, 이 특징을 고려한 기본 대응이다.

2. /s/ 뒤의 Narrow [p·t·k]

/s/ 뒤에서는 /p·t·k/가 숨이 거의 따라 나오지 않는 [p·t·k]로 실현된다. Narrow에서는 이 차이를 ㅍ·ㅌ·ㄱ으로 구별해 적는다. 이는 한국어의 된소리와 동일하다는 뜻이 아니라, 영어 발음에서 나타나는 숨의 동반 여부 차이를 표시하기 위한 것이다.

3. /tʃ/와 /t+s/

/tʃ/는 하나의 음소이므로 ㅈ으로 적고, /t+s/는 /t/와 /s/ 두 음소의 연쇄이므로 하나의 ㅈ으로 적는다. 두 값은 반드시 구별한다.

4. IPA—훈민바름 영어모음 대조표

IPA	핵	성격	대표 단어	정본 예
/i:/	ㅣ	장모음 · 두 모음 요소	fleece /fli:s/	엘리이스
/ɪ/	ㅣ	단모음	kit /kɪt/	킵
/ɛ/	ㅓ	단모음	dress /dres/	드레스
/æ/	ㅐ	단모음	trap /træp/	트랩
/ʌ/	ㅓ	단모음	strut /strʌt/	스트럿
/ʊ/	ㅜ	단모음	foot /fʊt/	풋
/u:/	ㅜ	장모음 · 두 모음 요소	goose /gu:s/	구우스
/ə/	·	약모음	comma /'kɒmə/	코마
/ɜ:/	ㅓ	RP 장모음	nurse (RP) /nɜ:s/	너어스
/ɑ/	ㅏ	GA 단모음	lot (GA) /lɑt/	랏
/ɑ:/	ㅏ	RP 장모음	palm /pɑ:m/	파아모
/ɒ/	ㅓ	RP 단모음	lot (RP) /lɒt/	롤
/ɔ/	ㅜ	GA 단모음	thought /θɔt/	올
/ɔ:/	ㅜ	RP 장모음	thought (RP) /θɔ:t/	오테
/i/	ㅣ	무강세 어말 /i/	happy /'hæpi/	해피

세 자모가 이 표의 핵심이다.

1. ㅣ (/ɪ/) — ㅣ 로 접으면 pick 과 peak 이 같아진다.
2. ㅜ (/ʊ/) — ㅜ 로 접으면 good 과 food 의 모음이 같아진다.
3. · (/ə/) — 영어에서 가장 흔한 모음이며, 강세 없는 자리를 표시한다.

RP SQUARE의 /ɛ:/=ㅓ ㅓ는 § 26의 모음 해례를 따른다.

5. IPA—훈민바름 영어복모음 대조표

뒤 모음 요소에는 읽기용 무음 초성 ㅇ을 붙인다. 아래 ‘구성’과 ‘핵 요소’는 음가 대응이며, ‘정본 예’는 이 조자 원칙을 적용한 단어 표시다.

IPA	구성	핵 요소	대표 단어	정본 예
/eɪ/	ㅈ + ㅣ	ㅈㅣ	face /feɪs/	페이ㅅ
/aɪ/	ㅈ + ㅣ	ㅈㅣ	price /praɪs/	프라이ㅅ
/ɔɪ/	ㅈ + ㅣ	ㅈㅣ	choice /tʃɔɪs/	쵸이ㅅ
/aʊ/	ㅈ + ㅞ	ㅈㅞ	mouth /maʊθ/	마 _후 ㅞ
RP /əʊ/	· + ㅞ	· ㅞ	goat /gəʊt/	구 _우 ㅞ
GA /oʊ/	ㅈ + ㅞ	ㅈㅞ	goat /gout/	고 _우 ㅞ
RP /ɪə/	ㅣ + .	ㅣ.	near /nɪə/	니 .
RP /ʊə/	ㅞ + .	ㅞ.	cure /kjʊə/	쿠 _후 .

goat 를 보면 RP는 · 로 시작하고 GA는 ㅈ 로 시작한다. IPA 에서 /əʊ/ 와 /oʊ/ 로 갈리는 차이가 훈민바름에서 그대로 눈에 보인다. 이것이 § 7 의 주제다.

6. 변환 갤러리 — 눈으로 보는 예

각론에 들어가기 전에, 이 체계가 실제로 무엇을 만드는지 한자리에서 본다. 아래 예는 모두 정보이며, 뒤 장들이 그 근거를 하나씩 풀어 준다.

6.1 기본 — 한 음절, 한 블록

단어	source IPA	정보	핵심
pick	/pɪk/	픽	초성·중성·종성 한 블록
pat	/pæt/	팻	종성 하나가 같은 블록의 종성 자리에
sing	/sɪŋ/	싱	/ŋ/ 은 종성 전용
fish	/fɪʃ/	페이 _△	/f/ = 0ㅍ
very	/ˈveri/	베이 ₀ 이 ₀	/v/ = 0ㅅ, /ɪ/ = 0어
judge	/dʒʌdʒ/	저 ₀ 아	/dʒ/ 는 한 음소
measure	/ˈmeɪʒə/	메이 ₀ 아	/ʒ/ = 0ㅈ

6.2 장모음 — 음절핵의 두 모음 요소

단어	source IPA	정보	핵심
food	/fu:d/	0ㅍ ₀ 우 ₀ ㅌ	뒤 요소는 같은 블록의 둘째 중성
soon	/su:n/	수 ₀ 우 ₀ ㄴ	/u : / 앞 /s/는 ㅅ
three	/θri:/	아 ₀ 어 ₀ 이	어두 자음군 + 장모음

6.3 단모음 ʊ — /u/ 는 ㅜ가 아니다

단어	source IPA	정본	대비
good	/ɡʊd/	꺠	↔ food /fu:d/ ㅍㅜ
foot	/fʊt/	풋	↔ boot /bu:t/ 부
soot	/su:t/	쑤	↔ suit /su:t/ 수

6.4 자음군 — 한 음절, 한 블록

단어	source IPA	정본	핵심
friend	/frɛnd/	ᄇᄇᄇ	초성군 + 음절핵 + 종성군
spin	/spɪn/	ㅅㅍㅍ	Broad. Narrow 는 ㅅㅍ
bank	/bæŋk/	뱅크	종성군 두 자모
milk	/mɪlk/	밀크	종성군 두 자모
play	/pleɪ/	플레이	C + /l/ 초성군
blue	/blu:/	블루	C + /l/ 초성군 + 장모음

6.5 종성 /s/ — 음소 순서대로 독립 표기

단어	source IPA	정본	핵심
pass (GA)	/pæs/	패스	팻 으로 합치지 않는다
bus	/bʌs/	버스	
this	/ðɪs/	디스	
boost	/bu:st/	부우스트	/s/ 뒤 종성은 음소 순서대로
task	/tɑ:sk/	타아스크	

6.6 /t+s/ 복합값과 /tʃ/ — 최소대립쌍

단어	source IPA	정보	핵심
cats	/kæts/	캐츠	두 음소 복합값
catch	/kætʃ/	캐 ^ㅅ	한 음소 파찰음
hits	/hɪts/	히츠	
abates	/əˈbeɪts/	ə 베 ^이 츠	복모음 + 복합값

이 대비는 /t+s/와 /tʃ/의 구조 차이를 한글에서 보존해야 하는 이유를 보여 준다. 훈민바름은 두 음소 연쇄 /t+s/를 츠으로, 한 음소 /tʃ/를 ㅅ으로 분리한다.

6.7 복모음

단어	source IPA	정보	핵심
bite	/baɪt/	바 ^이 ㅌ	ㅏ + ㅌ
say	/seɪ/	세 ^이	ㅓ + ㅌ
boy	/bɔɪ/	보 ^이	ㅓ + ㅌ
now	/naʊ/	나 ^우	ㅏ + ㅗ
choice	/tʃɔɪs/	오 ^이 스	복모음 + 종성군 /s/

6.8 Narrow — 표면형

단어	표면 IPA	정보	핵심
water (GA)	['wɔrə]	워 ^루 어	표면형 [ɹ] 보존 · 매핑 전 /ɹ/→/əɹ/ 정규화 · 모음 사이 /t/→[ɾ] 탄설음 ㄹ
stool	[stuːl]	스투우 ^ㄹ	/s/ 뒤 무기음 + dark L
milk	[mɪlk]	밀크	dark L

7. 영국식 RP와 미국식 GA는 어떻게 다른가

영어는 하나의 단일한 발음체계로 이루어진 언어가 아니다. 한국어에도 지역과 세대에 따라 발음 차이가 존재하듯 영어에도 여러 발음체계가 있으며, 정밀한 외국어 표기에서는 먼저 어떤 영어의 소리를 적는가를 밝혀야 한다.

그러나 이 장의 목적은 RP와 GA의 차이 자체를 설명하는 데 있지 않다. 핵심은 그 차이가 훈민바름 정보에서 실제 문자 차이로 보존되고 구별될 수 있는지를 확인하는 데 있다.

현용 한글의 통상 외래어 표기는 한국어 문자생활의 통일성과 관용을 우선하므로, 서로 다른 영어 발음도 같은 표기로 합쳐지는 경우가 많다. 반면 훈민바름은 source IPA와 언어모듈의 음운 대립을 자모값과 음절블록 구조에 대응시키므로, 모음의 질·길이·복모음의 이동·/ɹ/의 실현 여부와 같은 차이가 정보 자체의 차이로 나타난다.

예를 들어 같은 car라도 RP와 GA의 source IPA가 다르면 훈민바름 정보도 달라지고, go, pass, hot, near에서도 같은 일이 일어난다. 중요한 것은 어느 발음이 더 옳으냐가 아니라, 서로 다른 발음을 서로 다른 글자로 되살릴 수 있다는 점이다.

따라서 RP와 GA는 단순한 비교 대상이 아니라 훈민바름의 음운 변별력과 역변환 가능성을 시험하는 첫 두 참조모듈이다. 아래의 표준 어휘집합 대조는 두 방언의 차이를 나열하기 위한 표가 아니라, 그 차이가 훈민바름에서 얼마나 일관되게 문자화되는지를 검증하는 시험표다.

7.1 표준 어휘집합의 참조어 대조

아래는 표준 어휘집합의 이름을 참조어로 삼은 26개 사례다. FORCE 등 전체 집합을 망라하지 않는다. 대조 열은 해당 참조어의 전체 정보를 비교하며, 모든 화자의 변이를 대표하지 않는다. commA의 차이는 첫 모음이고 끝의 /ə/는 같다.

어휘집합	RP source IPA	RP 정보	GA source IPA	GA 정보	대조
KIT	/kɪt/	킵	/kɪt/	킵	같음
DRESS	/dres/	드레스	/dres/	드레스	같음
TRAP	/træp/	트랩	/træp/	트랩	같음
LOT	/lɒt/	롤	/lat/	랄	다름
STRUT	/strʌt/	스트럿	/strʌt/	스트럿	같음
FOOT	/fʊt/	풋	/fʊt/	풋	같음

기존 본문은 10월 9일 기준 · 연결 발음과 t 변이 갱신은 60쪽부터

어휘집합	RP source IPA	RP 정본	GA source IPA	GA 정본	대조
BATH	/bɑːθ/	바아 ^ㅅ	/bæθ/	배 ^ㅅ	다름
CLOTH	/klɒθ/	크로 ^ㅅ	/klɔθ/	크로 ^ㅅ	다름
NURSE	/nɜːs/	너어ㅅ	/nɜs/	— (기저모음 미결)	미결
FLEECE	/fliːs/	ᄇᆞ 리이ㅅ	/fliːs/	ᄇᆞ 리이ㅅ	같음
FACE	/feɪs/	ᄇᆞ이ㅅ	/feɪs/	ᄇᆞ이ㅅ	같음
PALM	/pɑːm/	파아ㅁ	/pam/	팜	다름
THOUGH HT	/θɔːt/	ㅅ ㅈ 오 ㅌ	/θɔt/	올	다름
GOAT	/gəʊt/	ㄱ 우 ㅌ	/gout/	고우 ㅌ	다름
GOOSE	/guːs/	구우ㅅ	/guːs/	구우ㅅ	같음
PRICE	/praɪs/	프 ㅍ라이ㅅ	/praɪs/	프 ㅍ라이ㅅ	같음
CHOICE	/tʃɔɪs/	쑈이ㅅ	/tʃɔɪs/	쑈이ㅅ	같음
MOUTH	/maʊθ/	마우 ^ㅅ	/maʊθ/	마우 ^ㅅ	같음
NEAR	/nɪə/	니 ㅎ	/nɪɹ/	닐	다름
SQUARE	/skweɪ/	스 ㅈ ㅅ ㅍ 에	/skweɪ/	스 ㅈ ㅅ ㅍ ㅅ	다름
START	/stɑːt/	스타아ㅌ	/staɪt/	스탁ㅌ	다름
NORTH	/nɔːθ/	노오 ^ㅅ	/nɔɪθ/	눔 ^ㅅ	다름
CURE	/kjʊə/	크후 ㅎ	/kjuɹ/	크후	다름
happY	/ˈhæpi/	해피	/ˈhæpi/	해피	같음
lettER	/ˈletə/	레 ㅌ	/ˈletə/	레 ㅌ	다름
commA	/ˈkɒmə/	코 ㅁ	/ˈkamə/	카 ㅁ	다름

※ CURE의 현재 정본 예는 /j/와 뒤따르는 모음값을 각각 보존한 명시형이다. C+/j/ 결합표기의 최종 판정을 뜻하지 않으며, 축약형·조자원리 확장형 등 다른 후보는 §16의 미결 과제로 비교한다.

26개 참조어의 정본 대조는 같음 12개, 다름 13개, 미결 1개다. GA NURSE의 기저 모음은 확정하지 않는다. 이는 선택한 예시의 내부 대조이며 전체 방언 차이의 전수조사가 아니다.

7.2 차이는 다섯 갈래로 정리된다

갈래 1 · rhoticity — 가장 큰 차이

GA는 어말과 자음 앞의 /ɹ/ 를 발음하고, RP는 발음하지 않는다. 훈민바름에서는 GA 에 종성 ㄹ이 나타나고 RP 에는 나타나지 않는다.

단어	RP source IPA	RP 정본	GA source IPA	GA 정본
car	/kɑː/	카아	/kɑɹ/	캬
four	/fɔː/	ᄇᄇ ᄇᄇ	/fɔɹ/	ᄇᄇ
here	/hɪə/	히 ᄇ	/hɪɹ/	히ᄇ
park	/pɑːk/	파아ᄇ	/pɑɹk/	팍ᄇ
short	/ʃɔːt/	ᄇᄇ ᄇᄇᄇ	/ʃɔɹt/	숏ᄇ
farm	/fɑːm/	ᄇᄇ ᄇᄇᄇ	/fɑɹm/	ᄇᄇ ᄇᄇ ᄇᄇ

현재 참조모듈에서 RP의 비권설 어말·자음 앞 /ɹ/ 비실현은 장모음이나 중심화된 모음과 함께 나타나는 경우가 많다. 그래서 car는 RP에서 카아(음절핵의 두 모음 요소), GA에서 ㄱ ᄇᄇ (/ɹ/ 종성 포함)으로 갈린다. 어느 쪽이든 한 음절이므로 한 음절블록이다. 두 발음의 길이와 rhoticity 차이가 정본 구조에 그대로 남는다.

약모음 뒤에서도 같은 일이 일어난다.

단어	RP source IPA	RP 정본	GA source IPA	GA 정본
better	/ˈbetə/	베ᄇ	/ˈbetɹə/	베ᄇ
water	/ˈwɔːtə/	ᄇᄇ ᄇᄇᄇ	/ˈwɔɹtə/	ᄇᄇᄇ
teacher	/ˈtiːtʃə/	티이ᄇ	/ˈtiːtʃɹə/	티이ᄇᄇ
mother	/ˈmʌðə/	머ᄇ	/ˈmʌðɹə/	머ᄇᄇ

better의 끝은 RP에서 ᄇ, GA에서 ᄇ이다. 두 발음을 모두 ‘베터’처럼 적는 통상 한글 병기만으로는 이 차이가 드러나지 않는다. 정밀 음가 기록에서는 이를 보존한다.

갈래 2 · GOAT — 시작점이 다르다

RP 의 /əʊ/ 는 약모음 ㅏ 에서 출발하고, GA 의 /ou/ 는 ㅓ 에서 출발한다. 끝점은 둘 다 ㅜ 로 같다.

단어	RP source IPA	RP 정본	GA source IPA	GA 정본
go	/gəʊ/	ㄱ우	/gou/	고우
low	/ləʊ/	ㄴ우	/lou/	로우
boat	/bəʊt/	ㅂ우ㅌ	/bout/	보우ㅌ
home	/həʊm/	ㅎ우ㅁ	/houm/	호우ㅁ
note	/nəʊt/	ㄴ우ㅌ	/nout/	노우ㅌ

이 대비는 ㅜ 판정이 있어야만 성립한다. ㅜ을 /ʊ/의 독립 자모값으로 두면 GOAT의 끝점이 /uː/와 구별된다. RP /əʊ/와 GA /ou/의 방언 차이는 시작점 ㅏ 와 ㅓ에서 드러나며, 두 모둘은 공통적으로 끝점 /ʊ/=ㅜ을 보존한다. § 4 에서 ㅏ·ㅜ 를 고집한 이유가 여기서 실제 효과로 나타난다.

갈래 3 · BATH — 집합이 갈라선 자리

bath, pass, dance, ask와 같은 단어들은 RP와 GA에서 서로 다른 모음으로 발음된다. RP에서는 palm의 모음과 같은 긴 /ɑː/가 나타나는 반면, GA에서는 trap의 모음과 같은 짧은 /æ/가 나타난다. 영어 음운론에서는 이러한 차이를 BATH 집합의 분화로 설명한다.

훈민바름에서는 이 차이가 그대로 정본에 나타난다. 예를 들어 pass는 RP /pɑːs/에서 파아ㅅ, GA /pæs/에서 패ㅅ으로 갈린다. 즉 단순히 ‘영국식과 미국식 억양이 다르다’는 설명에 그치지 않고, 서로 다른 모음과 길이 차이가 서로 다른 글자로 보존된다.

단어	RP source IPA	RP 정본	GA source IPA	GA 정본
pass	/pɑːs/	파아ㅅ	/pæs/	패ㅅ
bath	/bɑːθ/	바아ㅅ	/bæθ/	배ㅅ
dance	/dɑːns/	다아ㄴㅅ	/dæns/	댄ㅅ

단어	RP source IPA	RP 정보	GA source IPA	GA 정보
ask	/ɑ:sk/	아아스 ㅋ	/æsk/	애스 ㅋ
half	/hɑ:f/	하아 ^ㅁ	/hæf/	해 ^ㅁ

pass 는 RP 에서 파아스, GA 에서 패스 이다. 모음이 다를 뿐 아니라 음절핵의 구성까지 달라진다 — pass의 RP 정보는 한 음절블록의 음절핵에 두 모음 요소를 가진다.

갈래 4 · LOT·CLOTH·THOUGHT — 뒤쪽 모음의 재배치

RP는 짧은 ㅏ(/ɒ/)와 긴 ㅓㅓ(/ɔ:/)로 나뉘고, GA는 ㅏ(/ɑ/)와 ㅓ(/ɔ/)로 나뉜다. 경계선이 서로 다른 자리에 있다.

단어	RP source IPA	RP 정보	GA source IPA	GA 정보
hot	/hɒt/	홀	/hat/	할
stop	/stɒp/	ㅅ툵	/stap/	ㅅ탸
dog	/dɒg/	독	/dɔg/	독
long	/lɒŋ/	롱	/lɔŋ/	롱
caught	/kɔ:t/	코오ㅌ	/kɒt/	쿨

hot 은 RP 홀, GA 할 이다. 한 글자 안에서 중성만 바뀐다. dog 는 RP 독, GA 독 이다 — 같은 단어의 모음이 두 참조모듈에서 서로 다른 음가로 대응하기 때문이다.

갈래 5 · NEAR·SQUARE·CURE — RP와 GA의 핵 구조가 다르다

현재 RP 참조모듈은 NEAR /ɪə/와 CURE /ʊə/를 복모음으로 두고, SQUARE는 /ɛ:/ 장모음으로 정리한다. GA에서는 이 세 집합을 대체로 모음 + /ɪ/ 연쇄로 분석한다. 다른 RP 분석에서 SQUARE /eə/가 쓰일 경우에는 원문 IPA를 보존한 별도 변이로 관리한다.

기존 본문은 10월 9일 기준 · 연결 발음과 t 변이 갱신은 60쪽부터

단어	RP source IPA	RP 정보	GA source IPA	GA 정보
square	/skweɪ:/	스 ㄱ ㅁ ㅁ 에	/skwɛɪ/	스 ㄱ ㅁ ㅁ ㅁ
near	/niə/	니 ㅁ	/niɪ/	니 ㅁ
cure	/kjʊə/	크 후 ㅁ	/kjʊɪ/	크 후 ㅁ
air	/ɛ:/	에에	/ɛɪ/	엠펙

훈민바름에서는 이 차이가 핵과 종성의 구조로 드러난다. RP의 NEAR·CURE에는 음절핵의 두 번째 모음 요소가 오고, SQUARE는 /ɛ:/를 두 모음 요소를 가진 음절핵으로 적는다. GA에서는 대응 모음 뒤에 /ɪ/ 종성 ㅁ이 온다.

7.3 Narrow 에서만 갈리는 것 — GA 탄설음

GA에서는 강세 모음 뒤와 무강세 모음 앞 등의 환경에서 /t,d/가 [ɾ]로 실현될 수 있다. Broad의 기저 음소는 보존하고, Narrow에서 표면 이음을 구별한다.

단어	RP Broad	GA Broad	GA Narrow
water	ㅁ ㄴ ㅁ ㅁ ㅁ	ㅁ ㅁ ㅁ	ㅁ ㅁ ㅁ
better	베 ㅁ	베 ㅁ	베 ㅁ
city	씨 티	씨 티	씨 림

Broad 는 음소를 적고 Narrow 는 실현을 적는다. 따라서 Narrow 표기는 Broad 를 덮어쓰지 않고 덧붙인다.

7.4 이 장의 요점

	현용 한글의 통상 표기	훈민바름 (RP · GA)
car	카	카아 · ㄱ ㅅ
better	베터	베트 · 베툼
go	고	구 · 고우
pass	패스	파아스 · 패스
hot	핫	홀 · 할
near	니어	니우 · 님

이 장의 핵심은 RP와 GA가 서로 다르다는 사실 자체가 아니다. 이미 알려진 그 음운 차이를 훈민바름이 실제 문자 차이로 보존할 수 있다는 점이다.

현용 한글 표기에서는 car, go, pass, hot, near와 같은 낱말의 RP·GA 차이가 동일하거나 유사한 표기로 합쳐질 수 있다. 훈민바름에서는 선택한 언어모듈의 source IPA가 달라지면 그 차이가 모음 자모값, 장단, 복모음 구성, 종성 /ɹ/ 등의 구조 차이로 나타난다.

즉 훈민바름은 영어 발음을 한국어에 맞추어 하나로 평준화하는 표기법이 아니라, 서로 다른 원어 발음을 서로 다른 정본으로 보존하고 다시 구별해 읽을 수 있게 하는 표기체계를 지향한다.

RP와 GA의 대조는 그 가능성을 보여 주는 첫 시험이다. 이후 Australian English, Canadian English 등 다른 영어 발음체계가 추가되더라도 같은 원칙으로 독립 언어모듈을 구성할 수 있다.

8. 기본 원칙과 자모값·위치형

8.1 여덟 가지 기본 원칙

1. 표기의 출발점은 철자가 아니라 source IPA 다.
2. 대상 언어의 음소 또는 언어모듈이 지정한 음운 단위마다 하나의 자모값 ID 를 둔다.
3. 자모값은 새로 그리지 않고, 이미 등록된 한글 자모와 훈민정음의 조자 원리를 먼저 쓴다.
4. 하나의 음소는 위치에 따라 서로 다른 위치형을 가질 수 있으나, 자모값 ID 는 하나다.
5. 음절 첫머리의 모음 앞에 발음되는 초성 자음이 없으면 ㅇ 을 두어 훈민정음의 초·중·종성 블록 원리를 지킨다.
6. 하나의 영어 음절은 하나의 음절블록으로 적는다. 초성군과 종성군에는 자모가 둘 이상 올 수 있으며 음소 순서를 보존한다.
7. 어말 자음에 한국어 말음 중화를 적용하지 않는다. /k/→ㅋ, /t/→ㅌ, /p/→ㅍ 를 그대로 적는다.
8. 강세·성조·연결발화 정보는 분절 자모와 구별된 별도 층으로 다룬다.

8.2 자모값과 위치형

하나의 음소가 초성 자리와 종성 자리에서 다른 부호를 쓰는 일은 흔하다. 그러나 그것은 같은 자모값의 다른 얼굴이지 다른 음소가 아니다.

음소	자모값	초성형	종성형
/p/	JV-P	ㅍ	ㅍ
/t/	JV-T	ㅌ	ㅌ
/k/	JV-K	ㅋ	ㅋ
/f/	JV-F	ㅍ	ㅍ
/v/	JV-V	ㅍ	ㅍ
/s/	JV-S	ㅍ	ㅍ (종성군 안에서 음소 순서대로)
/z/	JV-Z	ㅍ	ㅍ
/ʃ/	JV-SH	ㅍ	ㅍ
/ɹ/	JV-R	ㅍ	ㅍ
/l/	JV-L	ㅍ	ㅍ

이 표의 위치형은 언어모듈과 자모값 ID를 통해 해당 자음 음소로 환원한다. 강제·운율과 원철자·출처까지 복원하려면 자모열 외의 구조화된 레코드를 함께 보존해야 한다.

9. 블록 문법

9.1 음절블록의 구조

한 블록 원칙과 손글씨 시범 조자법·글꼴 합성

제2본의 손글씨 정본은 한 영어 음절을 [] 한 쌍으로 묶어 하나의 음절블록으로 적는다. 복합 블록에서는 [] 안에 초성군→음절핵→종성군의 자모를 음소 순서대로 이어 적는다. 이것이 현 단계의 비디지털 손글씨 시범 조자법이며, 블록 수는 영어 음절 수와 일치한다.

제3본의 전산 직렬화는 이 내부 순서를 Unicode 옛한글 자모 연쇄로 보존하는 방식을 검토한다. KS X 1026-1과의 정합성 및 개정 필요 여부도 제3본에서 별도로 다룬다.

앞선 시험 기록은 pick의 합성과 bank·friend·boost·milk·play·task의 분리 표시를 보고했다. 이 관찰만으로 한 glyph 합성을 뜻하지는 않는다. glyph 수·표시 폭·음절블록 수는 구별하며, 환경을 명시한 재계측과 한계는 제3본에서 다룬다.

따라서 손글씨로 한 음절블록의 경계와 내부 순서를 기록하는 규칙은 제2본에 둔다. [] 안의 선형 시범 조자법을 넘어선 최종 압축 손글씨 배치형과, 같은 블록을 디지털 환경에서 하나의 자형으로 자동 합성하는 규칙·구현은 미결로 두며 후자는 제3본 전산본에서 다룬다.

음절블록은 초성군·음절핵·종성군의 세 자리로 이루어진다. 초성군에는 그 음절의 앞 자음이 모두, 음절핵에는 모음 요소가 모두, 종성군에는 뒤 자음이 모두 음소 순서대로 들어간다.

초성 중성 종성
ㅍ +] + ㅓ = 펍

9.2 초성이 없는 음절

모음으로 시작하는 음절에는 무음 초성 ㅇ을 둔다. 복모음·장모음의 뒤 모음에도 읽기용 무음 ㅇ을 붙인다(2026-09-17 저자 결정). 뒤 모음은 같은 음절핵의 요소이며 새 음절이 아니다. 이 표시용 ㅇ은 제3본의 저장 V 배열에는 더하지 않는다.

단어	source IPA	정본
ask	/æsk/	애스ㅋ
about	/ə'baʊt/	으 바우ㅌ
boy	/bɔɪ/	보이
food	/fu:d/	푸우드

9.3 C + /l/ 초성군

자음 뒤에 /l/이 오는 어두 자음군에서는 /l/을 ㄹ 한 번으로 적어 같은 음절블록의 초성군에 둔다. 앞 자음과 /l/의 음소 순서를 그대로 보존하며, 별도 중복 표시는 두지 않는다.

단어	source IPA	정본	구조
play	/pleɪ/	프 ㄹ ㅂ이	초성군 프·ㄹ ·핵 ㅂ+ ㅣ
blue	/blu:/	ㅂ ㄹ ㅡ우	초성군 ㅂ·ㄹ ·핵 ㅡ+ ㅡ
close	/klaʊz/	크 ㄹ ㅓ우 ㅈ	
flat	/flæt/	플 ㄹ ㅌ ㅌ	

따라서 play /pleɪ/는 초성군 프·ㄹ, 음절핵 ㅂ+ ㅣ로 분석하며 정본 시범 표기는 [프 ㄹ ㅂ이]이다.

9.4 종성 /s/ 규칙

1. 영어 /s/ 가 coda 위치에 놓이면 자모값은 ㅅ 계열의 종성값이다.
2. 앞 모음의 받침으로 합치지 않고, 손글씨 정본에서도 종성군 안에서 음소 순서에 맞추어 ㅅ을 독립하여 적는다.
3. 이 규칙은 /s/ 하나에만 적용하며 다른 마찰음·장애음으로 일반화하지 않는다.
4. 뒤에 다른 종성이 이어지면 같은 종성군 안에 음소 순서대로 이어 적는다.
5. 단, 등록된 /t+s/ 복합값 ㅌ 이 먼저 인식되는 환경에서는 복합값 규칙이 우선한다.

단어	source IPA	정본	금지형
pass (GA)	/pæs/	패 ㅅ	팻
bus	/bʌs/	버 ㅅ	븃
boost	/bu:st/	부우 ㅅ ㅌ	부웃 ㅌ
task	/tɑ:sk/	타아 ㅅ ㅋ	타앗 ㅋ

패스을 뺏처럼 합쳐 보이면 한국어 종성 중화에 따라 /t/로 읽힐 위험이 있다. 따라서 /s/는 같은 음절블록 안의 독립 요소로 드러낸다. Unicode 정준 합성 자체가 원 부호열의 복원 불가를 뜻하지는 않는다.

9.5 초성군과 종성군

한 음절의 자음은 모두 같은 블록의 초성군 또는 종성군에 들어간다. 자음군이라도 블록을 나누지 않으며 음소 순서를 바꾸지 않는다.

단어	source IPA	정본	블록
bank	/bæŋk/	뱅크	[뱅크]
milk	/mɪlk/	밀크	[밀크]
friend	/frend/	ᄇᆞᆫᆫᆞᆫ	[ᄇᆞᆫᆫᆞᆫ]
text	/tekst/	텍스트	[텍스트]

10. ㄹ 계열의 다중 대응

한국어 ㄹ에 가깝게 들릴 수 있는 영어 소리를 /l/, [ɫ], /ɹ/, [r]의 네 표기값으로 구별한다. Broad 음소와 Narrow 이름은 층위를 나누어 읽는다.

영어	자모	성격	예
/l/	ㄹ · ㄴ	치경 설측 접근음	light /laɪt/ 라이 _ㅌ
[ɫ]	ㄹ	연구개화 설측음 (dark L) · Narrow	milk [mɪɫk] 밀 _ㄹ
/ɹ/	어 · ㄹ	치경 접근음	right /raɪt/ ㄹ아이 _ㅌ
[r]	ㄹ	치경 탄설음 · GA Narrow	water ['wɔɹə] ㄹ로

light와 right를 보라. 현행 외래어 표기에서는 두 단어가 모두 ‘라이트’로 적힌다. 훈민바름에서는 라이_ㅌ와 ㄹ아이_ㅌ으로 첫 글자부터 갈린다.

최소대립쌍	현행 표기	훈민바름
light · right	라이트 · 라이트	라이 _ㅌ · ㄹ아이 _ㅌ
lice · rice	라이스 · 라이스	라이 _ㅌ · ㄹ아이 _ㅌ
collect · correct	컬렉트 · 커렉트	크렉 _ㅌ · 크렉 _ㅌ

11. /s/의 위치형

/s/는 하나의 자모값을 유지하면서 환경에 따라 세 위치형을 쓴다.

환경	부호	예
/u:, ʊ/ 앞 초성	ㅅ	soon /su:n/ 수우 _ㄴ
그 밖의 모음 앞 초성	ㅆ	sun /sʌn/ 썬
자음군 안	ㅅ	sky /skaɪ/ 스카이
종성	ㅅ (음소 순서대로)	pass /pæs/ 패ㅅ

이 위치형은 한국어 독자의 발음 근사를 돕기 위한 시범 선택이다. ㅅ·ㅆ을 영어 /s/와 음성적으로 동일하다고 보지는 않는다. 위치형이 달라도 자모값 ID는 같고, 역규칙에서는 /s/로 해석한다.

12. /t+s/ 복합값 츠

영어의 /tʃ/와 /t+s/는 비슷하게 들릴 수 있지만 구조가 다르다. /tʃ/는 하나의 파찰음 음소이고, /t+s/는 /t/와 /s/ 두 음소가 연속해서 나타나는 자음군이다. 훈민바름은 이 차이를 없애지 않고 정본에서 구별한다.

따라서 /tʃ/ 한 음소는 ㄷ으로 적고, /t+s/ 두 음소의 연쇄는 등록 복합값 츠으로 적는다. 츠은 두 음소를 한 자리에서 간결하게 표시하지만, 그 안의 /t/와 /s/라는 음소 순서는 그대로 보존한다.

그러므로 역변환할 때 ㄷ은 /tʃ/로, 츠은 /t/ + /s/로 되돌아간다. 예를 들어 catch /kætʃ/와 cats /kæts/는 소리 구조가 다르므로 훈민바름에서도 각각 캐ㄷ과 캐츠으로 명확히 구별된다.

	음소 수	자모	예
/tʃ/	1 (파찰음)	ㄷ	catch /kætʃ/ 캐ㄷ
/t+s/	2 (파열음 + 마찰음)	츠	cats /kæts/ 캐츠

복합값 츠은 종성군의 한 자리로 적으며, 종성 /s/ 규칙보다 우선한다.

단어	source IPA	정본
cats	/kæts/	캐츠
hits	/hɪts/	히츠
waits	/weɪts/	웨이츠
abates	/əˈbeɪts/	으베이츠
boots	/buːts/	부우츠

역변환은 츠을 /t/ + /s/ 두 음소로 되돌린다. 츠은 두 음소를 한 자리에서 간결하게 표시하지만, /t/와 /s/의 음소 순서는 그대로 보존한다.

13. Broad 와 Narrow

Broad 는 음소 대립을 기록하는 기본 정보이다. Narrow 는 이음 정보를 별도 속성으로 추가하며 Broad 를 덮어쓰지 않는다.

Narrow 기능	환경	자형	상태
무기 파열음	/s/ 뒤 [p t k]	ㅍ·ㅌ·ㄱ	시범
GA 탄설음	모음 사이 /t/ → [ɾ]	ㄹ	잠정 고정 · 글꼴 시험 필요
dark L	[ɫ]	ㄹ	잠정 고정
성문파열	[ʔ]	미정	미결
음절자음	[ŋ ɹ] 등	자모 + syllabic 속성	시범
동화·약화	연결발화 환경	표면형 필드	개별 규칙

Narrow는 Broad 정보 위에 실제 발음에서 나타나는 세부 조음 정보를 덧붙이는 층이다. 모든 세부 현상을 한꺼번에 적용하는 하나의 고정 형식이 아니라, 파열음에서 숨이 동반되는 정도, GA 탄설음, dark L, 성문파열, 음절자음, 약화, 동화, 연결발화 등 필요한 항목을 각각 선택하여 기록할 수 있다. 여러 항목을 함께 적용한 종합형은 구현 단계에서 하나의 묶음으로 제공할 수 있으나, 이는 정보의 필수 규칙이 아니라 사용 편의를 위한 선택사항이다.

14. 강세·장단·성조

영어에서 강세는 단순히 어떤 음절을 ‘세게 읽는가’를 표시하는 부가 정보가 아니다. 현대 표준 한국어에서는 영어와 같은 방식의 어휘 강세가 광범위한 음운 대립을 이루지 않지만, 영어에서는 한 낱말 안에서도 특정 음절이 다른 음절보다 두드러지게 발음되며, 그 차이가 모음의 길이와 음질, 음절의 약화, 말의 리듬과 음높이 움직임에까지 영향을 준다.

특히 강세를 받지 않는 음절의 모음은 원래의 모음값을 그대로 유지하지 않고 /ə/와 같은 약 모음으로 줄어드는 경우가 많다. 따라서 영어의 발음을 정확하게 되살리려면 자음과 모음의 종류뿐 아니라 어느 음절에 강세가 놓이는가도 함께 알아야 한다. 강세 위치가 잘못되면 개별 음소를 모두 맞게 발음하더라도 낱말 전체가 부자연스럽게 들리거나 다른 말처럼 인식될 수 있다.

강세는 낱말의 문법적 기능과 함께 달라지기도 한다. 예를 들어 일부 영어 낱말은 같은 철자를 쓰면서도 명사와 동사에서 강세 위치가 달라진다. record는 명사에서는 앞 음절에, 동사에서는 뒤 음절에 주강세가 놓이는 것이 대표적이다. 이처럼 철자가 같아도 강세가 달라지면 실제 발음 구조가 달라질 수 있으므로, 정밀한 발음 표기에서는 강세를 생략할 수 없다.

훈민바름은 이러한 운율 정보를 분절 자모와 뒤섞지 않고 별도의 구조 속성으로 보존한다. source IPA의 'ˈ'는 원문 정보로 그대로 남기고, 정본에서는 어느 음절이 제1강세·제2강세·무강세인지를 해당 음절블록의 운율 속성으로 기록한다. 따라서 같은 자모값을 유지하면서도 강세 위치를 별도로 보존하고 역으로 확인할 수 있다.

손글씨와 화면 표시에서는 이 구조 정보를 음절블록의 상대적인 크기 차이로 나타낼 수 있다. 강세가 강한 음절을 크게, 약한 음절을 작게 적으면 IPA의 강세기호를 따로 해독하지 않아도 낱말의 리듬이 글자의 크기만으로 직관적으로 드러난다. 훈민바름은 이처럼 발음의 분절적 차이뿐 아니라 강약의 구조까지 시각적으로 읽을 수 있는 표기를 지향한다.

제1강세 핵은 100%, 제2강세 핵은 100%, 무강세 핵은 75%로 표시한다. 같은 단어에 제2강세가 있을 때에만 제1강세 핵을 125%로 키운다. 뒤 모음 요소는 50%, 독립 자음은 75%이며 강세 배율을 중첩하지 않는다.

요소	정본 처리	표시 예
제1강세	별도 운율 속성	100% (제2강세가 있는 단어만 125%)
제2강세	별도 운율 속성	100% 크기
무강세	기본	75% 크기
장모음	음절핵의 두 모음 요소 + length 속성	뒤 요소 50% 크기

요소	정보 처리	표시 예
성조	언어모듈별 속성	영어 모듈에서는 미사용

분절음이 같을 때 강세만 달라지면 자모열은 같고 운율 속성만 달라진다. 분절음 검색과 강세 조건 검색을 구별하며, 강세를 포함한 역변환에는 운율 정보가 필요하다.

강세와 억양은 서로 관련되지만 같은 정보는 아니다. 강세는 낱말이나 발화 안에서 어떤 음절이 두드러지는지를 나타내고, 억양은 발화 전체에서 음높이가 어떻게 움직이는지를 나타낸다. 영어 참조모듈에서는 우선 강세와 장단을 구조화하고, 성조 언어에서는 같은 운율층을 언어모듈별 성조 정보로 확장할 수 있다. 억양과 연결발화의 세부 표시는 현재 정보의 분절 자모와 분리하여 별도의 운율·표면형 층에서 다룬다.

글자크기로 보는 영어강세

훈민바름은 영어의 강세를 부가적인 발음기호로 숨겨 두는 것이 아니라, 음절블록의 구조 정보로 보존하고 글자 크기의 차이로 직접 보이게 할 수 있다.

15. 어휘형과 연결발화형

사전에 실리는 형태(어휘형)와 실제 말소리(연결발화형)는 다르다. 훈민바름은 둘을 같은 층에 섞지 않는다.

층	내용	예
어휘형	사전 표제어의 Broad 정본	want to → 원 ㅌ · 투우
연결발화형	실제 발화의 표면형	wanna → 아 ·

연결발화형은 어휘형을 대체하지 않고 표면형 필드로 따로 기록한다. 학습자가 사전형과 실제 발화를 함께 볼 수 있게 하는 것이 목적이다.

16. 잠정 고정·시범·미결 목록

상태 표시

잠정 고정 — 현재 가표본의 기준값으로 사용하며, 이후 공개 검토와 검증에 따라 개정될 수 있다.

시범 — 원리 또는 필요성은 인정하되, 구체적 자형·경계·수치 등을 실제 적용과 사용자 시험을 위해 임시로 사용한다.

미결 — 복수의 합리적 대안이 남아 있어 최종 판정을 유보하고 비교·검증할 연구과제로 공개한다.

잠정적으로 고정된 항목 목록

항목	비고 · 적용 범위
자음 자모값	현재 RP·GA 참조모듈의 기본 음소-자모 대응값으로 사용한다. 위치에 따른 초성형·중성형은 같은 자모값의 위치형으로 본다.
공통 모음 자모값	RP·GA에서 공통으로 사용하는 기본 모음값. 특히 /ɪ/ = ㅣ, /u/ = 우, /ə/ = ·의 독립 대립을 보존한다.
1 영어 음절 = 1 훈민바름 음절블록	자모 수와 관계없이 원어의 음절 경계를 보존한다. 자음군·복합 음절핵도 같은 영어 음절이면 한 블록에 속한다.
복모음 구성요소 1:1	복모음을 하나의 한국어식 모음으로 환원하지 않고 source IPA의 두 모음값과 순서를 보존한다. 뒤 중성에는 읽기용 무음 초성 ㅇ을 붙인다.
복합 블록의 경계 [] 와 내부 음소 순서	손글씨에서는 한 영어 음절의 범위를 [] 로 표시하고 초성군→음절핵→중성군의 순서를 보존한다. 최종 압축 배치형은 미결이다.
중성 /s/의 독립 표기	/s/를 앞 모음의 현대 한글 받침으로 흡수하지 않고 원어 음소 순서에 따라 중성군의 독립 요소로 적는다.
/t+s/ 복합값 ㅈ	/tʃ/ 한 음소 ㅈ과 /t/+s/ 두 음소 연쇄를 구별한다. 역변환에서는 ㅈ → /t+s/로 되돌린다.
음절 첫머리의 무음 초성 ㅇ	발음되는 초성이 없는 음절 첫머리에 자리표지 ㅇ을 둔다. 뒤 중성의 읽기용 ㅇ도 무음이지만 새 음절을 뜻하지 않으며 저장용 초성군에는 더하지 않는다.
/j/ = ㅈ	/j/ 자체의 기본 자모값. /j/+V를 실제로 어떻게 조합할지는 미결이다.
/w/ = ㅍ	/w/ 자체의 기본 자모값. /w/+V의 합용·조자방식은 미결이다.
[r] = ㄹ	GA Narrow에서 탄설음 [r]을 나타내는 값. Broad /t,d/를 보존하고 표면 발음층에 추가한다.
[ɫ] = ㄹ	어말·자음 앞 dark L의 Narrow 값. Broad /l/의 자모값과 구별된 표면형이다.
rhotic vowel = vowel + /ɹ/	GA의 r-색채 모음을 별도 하나의 자모값으로 통째로 만들지 않고 모음값과 /ɹ/을 분리해 보존한다. 단 /ɜ/의 기저 모음 선택은 미결이다.

시범적으로 적용된 항목 목록

항목	비고 · 적용 범위
뒤 모음의 입기용 무음 초성 ㅇ	저자 결정(2026-09-17): 부우ㅅㅌ처럼 뒤 모음에 무음 ㅇ을 붙인다. 같은 핵의 요소이며 음소·음절 수와 저장 자모값은 바꾸지 않는다.
초성 /s/ 위치형의 모 음 경계	현재 /u:, ʊ/ 앞에서는 ㅅ, 그 밖의 모음 앞에서는 ㅅ을 쓰는 구별을 시험한다. 이 경계의 최종 범위는 사용자·지각 시험이 필요하다.
GA /ɔ/ = ㅓ	GA THOUGHT/CLOTH 계열의 /ɔ/를 RP /ɒ, ɔ:/와 구별하기 위한 현재 참조값. cot-caught merger 등 GA 내부 변이는 별도 모듈 판정이 필요하다.
/s/ 뒤 Narrow [p·t·k]의 ㅍ·ㅌ·ㅋ	숨이 거의 따라 나오지 않는 영어 [p·t·k]를 Broad ㅍ·ㅌ·ㅋ과 시각적으로 구별하기 위한 시험 표기다. 한국어 된소리와 동일하다는 뜻은 아니다.
음절자음	[ŋ, l] 등의 음절성을 자모값과 별도의 syllabic 속성으로 기록하는 방식을 시험한다. 세부 표시 방식은 아직 확정하지 않는다.
강세 크기값	제1·제2강세 핵 100%, 무강세 핵 75%. 제2강세가 있는 단어만 제1강세 125%를 적용한다. 세부 배치와 지각 효과는 후속 검증 대상이다.

미결 항목 목록

항목	비고 · 적용 범위
/j/+V 결합표기 및 역변환 규칙	ㅈ+모음 명시형, 기존 중성 활용형, 조자원리 확장형 등 복수 후보가 있어 판정을 유보한다.
/w/+V 결합표기 및 역변환 규칙	ㅍ+모음 명시형과 기존 합용중성 활용형 등을 비교 중이다. /wɪ, wʊ/ 등에서 기존 한글 중성과 음소 대립이 맞지 않는 문제가 있다.
반모음 결합의 Broad/Narrow 또는 정본/표시 차등화	음소값 보존형과 독자에게 익숙한 표시형을 서로 다른 층으로 분리할지 결정하지 않았다.
[ʔ] 자형	성문파열음의 Narrow 값은 필요성을 인정하지만 사용할 확장 한글 자형을 아직 배당하지 않았다.
/ɜ/의 normalized 기 저 모음	GA NURSE의 /ɜ/를 ㅓ+/ɪ/, ㅓ+/ɪ/ 또는 별도 모음값으로 분석할지 판정하지 않았다. §16.2의 공개 연구과제다.
복합 블록의 최종 압 축 손글씨 배치형	[] 안의 내부 음소 순서는 정했지만 여러 초성·중성·종성을 하나의 압축된 손글씨 자형으로 어떻게 배치할지는 정하지 않았다.
다중 초성·중성·종성 연쇄의 글꼴 합성	Unicode 부호열과 글꼴·셰이핑 환경에서 복합 블록을 자동 합성하는 방식은 제3본의 전산 과제로 남긴다.
뒤 중성의 세부 배치와 압축 자형	무음 ㅇ을 붙이는 입기 원칙은 저자가 결정했다. 같은 음절블록 안의 세부 배치와 최종 압축 자형은 후속 연구 대상으로 남긴다.

16.1 반모음 결합표기 — 미결 과제와 후보안

영어의 /j/와 /w/는 각각 ㅈ과 ㅍ을 기본 자모값으로 삼으며, 이 기저값은 현재 가표본에서 잠정 고정한다. 그러나 반모음이 모음과 결합할 때에는 음소값과 역변환의 명시적 보존, 한글 사용자의 가독성과 직관성, 훈민정음의 조자 원리가 서로 다른 표기방식을 제안할 수 있다.

현재 /j/+V에는 ① ㅈ과 해당 모음값을 각각 드러내는 음소값 보존형, ② ㅈ·ㅈ·ㅈ·ㅈ·ㅈ 등 익숙한 중성을 활용하는 직관형, ③ ㅈ→ㅈ, ㅈ→ㅈ 등 중성의 재출 관계를 현대적으로 확장 적용하는 조자원리 확장형을 후보로 둔다. 예를 들어 /jʊ/는 ㅈ+ㅈ, ㅈ, ㅈ의 여러 후보가 존재하며, 각각 음소 대립의 보존성과 즉독성에서 장단점이 다르다.

/w/+V 역시 ㅍ과 모음값을 보존하는 명시형과 ㅍ·ㅍ·ㅍ·ㅍ·ㅍ 등 합용중성을 활용하는 직관형을 비교할 수 있다. 다만 /wɪ/, /wʊ/ 등 주요 환경에서 기존 합용중성과 훈민바름의 모음 대립이 일치하지 않는 문제가 있어 현재 단계에서는 결합표기를 정하지 않는다.

따라서 /j/+V와 /w/+V의 결합표기는 모두 미결 과제로 둔다. 향후 음소 대립과 역변환 가능성, 한글 독자의 즉독성, 손글씨 구성, 조자 원리와의 정합성, RP·GA 및 다른 영어 방언에서의 적용성, 디지털 구현과 사용자 시험을 함께 비교한다. Broad/Narrow 층의 차등 적용과 정본·표시 프로파일의 분리 역시 해결 후보로 검토한다.

16.2 /ɜ:/ — 이 문서가 아직 답하지 못한 것

GA 의 NURSE 집합(nurse·bird·work)은 r-색채 모음 /ɜ:/ 를 가진다. 훈민바름은 rhotic 모음을 모음 + /ɹ/ 로 적는다는 원칙을 세웠고, 약모음 쪽 /ə/ 는 · + ㅅ으로 잠정 고정했다.

그러나 강세를 받는 /ɜ:/ 의 기저 모음을 무엇으로 볼 것인가, ·로 볼 것인가, ·로 볼 것인가, 별도의 자모값을 세울 것인가. 세 선택지가 각각 RP /ɜ : / 와의 대응 관계를 다르게 만든다.

GA NURSE는 완성 출력을 기대하는 양성 벡터에서 제외한다. 대신 source 보존과 미결 반환을 확인하는 부정·미결 벡터에 포함한다. 미결을 임의의 모음으로 채우지 않는다.

제2부 · 영어 음소 해례

영어 음소 해례의 목적은 표를 외우게 하는 데 있지 않다. 글자를 보아 입술·혀·입천장·목청의 움직임을 떠올리고, 손으로 쓴 정본을 다시 실제 소리로 되살릴 수 있게 하는 데 있다. 현대 음성학의 조음 위치·조음 방법·유성성·모음 공간을 사용하되, 설명은 한국어 화자가 거울과 측각으로 확인할 수 있도록 풀어 쓴다.

이 부의 RP와 GA는 각각 영국식·미국식 영어 전체를 대표한다는 뜻이 아니라, 현재 훈민바름 영어 모듈이 채택한 참조 발음체계다. 지역·세대·화자에 따른 변이는 별도 언어모듈과 Narrow 속성으로 관리한다. Australian English, Canadian English 등도 이후 독립 프로파일로 추가될 수 있다.

현대 해례체의 지위

이 부의 해례문은 『훈민정음』 해례본 원문을 인용한 것이 아니라, 그 설명 방식을 본떠 영어 음소의 조음 원리와 훈민바름 자모값을 밝힌 현대 의작이다.

현재 가표본의 기준값은 앞 장의 대응표와 상태표에 있으며, 해례는 그 값을 손과 입으로 이해하게 하는 설명층이다.

17. 발음기관과 해례 용어

해례 용어	주된 자리	뜻
양순	두 입술	두 입술을 맞대거나 가까이하여 낸다.
순치	아랫입술·잇니	아랫입술을 잇니 끝에 가볍게 대어 마찰시킨다.
치·치경	혀끝·잇나·잇잇몸	혀끝이나 혀날을 잇니 또는 잇잇몸에 대거나 가까이 둔다.
후치경	혀날·잇잇몸 뒤	혀날을 치경보다 뒤로 물려 넓은 통로를 만든다.
경구개	혀 앞·센입천장	혀 앞부분을 센입천장 가까이 올린다.
연구개	혀 뒤·여린입천장	혀 뒤를 올려 여린입천장과 막거나 가까이한다.
성문	목청	성대 사이의 열고 닫힘과 마찰로 낸다.
유성·무성	목 울림	유성은 목청이 울리고, 무성은 울리지 않는다.
파열·마찰·파찰	공기 흐름	완전히 막았다 터뜨리거나, 좁혀 비비거나, 막은 뒤 곧 마찰시킨다.
비음·접근음	코·넓은 통로	비음은 코로 공기를 보내고, 접근음은 마찰이 나지 않을 만큼 넓힌다.

18. 자음 총론

자음은 공기의 길을 어디에서, 어떤 방식으로 막거나 좁히는가에 따라 갈린다. 짝을 이루는 무성음과 유성음은 입과 혀의 자리는 같고 목청의 울림만 다르다. 한국어 화자는 자음 뒤에 — 나 으를 덧붙이지 말고, 원어의 자음군과 종성 순서를 그대로 지켜야 한다.

확인법 한 손을 목 앞에 대어 울림을 느끼고, 얇은 종이를 입 앞에 대어 바람의 세기와 지속을 본다. 코를 가볍게 막아 비음인지 확인하고, 거울로 입술의 닫힘과 둥글림을 살핀다.

19. 파열음

/p/

ㅍ

무성 양순 파열음

해례. 두 입술을 굳게 맞대어 숨을 잠시 가두었다가 한꺼번에 열어 내는 소리라. 목청은 울리지 아니한다.

현대 풀이. 영어 /p/는 특히 강세 음절의 첫머리에서 숨을 동반해 실현되는 경우가 많으며, /s/ 뒤에서는 그 숨이 크게 줄어든다. 종성에서는 터프림이 약하거나 들리지 않을 수 있으나 음소값은 /p/로 남는다.

입으로 익히기. 입술을 닫은 뒤 손바닥을 입 앞에 두고 pin의 첫소리에서 바람을 느껴 본다. spin에서는 바람이 훨씬 약해야 한다.

대표 예. pin /pɪn/, spin /spɪn/, cap /kæp/

/b/

ㅂ

유성 양순 파열음

해례. 내는 자리는 ㅍ과 같되 목청을 울려 소리를 더하는 소리라. 두 입술을 닫았다 열며, 울림이 ㅍ과의 차이를 이룬다.

현대 풀이. 영어 /b/는 특히 어말에서 완전한 유성이 약해질 수 있지만 Broad 음소는 /b/다.

입으로 익히기. 목을 만지고 bin의 첫소리에서 울림을 느껴 본다. pin과 번갈아 말해 입술 자리는 같고 목 울림만 달라지는지 확인한다.

대표 예. bin /bɪn/, cab /kæb/

/t/

ㅌ

무성 치경 파열음

해례. 혀끝이나 혀날을 윗잇몸에 대어 숨길을 막았다가 떼어 내는 소리라. 목청을 울리지 아니한다.

현대 풀이. 특히 강세 음절의 첫머리에서는 숨을 동반해 실현되는 경우가 많고, /s/ 뒤에서는 숨이 거의 따라 나오지 않는 [t̪]로 실현된다. GA의 모음 사이에서는 Narrow에서 [ɾ]로 실현될 수 있다.

입으로 익히기. 혀끝을 윗니가 아니라 그 뒤의 잇몸에 댄다. top과 stop을 비교하고, GA water에서는 한 번 텅기는 [ɾ]를 따로 느낀다.

대표 예. top /tɒp/, stop /stɒp/, cat /kæt/; GA water ['wɔɹəɪ]

/d/

ㄷ

유성 치경 파열음

해례. 나는 자리는 ㄷ과 같되 목청을 올리는 소리라. 혀끝이 윗잇몸을 막았다가 열리며 소리가 난다.

현대 풀이. GA에서는 강세 모음과 약세 모음 사이의 /d/도 [ɾ]에 가까워질 수 있다.

입으로 익히기. 목을 만진 채 den과 ten을 번갈아 말한다. 혀의 자리는 같고 올림만 달라야 한다.

대표 예. den /dɛn/, ladder /'lædər/, bad /bæd/

/k/

ㅋ

무성 연구개 파열음

해례. 혀 뒤를 여린입천장에 붙여 숨길을 막았다가 떼어 내는 소리라. 목청은 올리지 아니한다.

현대 풀이. 특히 강세 음절의 첫머리에서는 숨을 동반해 실현되는 경우가 많고, /s/ 뒤에서는 숨이 거의 따라 나오지 않는 [k]로 실현된다. 종성에서는 터프림이 약하거나 들리지 않아도 /k/의 음소값은 유지된다.

입으로 익히기. 혀 뒤쪽이 여린입천장에 닿아 막히는 느낌을 살핀다. key와 ski의 바람 차이를 비교한다.

대표 예. key /ki:/, ski /ski:/, back /bæk/

/g/

ㄱ

유성 연구개 파열음

해례. 나는 자리는 ㄱ과 같되 목청을 올리는 소리라. 혀 뒤가 여린입천장을 막았다가 열린다.

현대 풀이. 어말에서는 올림이 약해질 수 있으나 /k/와의 대립은 앞모음 길이와 폐쇄의 성질에도 드러난다.

입으로 익히기. go와 coat의 첫소리를 비교하며 목의 올림을 확인한다.

대표 예. go /gou/, bag /bæg/

20. 파찰음과 복합 자음값

/tʃ/

ㅈ

무성 후치경 파찰음

해례. 혀날로 윗잇몸 뒤를 막았다가 곧 좁은 틈으로 풀어 마찰시키는 소리라. 막음과 마찰이 한 음소로 이어진다.

현대 풀이. /t/와 /ʃ/를 따로 끊어 내는 것이 아니라 하나의 파찰음으로 낸다. 훈민바름 ㅈ는 /t+s/의 복합값 ㅈ과 구별한다.

입으로 익히기. chin의 첫소리를 길게 늘일 수 없고, 막힌 뒤 곧 /ʃ/ 같은 마찰이 이어지는지 확인한다.

대표 예. chin /tʃɪn/, match /mætʃ/

/dʒ/

ㄷʒ

유성 후치경 파찰음

해례. 내는 자리는 ㄷ과 같되 목청을 울리는 소리라. 막았다가 마찰시키는 두 동작이 한 소리로 이어진다.

현대 풀이. /tʃ/의 유성 짝이다. 종성 자형은 같은 HMBR 자모값의 위치형으로 다룬다.

입으로 익히기. 목을 만지고 gin과 chin을 번갈아 말한다.

대표 예. gin /dʒɪn/, judge /dʒʌdʒ/

/t+/s/

ㅌㅅ

치경 파열음+치경 마찰음
복합값

해례. 먼저 ㅌ의 막음을 이루고 곧바로 ㅅ의 마찰을 잇되, 두 음소의 차례를 잃지 않는 소리라. 이를 /tʃ/ 한 음소와 섞지 아니한다.

현대 풀이. 영어 어말 /ts/는 두 음소이나 훈민바름에서는 등록된 복합 자음값 ㅌㅅ으로 쓸 수 있다. 역변환에서는 반드시 /t/와 /s/ 두 음소로 되돌린다.

입으로 익히기. cats 끝을 catch와 비교한다. cats는 혀끝의 /t/ 뒤에 앞쪽 /s/ 마찰이 이어지고, catch는 더 뒤쪽의 /tʃ/ 한 소리다.

대표 예. cats /kæts/, hits /hɪts/, abates /əˈbeɪts/

21. 마찰음

/f/

ㅍㅍ

무성 순치 마찰음

해례. 아랫입술을 윗니 끝에 가볍게 대고 그 사이로 숨을 길게 흘려 마찰시키되 목청은 울리지 아니하는 소리라.

현대 풀이. 병서형 ㅍ을 기본으로 하고 연서형 ㅍ을 대안으로 둔다.

입으로 익히기. 입 앞의 종이가 계속 흔들리게 하되 목에는 울림이 없어야 한다. 입술을 완전히 닫으면 /p/가 되므로 삼간다.

대표 예. fan /fæn/, leaf /li:f/

/v/

ㅍㅍ

유성 순치 마찰음

해례. 내는 자리는 ㅍ과 같되 목청을 울려 소리를 더하는 소리라.

현대 풀이. 병서형 ㅍ이 기본이고 ㅍ은 역사적 대안형이다. ㅍ의 글리프가 병처럼 보이면 매핑이 아니라 글꼴을 점검한다.

입으로 익히기. 목을 만지고 fan과 van을 번갈아 말한다.

대표 예. van /væn/, very /ˈveri/, leave /li:v/

/θ/

ㅅ

무성 치 마찰음

해례. 혀끝을 윗니 사이 또는 윗니 뒤에 가볍게 두고 그 좁은 틈으로 숨을 흘리
되 목청은 울리지 아니하는 소리라.

현대 풀이. 혀를 세게 물지 않고 공기가 끊기지 않게 한다. /s/처럼 혀를 잇몸 안
에 숨기지 않는다.

입으로 익히기. 거울로 혀끝이 아주 조금 보이는지 확인하고 종이를 흔들어 본
다.

대표 예. thin /θɪn/, bath /bæθ/

/ð/

ㅅ

유성 치 마찰음

해례. 내는 자리는 ㅅ과 같되 목청을 울리는 소리라.

현대 풀이. 기능어에서 약하고 짧게 날 수 있으나 /d/로 막아 버리지 않는다.

입으로 익히기. 목을 만지고 thin과 then을 비교한다.

대표 예. then /ðɛn/, this /ðɪs/, breathe /briːð/

/s/

ㅅ·ㅅ·

중성군 ㅅ

무성 치경 마찰음

해례. 혀끝을 윗잇몸 가까이에 두고 가운데로 가는 좁은 홈을 내어 날카로운 바
람을 흘리는 소리라. 하나의 음소가 환경에 따라 여러 위치형을 취한다.

현대 풀이. /uː, ʊ/ 앞 ㅅ, 그 밖의 모음 앞 ㅅ, 자음 앞 ㅅ을 쓴다. 중성 /s/는 같은
블록 안에 음소 순서대로 독립 표기한다. 이 위치형 경계는 시범이다.

입으로 익히기. 이를 다물고 길게 sss를 내어 공기가 혀 가운데로 흐르는지 본
다. /ʃ/보다 혀가 앞에 있고 입술 둥글림이 적다.

대표 예. soon /suːn/, sun /sʌn/, sky /skaɪ/, pass /pæs/

/z/

ㅅ

유성 치경 마찰음

해례. 내는 자리는 /s/와 같되 목청을 울리는 소리라.

현대 풀이. /s/의 유성 짝이다. 어말에서는 울림이 일부 약해져도 마찰의 지속
과 앞모음 길이가 대립을 돕는다.

입으로 익히기. 목을 만지고 sip과 zip을 번갈아 말한다.

대표 예. zoo /zuː/, zip /zɪp/, buzz /bʌz/

/ʃ/

ㅅ

무성 후치경 마찰음

해례. 혀날을 윗잇몸 뒤로 물리고 넓은 틈으로 숨을 흘리며 입술을 조금 둥글리
는 소리라.

현대 풀이. /s/보다 조음점이 뒤이며 마찰음의 음색이 낮다.

입으로 익히기. shhh 하며 조용히 하라는 소리를 내고, sss와 혀의 앞뒤를 비교
한다.

대표 예. ship /ʃɪp/, wish /wɪʃ/

/ʒ/
ㅅ

유성 후치경 마찰음

해례. 내는 자리는 ㅅ과 같되 목청을 올리는 소리라.

현대 풀이. 영어에서는 단어 가운데에 흔하며 어두에는 드물다.

입으로 익히기. measure의 가운데를 길게 늘여 /ʒ/와 같은 자리에서 목이 울리는지 확인한다.

대표 예. measure /'meʒər/, vision /'vɪʒən/

/h/
ㅎ

무성 성문 마찰음

해례. 목청을 열어 숨을 내보내되 뒤따르는 모음의 입모양을 미리 갖추어 내는 소리라.

현대 풀이. 고정된 구강 조음점보다 뒤 모음에 의해 모양이 달라지는 성문 마찰음이다.

입으로 익히기. 손바닥에 따뜻한 숨을 내쉬며 hat, heat, who에서 뒤 모음에 따라 입모양이 달라지는지 본다.

대표 예. hat /hæt/, heat /hi:t/, who /hu:/

22. 비음

/m/
ㅁ

유성 양순 비음

해례. 두 입술을 닫고 입길을 막되 여린입천장을 내려 공기를 코로 보내며 목청을 올리는 소리라.

현대 풀이. 입술은 /b/와 같은 자리에 있으나 공기가 코로 흐른다.

입으로 익히기. 코를 막으면 소리가 막히는지 확인한다.

대표 예. man /mæn/, team /ti:m/

/n/
ㄴ

유성 치경 비음

해례. 혀끝을 윗잇몸에 대어 입길을 막고 공기를 코로 보내며 목청을 올리는 소리라.

현대 풀이. 혀 자리는 /d/와 같고 비강 통로만 열린다.

입으로 익히기. 코를 막은 채 no를 말하면 소리가 막혀야 한다.

대표 예. no /nou/, ten /ten/

/ŋ/

ㅇ

유성 연구개 비음

해례. 혀 뒤를 여린입천장에 붙이고 입길을 막은 채 공기를 코로 보내는 소리라.

현대 풀이. 영어에서는 주로 종성에 오며 sing의 끝에 원어에 없는 /g/를 덧붙이지 않는다. finger처럼 실제 /g/가 있는 단어는 별도다.

입으로 익히기. sing을 길게 늘여 혀끝이 아니라 혀 뒤가 막히는지 느낀다.

대표 예. sing /sɪŋ/, singer /ˈsɪŋər/, finger /ˈfɪŋɡər/

23. 접근음과 ㄹ 계열

/l/

ㄹ

유성 치경 설측 접근음

해례. 혀끝을 윗잇몸에 대되 혀의 양옆으로 공기를 흘려 내는 소리라.

현대 풀이. 음절 앞의 clear L과 종성의 dark L은 Broad에서는 /l/ 하나이며, [ɫ]은 Narrow에서 ㄹ로 구별한다. C+/l/에서도 /l/은 ㄹ 한 번으로 초성군에 적으며 앞 자음 뒤의 음소 순서를 그대로 보존한다.

입으로 익히기. 혀끝을 잇몸에 댄 채 양옆으로 공기가 흐르는지 확인한다.

대표 예. light /laɪt/, feel /fi:l/, play /pleɪ/

/ɹ/

ㄹᄇ. ㄹᄃ

유성 후치경 접근음

해례. 혀끝이나 혀날을 입천장에 닿지 않게 뒤로 모으거나 말아 올려, 가운데의 넓은 길로 울림을 내는 소리라.

현대 풀이. 영어 /ɹ/은 한국어 ㄹ처럼 혀가 닿아 튀지 않는다. onset은 ㄹᄇ, rhotic coda는 ㄹᄃ을 쓴다. RP의 비권설 위치에서는 철자 r가 있어도 surface /ɹ/이 없을 수 있다.

입으로 익히기. 혀끝이 어디에도 닿지 않은 채 red를 말하고, led와 비교한다.

대표 예. red /red/, right /raɪt/; GA car /kɑr/

[ɾ]

ㄹᄇ

유성 치경 탄설음

해례. 혀끝이 윗잇몸을 단 한 번 빠르게 스치고 곧 떨어지는 짧은 소리라.

현대 풀이. GA에서 강세 모음과 약세 모음 사이의 /t, d/가 흔히 [ɾ]로 실현된다. 이는 Broad /t, d/를 지우지 않는 Narrow 자형이다.

입으로 익히기. 한국어 모음 사이 ㄹ처럼 한 번만 가볍게 튕긴다. 길게 붙이면 /l/이 되고 막으면 /d/가 된다.

대표 예. GA water [ˈwɔtəɹ], city [ˈsɪɹi]

[ɫ]

르

유성 연구개화 설측 접근음

해례. 혀끝은 ㄹ의 자리에 두고 혀 뒤를 함께 올려 어두운 울림을 더하는 소리라.

현대 풀이. 영어의 종성·자음 앞 /l/에서 흔한 dark L의 Narrow 표시다.

입으로 익히기. feel 끝에서 혀끝은 잇몸에 두되 혀 뒤가 올라가 소리가 어두워지는지 느낀다.

대표 예. feel [fi:ɫ], milk [mɪɫk]

/w/

oo

유성 순-연구개 접근음

해례. 입술을 둥글게 오므리고 혀 뒤를 올렸다가 곧 뒤 모음으로 풀어 가는 소리라.

현대 풀이. /u/와 비슷한 입모양이지만 스스로 음절핵을 이루지 않고 다음 모음으로 빠르게 이동한다.

입으로 익히기. woo의 첫부분을 아주 짧게 하고 모음으로 즉시 넘어간다.

대표 예. we /wi:/, wood /wud/

/j/

o

유성 경구개 접근음

해례. 혀 앞을 센입천장 가까이 올렸다가 곧 뒤 모음으로 풀어 가는 소리라.

현대 풀이. 한국어의 반모음 y와 비슷하나 독립 모음 /i/가 아니다.

입으로 익히기. yes의 첫소리를 길게 끌지 말고 다음 모음으로 즉시 잇는다.

대표 예. yes /jes/, young /jʌŋ/

24. 모음 총론

모음은 공기의 길을 완전히 막지 않고 혀의 높이와 앞뒤 위치, 입술의 둥글림, 턱의 벌어짐, 길이와 긴장도로 갈린다. 모음표에서 위쪽은 혀가 높고 아래쪽은 혀가 낮으며, 왼쪽은 혀가 앞에 있고 오른쪽은 뒤에 있다. 같은 철자라도 RP와 GA의 음소값이 다를 수 있으므로 반드시 선택된 언어모듈의 IPA를 먼저 따른다.

손으로 익히는 법 거울로 턱의 벌어짐과 입술 둥글림을 보고, 한 모음에서 다른 모음으로 혀를 움직이며 차이를 느낀다. 단모음은 목표 자리에서 머물고, 복모음은 한 음절 안에서 시작점에서 끝점으로 이동한다.

25. 공통 핵심 모음

/ɪ/

근폐·근전설·평순 모음

해례. 혀를 입 앞쪽으로 올리되 /i:/만큼 높고 팽팽하게 하지 아니하며, 입술을 둥글리지 않고 짧고 느슨하게 내는 소리라.

현대 풀이. KIT 모음이다. /i:/를 단순히 짧게 한 것이 아니라 혀 높이와 긴장도가 다르다.

입으로 익히기. ship과 sheep을 번갈아 말하며 /ɪ/에서 턱과 혀가 조금 더 낮고 느슨한지 본다.

대표 예. ship /ʃɪp/, sit /sɪt/, pick /pɪk/

/i:/

폐·전설·평순 장모음

해례. 혀 앞을 높이 올리고 입술을 펴며 소리를 길고 팽팽하게 잇는 소리라.

현대 풀이. FLEECE 모음이다. 길이뿐 아니라 높은 혀 위치와 긴장된 음질이 중요하다.

입으로 익히기. 이를 살짝 드러내고 sheep을 길게 말하되 /j/를 앞에 덧붙이지 않는다.

대표 예. sheep /ʃi:p/, seat /si:t/

/ɛ/

중개·전설·평순 모음

해례. 혀 앞을 중간보다 조금 낮게 두고 턱을 열며 입술을 펴서 내는 소리라.

현대 풀이. DRESS 모음이다. 한국어 ㅔ보다 더 열릴 수 있다.

입으로 익히기. bed와 bad를 비교해 /ɛ/에서 턱이 /æ/보다 덜 벌어지는지 본다.

대표 예. bed /bed/, pet /pet/

/æ/

ㅏ

근개·전설·평순 모음

해례. 혀 앞을 낮추고 턱을 넓게 벌리되 입술을 둥글리지 않는 소리라.
현대 풀이. TRAP 모음이다. 한국어 ㅏ보다 대체로 더 낮고 입이 더 열린다.
입으로 익히기. 거울로 입이 넓게 벌어지는지 보고 man과 men을 비교한다.
대표 예. man /mæn/, pat /pæt/

/ʊ/

ㅜ

근폐·근후설·원순 모음

해례. 혀 뒤를 올리고 입술을 둥글리되 /u:/보다 힘을 덜고 짧고 느슨하게 내는 소리라.
현대 풀이. FOOT 모음이며 정본 자모값은 ㅜ이다. /u:/의 길이만 줄인 소리가 아니다.
입으로 익히기. full과 fool을 비교해 /ʊ/에서 혀와 입술의 긴장이 덜하고 소리가 짧은지 본다.
대표 예. foot /fʊt/, good /ɡʊd/

/u:/

ㅜㅜ

폐·후설·원순 장모음

해례. 혀 뒤를 높이 올리고 입술을 둥글려 소리를 길게 잇는 소리라.
현대 풀이. GOOSE 모음이다. 실제 화자에 따라 앞쪽으로 이동할 수 있으나 참조모듈에서는 /u:/ 대립을 보존한다.
입으로 익히기. fool을 길게 말하며 입술 둥글림과 긴장을 유지한다.
대표 예. food /fu:d/, soon /su:n/

/ə/

•

중설 약화 모음

해례. 혀와 입술에 힘을 빼고 입 한가운데에서 짧고 가볍게 나는 소리라.
현대 풀이. schwa는 주로 무강세 음절에 나타나며 주변 소리에 따라 음질이 달라진다.
입으로 익히기. about의 첫 음절을 세게 읽지 말고 아주 짧고 중앙에서 낸다.
대표 예. about /ə'baʊt/, control /kən'trəʊl/

/ʌ/

ㅏ

중개·중설~후설·평순 모음

해례. 혀를 낮고 가운데 또는 뒤쪽에 두고 입술을 둥글리지 않으며 힘주어 내는 소리라.
현대 풀이. STRUT 모음이다. RP와 GA에서 정확한 혀 위치는 다를 수 있으나 /ə/보다 강세가 있고 분명하다.
입으로 익히기. cup의 모음을 about의 schwa보다 분명하고 크게 낸다.
대표 예. cup /kʌp/, sun /sʌn/

26. RP의 모음 체계

/ɒ/



개·후설·원순 단모음

해례. 혀를 낮고 뒤에 두며 입술을 둥글려 짧게 내는 소리라.
현대 풀이. RP LOT 모음이다. GA에서는 흔히 /ɑ/와 대응한다.
입으로 익히기. RP hot에서 입술을 둥글리고 짧게 낸다.
대표 예. RP hot /hɒt/, lot /lɒt/

/ɔ:/



중개·후설·원순 장모음

해례. 혀 뒤를 중저 높이에 두고 입술을 둥글려 소리를 길게 이어 내는 소리라.
현대 풀이. RP THOUGHT 계열의 장모음이다. /ɒ/보다 혀가 높고 길다.
입으로 익히기. cot류의 /ɒ/와 caught류의 /ɔ:/를 길이와 음질로 대비한다.
대표 예. RP thought /θɔ:t/, port /pɔ:t/

/ɑ:/



개·후설·평순 장모음

해례. 혀를 낮고 뒤에 두며 입술을 둥글리지 않고 입을 넓게 열어 길게 내는 소리라.
현대 풀이. RP PALM·START 계열이다. 비권설 RP에서는 뒤의 철자 r가 별도 /ɹ/로 나지 않는다.
입으로 익히기. park를 길게 말하되 한국어 아보다 혀를 뒤로 둔다.
대표 예. RP palm /pɑ:m/, park /pɑ:k/

/ɜ:/



중설·평순 장모음

해례. 혀를 입 가운데에 두고 입술을 둥글리지 않은 채 길게 이어 내는 소리라.
현대 풀이. RP NURSE 모음이다. 비권설이므로 /ɹ/를 덧붙이지 않는다.
입으로 익히기. bird를 말할 때 혀끝을 말아 올리지 않고 중앙 모음을 길게 유지한다.
대표 예. RP bird /bɜ:d/, nurse /nɜ:s/

/ɛ:/



중개·전설~중설·평순 장모음

해례. 혀 앞을 중저 자리에 두고 입술을 펴며 길게 내는 소리라.
현대 풀이. 본 훈민바름 RP 참조모들은 SQUARE 계열을 /ɛ:/로 정리한다. 다른 RP 분석의 /eə/는 별도 모듈 변이로 관리한다.
입으로 익히기. square에서 첫 모음의 열린 음질을 유지하되 별도 /r/를 붙이지 않는다.
대표 예. RP square /skwɛ:/, shared /ʃɛ:d/

27. GA의 모음 체계

/ɑ/

ㅏ

개·후설·평순 모음

해례. 혀를 낮고 뒤에 두고 입술을 둥글리지 않아 내는 소리라.

현대 풀이. GA LOT·PALM 계열의 참조값이다. rhotic 환경에서는 뒤의 /ɹ/를 별도 자모값으로 잇는다.

입으로 익히기. GA hot을 말하며 입술을 둥글리지 않는다.

대표 예. GA hot /hɑt/, father /ˈfɑðər/

/ɔ/

ㅓ

중개·후설·원순 모음

해례. 혀 뒤를 중저 자리에 두고 입술을 둥글려 내는 소리라.

현대 풀이. cot-caught 대립을 유지하는 GA 모듈에서 THOUGHT 모음으로 쓰는 시범값이다. 대립이 없는 모듈에서는 /ɑ/와 합쳐질 수 있다.

입으로 익히기. cot과 caught를 구별하는 화자라면 caught에서 입술을 더 둥글리고 혀를 높인다.

대표 예. GA dog /dɔg/, caught /kɔt/

GA의 rhotic vowel 처리

GA의 rhotic vowel 가운데 /ɹ/는 source IPA를 보존한 뒤 normalized IPA 단계에서 /əɹ/로 분해한다. /ɜ/는 source IPA를 보존하되 기저 모음 선택이 아직 미결이므로 별도 모음 자모값을 확정하지 않는다.

source IPA /ˈwɒtɹ/ → normalized IPA /ˈwɒtəɹ/ → 훈민바름 정본의 순서다. 따라서 /ɹ/에는 별도의 모음 자모값을 두지 않고 /ə/ + /ɹ/의 결합으로 처리한다. /ɜ/ 역시 현재 가표본에서는 별도 모음 자모값을 확정하지 않은 상태다.

/ɹ/ → /əɹ/은 현재 기준이며, 강세를 받는 /ɜ/의 기저 모음 선택은 미결이다.

28. 복모음 총론: 한 음절 안의 이동

복모음은 두 음절을 나란히 읽는 것이 아니라 한 음절의 핵 안에서 혀와 입술이 시작점에서 끝점으로 움직이는 소리다. 구성요소의 자모값과 순서를 보존한다. 뒤 모음에는 읽기용 무음 초성 ㅇ을 붙인다(2026-09-17 저자 결정). 두 모음 요소는 함께 하나의 영어 음절핵을 이루며, 표시용 ㅇ은 음소·음절을 더하지 않는다. 아래 표는 표시용 ㅇ을 제외한 핵의 음가 구성이다.

모들	IPA	조음 이동	정본 핵	대표 단어
공통	/eɪ/	중전설 → /ɪ/	ㅈㅣ + ㅅ	say
공통	/aɪ/	저설 → /ɪ/	ㅈㅣ + ㅅ	bite
공통	/ɔɪ/	후설 원순 → /ɪ/	ㅈㅓ + ㅅ※	boy
공통	/aʊ/	저설 → /ʊ/	ㅈㅣ + ㅜ	how
RP	/əʊ/	중설 → /ʊ/	· + ㅜ	low
GA	/oʊ/	중후설 원순 → /ʊ/	ㅈㅓ + ㅜ	low
RP	/ɪə/	/ɪ/ → schwa	ㅣ + .	near
RP	/ʊə/	/ʊ/ → schwa	ㅜ + .	pure

※ 핵의 자모값은 그대로 보존하며 읽기용 조자법은 § 2.3과 § 16을 따른다.

29. 공통 복모음 해례

/aɪ/

ㅏ + ㅣ

저설→전설 상승 복모음

해례. 혀를 낮은 자리에서 시작하여 앞쪽의 /ɪ/ 자리로 올리며 입을 닫아 가는 소리라.

현대 풀이. 첫 요소는 모듈과 화자에 따라 앞·중앙·뒤로 조금 달라질 수 있으나 끝점 /ɪ/와의 이동을 보존한다.

입으로 익히기. I를 길게 말하며 입이 크게 열렸다가 좁아지는지 본다.

대표 예. bite /baɪ/, my /maɪ/, time /taɪm/

/ɔɪ/

ㅓ + ㅣ

후설 원순→전설 복모음

해례. 혀 뒤의 둥근 모음에서 시작하여 앞쪽 /ɪ/로 움직이며 입술의 둥글림을 풀어 가는 소리라.

현대 풀이. 현재 RP·GA 참조모듈에서는 /ɔɪ/의 첫 요소를 공통값 ㅓ로 적고, 끝점 /ɪ/는 ㅣ로 보존한다.

입으로 익히기. boy를 느리게 말하며 입술이 둥글었다가 펴지는지 본다.

대표 예. boy /bɔɪ/, choice /tʃɔɪs/

/eɪ/

ㅓ + ㅣ

전설 상승 복모음

해례. 혀 앞을 중간 높이에서 시작하여 /ɪ/ 자리로 올리며 턱을 조금 닫아 가는 소리라. 두 자리는 하나의 핵으로 이어진다.

현대 풀이. 첫 요소는 /ɛ/보다 높고 끝 요소는 완전한 /i:/가 아니라 느슨한 /ɪ/ 쪽이다. 정본은 ㅓ와 ㅣ의 값을 모두 보존한다.

입으로 익히기. say를 천천히 늘여 턱이 닫히고 혀가 올라가는 움직임을 느낀다. 세이처럼 두 음절로 끊지 않는다.

대표 예. say /seɪ/, late /leɪt/, day /deɪ/

/aʊ/

ㅏ + ㅜ

저설→후설 원순 상승 복모음

해례. 혀를 낮은 자리에서 시작하여 뒤쪽 /ʊ/로 올리고 입술을 점차 둥글리는 소리라.

현대 풀이. 끝 요소는 /u:/가 아니라 /ʊ/이므로 정본은 ㅜ를 쓴다.

입으로 익히기. how를 늘여 입이 열렸다가 둥글어지는 움직임을 거울로 본다.

대표 예. how /haʊ/, mouth /maʊθ/, now /naʊ/

30. RP 복모음 해례

/əʊ/

· + ㅜ
중설→후설 원순 복모음

해례. 힘을 뺀 가운데 모음에서 시작하여 뒤쪽 /ʊ/로 움직이며 입술을 둥글리는 소리라.

현대 풀이. GA /oʊ/와 시작점이 다르다. 편의형 ㅜ우로 합치지 않고 schwa · 와 /ʊ/ ㅞ의 1:1 값을 보존한다.

입으로 익히기. RP low를 천천히 말하며 첫소리가 ㅜ보다 중앙에서 시작하는지 느낀다.

대표 예. RP low /ləʊ/, go /gəʊ/, home /həʊm/

/ɪə/

! + .
전설→중설 중심 복모음

해례. 앞쪽의 느슨한 /ɪ/에서 시작하여 힘을 빼고 가운데 schwa로 옮겨 가는 소리라.

현대 풀이. RP NEAR 계열이다. GA에서는 대개 모음 + /ɪ/ 연쇄로 대응한다.

입으로 익히기. near를 비권설로 말하며 끝에서 혀를 말아 /r/를 만들지 않고 중앙으로 풀어 준다.

대표 예. RP near /nɪə/, here /hɪə/

/ʊə/

ㅞ + .
후설 원순→중설 중심 복모음

해례. 뒤쪽의 둥근 /ʊ/에서 시작하여 입술의 힘을 풀고 가운데 schwa로 옮겨 가는 소리라.

현대 풀이. RP CURE 계열의 참조값이다. 현대 화자에서는 다른 모음과 합쳐질 수 있으나 현재 모듈에서는 독립값을 둔다.

입으로 익히기. pure를 비권설로 말하며 첫 둥글림을 풀어 중앙 모음으로 끝낸다.

대표 예. RP pure /pjʊə/, cure /kjʊə/

31. GA 복모음과 vowel+/ɪ/

/ou/

ㅏ + ㅜ

중후설 원순→/u/ 복모음

해례. 혀 뒤의 중간 높이 등근 모음에서 시작하여 더 높고 느슨한 /u/ 자리로 옮기는 소리라.

현대 풀이. RP /əʊ/보다 시작점이 뒤쪽이고 둥글다. 정본은 ㅏ와 ㅜ를 보존한다.

입으로 익히기. GA low를 천천히 말하며 처음부터 입술이 둥글고 혀가 뒤에서 시작하는지 본다.

대표 예. GA low /lou/, go /gou/, home /houm/

GA의 NEAR·SQUARE·CURE 대응

GA에서는 RP의 /ɪə, eɪ 또는 eə, uə/에 대응하는 항목을 대개 모음 + /ɪ/ 연쇄로 분석한다. 훈민바름은 별도 rhotic vowel 자모를 만들지 않고 선택된 모음 정보 뒤에 /ɪ/ 자모값 ㅙ·ㅚ를 잇는다. 원문 사전이 /ɜ/를 쓰면 source IPA는 그대로 보존하고, normalized IPA 단계에서 /əɪ/로 분해한다. /ɜ/는 source IPA를 보존하되 기저 모음 선택이 아직 미결이므로 § 16.2의 연구과제로 둔다.

32. RP와 GA의 조음 차이 요약

항목	RP 참조모들	GA 참조모들
rhoticity	모음 뒤 /ɹ/을 일반적으로 실현하지 않음. linking R 등은 surface layer.	모음 뒤 /ɹ/을 실현하며 vowel+/ɹ/로 정본화.
LOT	/ɒ/	/ɑ/
THOUGHT	/ɔ:/	/ɔ/ 또는 cot-caught merged /ɑ/ 모들
GOAT	/əʊ/	/oʊ/
NEAR·CURE	/ɪə, uə/	vowel + /ɹ/
NURSE	/ɜ:/	모들 기저모음 + /ɹ/; /ɜ/ 원문은 보존
flapping	기본 참조규칙 아님	/t,d/의 [ɾ] Narrow 가능

제7장에서 본 RP·GA의 차이는 조음 관점에서도 같은 다섯 갈래로 묶인다. rhoticity, GOAT의 시작점, BATH의 어휘집합 분기, LOT·CLOTH·THOUGHT 계열의 후설모음 재배치, 그리고 RP의 NEAR·CURE 및 SQUARE와 GA의 vowel+/ɹ/ 대응이다. 이 표는 그 차이를 빠르게 되짚는 요약이며, 개별 음가의 조음은 앞 절의 해례를 따른다.

33. 발음 점검과 대립 연습

해례는 설명으로 끝나지 않고 대립을 통해 입과 귀에 새겨야 한다. 아래 단어쌍은 같은 자리를 유지하면서 한 요소만 바꾸어 연습한다. 지역 모듈에 따라 일부 대립은 합쳐질 수 있으므로 RP·GA 선택을 먼저 확인한다.

연습 축	단어쌍·연쇄
무성·유성	fan-van /f-v/ · thigh-thy /θ-ð/ · sip-zip /s-z/
조음점	sip-ship /s-ʃ/ · ten-ken /t-k/ · light-right /l-r/
비음	sin-sing /n-ŋ/ · sum-sun /m-n/
고모음	ship-sheep /ɪ-i:/ · full-fool /ʊ-u:/
전설 저모음	bed-bad /ɛ-æ/
RP 후설	cot-caught /ɒ-ɔ:/ · cut-cart /ʌ-ɑ:/
GA 후설	cot-caught /ɑ-ɔ/ (대립 모듈에서만)
복모음 이동 (GA)	late-light /eɪ-aɪ/ · boy-buy /ɔɪ-aɪ/ · boat-bout /oʊ-au/
파찰·연쇄	catch-cats /tʃ-t + s/

이 부의 잠정 내용

복모음은 /eɪ, aɪ, ɔɪ, aʊ/, RP /əʊ, ɪə, ʊə/, GA /oʊ/를 포함한다. 각 복모음은 하나의 영어 음절핵이며 구성요소의 자모값을 보존한다. 뒤 모음에 읽기용 무음 ɹ을 붙여도 음소·음절은 늘지 않는다. GA의 rhotic 항목은 vowel + /ɹ/로 분석하되 /ɜ/의 기저 모음은 미결이다.

34. 맏음말 — 훈민바름 정본의 강점

이 책에서 제시한 훈민바름 정본의 목적은 영어 발음을 한글과 비슷한 소리로 대략 옮기는 데 있지 않다. 그보다 중요한 목표는 원어의 음운 차이를 가능한 한 잃지 않으면서, 한글의 자모와 음절블록이라는 익숙한 구조 안에서 사람이 직접 적고 다시 읽을 수 있게 하는 것이다. 이 책에서 살펴본 영어 RP·GA 참조모듈은 그 가능성을 시험하는 첫 사례다.

첫째, 훈민바름은 현용 한글 표기에서 하나로 합쳐지기 쉬운 소리의 차이를 문자 차이로 보존한다. /ɪ/와 /i:/, /ʊ/와 /u:/, /l/과 /ɹ/, /tʃ/와 /t+s/처럼 영어에서 구별되는 음가는 서로 다른 자모값이나 구조로 나타난다. RP와 GA처럼 같은 영어 안의 서로 다른 발음체계도 하나로 평준화하지 않는다. car, go, pass, hot, near에서 확인했듯 source IPA가 달라지면 그 차이가 정본에도 드러난다. 즉 훈민바름의 강점은 발음 차이를 설명하는 데 그치지 않고, 그 차이를 실제로 눈에 보이는 글자의 차이로 남길 수 있다는 데 있다.

둘째, 훈민바름은 원어의 음절구조를 보존한다. 영어 한 음절을 훈민바름 한 음절블록에 대응시키고, 그 안에 초성군·음절핵·종성군을 음소 순서대로 담는다. 따라서 friend, bank, boost처럼 여러 자음이 이어지는 낱말도 한국어식으로 임의의 모음을 삽입하거나 원래 하나인 음절을 여러 음절로 나눌 필요가 없다. 복모음과 장모음 역시 하나의 음절핵 안에서 구성요소와 길이를 보존한다. 한글의 음절블록 원리를 버리는 것이 아니라 그 조자 범위를 원어의 음절구조에 맞게 확장하는 것이 정본의 기본 발상이다.

셋째, 하나의 고정된 ‘영어 발음’을 강요하지 않고 언어모듈을 통해 발음체계의 다양성을 수용할 수 있다. 이 책에서는 첫 참조범위로 RP와 GA를 나누었지만, 같은 원리는 Australian, Canadian, Irish, Scottish English 등 다른 영어 발음체계에도 적용할 수 있다. 나아가 각 언어의 음운체계와 source IPA, 실제 화자 검토가 갖추어진다면 다른 언어 역시 독립 모듈로 연구할 수 있다. 훈민바름 정본은 특정 영어 억양 하나를 한글로 고정하는 체계가 아니라 어떤 원어의 어떤 발음을 기록하는지를 먼저 밝히고 그 차이를 보존하는 구조를 지향한다.

넷째, 사전적 음소와 실제 말소리를 한 층에 뒤섞지 않는다. Broad 정본은 음소 대립을 안정적으로 기록하고, Narrow에서는 GA 탄설음, dark L, 성문파열, 음절자음, 약화·동화와 같은 실제 발음의 세부 정보를 필요에 따라 덧붙일 수 있다. 어휘형과 연결발화형 역시 분리한다. 따라서 기준이 되는 발음 구조를 잃지 않으면서도 실제 화자의 발화에 가까운 정보를 단계적으로 추가할 수 있다.

다섯째, 훈민바름은 자음과 모음뿐 아니라 강세와 같은 운율 정보도 글자로 직관화할 가능성을 가진다. 영어에서는 강세가 모음의 음질과 약화, 리듬, 낱말의 인식에 큰 영향을 주며 같은 철자라도 품사에 따라 강세 위치가 달라지기도 한다. 훈민바름은 이를 자모 자체와 섞지 않고

별도 운율 속성으로 보존하면서, 손글씨와 화면에서는 음절블록의 상대적인 크기 차이로 나타내는 방식을 시험한다. 이는 IPA의 강세기호를 읽지 못하는 사용자도 낱말의 강약 구조를 글자의 크기로 직접 볼 수 있게 하는 교육적 가능성을 보여 준다.

여섯째, 정본은 디지털 기술에 종속되지 않는다. 현재 판정이 끝난 범위에서는 손과 종이와 IPA 만으로도 원리를 적용할 수 있다. 동시에 자모값 ID, 음소 순서, 음절블록, 운율 속성과 같은 구조를 갖기 때문에 이후 전산본에서는 Unicode 직렬화·검색·역변환·글꼴 합성으로 확장할 수 있다. 사람이 읽고 쓰는 정본을 먼저 세우고, 기계는 그것을 보존하고 처리하도록 하는 순서다.

이러한 특징을 하나로 묶으면 훈민바름 정본의 핵심은 ‘더 많은 글자를 쓰는 것’이 아니라 ‘더 많은 원어 정보를 잃지 않는 것’이라고 할 수 있다. 현용 외래어 표기가 한국어 문자생활에서 통일된 관용 표기를 제공하는 데 목적이 있다면, 훈민바름은 교육·사전·학술·접근성·언어기술과 같이 원어의 발음 구조를 더 정밀하게 남겨야 하는 영역에서 그와 다른 역할을 시험한다.

다만 이 문서에서 제시한 모든 조자법과 자모값이 최종적으로 확정된 것은 아니다. 복합 음절 핵의 최종 조자법, 반모음 결합, 일부 Narrow 자형, 압축 블록 배치와 글꼴 합성 등에는 여전히 시범 또는 미결 항목이 남아 있다. 그럼에도 이 제2본이 보여 주려 한 것은 완성된 규범의 선언이 아니라, 한글의 역사적 자원과 조자 원리를 이용해 원어의 음운구조를 훨씬 더 세밀하게 보존하는 정본이 실제로 구성될 수 있다는 가능성이다.

우리 글로 적고, 원래 소리로 되살린다. 그것이 이 해례본이 시험한 훈민바름 정본의 출발점이다.

부록 A. 손글씨 작성 규칙

1. 옛한글 자모는 현대 완성형으로 바꾸지 말고 자모 자체의 형태를 쓴다.
2. ㅍ+ ㅣ+ ㅅ은 한 블록으로 적는다. 외형이 현대 픽과 비슷해도 중성은 ㅣ가 아니라 ㅣ이다.
3. 초성군과 종성군의 자모는 같은 음절블록 안에 음소 순서대로 둔다. 자음군이라도 블록을 나누지 않는다.
4. 복모음·장모음의 뒤 모음에는 읽기용 무음 초성 ㅇ을 붙인다(2026-09-17 저자 결정). 부우ㅅㅌ의 우는 같은 음절블록의 뒤 중성 요소이며 새 음절이 아니다. 표시용 ㅇ은 음소 대응과 저장 부호열에 더하지 않는다.
5. C+/l/에서는 /l/을 ㄹ 한 번으로 적어 같은 음절블록의 초성군에서 앞 자음 뒤에 둔다.
6. 강세는 자모형을 바꾸지 않고 선택된 운율 표시방식으로만 나타낸다.

부록 B. 현재 가표본 판정 요약

판정	내용
/v/	ㅁ 기본, ㅅ 대안
/f/	ㅍ 기본, ㅍ 대안
종성 /s/	같은 블록 안에서 음소 순서대로 독립 표기
/t+s/	복합 자음값 ㅅ; /tʃ/의 ㅅ과 구별
/u/	ㅜ U+118D 자형
[ɹ]	ㄹ
rhotic vowel	별도 모음값 없이 vowel + /ɹ/
영국 모듈	현 단계 RP 유지; MLE는 독립 모듈로 연구·검토한다는 방침만 잠정 고정, 음소목록·사전·참조 화자·매핑은 미확정
[ɫ]	ㅌ (dark L Narrow)
C+/l/	/l/은 초성군에 ㄹ 한 번; 중복 표기 없음
/s/ 위치형	/u:, u/ 앞 ㅅ · 그 밖의 모음 앞 ㅅ · 자음 앞 ㅅ · coda 종성군 ㅅ (경계는 시범)
SQUARE	현 RP 참조모듈 /ɛ:/, /eə/ 분석은 별도 변이로 보존
복모음·장모음 둘 째 중성	구성요소 1:1과 한 음절핵 보존 · 뒤 중성에 읽기용 무음 ㅇ을 붙임(저자 결정) · 세부 배치와 최종 압축 자형은 후속 과제

GA의 t 변이를 구별하는 기준

2026년 10월 12일 개정 부록. 제7.3절과 제13·15절의 Broad와 Narrow 구분을 유지하면서, t 변이의 조음법과 웹앱 지원 범위를 보충한다. / /는 원형 음소 표기, []는 선택한 세부 실현이다. 여기의 예시는 가능한 발음이며 유일한 표준 발음이 아니다.

현상	IPA	조음의 차이
탈락	∅	별도의 t 폐쇄·파열을 생략한다.
성문 대치	[ʔ]	성문에서 흐름을 잠시 제한한다.
성문 강화	[ʔt]	성문 조절과 치경 폐쇄가 함께 있다.
미파열	[t̚]	혀의 폐쇄를 유지하며 뚜렷한 파열을 내지 않는다.
탄설음	[ɾ]	혀끝이 치경에 짧게 한 번 닿는다.

미파열은 탈락이 아니며, [ʔ]도 단순한 무음 표시가 아니다. 실제 성문음화에는 불규칙한 성대 진동과 다양한 폐쇄 정도가 포함된다. 성문 강화와 미파열이 함께 나타날 수도 있다. 앱에서 같은 t에 한 선택만 허용하는 것은 학습형을 분명히 비교하기 위한 인터페이스 정책이지 생리적으로 모든 현상이 배타적이라는 뜻은 아니다.

혀와 목청으로 읽는 조음 안내

/t/의 기본 조음에서는 혀끝·혓날로 윗잇몸 부근을 막았다가 푼다. 탈락 연습에서는 그 별도 동작을 더하지 않고 앞뒤 소리로 이동한다. 미파열 연습에서는 혀의 막음을 유지하되 “트” 같은 모음이나 큰 파열을 덧붙이지 않는다.

[ʔ]는 영어 uh-oh의 가운데처럼 목청에서 짧게 끊기는 느낌으로 연습한다. 목을 세게 조이지 않는다. [n]은 혀의 치경 접촉을 유지하면서 코로 내는 소리가 음절핵이 된 것이다. 별도 모음 없이 한 음절을 이룬다. 이는 훈민정음의 조음 기관 중심 읽기를 활용한 현대 음성학적 설명이며 옛 문헌이 영어 변이 규칙을 직접 규정했다는 주장은 아니다.

[ɾ]는 탈락과 구분한다. GA에서 강세 모음 뒤, 다음 무강세 모음 앞 등의 환경에서 나타날 수 있다. 이 부록은 그 대비를 설명하지만 현재 앱에서 flapping을 자동 생성하지 않는다. [ʔ]의 자형은 여전히 미정이며 [n]의 시범 표기도 정본으로 확정하지 않는다.

적용 조건과 보수적인 추천

환경과 예	선택 가능한 연습형	앱의 추천
자음군 끝 t/d 뒤 자음 next day	t/d 탈락 [nɛks deɪ]	빠른 회화, 강조·쉽 없을 때
n 뒤 t와 다음 무강세 모음 winter /'wɪntəɪ/	t 탈락 ['wɪnəɪ]	GA 빠른 회화, 강조 없을 때
무강세 단어말 tən button /'bʌtən/	['bʌʔn]	GA 보통·빠른 회화
자음 앞 t get me	[gɛʔ mi:] 또는 [getʰ mi:]	[ʔ]은 보통·빠른 회화의 일부 환경
쉽 앞 단어말 t get	[gɛʔ] 또는 [getʰ]	기본 추천 없음, 수동 선택

자음군 탈락은 양옆이 자음일 때 제안하며, 이번 앱은 다음 h/j 앞을 제외한다. n 뒤 t 탈락은 다음 음절에 강세가 없다고 확인할 수 있을 때만 제안한다. 다음 모음에 강세가 있거나 강세 정보가 부족하면 제안하지 않는다. /tən/ 어말은 별도의 성문음화·음절성 비음 선택으로 다룬다.

winter류에서 t가 들리지 않는 결과는 화자에 따라 비음 탄설음이나 n/t 동화로도 분석될 수 있다. 앱의 ∅는 그 모든 미세한 조음을 기술하는 분석이 아니라 비교를 위한 단순화된 표면형이다. button의 [ʔn] 선택은 t 성문음화뿐 아니라 /ə/의 약화·생략과 n의 음절성도 함께 표시한다.

성문음화 후보는 모음·공명음 뒤 t가 자음이나 쉽 앞에 있을 때, 또는 무강세 어말 /tən/에서 제시한다. m/n/l 앞과 button형은 보통 회화에서도 추천하고, 그 밖의 자음 앞은 빠른 회화에서 추천한다. 미파열과 쉽 앞 성문음화는 수동 선택으로 남긴다. 강조한 단어에서는 t 변이를 기본 추천하지 않으며 t 탈락은 차단한다.

이 추천 속도, 우선순위와 범위는 앱의 편집 결정이다. 연구에서 GA 모든 화자에게 동일한 규칙이 입증된 것으로 읽지 않는다. 성문음화가 GA에만 있다는 뜻도 아니다. 이번 자동 후보 생성 범위를 GA로 제한한 것이다.

비교 연습과 참고 근거

한 번에 한 가지 차이를 듣고 만들기

- ① GA에서 button을 변환하고 연음화 버튼을 누른다. 원형 /ˈbʌtən/과 [ˈbʌʔn]을 번갈아 연습한다. 목청에서의 끊김과 모음 없는 [n]을 따로 의식한다. 앱의 훈민바름 옆에 원형 참조가 표시되는 이유도 확인한다.
- ② GA에서 winter를 입력하고 연음화 버튼을 누른 뒤 빠른 회화로 바꾼다. 추천된 t 탈락을 켜다가 끈다. 별도의 /t/ 폐쇄·파열을 추가하는지 비교한다. 다음 모음에 강세를 준 별도 예에는 이 단순화 규칙을 적용하지 않는다.
- ③ get me를 입력하고 연음화를 누른 뒤 [ʔ]과 [t]을 각각 선택한다. 앞 선택은 해제된다. 혀의 폐쇄와 성문 조절이 어떻게 다른지 설명한다. 씬을 설정하면 단어 사이 연결은 끊기지만 씬 앞 t 변이는 수동으로 비교할 수 있다. ④ next day에서 자음군 탈락을 비교한다.

출처와 검토 범위

Seyfarth & Garellek (2020), Physical and phonological causes of coda /t/ glottalization in the mainstream American English of central Ohio, Laboratory Phonology 11(1). 미국 영어의 어말·음절말 t 성문음화와 변이를 다룬다. 중앙 오하이오 Buckeye 자료의 40명 표본이므로 모든 GA 화자의 빈도로 일반화하지 않는다.
pages.ucsd.edu/~mgarellek/files/Seyfarth_Garellek_2020_LabPhon.pdf

Pronuncian, When /t/ sounds like /d/ or a glottal stop. t 실현 조건과 교육 예시. 이 자료의 quick d 표기를 본문에서는 [r]와 구별하여 설명했다. pronuncian.com/when-t-doesnt-sound-like-t/

Baruch College, Deletion. 자음 사이 t/d 탈락과 h 앞 예외. tfcs.baruch.cuny.edu/deletion/ · Mark Liberman의 Language Log 해설은 nt와 음절성 n의 차이 및 비음 탄설음 가능성을 보충한다.
languagelog.ldc.upenn.edu/nll/?p=62989

Education University of Hong Kong, Phonetics and Phonology §3.2는 치경 폐쇄와 음절말 미파열을 설명한다. corpus.eduhk.hk/english_pronunciation/index.php/3-2-acoustic-aspects-of-consonants/ · 국제음성학회 IPA 도표는 ʔ를 가청 파열 없음으로 정의한다. internationalphoneticassociation.org/content/ipa-diacritics

검토일 2026-10-12. 위 문자 본문·IPA를 확인했으며 제공 오디오를 청취 검증하지 않았다. button·winter의 원형은 Collins 영어사전과 대조했다. 일반 사전 전수 검수, 발음 교육 효과 평가, 완전한 Narrow 음성 전사는 이번 갱신의 범위에 포함하지 않는다.

외국어 표기법 연구안

훈민바름 전산본

정본을 잃지 않고 저장하고 나르고 다시 읽는 규격

영어확장판 제3본 · Unicode 직렬화 · 글꼴과 셰이핑 · 국가표준 정합성

2026년 10월 12일 개정판 · 연구·검토용

Park Il Hwan

연구·검토용 비구속 문서

우리 글, 바른 소리

훈민바름 영어확장판 4본의 구성

훈민바름 영어확장판은 하나의 체계를 네 역할로 나눈다. 선언본은 뜻과 공공적 필요를 밝히고, 해례본은 손으로 기록할 수 있는 정본의 구조와 발음 원리를 풀며, 전산본은 정본의 디지털 보존·처리 규격을 세우고, 구현본은 웹 참조구현과 시험으로 실제 작동을 검증한다.

문서	핵심 내용	답하는 질문
제1본 선언본	역사·문화·사회·정책·교육·접근성·AI·표준화와 세계화 비전	왜 필요한가
제2본 해례본	영어 RP·GA 정본, 음가-자모 대응과 손글씨 시범 조자법	어떤 소리를 어떻게 적고 읽는가
제3본 전산본	구조 데이터·Unicode·글꼴·셰이핑·검색·복사·역변환	정본을 어떻게 보존하고 처리하는가
제4본 구현본	웹 참조구현·실제 시험·감사·공개 배포와 오류 제보	규격이 실제로 어떻게 작동하는가

이 문서의 지위와 명칭

‘훈민바름’은 옛한글·확장 한글 자모를 활용한 「외국어 표기법」의 가능성을 연구·개발하는 동안 사용하는 내부 코드명이다. 공적 규범의 명칭을 선점하지 않는다. 현재 훈민바름은 하나의 가표본(working specimen) 연구이자 참조 실험(reference experiment)이며, 이 체계를 그대로 채택하도록 요구하는 규범이 아니다.

첫 참조범위는 영어 RP·GA다. 다른 영어 방언은 독립 프로파일로, 다른 언어는 각자의 원문·source IPA·음운체계와 화자 검토를 갖춘 독립 모듈로 연구한다. 영어의 자모 대응을 다른 언어에 자동으로 상속하지 않는다.

현재 개발·구현 상태

제1-4본은 연구·검토용 판본이다. 이 제3본은 재현 가능한 사전 전수조사와 현재 참조구현 상태를 반영한 교정본이다. 제4본과 digital conversion engine, web app은 계속 개발·시험하며, 규격 제안은 최종 표준 채택이나 전체 구현 완료를 뜻하지 않는다.

문서 근거: [I1] 제1본 2쪽, [I2] 제2본 2·6쪽, [I3] 커버레터 2-3쪽.

이 문서의 읽는 순서와 근거

먼저 제1부에서 무엇을 보존하는지 정하고, 제2부에서 그 구조를 어떤 부호열로 운반하는지 살핀다. 제3·4부는 표시와 정보 회수를 다루며, 제5부는 이 기술 규격을 국가표준 정합성과 개정 제안으로 연결한다. 제6부와 부록은 구현자가 확인할 항목을 제공한다.

표지	이 문서에서의 뜻
잠정 고정	현 가표본의 기준값. 공개 검토와 시험에 따라 개정 가능
시범	비교·사용자 시험을 위한 잠정 형식·수치·절차
미결	최종 판정을 유보한 공동 연구 과제

근거를 구분하는 방식

제1·2본의 정책적 지위와 정본 규칙은 [I1]·[I2]로 인용한다. 이전 전산본과 대화 기록은 변경 경위를 확인하는 내부 자료 [I4]·[I5]로 구분한다. Unicode·OpenType·국가표준 관련 설명에는 외부 문헌 [E1]–[E6]를 붙인다. mappingUnits, 오류 상태와 교환 봉투의 세부 필드는 이번 개정에서 제안하는 전산 설계다.

셰이핑 폭은 2026-09-08 작성 환경의 계측값을 유지한다. 사전 전수조사는 저장소의 data/dict-rp.json·data/dict-ga.json과 tools/cluster_census.mjs로 다시 실행해 §21의 집계가 재현됨을 확인했다. 사전판·원출처·라이선스의 별도 확인이 필요하다는 한계는 그대로 남긴다.

열람 한계

KS X 1026-1:2007은 한국 측이 공개한 영문 번역본의 관련 조항을 확인했다. 번역본은 국문 정식 표준이 아니며 국문판이 최종 권위임을 스스로 밝힌다. 현행 국문판의 조문·개정 이력은 이번 집필에서 확정하지 않았다. 개정 제안은 이 한계를 명시한 연구안이다.

이번 개정의 범위

2026-08-14의 41쪽 전산본이 가진 6부·28절·부록 A–D의 조직을 유지한다. 데이터 모델, 원문 보존, 직렬화와 역변환의 한계, 동치 판정, 계측 근거와 국가표준 논증을 현재 제2본에 맞춰 다시 썼다. 웹앱의 현재 출력과 규격 사이의 차이는 제4본의 감사 대상이다.

차례

1. 4본 안에서 전산본의 지위	5
2. 다섯 층위 모델	5
3. 정본 데이터 모델	6
4. 자모값 등록부 · 자음	9
5. 직렬화의 기본 원칙	12
6. 자리별 직렬화 규칙	13
7. 직렬화 갤러리	14
8. 정규화와 동치	16
9. 금지되는 연쇄와 오류 검출	17
10. 표시층의 지위	18
11. 글꼴 요구사항	18
12. 세이핑 요구사항	19
13. 현재 환경에서 무엇이 합성되는가 · 재측정	20
14. 대체 표시	22
15. 검색	23
16. 복사와 붙여넣기	24
17. 역변환 · 무엇을 되찾는가	25
18. 교환 형식	27
19. 이미 표준 안에 있는 것	28
20. KS X 1026-1에서 확인한 조항	29
21. 훈민바름 연쇄가 막히는 지점	30
22. 두 갈래의 해법과 그 비용	31
23. 개정 제안	32
24. 국제표준과의 관계와 이행 경로	34
25. 적합성 등급	35
26. 시험 벡터와 27. 버전·변경 관리	36
28. 잠정 고정·시범·미결 목록	37
맷음말 · 부호는 정본을 섬긴다	38
부록 A. 복합 음절핵과 코드포인트 식별	39
부록 B. 시험 벡터와 기대 판정	41
부록 C. 측정 방법과 재현 절차	42
부록 D. 용어 대조	44
참고문헌 및 작성 근거 · 내부 자료	45

제1부 · 정보를 계산 가능하게 만드는 층

1. 4본 안에서 전산본의 지위

제2본이 세운 정보는 사람이 손으로 적을 수 있는 자모값과 배열 원리다. Unicode 부호열과 글꼴은 이를 저장하고 나르는 수단이다. 디지털 장치가 어떤 자형을 잘 그린다는 이유로 음가를 바꾸거나, 한 음절의 자음군을 여러 음절로 나누지 않는다.

동시에 제2본이 시범·미결로 남긴 조자법을 전산본에서 확정 규칙으로 승격하지 않는다. 현재 저장형의 무초성 중성 연쇄를 부호열로 시험할 수는 있지만, 그것이 복합 음절핵의 최종 손글씨 조자법을 결정하지는 않는다.

2. 다섯 층위 모델

층	이름	보존·처리하는 정보
1	canonicalModel	음운 토큰·mappingUnits·자모 요소·음절 경계·운율·판정 상태
2	unicodeSerialization	지원되는 정보 구조의 결정적 조합형 자모 연쇄
3	displayRuns	블록별 표시 명령, 상대 크기, 대체 표시 상태
4	displayText	글꼴·조판 환경에 따라 나타나는 파생 문자열
5	fallbackText	합성 불가·미배당 시 제공하는 선형 표기·IPA·설명

다섯 층 앞에는 원문과 source IPA를 보관하는 source 레코드가 있다. source에서 normalized IPA를 만들고 음운 토큰을 분석한 뒤 층 1을 구성한다. 이 변환에는 언어모듈·정규화 규칙·음절분석의 버전과 근거를 남긴다.

같은 입력만으로는 재현 조건이 충분하지 않다. 같은 source, 언어모듈, registry, 정규화·음절분석 규칙과 직렬화 프로파일 버전이 함께 고정되어야 같은 결과를 재현할 수 있다.

층 1과 source가 보존의 권위다. 층 2는 구조의 일부를 옮긴 파생 형식이며, 층 3–5는 이를 읽게 하는 출력이다. 표시 실패는 표시 상태를 바꿀 뿐 source나 음운 분석을 바꾸지 않는다.

[I1] 2·29–30쪽; [I2] 6·8–10·26쪽. 세부 데이터 필드: 본 개정 설계안.

3. 정보 데이터 모델

3.1 원문과 정규화 결과를 함께 보존한다

source IPA는 출처가 제공한 문자열 그대로 보존한다. normalized IPA는 선택된 언어모듈이 매핑 직전에 사용하는 분석형이다. 정규화는 철자를 고치는 작업과 음운 분석을 구분하며, 변경 규칙과 원문상의 위치를 기록한다. Unicode 정규화와 IPA 분석 정규화도 서로 다른 절차다.

예를 들어 GA source /'wɒtə/는 해당 규칙에 따라 normalized /'wɒtə/로 분석한다. source의 /ə/를 삭제하지 않고 변환 이력으로 연결한다. /ɜ/의 기저 모음은 제2본에서 미결이므로 임의로 /ə/나 /ʌ/로 바꾸지 않는다.

3.2 음운 토큰과 자모 요소를 분리한다

phonemeTokens는 선택된 모듈이 인정한 음운 단위의 순서다. mappingUnits는 이 토큰들과 정보 자모값의 대응 관계다. 하나의 음운 단위가 여러 자모 요소를 가질 수 있고, 여러 음소가 하나의 등록 복합 자모값을 가질 수 있다. 따라서 음운 토큰 수와 자모 요소 수를 같다고 가정하지 않는다.

음운 입력	자모값 배열	관계
/p/	[JV-P]	1:1
/u:/	[JV-U, JV-U]	1:N
/eɪ/	[JV-E, JV-KIT]	1:N
/t/ + /s/	[JV-TS]	N:1

각 mapping unit은 입력 token_ids의 순서, 출력 jamo_values의 순서, 적용 rule_id와 판정 상태를 가진다. 장모음은 기존 모음 요소를 반복하고 length=long을 보존한다. 복합값 ʌ은 /t/와 /s/ 두 토큰을 남긴다. 역변환은 이 관계와 등록된 역규칙을 함께 따른다.

3.3 /l/과 무음 초성

play /pleɪ/의 /l/은 같은 초성군에서 /p/ 뒤에 한 번만 기록한다. 초성군 자모값은 [JV-P, JV-L]이다. 모음으로 시작하는 음절의 ɔ은 음소가 아닌 자리표지이며, JV-ZERO와 kind=orthographic_zero로 기록한다. Unicode 채움 문자와 구별한다.

[I2] 7·25-27·30·37·58-59쪽; [I5] Master 3, 137-140쪽. 식별자 철자와 필드 구성은 시범 스키마.

3. 정보 데이터 모델 · 음절과 상태

3.4 블록은 자모가 아니라 선택된 음절분석을 따른다

한 영어 음절은 하나의 훈민바름 음절블록에 대응한다. 각 블록은 onset[], nucleus[], coda[]를 가진다. 배열 원소는 mapping unit과 그 출력 요소의 번호를 가리킨다. 장모음 /u:/의 두 중성 요소는 동일한 음운 토큰에 연결되지만, 각각 구별되는 출력 요소다.

영어의 음절분석은 철자열이나 자모 종류만으로 새로 추정하지 않는다. 출처 사전의 경계, 언어מוד의 분석 규칙 또는 검토자의 판정을 syllabification에 기록한다. 둘 이상의 합리적 분석이 있으면 선택된 분석과 후보를 구분한다. 부호열에서 읽는 경계는 이미 결정된 분석을 회수하는 경계다.

3.5 완성·시범·미결을 데이터로 표현한다

필드	예시 값	용도
status	complete / partial / unresolved	이 레코드의 구조 완성 상태
rule_status	provisional / trial / unresolved	개별 규칙의 잠정 고정·시범·미결
serialization.status	complete / unavailable	평문 전체를 만들 수 있는지
render_status	composed / linear / missing	화면 표시의 상태
issues[]	issue_id, token_ids, reason	실패 위치와 원인을 보존

구조가 유효하더라도 중성 위치형 부호가 미배당이면 평문은 unavailable일 수 있다. 반대로 평문을 생성할 수 있어도 최종 압축 자형은 아직 만들지 못할 수 있다. 미결 항목이 포함된 레코드를 단순히 성공 또는 실패 한 값으로 압축하지 않는다.

3.6 Narrow와 운율

Broad 구조를 먼저 보존하고 Narrow는 underlying token, surface phone, feature_id, 환경 규칙과 함께 별도 층에 추가한다. GA [ɾ]와 dark L [ɫ]은 Broad의 /t,d/와 /l/을 덮어쓰지 않는다. [ʔ] 자형과 음절자음의 최종 처리처럼 미결인 항목에는 완성된 자모열을 만들어 넣지 않는다.

stress는 primary·secondary·none·unknown을 구분한다. 길이는 핵의 음운 속성이다. 제1강세·제2강세 핵은 100%, 무강세 핵은 75%이며 같은 단어에 제2강세가 있을 때만 제1강세를 125%로 표시한다. 강세기호 없는 단음절을 자동으로 primary로 판정하지 않는다.

[I2] 31–37쪽. 상태 필드와 unknown 처리: 본 개정 설계안.

3. 정보 데이터 모델 · 교환 예시

다음은 boost의 핵심 구조를 보이는 유효한 JSON 예시다. 문서 예시이므로 source의 사전 출처는 비워 두며, 실제 사전 레코드에는 부록 C의 출처 정보를 채워야 한다. 버전 문자열은 이 초안의 시범 식별자다.

```
{
  "schema_version": "draft-2026-09-08",
  "module_id": "en-RP", "registry_version": "draft-2026-09-08",
  "source": {"spelling": "boost", "ipa": "/bu:st/",
    "provenance": {"kind": "document-example", "ref": "I2:p9"}},
  "normalized_ipa": "/bu:st/",
  "normalization_rules": [],
  "canonicalModel": {
    "status": "complete",
    "phonemeTokens": [
      {"id": "t1", "ipa": "b"}, {"id": "t2", "ipa": "u:"},
      {"id": "t3", "ipa": "s"}, {"id": "t4", "ipa": "t"}],
    "mappingUnits": [
      {"id": "m1", "token_ids": ["t1"], "jamo_values": ["JV-B"],
        "rule_id": "consonant-b"},
      {"id": "m2", "token_ids": ["t2"],
        "jamo_values": ["JV-U", "JV-U"], "length": "long",
        "rule_id": "long-u", "rule_status": "trial"},
      {"id": "m3", "token_ids": ["t3"], "jamo_values": ["JV-S"],
        "rule_id": "coda-s"},
      {"id": "m4", "token_ids": ["t4"], "jamo_values": ["JV-T"],
        "rule_id": "consonant-t"}],
    "blocks": [{"id": "b1", "stress": "unknown",
      "onset": [{"m1", 0}], "nucleus": [{"m2", 0}, {"m2", 1}],
      "coda": [{"m3", 0}, {"m4", 0}]}]
  },
  "serialization": {"profile": "hmbr-jamo-trial-2026-09-08",
    "status": "complete", "normalization": "NFD",
    "codepoints": [4359, 4462, 4462, 4538, 4544]}
}
```

자리 참조 ["m2", 0]은 m2의 첫 자모 요소를 가리킨다. 두 중성 요소를 모두 소비한 뒤 같은 출력 요소를 다른 자리에 다시 사용하지 않는다.

위 예시의 source와 canonicalModel, serialization을 합친 전체 레코드가 보존 대상이다. serialization만 따로 복사하면 source와 강세·출처가 함께 전달되는 것은 아니다.

4. 자모값 등록부 · 자음

등록부는 문서·구현·시험이 공유하는 자모값 ID와 위치형의 연결표다. 음가-자모 대응은 제2본을 따르고, 전산본은 위치별 Unicode 부호와 처리 상태를 명시한다. 자형만 닮았다는 이유로 부호를 대체하지 않는다.

IPA	자모값 ID	초성형	종성형
/p/	JV-P	U+1111	U+11C1
/b/	JV-B	U+1107	U+11B8
/t/	JV-T	U+1110	U+11C0
/d/	JV-D	U+1103	U+11AE
/k/	JV-K	U+110F	U+11BF
/g/	JV-G	U+1100	U+11A8
/f/	JV-F	U+114B	U+11F4
/v/	JV-V	U+1144	U+11E6
/θ/	JV-TH	U+1145	미배당
/ð/	JV-DH	U+1142	미배당
/s/	JV-S	U+1109 / U+110A	U+11BA
/z/	JV-Z	U+1136	U+D7EF
/ʃ/	JV-SH	U+1140	U+11EB
/ʒ/	JV-ZH	U+1146	미배당

/s/의 초성 위치형은 /u:, ʊ/ 앞 ㅅ, 그 밖의 모음 앞 ㅆ, 자음군 안 ㅅ이라는 현재 시범 경계를 사용한다. 경계의 최종 범위는 사용자·지각 시험 대상이다. 종성형은 U+11BA이며, 표시층에서 한국어식 발음 중화를 적용하지 않는다.

등록부의 상태는 두 축으로 관리한다. 음가-자모값의 판정과 해당 위치형의 Unicode 배당 상태는 서로 다르다. 종성 부호가 미배당이라는 말은 해당 음소나 손글씨 자모값 자체가 없다는 뜻이 아니다.

[I2] 11-12·25·29·36·59쪽; [I4] 9-11쪽. 코드포인트 이름은 부록 A의 처리 원칙에 따라 확인한다.

4. 자모값 등록부 · 자음과 특수값

IPA·기능	자모값 ID	초성형	종성형
tʃ	JV-CH	U+1149	미배당
dʒ	JV-JH	U+1148	미배당
t+s	JV-TS	해당 없음	U+11BE
h	JV-H	U+1112	해당 없음
m	JV-M	U+1106	U+11B7
n	JV-N	U+1102	U+11AB
ŋ	JV-NG	해당 없음	U+11BC
l	JV-L	U+1105	U+11AF
ɹ	JV-R	U+A976	U+D7DB
j	JV-Y	U+1159	해당 없음
w	JV-W	U+1147	해당 없음
[ɾ]	JV-FLAP	U+111B	해당 없음
[ɬ]	JV-DARKL	해당 없음	U+11D0
무음 ◯	JV-ZERO	U+110B	해당 없음

JV-TS는 두 음소 /t+s/에 대응하는 종성 복합값이다. 이 프로파일에서는 초성 자리에 적용하지 않는다. /tʃ/ 한 음소의 JV-CH와 구분한다. 어두 /j/, /w/의 기저값은 보존하되 모음과의 결합·축약형은 미결이므로 명시형 예시를 최종 결합규칙으로 선언하지 않는다.

4.1 미배당 위치형의 처리

/θ, ð, ʒ, tʃ, dʒ/의 종성은 이 초안에 계승한 등록부에서 전용 부호를 확정하지 않았다. 구조형은 해당 음소·자모값·coda 위치를 그대로 저장한다. 평문 전체는 unavailable로 두고, 누락 음소와 이유를 issues[]에 기록한다.

초성 부호를 종성 자리에 끼워 넣거나, filler를 붙여 임시 평문을 정상 직렬화처럼 내보내지 않는다. 그렇게 하면 L/V/T 순서에 의한 경계 판별과 위치별 역변환의 전제가 무너진다. 부호 후보를 시험하려면 별도 실험 프로파일과 버전을 사용한다.

[I2] 30–37쪽; [I4] 10–11쪽. 미배당 시 반환 상태: 본 개정 설계안.

4. 자모값 등록부 · 모음

음가·요소	ID	부호	적용
i	JV-I	U+1175	단모음·장모음의 구성 요소
ɪ	JV-KIT	U+119D	KIT 대립 보존
ɛ / e	JV-E	U+1166	DRESS와 FACE 구성 요소
æ	JV-AE	U+1162	TRAP
ɑ / a	JV-A	U+1161	GA LOT·PALM 및 복모음 요소
ɒ / o	JV-O	U+1169	RP LOT, GOAT·CHOICE 요소
ɔ	JV-AW	U+1182	GA THOUGHT 시범값
u	JV-U	U+116E	장모음의 구성 요소
ʊ	JV-FOOT	U+118D	FOOT 및 복모음 끝 요소
ʌ	JV-EO	U+1165	STRUT·RP NURSE 구성 요소
ə	JV-SCHWA	U+119E	약모음·RP GOAT 구성 요소

JV-KIT·JV-FOOT 등은 이번 시범 등록부의 식별자다. 식별자 철자 자체가 새 음가 판정은 아니다. 여러 음운 환경이 같은 기본 자모 요소를 쓰는 경우, 그 관계는 개별 mapping unit의 rule_id로 구분한다.

한 자모 부호를 거꾸로 읽는 것만으로 원래 음운 단위를 항상 알 수는 없다. U+1169는 RP /ɒ/의 단모음, GA /oʊ/의 첫 요소, 공통 /ɔɪ/의 첫 요소로 쓰인다. U+1166 역시 /ɛ/와 /eɪ/의 첫 요소에 쓰인다. 역변환은 모듈과 전체 음절핵 규칙을 필요로 한다.

4.2 장모음과 복모음

장모음과 복모음은 단일 부호 이름에서 분해하지 않고 제2본의 구성표에서 명시적으로 가져온다. /u:/는 [JV-U, JV-U], /ʊ/는 [JV-FOOT]이며 둘은 다르다. RP /ɔ:/는 [JV-O, JV-O], GA /ɔ/는 [JV-AW]다. 전체 대응표는 부록 A에 둔다.

등록부에는 codepoint, Unicode 이름, Hangul_Syllable_Type, 허용 위치, module_id, 규칙 상태와 근거를 함께 둔다. 등록되지 않은 음가를 가장 비슷한 자모로 추정하여 채우지 않는다.

[I2] 13-14·35-37·50-53쪽. 자모 ID와 registry 구조: 본 개정 설계안.

제2부 · Unicode 직렬화 규격

5. 직렬화의 기본 원칙

한 음절블록 = 한 자모 연쇄

지원되는 Broad 블록을 L+ V+ T* 형태로 직렬화한다. L은 하나 이상, V는 하나 이상, T는 없거나 하나 이상이다. 이것은 현 시범 조자법을 운반하는 전산 프로파일이며 최종 압축 자형의 완성을 뜻하지 않는다.

음가를 가진 초성군이 비어 있으면 JV-ZERO의 U+110B를 앞에 둔다. 뒤 중성의 저장 부호열에는 ◦을 넣지 않는다. 2026-09-17 저자 결정으로 읽기용 표시에만 무음 ◦을 붙인다. 무음 초성은 U+115F와 다르다.

층 2에는 이 프로파일의 등록부가 허용한 조합형 자모만 사용한다. 현대 완성자, 호환 자모, filler와 임의 제어문자는 허용하지 않는다. 이는 훈민바름의 저장 기준형을 정한 것이며, 해당 문자들이 Unicode에서 잘못된 문자라는 뜻은 아니다.

5.1 직렬화 단위와 지원 범위

직렬화 단위는 하나의 어휘 레코드다. 단어 안의 블록 사이에는 구분 부호를 넣지 않는다. 여러 단어·문장·구두점을 교환할 때는 레코드 배열이나 별도 문서 구조를 사용한다. [], 세로줄과 공백은 설명용 표에서 경계를 보일 수 있으나 정본 자모열에는 포함하지 않는다.

nucleus[]가 없는 음절자음 구조는 이 모음핵 기반 평문 프로파일의 지원 범위 밖이다. source와 syllabic 속성은 구조형에 남기고, 음절핵에 임의의 모음을 추가하지 않는다. 미배당 coda가 있으면 어휘 전체의 직렬화 상태를 unavailable로 돌려준다.

5.2 경계 판별

L을 읽는 동안 같은 초성군에 더하고, V가 나오면 음절핵을 연다. V 다음의 V는 같은 핵, T는 종성군이다. V 또는 T 다음의 L은 다음 블록의 시작이다. T 다음의 V, 첫 위치의 T, V 없이 끝나는 L 연쇄는 오류다. 이 판별은 부호열에 보존된 경계를 읽는 절차이며 영어 음절을 새로 분석하는 알고리즘은 아니다.

[I2] 8-10·26·36쪽; [E1] §3.12 D134. 프로파일 제한은 본 연구의 설계다.

6. 자리별 직렬화 규칙

6.1 초성군

초성군의 각 요소를 등록부의 L 위치형으로 순서대로 출력한다. friend의 초성군은 U+114B U+A976, play는 U+1111 U+1105다. /l/을 한 번만 출력하며 자음 사이에 모음을 넣지 않는다. /s/의 위치형 선택은 해당 모듈의 명시된 시범 규칙을 따른다.

6.2 음절핵

단모음은 하나의 V를, 장모음·복모음은 구성표에 등록된 V 배열을 출력한다. /eɪ/는 U+1166 U+119D, /aʊ/는 U+1161 U+118D다. 코드포인트 이름에 U-U나 O-O가 들어 있다고 하여 같은 문자의 두 번 반복과 동치로 해석하지 않는다.

6.3 종성군과 /s/

종성군은 T 위치형을 음소 순서대로 출력한다. bank는 U+11BC U+11BF, text는 U+11BF U+11BA U+11C0이다. /s/는 같은 음절블록의 독립된 종성 요소로 보존한다. 디지털 표시에서는 제2본의 선형 시범 원칙에 따라 패스처럼 /s/를 드러내며, 한국어 완성자 팻을 정상 HMBR 표시로 대체하지 않는다.

평문 pass(GA)의 U+1111 U+1162 U+11BA는 Unicode NFC에서 팻으로 합성될 수 있다. 그 부호상의 정준 동등성과 훈민바름이 요구하는 학습용 표시 규칙은 서로 다른 문제다. 저장에서는 기준형을 지키고 표시에서는 자모값과 /s/의 가시성을 지킨다.

6.4 /t+s/ 복합값

같은 종성군 안에서 인접한 /t/와 /s/가 등록된 복합값 규칙에 해당하면 JV-TS를 먼저 적용하고 U+11BE를 출력한다. cats는 /k æ t s/ 네 음소와 세 자모 요소를 가진다. 뒤따르는 자음이 있으면 그 요소를 이어 기록한다. 음절·어휘 경계를 가로질러 결합하지 않는다.

```
{ "id": "m3", "token_ids": ["t3", "t4"],
  "jamo_values": ["JV-TS"], "rule_id": "coda-ts",
  "reverse_ipa": ["t", "s"] }
```

/tʃ/를 /t/와 /ʃ/로 쪼개 이 규칙에 넣지 않는다. tokenizer가 인정하는 파찰음 표기 변이는 원문과 정규화 이력을 남기고 하나의 음소로 분석한다.

[I2] 25–30·58–59쪽; [E2] 정준 동등성. /s/ 표시와 직렬화의 역할을 분리한다.

7. 직렬화 갤러리

아래는 제2본의 예시를 현 시범 프로파일의 코드포인트 열로 옮긴 것이다. 코드포인트 사이의 공백은 읽기용이며, 실제 문자열에는 없다. 모든 행은 한 영어 음절이며 한 훈민바름 음절블록이다.

낱말	IPA	코드포인트 (각 항 U+)
pick	/pɪk/	1111 119D 11BF
pat	/pæt/	1111 1162 11C0
sing	/sɪŋ/	110A 119D 11BC
bank	/bæŋk/	1107 1162 11BC 11BF
milk	/mɪlk/	1106 119D 11AF 11BF
text	/tɛkst/	1110 1166 11BF 11BA 11C0
spin	/spɪn/	1109 1111 119D 11AB
friend	/frɛnd/	114B A976 1166 11AB 11AE
three	/θri:/	1145 A976 1175 1175
play	/pleɪ/	1111 1105 1166 119D
boy	/bɔɪ/	1107 1169 119D
boost	/bu:st/	1107 116E 116E 11BA 11C0
food	/fu:d/	114B 116E 116E 11AE

장모음·복모음의 두 V는 같은 핵의 저장 요소다. 읽기용 표시에는 뒤 V 앞에 무음 ㅇ을 붙인다. 이 표시용 자리는 음소·음절 수나 저장 부호열을 바꾸지 않는다. /s/의 위치형 경계는 시범이다.

[I2] 7-17·26-30쪽. 실제 표시 자형과 부호열은 별개로 검증한다.

7. 직렬화 갤러리 · 복합값과 모듈 대조

날말·모듈	normalized IPA	코드포인트 (각 항 U+)
cats RP·GA	/kæts/	110F 1162 11BE
boots RP·GA	/bu:ts/	1107 116E 116E 11BE
blue RP·GA	/blu:/	1107 1105 116E 116E
pass RP	/pɑ:s/	1111 1161 1161 11BA
pass GA	/pæs/	1111 1162 11BA
car RP	/kɑ:/	110F 1161 1161
car GA	/kɑɪ/	110F 1161 D7DB
go RP	/gəʊ/	1100 119E 118D
go GA	/goʊ/	1100 1169 118D
ask RP	/ɑ:sk/	110B 1161 1161 11BA 11BF
about RP·GA	/əˈbaʊt/	110B 119E 1107 1161 118D 11C0
very RP·GA	/ˈvɛɹi/	1144 1166 A976 1175

세로줄은 이 표에서만 쓰는 편집 기호다. about과 very는 각각 두 블록이며 실제 부호열에는 세로줄이 없다. very의 출처형 /vɛri/를 쓰는 자료는 source에 그대로 두고, 본 표에서는 분석에 사용하는 /ɹ/를 명시했다.

7.1 미결 레코드도 버리지 않는다

catch /kætʃ/는 source와 /tʃ/ 토큰, JV-CH와 coda 위치를 구조형에 보존할 수 있다. 다만 이 등록부에서 해당 종성 부호가 미배당이므로 완성 평문을 제공하지 않는다. nurse(GA)의 /ɜ:/는 기저 모음이 미결이므로 normalized IPA와 mapping 결과를 확정하지 않는다.

예시의 위계

제2본의 § 9.3·부록 A와 현 교정본 예시에 따라 ‘초성군의 ㄹ 한 번’을 적용한다. 이전 예시의 외관을 역으로 분석해 폐기된 중복 규칙을 되살리지 않는다.

[I2] 20–24·27·30·37·58쪽; [I5] Master 3, 138쪽.

8. 정규화와 동치

8.1 기준형은 NFD 조합형이다

이 프로파일은 등록된 조합형 자모 연쇄를 저장 기준형으로 삼는다. 기준형의 코드포인트 일치와 Unicode의 정준 동등성을 구분한다. 전자는 규칙이 고른 직렬화가 같은지, 후자는 Unicode 정규화 기준에서 같은 문자열인지를 검사한다.

입력·처리	본 프로파일의 판정
규격에 맞는 NFD 자모열	기준형 후보, registry·배열 규칙 추가 검사
NFC로 합성된 완성자 포함 입력	원입력을 보존한 뒤 NFD로 분해하고 검증하여 수용 가능
호환 자모·NFKC/NFKD 적용 입력	출처와 변환 이력을 확인. 원래 위치를 추정하여 자동 확정하지 않음
U+118D와 U+116E U+116E	서로 다른 값. 동치 치환 금지
U+1182와 U+1169 U+1169	서로 다른 값. 동치 치환 금지

NFC는 지원되는 일부 자모를 현대 완성자로 합칠 수 있으나 그 사실만으로 음운 정보가 소실됐다고 단정하지 않는다. 이 문서의 시험 자모열은 NFC를 거친 뒤 NFD로 되돌리면 원래 코드포인트 열이 복원됐다. 그러므로 수신 어댑터는 NFD 후 검증할 수 있다. 다만 source·강세·규칙 이력은 처음부터 평문에 없었으므로 정규화로 복원할 수 없다.

8.2 프로파일 의미 동치

등록 겹낱자와 낱자 연쇄의 의미 동치는 Unicode 이름이나 모양으로 자동 결정하지 않는다. 해당 모듈·버전이 동일한 값과 역규칙을 명시한 경우만 별도 별칭 관계로 둘 수 있다. 이 관계를 Unicode 정규화라고 부르지 않으며 UAX #15의 알고리즘을 바꾸지 않는다.

현재 기준형에는 겹낱자 별칭의 포괄적 치환을 도입하지 않는다. 따라서 등록부의 허용 목록에 없는 복합 부호를 받아 임의로 분해하지 않는다. 특히 /ʊ/와 /u:/의 대립을 지키기 위해 U-U라는 이름을 분해 지시로 사용해서는 안 된다.

[E2] UAX #15 §§1-2-6. 정준 정규화와 연구 프로파일의 의미 규칙은 구분한다.

9. 금지되는 연쇄와 오류 검출

검사기는 입력 오류를 발견하면 원입력과 위치를 보존한다. 미배당이나 미결을 다른 음가로 바꾸어 성공 처리하지 않는다. 수신 어댑터의 정규화와 기준형 검사를 구분하여, 고쳐 받은 입력이 원래부터 기준형이었던 것처럼 기록되지 않게 한다.

오류 ID	검출 대상	처리
E-CHAR	기준형의 완성자·호환 자모·filler·제어문자	원입력 보존, 수신 어댑터 여부 확인
E-ORDER	첫 T, T 다음 V, V 없는 블록	해당 위치의 직렬화 거부
E-REGISTRY	미등록 자모값·미배당 위치형	구조 보존, 평문 unavailable
E-REFERENCE	없는 token_id·mapping id·요소 번호	구조형 검증 실패
E-COVERAGE	입력 음소 누락·중복 소비, 출력 요소 누락	mapping 관계 재검토
E-BOUNDARY	구조형 경계와 자모열 경계 불일치	전체 레코드 일관성 실패
E-AMBIGUOUS	둘 이상의 역규칙 후보	후보 반환, 자동 단일화 금지
E-UNRESOLVED	/ʒ/·음절자음 등 미결 단계	원문·미결 사유 보존

9.1 구문만으로는 잡지 못하는 오류

음절핵 도중에 잘못 삽입된 L은 자모열만 보면 새 음절의 시작처럼 보일 수 있다. 따라서 정규식 검사만으로 구조가 맞다고 선언하지 않는다. 원 canonicalModel의 블록 수·자리 순서·토큰 대응과 다시 대조해야 한다.

한 mapping unit이 여러 토큰을 소비할 수 있지만, 같은 입력 토큰을 관계없는 두 mapping unit에서 중복 소비하지 않는다. 출력 요소 참조도 정확히 한 자리에 속한다. 무음 조성처럼 음소를 소비하지 않는 요소는 예외의 종류를 명시해야 한다.

9.2 성공 조건

complete 평문을 내보낼 때는 모든 블록과 모든 위치형이 직렬화 가능해야 한다. 부분 결과는 완성 평문 필드에 쓰지 않는다. 표시용 설명에서는 이미 처리한 부분과 미결 부분을 함께 보여 줄 수 있으나, 그 전체를 유효한 정보 문자열로 표시하지 않는다.

오류 ID와 검증 절차: 본 개정의 시범 규칙. 제2본의 미결사항을 자동 보정하지 않는다.

제3부 · 글꼴과 세이핑

10. 표시층의 지위

표시층은 정본을 읽게 만드는 층이다. 부호열의 글리프가 화면에 나타난다는 것, 음절블록이 하나의 표시 영역으로 묶인다는 것, 자모값이 식별된다는 것은 서로 다른 조건이다. 한 조건의 성공으로 다른 조건까지 충족했다고 판단하지 않는다.

displayRuns는 block_id, 자모 요소 참조, 표시 프로파일, 강세값과 render_status를 가진다. displayText는 그 명령을 문자열로 나타낼 수 있을 때의 파생 결과다. SVG나 canvas로 그리는 경우 문자열만으로는 실제 배치를 모두 설명할 수 없으므로, 자모 참조와 접근성 텍스트를 함께 유지한다.

11. 글꼴 요구사항

검사	확인할 것	그것만으로 알 수 없는 것
cmap	요구 코드포인트가 글리프에 연결되는가	결합 후 올바른 모양과 위치
글리프	/v/·/f/·KIT·FOOT 등 자형이 구별되는가	임의 복합 연쇄의 조판 성공
GSUB·GPOS	필요한 치환·배치 기능과 lookup	해당 연쇄에 실제 적용되는지
글꼴 로드	실제 선택된 파일과 버전	다른 환경의 동일 결과
라이선스	사용·수정·배포 조건	정본 규칙의 타당성

Microsoft의 Hangul 설명은 등록 조합을 처리하는 전처리와 ljmo·vjmo·tjmo 등의 적용을 설명한다. 기능 태그가 존재한다는 사실만으로 모든 L+ V+ T* 연쇄의 배치가 보장되는 것은 아니다. 실제 엔진·글꼴·적용 기능을 함께 시험한다.

11.1 이번 시험 글꼴

이번 재측정에는 NanumMyeongjoOTF-YetHangul.otf, Version 1.0.0.9 Build 20140917을 사용했다. 글리프 수는 26,030개이며 GSUB에는 aalt·ccmp·liga·ljmo·tjmo·vjmo가 있다. 이전 보고의 28,844개와 다른 파일이므로, 이번 결과를 이전 글꼴 파일의 동일 재현이라고 부르지 않는다.

실제 서비스에서는 글꼴 다운로드 성공과 폰트 선택, 코드포인트 범위, 필요한 위치형의 표시를 확인한다. 다른 글꼴의 자동 대체로 우연히 보인 결과와 목표 글꼴 자체의 지원을 구분한다.

[E3] Microsoft Hangul 설명. 파일 정보와 수치: 2026-09-08 직접 측정, 부록 C.

12. 셰이핑 요구사항

12.1 보존해야 할 것

셰이핑은 블록 안 자모의 순서와 음가 식별을 지키면서 표시 영역을 구성해야 한다. 블록 폭이나 모양을 맞추기 위해 /ʊ/를 /u:/의 부호로 바꾸거나 /t+s/를 /tʃ/로 읽게 하지 않는다. 여러 글리프로 한 블록을 그릴 수도 있으므로 글리프 수 하나를 합격 기준으로 삼지 않는다.

최종 압축 손글씨 배치형은 제2본의 미결 과제다. 그 상태에서 전산본이 임의의 압축 모양을 최종 정본으로 고정할 수는 없다. 현 단계에서는 블록 경계와 내부 순서가 보이는 선형 시범 표시를 지원하고, 압축 배치 후보는 별도 프로파일에서 비교한다.

12.2 세 종류의 경계

경계	결정 기준	시험
음운 음절 경계	선택된 언어모듈·사전·검토	canonicalModel과 source 분석 대조
텍스트 처리 경계	Unicode 분절 규칙과 앱의 명시적 조정	선택·삭제·커서 이동·줄바꿈
실제 자형 영역	글꼴·셰이핑·배치 프로파일	페이지·화면의 자모 식별과 배치

UAX #29의 GB6–GB8은 한글 연쇄 내부의 그래핌 경계를 다룬다. 이 경계는 원어의 음운 분석을 대신하지 않으며, 여러 자모를 한 네모로 배치하는 기능 그 자체도 아니다. 앱은 실제 커서와 삭제 동작을 따로 시험한다.

12.3 폭 측정의 역할

한 음절의 기준 폭으로 나눈 상대 폭은 분산 표시를 탐지하는 보조 지표다. 약 1자리 폭이라고 해도 자모가 겹쳐 안 보이거나 다른 음가처럼 보이면 실패다. 최종 표시 판정에는 자모 누락·겹침·읽는 순서·/s/ 가시성·모음 대립을 시각적으로 확인하는 절차가 필요하다.

표시 후보가 성공하더라도 canonicalModel과 기준 자모열은 유지한다. 프로파일 이름과 버전을 바꾸지 않은 채 새로운 축약 모양을 배포하지 않는다.

[I2] 26·36쪽; [E4] UAX #29 §3.1.1, GB6–GB8. 표시 요구: 본 연구의 제안.

13. 현재 환경에서 무엇이 합성되는가 · 재측정

2026-09-08, HarfBuzz 8.3.0과 §11.1의 OTF 파일을 사용해 문자열의 가로 advance 합계를 측정했다. 글꼴 좌표계는 1,000 units/em이며 기준 문자열 가(U+AC00)의 advance는 950이다. 상대 폭은 시험 문자열의 advance를 950으로 나눈 값이다.

입력	부호 수	글리프 수	advance	상대 폭
pick	3	3	973	1.0242
bank	4	2	1923	2.0242
boost	5	4	3869	4.0726
spin	4	4	1946	2.0484
milk	4	4	1946	2.0484
friend	5	5	2919	3.0726
play	4	3	2896	3.0484
boy	3	2	1923	2.0242
cats	3	1	950	1.0000

13.1 확인된 결과

pick은 세 글리프를 가지면서 약 1.02자리, cats는 한 글리프로 1자리 폭을 차지했다. 반면 bank·milk·spin·boy는 약 2자리, friend·play는 약 3자리, boost는 약 4자리다. 이 환경에서는 여러 자모를 포함한 연쇄가 하나의 압축된 표시 영역으로 모이지 않았다.

이는 이번 입력과 글꼴·엔진 조합에서의 폭 측정 결과다. 모든 조합형 자모, 다른 글꼴·운영체제·브라우저를 대표하는 수치가 아니다. 또한 상대 폭은 가독성·음성학적 타당성·교육 효과의 점수가 아니다.

13.2 이전 보고와의 관계

이전 41쪽 전산본의 그림과 보고도 비슷한 분산 표시를 보여 주지만, 글꼴 파일과 엔진 버전의 동일성이 입증되지 않았다. 따라서 새 측정값과 이전 보고를 혼합해 하나의 실험처럼 제시하지 않는다. §21의 사전 전수조사는 셰이핑 실험과 별개로 현재 저장소의 코드와 자료에서 재현했다.

이번 직접 측정, 글꼴 해시·실행 조건·재현 절차는 부록 C. 이전 보고: [I4] 17–20쪽.

13. 현재 환경에서 무엇이 합성되는가 · 대조

같은 환경에서 일부 자모 연쇄를 등록된 복합 부호로 바꾼 대조 입력도 측정했다. 이 입력은 글리프 처리의 차이를 살피는 실험용이며, 훈민바름 정본의 동치 표기로 승인한 것이 아니다.

대조 이름	실험용 부호열 (각 항 U+)	상대 폭
bank	1107 1162 11EF	1.0242
boost	1107 118D D7F1	1.0242
spin	113A 119D 11AB	1.0242
milk	1106 119D 11D8	1.0242
friend	114B A976 1166 11C6	2.0484

bank 대조에서는 종성 두 부호 대신 U+11EF, spin에서는 초성 두 부호 대신 U+113A, milk에서는 종성 두 부호 대신 U+11D8을 넣었다. friend는 겹종성 U+11C6만 적용했으므로 초성군 부분의 분산이 남았다.

boost 대조의 한계

U+118D는 이 연구에서 /ʊ/의 부호다. U+116E U+116E인 /u:/를 U+118D로 바꾼 대조 입력은 boost의 정상 표기가 아니다. 약 1자리 폭이 나왔다는 사실은 /u:/를 손실 없이 표현했다는 증거가 아니다.

13.3 이 실험으로 말할 수 있는 것

이 환경은 일부 등록 복합 부호가 포함된 입력을 약 1자리 폭으로 처리했다. 이는 동일한 엔진에서도 입력 부호의 구성이 표시 결과에 영향을 준다는 근거다. 그러나 임의의 미등록 조합을 현재 글꼴이 이미 그릴 수 있다거나, 셰이핑의 유일한 원인이 부호 배당이라는 결론은 나오지 않는다.

글꼴의 lookup, 엔진 전처리, 클러스터 인식과 최종 배치는 함께 검토해야 한다. 새 문자 부호를 배당하는 일과 기존 글꼴에 그 글리프를 설제하는 일도 별개다. 부호를 등록하면 즉시 모든 글꼴이 지원한다는 주장은 성립하지 않는다.

13.4 다음 시험

최종 배치 후보마다 원 자모의 식별과 읽기 순서를 확인하고, 원본 자모열 유지 여부와 복사 결과를 검사한다. 다른 글꼴·브라우저·운영체제의 시험은 제4본에서 환경별로 축적한다.

대조 입력과 폭: 이번 직접 측정. 의미 판정은 [I2]의 모음 대립과 §8의 동치 원칙을 따른다.

14. 대체 표시

대체 표시는 표시 실패 때도 source와 정본 구조를 읽을 수 있게 하는 출력이다. 블록을 합성하지 못했다고 영어 음절을 새로 나누지 않는다. 여러 자모가 선형으로 보이더라도 한 음절의 범위를 [] 로 묶고 IPA를 함께 제시한다.

상태	출력 원칙	기록
복합 블록 압축 불가	블록별 선형 자모·음소 순서 표시	render_status=linear
특정 글리프 없음	코드포인트·자모값·IPA 병기	missing_glyph와 해당 요소
종성 부호 미배당	구조와 IPA 설명, 완성 평문 없음	serialization.status=unavailable
원문 분석 미결	source IPA와 미결 사유 표시	status=unresolved

boost는 한 블록 [부우ㅅㅌ] 으로 표시한다. 2026-09-17 저자 결정에 따라 뒤 모음에 무음 ㅇ을 붙여 읽기 쉽게 한다. 우는 같은 /u:/의 뒤 요소이며 새 음절이 아니다. 저장형 L V V T T와 /b u: s t/는 그대로 보존한다.

14.1 정상 표시와 시범 표시

단순 블록이 잘 합성되면 composed로 기록할 수 있다. 복합 블록의 선형 시범형은 linear로 알린다. 미결인 압축 모양은 experimental 표시 프로파일로 구분한다. ‘옛한글 지원’이라는 한 문구로 이 상태들을 합쳐 알리지 않는다.

14.2 접근성

자모의 모양만으로 정보를 전달하지 않는다. source spelling·선택한 발음 모듈·IPA·블록 수·강세를 텍스트로 제공하고, 장모음·복모음은 핵의 구성과 길이를 설명할 수 있어야 한다. 색상이나 글자 크기만으로 상태나 강세를 전달하지 않는다.

화면낭독기가 훈민바름 문자열을 정확히 발음한다고 가정하지 않는다. 원문·IPA 읽기, 음성자료 연결과 구조 설명을 별도로 시험한다. 단순 선형 표시를 복사하려는 사용자의 선택은 허용하되, 정본 교환용 복사와 구별하여 어떤 형식이 전달되는지 알린다.

[I1] 접근성 원칙; [I2] 9-10·26쪽. display/fallback 세부 요구는 본 개정 설계안.

제4부 · 검색·복사·역변환

15. 검색

검색은 목적에 따라 키를 나눈다. 부호열 완전일치, 같은 음소 찾기, 같은 블록 구조 찾기는 서로 다른 질문이다. 어느 키든 모듈·등록부·규칙 버전을 포함하여 다른 분석을 우연히 합치지 않는다.

검색 종류	키	보존하는 구별
원문 검색	spelling, source language/script	원철자·원문 문자
출처 IPA 검색	source.ipa + source id	사전 표기 관행
음운 검색	normalized tokens + module/version	선택된 음운 대립
자모값 검색	mapping 결과의 자모 요소열	자모값·출력 순서
정본 부호열 일치	profile/version + NFD codepoints	기준 직렬화
구조 검색	onset/nucleus/coda와 prosody	음절·자음군·핵·강세
표시 별칭 검색	승인된 alias + profile/version	검색 편의, 별칭 출처

soon /su:n/의 자모값 키는 JV-S, JV-U, JV-U, JV-N이다. 여기에는 장모음 구성요소가 둘 있지만 음운 키에는 /u:/ 하나가 있다. 이 차이를 유지해야 ‘자모가 네 개인 형태’와 ‘음소가 세 개인 형태’를 각각 검색할 수 있다.

15.1 동음어와 동형 자모열

같은 발음을 가진 원철자는 여러 개일 수 있다. 따라서 자모열에서 단어 철자를 유일하게 복원한다고 선언하지 않는다. 또한 서로 다른 언어모듈이 같은 자모열을 다른 분석으로 사용할 수 있으므로 module_id 없는 검색 결과에는 분석 범위를 표시한다.

15.2 정규화와 검색 인덱스

입력 원문은 보존하고 별도의 검색 기준키를 만든다. NFC 입력을 NFD로 변환해 검색할 수 있으나, 원 source IPA를 그 과정에서 덮어쓰지 않는다. 표시 별칭은 별칭 검색에만 쓰고 정본의 의미 동치로 승격하지 않는다. 새로운 registry 버전은 기존 키를 재생성할 수 있어야 한다.

[I1] 원문·발음·데이터 연결 원칙; [E2] 정규화. 검색 키 분리본 개정의 시범 규칙.

16. 복사와 붙여넣기

복사는 사용자가 선택한 정보 수준을 전달해야 한다. 화면 문자열만 보내는 복사, 정본 자모열 복사, source와 구조 전체를 보내는 복사는 보존 범위가 다르다. 구현은 이를 명확히 구별한다.

형식	답는 것	한계
text/plain	지원되는 기준 자모열	source·강제·버전이 스스로 따라가지 않음
JSON 파일·지원되는 구조 형식	source·canonicalModel·규칙 버전·직렬화	수신 측의 스키마 지원 필요
text/html	표시 마크업·IPA·기준 자모열·버전 표시	수신 앱이 속성을 제거할 수 있음
읽기용 표시 복사	블록 경계·선형 표기·설명	정본 교환 형식으로 재해석하지 않음

16.1 기본 복사

평문을 생성할 수 있는 레코드의 정본 복사는 기준 자모열을 제공한다. 가능한 환경에서는 구조 형식을 함께 제공하되, 임의 MIME 형식을 모든 클립보드와 앱이 지원한다고 전제하지 않는다. 구조형 파일 다운로드를 위한 별도 경로다.

평문이 unavailable인 레코드에는 완성 자모열을 제공하는 것처럼 보이는 단추를 두지 않는다. 대신 구조형 내보내기 또는 IPA·미결 설명 복사를 제공한다. 표시 문자열에서 빠진 음소가 복사 과정에서도 조용히 사라지지 않게 한다.

16.2 부분 선택과 붙여넣기

정본 블록 복사는 선택 범위를 블록 경계까지 넓힐 수 있으며 그 동작을 사용자에게 알려야 한다. 자모 요소만 필요한 고급 복사는 fragment임을 명시한다. 조각을 완성 음절로 읽지 않는다.

붙여넣은 평문은 원입력을 보존한 뒤 NFD 변환과 프로파일 검사를 거친다. 형식·버전을 알 수 없거나 역규칙이 여러 개면 후보로 남긴다. HTML의 보조 속성이 사라졌다면 구조 전체를 복구했다고 선언하지 않는다. 복사 왕복은 실제 수신 앱을 포함해 시험한다.

클립보드 형식은 교환 설계 요구다. 실제 브라우저·앱 지원과 시험 결과는 제4본에서 보고한다.

17. 역변환 · 무엇을 되찾는가

역변환은 정보 수준별로 정의한다. source 레코드에서 원문을 다시 읽는 것, canonicalModel에서 선택된 음운 분석을 회수하는 것, 평문에서 가능한 분석을 구하는 것은 서로 다른 작업이다.

경로	되찾을 수 있는 것	되찾지 못하는 것
전체 구조 레코드	보존한 source, 토큰, mapping, 강세와 버전	원래 기록하지 않은 정보
기준 자모열 + 모듈·역규칙	지원 범위의 자모 구조와 유일한 음운 후보	출처 표기·강세·분석 이력·원철자
기준 자모열만	부호와 L/V/T 구조	모듈을 모르면 음운 해석을 확정할 수 없음
표시·대체 문자열	보이는 정보의 참고적 해석	완전한 정보 복원 보장 없음

17.1 구조형 왕복

JSON을 쓰고 다시 읽었을 때 배열 순서, 문자열, 필드의 null과 생략, 상태값과 참조 관계가 보존되는지 검사한다. JSON 키의 직렬화 순서가 바뀐 것과 데이터 의미가 바뀐 것은 구분한다. source IPA의 원문 복원은 보관된 source.ipa를 읽는 절차다.

17.2 평문 왕복의 전제

부호열을 블록으로 나누고 위치형을 자모 요소로 되돌린 뒤, 모듈의 전체 음절핵·복합값 역규칙을 적용한다. 등록부가 단일 부호마다 하나의 자모값을 가리킨다는 조건만으로 충분하지 않다. 장모음·복모음은 여러 부호를 함께 읽으며 /t+s/는 한 부호에서 두 음소를 회수한다.

지원 범위에서 하나의 자모 구조가 둘 이상의 normalized 음운 분석을 허용하면, 반환값은 ambiguous와 후보 목록이다. 임의로 첫 후보를 고르지 않는다. 이 초안의 시험 예시가 모두 유일하더라도 영어 전체의 유일 역변환을 입증한 것은 아니다.

[I2] 13-14·30·37·50-53쪽의 대응에서 도출한 전산 요구. 완전 보존은 구조 레코드를 전제로 한다.

17. 역변환 · 복합 대응과 제한

17.3 boost와 cats

입력	구조 회수	음운 회수
boost의 L V V T T	B / U,U / S,T	/b/ + /u:/ + /s/ + /t/
cats의 L V T	K / AE / TS	/k/ + /æ/ + /t/ + /s/
play의 L L V V	P,L / E,KIT / Ø	/p/ + /l/ + /eɪ/

boost의 두 U는 long-u 역규칙으로 /u:/ 하나가 된다. 이를 /u/ 두 음소라고 읽지 않는다. cats의 TS는 coda-ts 규칙으로 /t/와 /s/ 두 음소가 된다. play의 초성 L은 한 번만 회수한다.

17.4 같은 기본 요소, 다른 음절핵

U+1169 U+119D는 공통 CHOICE 규칙에 따라 /ɔɪ/다. U+1169 U+118D는 GA GOAT 규칙에 따라 /oʊ/다. 단독 U+1169가 RP에서 /ɒ/를 나타낸다고 하여 모든 연쇄의 첫 요소를 /ɒ/로 읽지 않는다. 역규칙은 부록 A의 전체 핵 단위로 조회한다.

U+1165 U+1165는 현 RP NURSE 규칙에서 /ɜ:/다. U+1165 단독의 /ʌ/를 두 번 출력해 해결하지 않는다. 같은 원리로 모듈의 음운 단위와 자모 구성요소를 분리해야 한다.

17.5 복원하지 않는 정보

about의 평문에는 강세기호가 없으므로 /ə'baʊt/의 제2음절 강세를 평문만으로 되찾았다고 말할 수 없다. 규칙으로 음소 분절열 /əbaʊt/를 회수할 수 있더라도 source의 기호 배치와 출처는 구조 레코드가 보존한다.

Narrow [ɹ]가 보이더라도 기저 음소가 /t/인지 /d/인지는 surface phone만으로 정해지지 않을 수 있다. underlying token을 저장해야 한다. 연결발화 표면형 역시 사전형을 자동으로 유일하게 복원하는 자료가 아니다.

왕복 적합성의 보고 단위

원문 문자열, normalized 음운 토큰, 자모 구조, 코드포인트 열, 표시 결과를 각각 검사한다. 서로 다른 정보 수준을 ‘무손실’ 한 단어로 묶지 않는다.

18. 교환 형식

완전 보존의 기본 교환 단위는 `source`와 `canonicalModel`을 포함하는 구조 레코드다. 직렬화는 전체 정보가 아니라 함께 제공하는 파생 값이다. 받는 쪽은 해당 버전을 지원하는지 확인하고, 구조에서 다시 만든 자모열이 함께 온 직렬화와 일치하는지 검증한다.

필수 묶음	내용
식별·버전	<code>schema_version</code> , <code>module_id</code> , <code>registry_version</code> , 규칙·프로파일 버전
<code>source</code>	원문, <code>source IPA</code> , 언어·문자·출처와 이용조건
<code>analysis</code>	<code>normalized IPA</code> , 정규화 이력, 음절분석과 판정 근거
<code>canonicalModel</code>	음운 토큰, <code>mappingUnits</code> , 블록, 운율·Narrow와 상태
<code>serialization</code>	상태, <code>profile</code> , 기준 <code>codepoints</code> 또는 <code>unavailable</code> 사유
<code>validation</code>	검사 버전, 오류·미결 목록, 검사 범위

18.1 코드포인트 배열과 문자열

`codepoints[]`는 Unicode 스칼라값을 정수 배열로 담는다. 이 프로파일은 허용된 자모만 포함하므로 각 값이 `registry`와 L/V/T 자리 규칙을 충족해야 한다. `text` 필드를 함께 두면 `codepoints`로 만든 문자열과 같은지 검사한다.

정규화를 강제하는 전송 환경에서는 숫자 배열 또는 명시된 이스케이프 전송을 쓸 수 있다. JSON이라는 포장만으로 내부 문자열의 정규화가 자동 방지되는 것은 아니다. 수신 측에서 재구성 후 검증해야 한다.

18.2 불일치와 버전 호환

구조형과 함께 온 자모열이 다르면 수신자는 한쪽을 조용히 정본으로 채택하지 않는다. 원 레코드와 불일치를 보존하고, 선택한 `registry`로 재생성한 결과를 별도 후보로 기록한다. 알 수 없는 필드를 버리는 구현은 완전 왕복을 주장하지 않는다.

교환 과정의 `content hash`는 전송 손상을 확인하는 수단이다. 그것이 음운 분석의 타당성이나 출처 라이선스를 입증하지는 않는다. 동일 레코드의 바이트 해시가 필요하다면 JSON 직렬화 방식까지 고정해야 한다.

교환 봉투와 필드명은 이번 개정의 시범 설계다. 현재 웹앱이 이미 이 형식을 사용한다고 주장하지 않는다.

제5부 · 국가표준·국제표준 정합성

19. 이미 표준 안에 있는 것

정밀 표기층을 연구하려면 문자부호, 음절의 구조, 표시 기술과 적용 규범을 나누어 살펴야 한다. 문자가 Unicode에 있다는 사실은 표현의 출발점이다. 모든 언어의 음가 대응이나 모든 환경의 상호운용성이 이미 해결됐다는 뜻은 아니다.

19.1 Unicode 음절블록과의 구조적 정합성

Unicode 17.0 §3.12 D134는 표준 한국어 음절블록을 하나 이상의 L, 하나 이상의 V, 0개 이상의 T가 이어지는 연쇄 또는 그와 정준 동등한 연쇄로 정의한다. 본 프로파일의 L+ V+ T* 구조는 이 정의의 조합형 구조와 정합한다. 다만 Unicode의 정의가 훈민바름의 영어 음운 분석과 음가 배당을 승인하는 것은 아니다. [E1]

또한 이 정의는 실제 글꼴이 임의의 복합 연쇄를 읽기 좋은 한 블록으로 그린다는 보증이 아니다. 음절 경계, 자형 합성과 복사·검색의 상호운용성은 별도로 검증해야 한다.

19.2 사용 범위

Unicode 블록	범위	본 프로파일
Hangul Jamo	U+1100–U+11FF	등록부의 L/V/T 위치형만 허용
Hangul Jamo Extended-A	U+A960–U+A97F	등록된 초성 위치형
Hangul Jamo Extended-B	U+D7B0–U+D7FF	등록된 중성·종성 위치형
Hangul Syllables	U+AC00–U+D7A3	수신 어댑터에서 NFD 후 검사
Hangul Compatibility Jamo	U+3130–U+318F	설명·날자 표시용, 기준형에는 배제

범위 전체를 허용 문자 목록으로 쓰지 않는다. 미할당 코드포인트, 미등록 위치형과 filler는 별도 검사한다. Unicode 이름은 문자 식별을 위한 자료이며 HMBR 음가를 자동으로 계산하는 규칙이 아니다.

표준의 역할 구분

Unicode 적합성, KS 적용 프로파일과의 정합성, HMBR 규칙 적합성, 실제 표시 적합성은 각각 판정한다. 어느 한 층의 성공이나 제약을 나머지 모든 층의 결론으로 확대하지 않는다.

20. KS X 1026-1에서 확인한 조항

본 절의 직접 확인 자료는 한국 측 국제회의 제출문서 N3422에 실린 KS X 1026-1:2007 영문 번역본이다. 이 자료는 국문 정식 표준이 아니며, 국문판과 차이가 있으면 국문판이 최종 권위임을 밝힌다. 아래 절 번호와 요지는 이 번역본 기준이다. [E5]

번역본의 절	확인한 요지	검토할 관계
§5.1, §5.3(1)	복합 낱자를 단순 낱자 부호 여러 개의 연쇄로 나타내는 것을 제한	복수 L·V·T를 사용하는 확장 프로파일
§5.2, §5.3(3)	현대 한글 음절의 외부 교환은 완성형 사용	일관 조합형 기준 저장 방식
§5.3(2)	완성 음절과 조합형 낱자를 합쳐 다른 블록으로 표현하는 것을 제한	NFC 입력·NFD 기준형과의 관계
§6	정규화와 별도 권고 알고리즘	UAX #15와 적용 프로파일의 구분

20.1 연구 질문의 범위

이 조항들을 외국어 정밀 표기 프로파일에도 그대로 적용할 경우, 하나의 음절 안에서 여러 자모 요소를 보존하는 방식을 어떻게 수용할지 검토가 필요하다. 특히 등록된 복합 부호의 사용 관행과, 같은 기본 요소의 연쇄에 독립된 음가를 부여하는 연구 프로파일 사이에는 명시적 조정이 필요하다.

현행 정보교환 프로파일이 허용하는 범위만 적용하면, 이 연구가 요구하는 복수 자모 요소의 독립 음가와 순서를 완전하게 표현할 길이 없다. 그러나 그것이 곧 법적 금지를 뜻하지는 않는다. 실제 구속력·기관별 채택·현행 국문판 적용은 별도 확인 대상이며, 이 절은 정보교환 프로파일의 표현 범위와 정합성에 한정한다.

20.2 개정 전에 확인할 사항

현행 국문 정식본과 개정·확인 이력을 대조하고, 정밀 외국어 표기층을 기존 정의에 포함시킬지 부속 프로파일로 돌지 검토한다. 국제표준의 정규화 정의를 바꾸지 않으면서 허용 표현과 교환 조건을 명시할 수 있는지 확인한다. 그 결과에 따라 본문 개정, 부속서 또는 독립 응용 프로파일 가운데 적절한 형식을 선택한다.

[E5] PDF 5쪽의 번역본 지위, 12~14쪽의 관련 조항. 개정 방향과 적용 해석은 본 연구의 제안이다.

21. 훈민바름 연쇄가 막히는 지점

21.1 확실한 사례부터 분리한다

첫째는 자모값·위치형의 미배당이다. /θ/ 등의 종성형 문제는 현재 연구 등록부의 과제이므로, 이를 모두 국가표준의 제약으로 돌리지 않는다. 둘째는 구성 자모가 있어도 원하는 연쇄의 표시와 교환이 충분히 지원되지 않는 경우다. 셋째는 복합 부호와 구성 연쇄를 자동 동치로 취급할 때 음운 대립이 사라지는 경우다.

§13은 둘째 문제를 한 환경에서 직접 측정했다. /ʊ/와 /u:/, GA /ɔ/와 RP /ɔ:/의 부호 대비는 셋째 문제를 설명한다. 이 사례들은 대규모 사전의 비율을 확정하지 않아도 프로파일의 동치 규칙을 신중하게 설계해야 한다는 근거가 된다.

21.2 사전 전수조사의 재현 결과

자리·모듈	종류 총계	부호 있음	부호 없음	미배당	복합값 /t+s/
초성 RP	207	16	190	1	0
초성 GA	123	11	112	0	0
종성 RP	447	38	326	82	1
종성 GA	303	31	208	63	1

2026-09-16 저장소의 두 사전과 tools/cluster_census.mjs로 다시 계산했다. RP는 초성군 207종(부호 있음 16·없음 190·미배당 1), 종성군 447종(38·326·82·복합값 /t+s/ 1)이다. GA는 초성군 123종(11·112·0), 종성군 303종(31·208·63·복합값 /t+s/ 1)이다. 종성 표의 과거 소계가 1종 작았던 것은 /t+s/를 구성 연쇄가 아니라 복합값 ㅈ으로 배당해 세 범주 대조에서 제외했기 때문이다.

재현 결과와 해석 범위

완전 해소에 필요한 새 겹낱자는 RP 516종, GA 320종이며 빈도 상위 60종의 누적 해소율은 각각 90.4%, 91.8%다. 2요소 음절핵은 14종이고, 코드포인트 이름상 후보는 3종이지만 U+118D와 U+1182는 이미 각각 /ʊ/와 GA /ɔ/에 배당돼 자동 동치로 쓸 수 없다. 이 수치는 현 사전과 전처리 규칙의 재현 결과이며 영어 전체의 영구적 하한은 아니다.

어두·어말 추출은 고유명사·차용어·오류·음절자음·구두점과 각 사전의 RP/GA 표기 관행에 영향을 받는다. 따라서 결과에는 사전 파일, 실행 코드, 분모·제외 조건과 실행일을 함께 인용한다. 저장소에 포함된 사전의 원출처·판·재배포 조건은 별도 확인 대상으로 남긴다.

재현 자료: data/dict-rp.json, data/dict-ga.json, js/registry.js, tools/cluster_census.mjs. 실행: node tools/cluster_census.mjs. 확인일: 2026-09-16. 기존 보고: [I4] 25–28·35–40쪽.

22. 두 갈래의 해법과 그 비용

22.1 갈래 A · 필요한 복합 부호를 검토한다

필요한 조합에 독립 부호를 배당하는 길은 문자 부호화 제안, 용례·구별 필요성 검토와 글꼴·엔진 지원을 함께 요구한다. 기존 부호를 활용할 수 있어도 해당 부호의 자형·문자적 정체성과 HMBR 의미의 적합성을 검토해야 한다.

새 부호가 등록되었다고 기존 글꼴이 즉시 그릴 수는 없다. §13의 대조는 이미 있는 일부 복합 부호를 처리한 결과다. 아직 없는 부호에 대한 구현 비용이 없다는 증거로 사용할 수 없다.

22.2 갈래 B · 낱자 연쇄를 보존하는 프로파일

구성 자모열을 기준형으로 유지하고 블록 경계·교환·표시 규칙을 별도로 정하는 길이다. 연쇄의 음운 대응과 직렬화 구조를 명시하므로, 새 언어를 추가할 때 공통 스키마와 처리 도구를 재사용할 수 있다. 다만 새 음가 배당, 언어별 분석과 배치·사용자 시험은 여전히 필요하다.

비교 항목	갈래 A · 복합 부호	갈래 B · 구성 연쇄
추가 절차	문자 부호화·글꼴·엔진 검토	프로파일·배치·상호운용성 구현
확장 비용	새 조합의 문자 식별 검토	새 모듈의 분석·registry·표시 시험
기존 환경	기존 부호의 일부는 지원	지원되지 않는 연쇄는 fallback
의미 동치	자동 인정하지 않음	자동 인정하지 않음
공통 조건	음가·순서·출처와 버전 보존	음가·순서·출처와 버전 보존

22.3 이 연구의 권고

구성 연쇄와 구조형을 보존하는 갈래 B를 우선 연구한다. 갈래 A는 독립된 문자적 필요성과 음가 보존이 입증된 항목에 한해 검토한다. 어느 경로든 단일 복합 부호와 연쇄의 동일성은 명시적 registry 관계로만 허용하며, Unicode 정규화 알고리즘을 바꾸지 않는다.

갈래 비교와 권고는 본 연구의 설계·정책 제안이다. 새 부호의 등록 가능성이나 일정은 확정하지 않는다.

23. 개정 제안

제1본이 제안한 것은 일반 외래어 표기의 기능과 통일성을 유지하면서, 원어 음가를 정밀하게 기록할 분야에 별도의 연구·적용 공간을 마련하자는 것이다. 아래 제안은 그 공간에서 데이터를 저장하고 교환하는 기술적 조건을 다룬다.

과제	제안 방향	보존할 조건
구성 연쇄의 표현	확장 프로파일에서 복수 초·중·종성 연쇄를 명시적으로 허용	원래 음가·순서·블록 경계
기준형의 선택	일반 조합형 교환 프로파일과 수신 정규화 절차를 문서화	Unicode 정준 동등성
의미 동치의 범위	registry가 승인한 경우만 프로파일 별칭 허용	서로 다른 음운 대립
표시·편집 요구	블록 처리·대체 표시·복사·접근성 시험을 포함	미지원 상태의 투명성

23.1 적용 범위

확장 프로파일은 교육·사전·학술·고유명사·접근성·언어기술 등 원어 정보 보존이 필요한 용도를 대상으로 한다. 일반 한국어 문서에 새로운 표기나 조판을 강제하지 않는다. 또한 훈민바름의 특정 자모 배당을 국가표준의 유일한 값으로 채택하도록 요구하지 않는다.

23.2 표준과 응용 규칙을 분리한다

국가표준 또는 부속 프로파일에는 허용 문자열·교환 조건·상호운용성·검증 방법을 중심으로 최소 공통 요구를 둔다. 특정 언어의 음가·자모 대응과 교육용 표시 수치는 언어모듈·응용 프로파일에서 다룬다. 기술중립성을 지키기 위해 특정 글꼴이나 특정 앱을 적합성의 유일한 기준으로 삼지 않는다.

본 제안은 일반 표기를 흔들지 않고 정보 보존이 필요한 영역의 적용 범위를 합리화하는 방향이다. 개정 형식과 대상 조항은 현행 국문판의 확인 및 공공 검토를 거쳐 정한다.

[I1] 23-24·29-33쪽의 공공성·단계적 표준화 원칙을 전산 규칙으로 구체화한 제안.

23. 개정 제안 · 문언 예시

아래 문언은 채택된 표준 조항의 인용이 아니라, 공개 검토를 위한 연구자의 예시안이다. 현행 국문판의 절 번호와 문체에 맞추는 작업은 별도로 남긴다.

확장 표기 프로파일의 자모 연쇄와 처리

- ① 정밀 음가 기록을 위한 확장 표기 프로파일은 해당 프로파일이 정한 자모값의 순서와 음절 경계를 보존하기 위하여, 하나의 음절블록 안에 복수의 초성·중성·종성 조합형 자모를 차례대로 표현할 수 있다.
- ② 프로파일은 허용 자모, 기준 직렬화, 경계 판별, 버전 식별과 수신 시 정규화 절차를 명시하여야 한다. Unicode의 정규화 정의를 변경하는 방식으로 별도의 문자값을 정하여서는 안 된다.
- ③ 단일 복합 부호와 구성 낱자 연쇄의 의미 동치 여부는 부호의 이름이나 자형만으로 정하지 않는다. 해당 프로파일이 두 표현에 같은 값과 역규칙을 명시적으로 부여하고 음운 대립의 보존을 검증한 경우에만 프로파일 내부의 동치로 인정한다.
- ④ 처리계는 하나의 정보 음절블록을 정보처리 단위로 보존하고, 표시 프로파일에 맞춰 구성 요소를 식별 가능하게 나타내야 한다. 압축 합성을 지원하지 못하면 블록 경계·선형 표시·IPA 또는 대체텍스트를 제공하고 지원 상태를 밝혀야 한다.
- ⑤ 정보 교환 형식은 표시 결과와 분리하여 보존하여야 한다. 표시를 위한 치환이나 축약은 원문, 자모값, 음소 순서와 규칙 버전을 덮어써서는 안 된다.
- ⑥ 본 프로파일의 적용은 일반 한국어 문서의 표기와 교환 방식을 일률적으로 변경하지 않으며, 다른 연구 표기체계에도 같은 검토 조건을 적용한다.

문언의 효과와 한계

①은 연쇄 표현의 경로, ②는 교환 조건, ③은 음가 대립의 보호, ④-⑤는 실제 사용자에게 전달되는 정보의 보존을 담당한다. 이 조항들을 제안한다고 해서 글꼴 합성과 음가 배당이 자동으로 완성되는 것은 아니다. 적합성 시험과 지원 범위를 함께 공개해야 한다.

24. 국제표준과의 관계와 이행 경로

24.1 2009년 의견서의 의미

미국과 Unicode Consortium의 2009년 공동 의견서 L2/09-151은 KS X 1026-1:2007을 ISO/IEC 10646에서 참조하려는 제안에 대해, 영문판 제6절의 별도 정규화 권고와 Unicode·ISO/IEC 10646의 정합성 우려를 제기했다. 국문 원문을 검토하지 못했다는 한계와, 정규화가 의무가 아닌 상황의 교환 권고라면 다른 해석이 가능하다는 점도 함께 밝혔다. [E6]

따라서 이 자료는 당시 정규화 논점이 실제로 제기되었다는 역사적 근거다. 그 우려가 2026년에도 동일한 문언으로 남아 있다거나 이미 법적 위반이 확정됐다는 증거로 사용하지 않는다.

24.2 단계적 이행

단계	산출물	다음 단계의 조건
공개 응용 프로파일	정본·registry·스키마·시험 벡터	독립 구현과 오류 제보가 가능함
기술보고서	환경별 계측·왕복·접근성 결과	재현 자료와 부정적 결과 공개
단체·공공 검토	최소 공통 규격과 용도별 적용 범위	음운·교육·문자·정보기술 검토
국가표준 검토	현행 국문판 대조와 개정·부속서안	운용 근거와 상호운용성 검증

제3본은 응용 프로파일과 검증 요구를 세우는 단계다. 제4본은 구현 결과와 그 한계를 기록한다. 연구 규격의 공개는 상위 표준 개정을 기다리지 않고 진행할 수 있으나, 표준 적합성을 선언할 때는 적용 판본·프로파일·예외를 밝혀야 한다.

24.3 공개 자료의 안내

공개 저장소 주소는 <https://github.com/ivanpark/hmbr> 이다. 제3본·제4본, 웹 참조구현과 시험자료의 최신판은 그 공개 경로에서 관리한다. 현재 배포 파일과 웹앱 동작의 감사 결과는 제4본과 2026-10-09 네 본 상태 보고서를 참조한다. 이 전산본은 설계 규격과 문서 내부의 일관성을 다룬다.

[I1] 29-31쪽; [I3] 2-3쪽; [E6] 2009년 공동 의견서 전문. 일정과 현재 배포 상태는 별도 확인 대상.

제6부 · 적합성과 판정

25. 적합성 등급

적합성은 선언한 모듈·프로파일·지원 범위 안에서 판정한다. 다음 세 등급은 시범 분류다. 상위 등급은 하위 요구를 포함하지만, 범위 밖 미결 항목까지 지원한다는 뜻은 아니다.

등급	요구	필수 보고
1 · 보존 적합	source와 구조형 보존, 유효한 참조 관계, 직렬화 가능 여부의 정확한 판정	미배당·미결 항목과 지원 범위
2 · 표시 적합	1등급 + 블록·음가 식별, 미지원 표시와 접근성 정보	압축 합성 범위와 linear/fallback 범위
3 · 왕복 적합	2등급 + 선언 범위의 구조형 왕복·평문 역변환·검색·복사 검증	복원 정보 수준, 후보·실패 및 실제 수신 환경

선형 대체 표시를 정확히 제공하는 구현은 해당 범위의 표시 요구를 충족할 수 있다. 그러나 이를 복합 블록의 압축 합성 성공이라고 부르지 않는다. 보존 적합에서도 미결 레코드를 조용히 버리거나 음소를 바꾸면 실패다.

25.1 적합성 선언의 최소 항목

규격·schema·registry·모듈 버전, 입력 데이터셋과 범위, 지원·제외 항목, 시험 결과, 글꼴·엔진·플랫폼 버전, 검사 일자와 재현 자료를 함께 공개한다. ‘모든 테스트 통과’라는 문구에는 어떤 테스트와 어떤 분모인지를 붙인다.

25.2 서로 다른 검증을 섞지 않는다

이 문서를 작성하면서 JSON과 예시 코드포인트를 검사한 결과는 문서의 내부 일관성 검사다. 그것이 웹앱의 3등급 적합성이나 전체 영어의 왕복 성공을 의미하지 않는다. 음성학적 타당성, 학습 효과와 사회적 수용성 역시 별도의 검증이다.

미결 항목을 처리하지 못한다는 사실보다, 처리하지 못한 범위를 숨기거나 다른 음가로 대체하는 것이 더 큰 보존 실패다. 적합성 보고는 성공과 함께 범위와 실패를 읽을 수 있게 해야 한다.

26. 시험 벡터와 27. 버전·변경 관리

26.1 시험 묶음

묶음	확인할 내용	예시
기본·자음군	L/V/T 순서·블록 수·자모 순서	pick, bank, friend, text
복합 음절핵	1:N 대응과 단모음 대립	boost, boy, food, full/fool 핵
복합값	N:1 대응과 두 음소 회수	cats, boots
일반 초성군	/l/ 한 번 기록	play, blue
모듈 대조	RP·GA의 서로 다른 기준형	pass, car, go
미결·미배당	원문 보존과 거부 상태	/ɜː/, catch의 종성 /tʃ/
정규화	NFC 수신과 NFD 기준형 회수	pat, bank, pass(GA)
구조 무결성	누락·중복·없는 참조·경계 불일치 검출	토큰·mapping·블록 교차 검사

부록 B의 양성 벡터와 음성 벡터를 함께 사용한다. 양성 예시만 통과한 것을 오류 검출의 증거로 삼지 않는다. 아주 짧은 자모열의 성공을 사전 전수조사 성공으로 확대하지 않는다.

27.1 버전의 대상

문서판, schema, registry, 언어모듈, 직렬화 프로파일, 표시 프로파일과 시험셋의 버전을 각각 기록한다. 문서 설명 수정과 음가 배당 변경은 영향이 다르다. 미결이 잠정 고정으로 바뀌더라도 기존 출력·역해석이 달라지면 호환성 변경으로 취급한다.

27.2 이행 원칙

구판 데이터는 source에서 새 분석을 다시 만들 수 있는지 먼저 확인한다. 원 source를 보존하고 이전 규칙과 새 규칙의 차이를 기록한다. 구판의 filler나 중복 위치를 단순 삭제하는 방식으로 이행하지 않는다. source가 없어 음소를 확정할 수 없으면 수동 검토 대상으로 남긴다.

공개 배포에는 변경 이유, 영향받는 레코드, 새 시험과 이전판 읽기 절차를 포함한다. 여기서 제안한 draft-2026-09-08 식별자는 제품의 정식 배포 버전 번호가 아니다.

28. 잠정 고정·시범·미결 목록

28.1 제2본에서 받는 기준

한 영어 음절 = 한 음절블록, 내부 음소 순서, /l/ 한 번, /tʃ/와 /t+s/의 구별, /ɪ/·/ʊ/·/ə/의 독립 대립, source IPA 보존을 따른다. 자모값의 음가를 이번 전산 설계의 편의를 위해 바꾸지 않는다.

상태	이 개정에서의 처리
잠정 고정 원칙	음운 토큰과 자모 요소 분리; 1:1·1:N·N:1 보존; 층 1과 source의 권위; 표시와 저장 분리
시범 전산 프로파일	L+ V+ T* 기준형; NFD 수신·저장 절차; mappingUnits 필드; 오류 코드; 교환 형식; 적합성 등급
시범 표시	복합 핵의 뒤 모음에 읽기용 무음 초성 ㅇ을 붙이는 선형 표시(저장 V 배열에는 미포함), /s/ 위치형 경계, GA /ɔ/ 값, 강세 상대 크기.
미결 자모·분석	/ɜ/ 기저 모음, /j/+V·/w/+V 최종 결합, [ʔ], 음절자음 최종 처리
미결 전산·표시	미배당 종성 부호, 복합 압축 자형, 동치 별칭의 개별 승인, 다중 환경 적합성

이 상태표는 초안의 적용 범위를 밝히는 것이다. 새 전산 필드와 프로파일이 이미 저자의 최종 승인을 받았다는 뜻이 아니다. 제2본이 시범·미결로 둔 항목은 이 문서에서도 해당 상태를 유지한다.

28.2 제4본으로 넘기는 것

제4본은 이 문서의 registry와 schema를 실제 데이터·코드로 옮기고, source 정규화와 음절분석, 직렬화·역변환, 표시·복사·검색·접근성을 시험한다. 사전 전수조사는 출처·라이선스·원파일·실행 버전과 함께 다시 수행한다. 앱 화면과 실제 동작의 차이는 그 구현본에서 보고한다.

그 결과가 현재의 조자법·음가 배당을 다시 검토해야 함을 보이면 제2본의 공개 판정 절차로 되돌린다. 기술 구현이 정본을 몰래 바꾸거나, 정본의 미결사항을 구현 성공으로 덮지 않는다.

[I2] 35–37·58–59쪽; [I5] Master 3, 137–140쪽. 새로운 전산 항목은 검토용 설계안.

맷음말 · 부호는 정본을 섬긴다

훈민바름 전산본의 첫 과제는 더 많은 글자를 화면에 보여 주는 일이 아니다. 어떤 원문의 어떤 소리를 어떤 분석과 규칙으로 기록했는지 잃지 않는 일이다. 정본의 자모값, 원어의 음운 단위, Unicode 부호열과 화면 자형을 나누는 이유도 여기에 있다.

하나의 /u:/가 두 모음 요소로 적히고, /t/와 /s/ 두 음소가 하나의 복합값으로 적힐 수 있다. 이 관계를 데이터에 담지 않으면 화면은 그럴듯해도 원래 소리를 다시 읽는 과정에서 정보가 달라진다. 음운 토큰과 자모 요소를 분리하는 모델은 이런 차이를 감추지 않기 위한 장치다.

한 영어 음절을 한 블록으로 보존하는 일도 같은 문제다. 영어 자음군과 복합 음절핵이 현재 글꼴에서 여러 자리로 나뉘어 보이더라도, 그 표시 사정 때문에 원어에 없는 음절을 만들어서는 안 된다. 압축된 최종 자형이 아직 미결이라면 블록 경계와 내부 순서, IPA를 통해 현재의 한계를 정직하게 보여 주어야 한다.

국가표준의 검토 역시 같은 태도를 요구한다. 국제 문자부호의 구조적 가능성, 국가표준의 적용 범위, 개별 글꼴과 엔진의 구현 한계를 하나의 원인으로 뭉뚱그리지 않는다. 직접 확인한 조항과 측정값을 밝히고, 아직 확보하지 못한 원자료와 미결사항도 함께 남긴다. 그래야 연구가 다른 사람의 검토와 재구현으로 이어질 수 있다.

이 연구가 요청하는 것은 특정 표기체계의 독점적 채택이 아니다. 일반 외래어 표기의 역할을 지키면서, 원어 음가를 정밀하게 보존할 필요가 있는 분야에서 한글의 역사적 자원과 조자 원리를 공개적으로 시험할 수 있는 공간이다. 그 공간에는 훈민바름뿐 아니라 다른 연구와 대안도 함께 들어올 수 있어야 한다.

정본은 사람이 적고 읽는 원리를 세우고, 전산본은 그 원리를 잃지 않게 보존하며, 구현본은 실제로 얼마나 지켰는지 밝힌다. 세 단계가 서로의 한계를 드러내고 수정할 수 있을 때, 가표본은 고정된 문자 목록을 넘어 계속 검증하고 발전시킬 수 있는 연구가 된다.

우리 글로 적고, 기록한 소리 정보를 다시 되찾는다.

부호는 정본을 섬기고, 구현은 그 보존을 증명해야 한다.

부록 A. 복합 음절핵과 코드포인트 식별

아래 표는 음운 단위와 저장 자모 요소열의 대응이다. 표시용 무음 ㅇ은 포함하지 않는다.

읽기용 표시는 뒤 요소에 ㅇ을 붙이며, 제2본의 음가 대응과 이 표의 부호열은 유지한다.

음운 단위	모듈	자모 요소 (각 항 JV-)	부호열 (각 항 U+)
/i:/	공통	I I	1175 1175
/u:/	공통	U U	116E 116E
/ɑ:/	RP	A A	1161 1161
/ɔ:/	RP	O O	1169 1169
/ɜ:/	RP	EO EO	1165 1165
/ɛ:/	RP	E E	1166 1166
/eɪ/	공통	E KIT	1166 119D
/aɪ/	공통	A KIT	1161 119D
/ɔɪ/	공통	O KIT	1169 119D
/aʊ/	공통	A FOOT	1161 118D
/əʊ/	RP	SCHWA FOOT	119E 118D
/oʊ/	GA	O FOOT	1169 118D
/ɪə/	RP	KIT SCHWA	119D 119E
/ʊə/	RP	FOOT SCHWA	118D 119E

GA NURSE /ɜ/는 이 완성 대응표에 넣지 않는다. /ɝ/는 source에서 normalized /əɪ/로 분석하는 규칙을 따르며, 모음 하나의 복합 nucleus로 등록하지 않는다. /ɪ/은 그 음절 분석에 맞는 자음 위치에 기록한다.

A.1 허용 부호의 이름 검사

registry의 codepoint와 Unicode 공식 이름, Hangul_Syllable_Type을 대조한다. 코드포인트 이름의 문자열 유사성은 음가 대응이나 프로파일 동치의 근거로 사용하지 않는다. 별칭과 역사적 이름이 필요한 문자는 해당 Unicode 판본의 주석도 함께 확인한다.

[I2] 13-14-22-23-35-37-50-53쪽. 구성 요소 ID는 이 개정의 시범 식별자.

부록 A. 핵심 문자 식별표

부호	단독 자모	Unicode 공식 이름
U+119D	ㅣ	HANGUL JUNGSEONG I-ARAEA
U+118D	ㅍ	HANGUL JUNGSEONG U-U
U+1182	ㅍ	HANGUL JUNGSEONG O-O
U+119E	.	HANGUL JUNGSEONG ARAEA
U+1144	ㅐ	HANGUL CHOSEONG IEUNG-PIEUP
U+114B	ㅐ	HANGUL CHOSEONG IEUNG-PHIEUPH
U+A976	ㅐ	HANGUL CHOSEONG IEUNG-RIEUL
U+D7DB	ㅐ	HANGUL JONGSEONG RIEUL-YESIEUNG
U+111B	ㅐ	HANGUL CHOSEONG KAPYEOUNRIEUL
U+11D0	ㅐ	HANGUL JONGSEONG SSANGRIEUL
U+11BE	ㅐ	HANGUL JONGSEONG CHIEUCH
U+1149	ㅐ	HANGUL CHOSEONG IEUNG-CHIEUCH
U+115F		HANGUL CHOSEONG FILLER
U+1160		HANGUL JUNGSEONG FILLER
U+110B	ㅐ	HANGUL CHOSEONG IEUNG

단독 자모 열은 코드 식별용이다. 독립 음절블록이나 새 정본 예시가 아니다. 표에서 글리프를 홀로 보여 주는 일과 기준형 자모열을 직렬화하는 일은 구분한다.

특히 구별할 두 쌍

U+118D (FOOT /ʊ/) ≠ U+116E U+116E (/u:/)

U+1182 (GA /ɔ/) ≠ U+1169 U+1169 (RP /ɔ:/)

어느 쪽도 이름이나 모양만으로 자동 치환하지 않는다.

이름과 코드값은 이번 실행 환경의 Unicode 데이터에서 대조했다. 확인 대상은 오래전부터 배당된 본 표의 문자들이며 최신판 전체의 변경 검사는 아니다.

부록 B. 시험 벡터와 기대 판정

§7의 25개 직렬화 예시는 양성 벡터다. RP·GA 공통 표시는 두 모듈에 같은 예시 부호열을 적용한다는 뜻이다. 아래 항목은 그 예시만으로 드러나지 않는 보존 실패와 미결 처리를 추가로 검사한다.

시험	입력·조건	기대 판정
B-01	boost /u:/의 mapping	음운 토큰 1개, 자모 요소 2개
B-02	cats의 /t+s/	음운 토큰 2개, JV-TS 1개
B-03	play·blue	초성군의 JV-L이 1개
B-04	FOOT 부호와 U U 연쇄	두 값의 동치 비교는 거짓
B-05	GA AW와 RP O O 연쇄	두 값의 동치 비교는 거짓
B-06	about 평문만 역변환	강세·출처를 복원했다고 선언하지 않음
B-07	catch 중성 JV-CH	구조 보존, 평문 unavailable
B-08	GA /3~/	source 보존, 기저 모음 unresolved
B-09	음절자음만 있는 nucleus	모음 삽입 없이 지원 범위 밖으로 보고
B-10	filler가 포함된 기준형	E-CHAR
B-11	T 다음 V 또는 V 없는 L	E-ORDER
B-12	없는 mapping 요소 참조	E-REFERENCE
B-13	같은 입력 음소의 중복 소비	E-COVERAGE
B-14	NFC(bank)를 수신	원입력 보존, NFD 후 원 코드열 회수
B-15	모듈·버전 없는 평문	필요 정보 요청 또는 후보 상태 유지

표의 기대 판정은 구현 요구이며 이 책을 작성한 것만으로 웹앱에서 모두 통과한 것은 아니다. 문서 내부 검사는 부호열·경계·대립·JSON 예시의 일관성을 확인한다. 사용자 인터페이스와 실제 복사·낭독 동작은 제4본의 시험 결과를 필요로 한다.

부록 C. 측정 방법과 재현 절차

C.1 이번 세이핑 계측의 식별 정보

항목	값
측정일	2026-09-08
글꼴	NanumMyeongjoOTF-YetHangul.otf
파일 판	Version 1.0.0.9 Build 20140917
글꼴 좌표	unitsPerEm = 1000
엔진	HarfBuzz 8.3.0
방식	HarfBuzz C API, OpenType font functions
버퍼	UTF-8 입력, segment properties 자동 추정
추가 기능	사용자 지정 feature 없음, 엔진 기본값
폭 기준	가 U+AC00의 advance 합 = 950
정규화	입력 코드포인트 열을 직접 구성

글꼴 파일 SHA-256:

```
c2e3217e63456a8cdcfce5670692d5ec  
78d7b1d1396f6968bef70ddb72dc1bb
```

위 해시는 줄바꿈 없이 이어 붙인 64자리 값이다. 시험 입력의 정확한 코드포인트는 §7과 §13의 대조표에 있다. 동일한 글꼴 이름이어도 파일 해시가 다르면 같은 실험 조건으로 취급하지 않는다.

C.2 재현 단계

동일한 해시의 글꼴을 로드하고 `hb_face_create`와 `hb_font_create`로 font를 만든다.

`hb_ot_font_set_funcs`를 적용하고 font scale을 (1000,1000)으로 둔다. 각 코드포인트 열을 UTF-8로 만들고 새 buffer에 넣어 `hb_buffer_guess_segment_properties`를 호출한다. 사용자 지정 feature 없이 `hb_shape`를 실행한다.

`hb_buffer_get_glyph_positions`에서 `x_advance`를 합하고 기준값 950으로 나눈다. 글리프 수, 코드포인트 열, advance 합과 비율을 함께 저장한다. 자형 판정에는 실제 화면 또는 페이지 렌더를 별도로 확인하며, 1자리 폭만으로 정상 HMBR 표시를 선언하지 않는다.

이번 환경의 직접 계측. Linux 공유 라이브러리 HarfBuzz C API를 Python ctypes에서 호출했다.

부록 C 계속 · 사전 전수조사 재현 기록

C.3 재현 환경과 명령

묶음	기록할 정보
자료 식별	사전명·판·날짜·원파일·해시·표제어 수·변이 처리
이용조건	출처·권리자·라이선스·재배포 허용 범위
발음 모듈	RP·GA 표기 관행, 비권설·rhotic·길이·강세 처리
전처리	기호 치환·IPA 분해·음절자음·제외 조건과 로그
매핑	registry 버전, 위치형, /t+s/ 소비 규칙
범주	배당된 구성값 / 미배당 구성값 / 분석 미결 / 오류
계수	자음군 종류 수·출현 수·표제어 수를 분리
재현성	실행 코드 버전·환경·명령·원 집계표

저장소 루트에서 `node tools/cluster_census.mjs`를 실행한다. 스크립트는 `data/dict-rp.json`과 `data/dict-ga.json`을 읽고 `js/registry.js`의 현재 배당을 적용한다. 결과에는 표제어·해석 수, 조성·종성 군집, 부호 있음·없음·미배당, 복합값 /t+s/, 새 겹낱자 수, 상위 60종 누적률과 2요소 음절핵을 함께 출력한다.

C.4 이번 문서의 내부 검사와 사전 재현 범위

§7의 25개 예시에서 허용 자모와 L/V/T 배열, 블록 경계를 검사한다. 각 자모열의 NFD 안정성과 NFC 뒤 NFD 복원을 검사하고, FOOT/U U 및 AW/O O 대립을 확인한다. §3 JSON의 참조와 boost 직렬화가 대응하는지도 검사한다.

이 검사는 예시의 내부 일관성과 현재 사전 집계 of 재현성을 확인한 것이다. 독립 변환 엔진의 영어 전범위 발음 정확도나 사전 원출처의 권리 상태를 대신하지 않는다. §13의 세이핑 계측과 §21의 사전 집계는 서로 다른 자료·명령·판정 범위를 가진다.

C.5 공개 기록과 인용 조건

제4본의 적합성 보고에는 성공·실패·미결 결과와 전체 분모를 함께 남긴다. §21 수치를 인용할 때에는 저장소 커밋, 사전 파일, `tools/cluster_census.mjs`와 실행일을 함께 적는다. 사전판·라이선스가 확정되기 전에는 이 결과를 영어 전체의 절대적 규모나 법적 결론으로 확대하지 않는다.

부록 D. 용어 대조

이 문서의 용어	이름	뜻
원문 IPA	source IPA	출처의 기호와 문자열을 보존한 IPA
정규화 IPA	normalized IPA	언어모듈이 매핑에 사용하는 분석형
음운 토큰	phoneme token	모듈이 정한 음운 단위
대응 단위	mapping unit	음운 토큰열과 자모 요소열의 관계
자모값 ID	jamo value ID	등록부의 자모 요소·복합값 식별자
음절블록	syllable block	선택된 영어 음절 하나를 담는 정본 단위
초성군	onset[]	블록 앞 자음 요소. 없으면 무음 초성 사용
음절핵	nucleus[]	한 음절을 이루는 모음 구성 요소
종성군	coda[]	블록 뒤 자음 요소의 순서 있는 배열
위치형	positional form	같은 자모값의 자리별 코드·모양
직렬화	serialization	구조를 정해진 순서의 교환 표현으로 변환
셰이핑	shaping	문자열을 글리프 선택·배치로 처리
기준형	profile canonical form	본 응용 프로파일이 선택한 저장형
정준 동등성	canonical equivalence	Unicode가 정한 문자열의 동등성
의미 동치	profile equivalence	모듈·버전이 승인한 의미 관계
대체 표시	fallback	미지원 환경에서 정보 보존을 돕는 출력

‘정본’이라는 연구상의 권위와 Unicode의 canonical이라는 기술 용어를 혼동하지 않는다. 정본의 음가 규칙은 제2본을 따르고, Unicode 정준 동등성은 Unicode 규칙을 따른다. 이 책의 기준형은 응용 프로파일이 선택한 보존 형식이다.

참고문헌 및 작성 근거 · 내부 자료

아래 자료는 이번 개정에서 직접 열람한 근거다. 페이지는 원본 PDF의 물리적 쪽수 기준이며, 본문에 인쇄된 차례 번호가 다른 경우 PDF 쪽수를 우선한다.

[I1] 제1본 선언본

Park Il Hwan, 「훈민바름 영어확장판 · 제1본 선언본」, 제출용, 원표지 날짜 2026-08-14, 36쪽. 2026-09-08 제공본에서 출발했으며 현행 기준은

HMBR_Book1_Seoneonbon_20261009.pdf(2026-10-09 교정판)다. 문서 지위·정밀 표기층·공공성·표준화·한계의 표현 기준.

[I2] 제2본 해례본

Park Il Hwan, 「훈민바름 영어확장판 · 제2본 해례본」, 제출용, 원표지 날짜 2026-08-14, 59쪽. 현행 기준은 HMBR_Book2_Haeryebon_20261009.pdf(2026-10-09 교정판)다. 최초

업로드의 ‘정본’ 대신 문서 내부 명칭인 ‘해례본’을 따른다. 음가 배당과 현 시범 조자법·상태의 기준.

[I3] 제안서와 커버레터

「규제합리화 대국민 공모전 제안서 작성본(1)」, 1쪽; 「심사위원 여러분께(3)」, 4쪽. 정밀 표기층의 적용 범위 합리화와 4본의 역할·공개 계획을 확인했다.

[I4] 이전 전산본과 작성 보고

「훈민바름 영어확장판 · 제3본 전산본 260807」, 12쪽.

HMBR_BOOK3_JEONSANBON_A4.docx와 HMBR_BOOK3_PROOF.pdf, 2026-08-14, 41쪽. HMBR_BOOK3_REPORT.md, Claude, 2026-08-14. 이 가운데 41쪽 후속판의 조직과 판형을 개정 기반으로 삼았다. 과거 사전 집계는 2026-09-16 저장소의 스크립트로 재현했다.

[I5] 이전 대화 기록

「바르미 Master 3 / 공모전 2」, 140쪽; 「바르미 Master 2 / 공모전」, 623쪽. 전산본 관련 결정과 마지막 수정 지시를 검색·대조했다. 특히 Master 3의 132-140쪽에서 후속 원고 평가, mapping unit, /I 한 번과 의미 동치 수정 지시를 확인했다. 대화에 언급된 과거 첨부물 전부를 복원한 것은 아니다.

이번 개정의 JSON 필드, 오류 코드, 교환·역변환 조건과 적합성 요구는 위 자료를 전산적으로 정합화하기 위한 설계안이다. 원 저자의 새로운 음성학 판정을 대신하지 않는다.

참고문헌 및 작성 근거 · 외부 자료

[E1] Unicode 17.0.0, Chapter 3

§3.12 Conjoining Jamo Behavior, D133–D134. 음절블록 정의와 한글 정준 분해·합성.

<https://www.unicode.org/versions/Unicode17.0.0/core-spec/chapter-3/>

[E2] UAX #15: Unicode Normalization Forms

정준·호환 동등성, NFC·NFD·NFKC·NFKD의 구분.

<https://www.unicode.org/reports/tr15/>

[E3] Microsoft, Developing OpenType Fonts for Korean Hangul Script

등록 조합 전처리, Uniscribe 설명, ccmp·ljmo·vjmo·tjmo. 모든 엔진의 현재 동작을 대표하지 않는다.

<https://learn.microsoft.com/en-us/typography/script-development/hangul>

[E4] UAX #29: Unicode Text Segmentation

Unicode 17.0.0, Revision 47. GB6–GB8과 한글 분절.

<https://www.unicode.org/reports/tr29/tr29-47.html>

[E5] N3422, KS X 1026-1:2007 영문 번역본

한국 측 공개 제출 자료. PDF 5쪽의 지위 표기와 12–14쪽 관련 조항. 국문 정식본을 대신하지 않는다.

<https://www.unicode.org/L2/L2008/08225-n3422.pdf>

[E6] L2/09-151 / N3633

US / Unicode Consortium Comments on N3611, 2009-04-21, 2쪽. 당시 정규화 우려와 해석 한계를 확인했다.

<https://www.unicode.org/L2/L2009/09151-us-utc-cmt-n3611.pdf>

외부 자료 열람일: 2026-09-08. 법령의 구속력, 현행 KS 국문판과 저장소의 현재 배포 상태는 이 문서의 확인 범위 밖이다.

원형 토큰과 t 표면 주석의 보존

2026년 10월 12일 개정 부록. 제3.6절과 모델·교환 논의에 연결 발음 연습층의 실제 구현을 보충한다. 기존 Broad 정보, 음가 배당, 블록 구조와 NFD 직렬화는 변경하지 않는다. Narrow 전체 스키마의 완성을 뜻하지 않는다.

서로 다른 세 층

층	보존하거나 생성하는 정보
원형 records	입력 IPA, 출처, phonemeTokens, canonicalModel, 정보 직렬화
파생 records	지원 음소만 달라진 약형·동화·탈락을 기존 엔진으로 변환한 연습형
표면 surface_records	[ʔ], [ŋ], [tʰ] 등을 포함할 수 있는 IPA와 훈민바름 참조 상태

표면 IPA가 기본 엔진에 없는 세부 음을 포함해도 그 기호를 원형 자모열에 끼워 넣지 않는다. button을 [ˈbʌʔŋ]으로 선택하면 원형 /bʌtən/ 레코드를 유지하며 display_status는 broad-reference다. winter의 t 탈락처럼 지원 음소만 남는 경우에는 연습용 파생 훈민바름을 만들되 원형 records를 별도로 보존한다.

t 주석의 필드

필드	의미
word_index	원형 단어 배열의 0부터 시작하는 위치
token_ids	해당 원형 canonicalModel.phonemeTokens의 ID
underlying_ipa	영향받은 원형 토큰의 IPA
surface_ipa	선택한 실현. 탈락이면 빈 문자열
feature_id	ntdrop, glottal 또는 unreleased
environment / source	적용 환경 설명과 참고 자료 URL

button의 /tən/을 [ʔŋ]으로 바꾸는 선택은 원형 tən 토큰 세 개를 가리킨다. 표면 [ŋ]을 새 Broad 음소로 오인하지 않는다. 겨냥한 원형 토큰이 남으므로 선택을 해제할 때 추측으로 복원할 필요가 없다. 이 필드는 이번에 추가한 t 변이에 제공하며 기존 연결 규칙 전체의 토큰 수준 주석을 완성한 것은 아니다.

내보내기와 일관성 규칙

연음화 패널을 연 뒤 전체 기록 저장을 누르면 원형 records에 connected_speech를 추가한다. 일반 변환만 한 결과에는 이 필드를 넣지 않는다. version은 connected-speech-t-variants-20261012다. accent, pace, emphasis, pauses, candidates, applied, 원형과 선택형 IPA, 파생 records 및 surface_annotations와 surface_records를 담는다. GA 단어 하나라도 t 후보가 있으면 이 연습층을 저장할 수 있다.

surface_records 필드	해석
ipa	선택 후 표면 문자열
hmbr_reference_ipa	현재 훈민바름이 참조하는 IPA
display_status	derived 또는 broad-reference
display_notes	미정 자형·미파열 주석 등 화면 표시의 이유

derived는 선택 후의 지원 가능한 IPA로 표시할 수 있다는 뜻이며 항상 원형과 다른 문자열이라는 뜻은 아니다. broad-reference이면 훈민바름만 보고 표면의 [ʔ][n]까지 확정해 읽어서는 안 된다. 표면 문자열에는 원형에서 유지한 음절 점이 남을 수 있으며 완전한 재음절화·운을 분석을 뜻하지 않는다.

충돌과 되돌리기

같은 원형 t의 탈락·성문음화·미파열은 하나의 선택 그룹으로 묶는다. 수동 선택에서는 마지막 선택이 앞 선택을 대체한다. 빠른 회화에서 자음군 탈락과 성문음화가 모두 추천될 수 있는 경우 성문음화의 추천 우선순위를 높인다. 이는 인터페이스의 단순화이며 복합 조음 자체를 금지하는 음성학 규칙이 아니다.

쉽은 해당 단어 경계의 연결 부호를 막는다. 쉽 앞 성문음화·미파열은 별도의 수동 선택이다. 입력이나 발음 기준을 바꾸면 이전 선택을 버린다. 원형 자모열 복사는 기존 원형만 복사한다. 표면 IPA나 연결 부호를 정본의 무손실 직렬화로 바꾸지 않는다.

검증 범위와 음성 연계

새 시험은 원형 불변, 토큰 ID의 존재, 세 토큰을 가리키는 button 주석, 강제 조건, 쉽, 선택 충돌, 실제 파생 자모의 변화를 검사한다. 역변환은 surface_ipa만 보고 원형을 추정해서는 안 된다. [r]에서 t/d를 복구하기 어려운 기존 문제와 동일한 보존 원칙을 따른다.

다음 TTS 단계는 이 저장된 선택을 음성 요청에 연결하고 반환된 발음이 요청과 일치하는지 별도로 확인해야 한다. 현재 내보내기 구조에는 합성된 오디오나 발음 일치 보증을 넣지 않는다. 외부 JSON 가져오기와 완전한 Narrow 변환도 별도 과제로 남긴다.

훈민바름 구현본

웹 참조구현·시험·사용 안내

영어확장판 제4본 · 바르미 웹 공개 참조구현 사용해례

정본의 원리를 실행하고, 보존한 정보와 남은 한계를 확인하는 방법

2026년 10월 12일 개정판 · 연구·검토용

Park Il Hwan

연구·검토용 비구속 문서

우리 글, 바른 소리

훈민바름 영어확장판 4본의 구성

훈민바름 영어확장판은 하나의 체계를 네 역할로 나눈다. 선언본은 뜻과 공공적 필요를 밝히고, 해례본은 정본의 구조와 발음 원리를 풀며, 전산본은 정본의 디지털 보존·처리 규격을 세우고, 구현본은 실제 프로그램과 시험으로 그 작동을 확인한다.

문서	담당 범위	이 책과의 관계
제1본 선언본	뜻과 미래, 공공적 필요	목적과 연구 지위를 받는다.
제2본 해례본	영어 RP·GA 음가와 블록 문법	무엇을 보존할지 정한다.
제3본 전산본	구조·직렬화·표시·교환 규격	구현과 시험의 기준이다.
제4본 구현본	웹 사용·실행·시험·공개 기록	실제 구현 범위와 차이를 밝힌다.

이 문서의 지위

‘훈민바름’은 옛한글·확장 한글 자모를 활용한 「외국어 표기법」의 가능성을 연구·개발하는 동안 사용하는 내부 코드명이다. 현재 영어확장판은 가표본(working specimen) 연구이자 참조 실험(reference experiment)이며, 그대로 채택하도록 요구하는 공적 규범이 아니다.

이 책은 2026-08-07 제4본의 12개 장과 부록 A·B를 계승하며 2026-09-16·17 교정과 2026-10-09 후속 갱신을 반영한다. 본문의 9월 16일 시험 기록은 당시 기준판의 결과이며 최신 코드판과 검증 기록은 부록 D를 따른다. 구판의 규칙이나 시험 수치를 최신 구현의 사실로 옮기지 않는다.

읽는 길

처음 쓰는 독자는 §3과 §7부터, 수업에 활용하는 교사는 §6·7·8부터 읽을 수 있다. 개발자는 §4·5·9·10과 부록의 재현 조건을 함께 확인한다. 공개 저장소에는 문서와 코드가 함께 있으며, 실제 웹사이트의 서비스 상태는 저장소 파일 공개와 별도로 확인한다.

현재 구현은 제3본 기준의 Broad 시범 변환기다. 전체 규격의 적합성 인증, 전 사전 발음 검수 완료 또는 교육 효과 입증을 뜻하지 않는다.

출처: 제1본 2쪽, 제2본 2:25–37쪽, 제3본 2:37쪽. 문서판과 코드판은 부록 C 참조.

차례

항목	쪽
머리말 · 1. 구현본의 목적과 권위 관계	4
2. 현재 참조구현의 구성	5
3. 사용자가 거치는 기본 흐름	6
계속 · 복사·저장·입력 오류의 처리	7
4. 변환 파이프라인	8
5. 출력층과 화면 표시	9
6. RP·GA와 Broad·Narrow 제어	10
7. 핵심 시험어로 읽는 구현	11
계속 · 교사를 위한 사용례	12
8. 글꼴·셰이핑과 대체 표시	13
9. 자동시험·전수감사·재현성	14
계속 · 개발자가 실행하는 재현 절차	15
10. 현 구현과 전산 규격의 정합화 과제	16
11. 공개 배포와 버전 확인	17
계속 · 오류 제보와 문제 해결	18
12. 향후 다언어 구현 원칙	19
부록 A. 적합성 체크리스트	20
부록 B. 핵심 시험 벡터	21
부록 C. SOURCE AUDIT와 작성 근거	22
부록 C 계속 · 판본 차이와 후속 갱신	23
부록 D. 2026-10-09 구현 상태 갱신	24

쪽수는 이 개정판의 PDF 물리적 쪽수를 따른다. ‘현재 구현’은 부록 D의 기준판을 가리킨다. §8.3.9와 부록 C의 2026-09-16 기록은 당시 기준판의 결과다. 이후 main이 변경되면 해당 커밋과 검증 기록을 함께 확인한다.

머리말 · 구현은 규격을 대신하지 않는다

훈민바름 웹앱의 역할은 원어의 소리와 정본의 구조가 디지털 환경에서 얼마나 충실하게 보존되는지 드러내는 데 있다. 입력한 IPA가 어떤 음운 단위로 분석됐는지, 그 단위가 어떤 자모값과 위치형에 연결됐는지, 화면과 복사 결과가 무엇을 담고 무엇을 담지 못하는지를 확인할 수 있어야 한다.

화면이 그럴듯하다고 정본이 옳은 것은 아니다. 부호열이 규칙에 맞는다고 사전의 발음이 검수된 것도 아니다. 구현본은 이 차이를 구체적인 입력과 실행 결과로 설명한다. 미결을 미결로 남기는 결과도 연구 기록의 일부다.

1. 구현본의 목적과 권위 관계

1.1 규칙을 받는 방향

음가·자모값·손글씨 블록 문법은 제2분을 따른다. 구조 데이터와 Unicode 기준형은 제3분의 시범 규격을 따른다. 앱이 구현의 편의를 위해 고른 음절 추정법이나 화면 배치는 별도의 구현 선택으로 밝힌다. 이런 선택이 제2분의 새로운 최종 판정이 되는 것은 아니다.

예를 들어 현재 앱은 음절 경계가 없는 다중 모음핵 입력에 최대 초성 추정 휴리스틱을 사용한다. 이때 검토 필요 상태를 반환한다. 반면 IPA에 점으로 주어진 경계는 우선 보존한다. 같은 단어라도 분석 경계가 달라질 수 있는 이유를 이 기록에서 확인한다.

1.2 구현이 밝혀야 할 네 가지

첫째, 어떤 입력과 모듈을 사용했는가. 둘째, 어떤 규칙과 버전이 그 입력을 처리했는가. 셋째, 원문·강세·대응 관계가 어느 출력에 남아 있는가. 넷째, 아직 구현하지 못했거나 판정하지 않은 부분은 무엇인가.

구현과 문서가 다르면 입력, 등록부, 분석, 직렬화, 표시의 순서로 차이를 찾는다. 승인된 규칙 변경은 문서·코드·시험 벡터에 함께 반영한다. 기존 결과를 새 규칙의 정답으로 삼아 시험만 바꾸지 않는다.

출처: 제3분 3-8:37-38쪽, 새 앱 js/engine.js. 음절 추정법은 이번 웹앱의 구현 선택.

2. 현재 참조구현의 구성

새 앱은 정적 HTML·CSS·JavaScript와 첨부 JSON 사전, 옛한글 웹글꼴로 구성된다. 변환을 위해 API 키나 외부 AI 호출을 사용하지 않는다. 브라우저가 필요한 사전 파일을 읽고, 변환 엔진이 그 IPA를 처리한다. 영어 철자에서 새 발음을 생성하는 모델은 포함하지 않는다.

파일·디렉터리	역할
index.html, css/style.css	입력·발음 기준·결과·문서 링크
js/app.js	사용자 입력, 결과 표시, 복사·JSON 저장
js/dict.js	문서 예시·강세 교정 자료·사전 원문 조회
js/registry.js	자모값·위치형·모음 구성값
js/engine.js	IPA 분석, 대응, 직렬화, 구조 검사
data/examples.json	출처를 표시한 문서 예시
data/ipa-corrections.json	일부 강세 교정 자료; 전체 발음 검수와 구분
data/dict-ga.json, dict-rp.json	기존 ZIP에서 계승한 발음 사전
fonts/	나눔명조·나눔바른고딕 옛한글 WOFF
tests/, tools/	자동시험, 정적 검사, 배포 파일 생성
docs/, legacy/	문서 및 비교용 이전 앱

2.1 사전 자료의 지위

첨부 GA 사전은 125,004개, RP 사전은 135,495개 항목을 담는다. 이는 파일의 항목 수이며 고유한 어휘 수나 검수된 발음 수를 뜻하지 않는다. 기존 자료는 GA를 CMU 계열, RP를 2023년 일부 기계 생성·검수 전 자료로 설명한다. 원 사전의 모든 IPA와 변이를 이번 개정에서 새로 검수하지 않았다.

문서 예시·일부 강세 교정 자료·일반 사전 조회는 화면의 출처 표지로 구별한다. ‘문서 예시’는 책에 수록된 참조 입력이라는 뜻이며 전체 사전의 품질 보증이 아니다. 일부 강세 교정도 전체 IPA 검수 완료란 뜻하지 않는다. 수업용 새 어휘는 교사가 실제 source IPA와 선택 모듈을 먼저 확인한다.

2.2 이전 앱의 위치

legacy/index.html은 v2.0 계열의 비교용 화면이다. 원 코드의 Broad·Narrow·연음·조음 안내를 보존 하되 현재 제3본의 적합 출력과 구분한다. 구버전의 동작 확인은 구버전 자체의 회귀시험이며 새 규격의 통과 증거가 아니다.

출처: 앱 README, js/dict.js, data/ 및 legacy/README.md. 항목 수는 이번 원파일 직접 계수.

3. 사용자가 거치는 기본 흐름

3.1 영어로 찾기

‘영어로 찾기’를 선택하고 단어나 짧은 문장을 입력한다. ‘발음 기준’에서 **RP** 또는 **GA**를 선택한 뒤 ‘훈민바름으로 보기’를 누른다. 입력란에서는 **Ctrl+Enter** 또는 **Command+Enter**로도 변환할 수 있다. 처음에는 ‘책 속의 예시’에 있는 **boost, play, friend, cats, about**를 차례로 확인하는 편이 결과 구조를 익히기 쉽다. 영어 조회는 선택한 모듈의 문서 예시, 일부 강세 교정 자료, 첨부 사전의 순서로 찾는다. 일반 문장은 어휘별 결과의 배열이며 연결발화를 자동 재현한 결과가 아니다.

3.2 IPA 직접 입력

‘IPA 직접 입력’을 선택하고 **Broad IPA**를 넣는다. 전체를 **//**로 감싸거나 기호 본문만 입력할 수 있다. 공백은 단어 사이, 점(.)은 음절 사이 경계로 사용한다. 예를 들어 **/ə.'bʌʊt/**는 두 음절과 둘째 음절의 제1강세를 명시한다.

입력	확인할 사항
/bu:st/	한 모음핵, 한 음절블록
/ə.'bʌʊt/	두 블록, 둘째 블록의 제1강세
/pɪk kætʃ/	두 단어, 두 개의 레코드
/vɛʃi/	음절 경계 자동 추정과 검토 필요 표시
/kætʃ/	종성 위치형 미배당, 직렬화 미결

3.3 결과를 읽는 순서

먼저 원문 IPA와 출처 표지를 확인한다. 그다음 **[]** 한 쌍을 한 음절블록으로 읽는다. 자음군이나 긴 모음이 여러 요소로 보이더라도 같은 괄호 안이면 하나의 영어 음절이다. 끝으로 **IPA · 자모 대응 · 코드포인트**를 열어 입력 음소와 자리별 부호를 대조한다.

‘시범 변환’은 표시 완료의 상태다. ‘검토 필요 · 시범 변환’이면 자동 음절 추정 또는 미검수 사전 등의 조건을 함께 살핀다. ‘미결 항목 있음’이면 그 아래 사유를 먼저 읽는다.

출처: index.html, js/app.js, js/dict.js, js/engine.js. 화면 라벨과 동작은 코드에서 확인했으며 브라우저 조작 검증과 구분한다.

계속 · 복사·저장·입력 오류의 처리

3.4 자모열 복사

‘자모열 복사’는 NFD 조합형 자모열을 복사한다. 화면의 괄호, 크기, 출처 표지는 포함하지 않는다. 단어 하나의 기준형 안에는 편집용 공백이 없다. 여러 단어를 함께 복사할 때만 단어별 자모열을 공백으로 연결한다.

전체 결과에 미결 단어가 있으면 전체 자모열 복사 버튼을 비활성화한다. 누락된 음소를 빼고 나머지만 완성 결과처럼 복사하는 일을 막기 위한 동작이다. 자동 클립보드 쓰기가 실패하면 수동으로 선택해 복사할 입력란을 표시하도록 구현돼 있다. 이 경로의 실제 브라우저별 동작은 별도 시험 대상이다.

3.5 전체 기록 저장

‘전체 기록 저장’은 `hmb-r-records.json`을 내려받도록 구현돼 있다. 원입력, 선택 모듈, 원문 IPA, 정규화 이력, 강세, 음소와 자모의 대응, 직렬화 상태와 미결 사유가 기록된다. 수정·재검토·수업 자료 관리에는 평문만 보관하는 것보다 이 레코드가 적합하다.

저장 파일명은 고정이므로 내려받은 후 단어·모듈·검토일을 포함하는 이름으로 구별해 보관한다. 현재 화면에는 이 JSON을 다시 불러오는 가져오기 기능이 없다. ‘저장 가능’과 ‘앱에서 재열기 가능’을 혼동하지 않는다.

3.6 입력 오류와 미결의 차이

영어 조회에서 사전에 없는 단어가 하나라도 있으면 문장 전체의 변환 결과를 만들지 않고 해당 단어를 알린다. IPA를 알고 있다면 직접 입력으로 바꿀 수 있다. 이때도 모듈·음절 경계·출처를 별도로 확인한다.

IPA 직접 입력은 단어별 레코드를 만든다. `/pIk kætʃ/`처럼 일부가 미결이면 가능한 단어와 미결 단어를 함께 보여 주며 전체 JSON에는 두 상태가 모두 남는다. 한 번에 100단어, 전체 입력 5,000자, 단어별 IPA 1,000자의 제한이 있다.

새 입력을 편집하면 이전 결과와 저장 가능 상태를 무효화한다. 입력문과 오래된 결과가 다른데도 같은 결과로 복사되는 일을 줄이기 위한 처리다.

출처: `js/app.js`의 `doConvert-invalidate`·복사·저장 처리, `js/engine.js`의 `convertPhrase-convertWord`.

4. 변환 파이프라인

4.1 원문에서 음절분석까지

영어 입력이면 사전 조회로 source IPA를 얻는다. IPA 직접 입력이면 사용자의 기호열을 원문으로 받는다. 이후 기호 별칭과 길이 부호를 처리하면서 바꾸기 전후와 규칙 식별자를 `normalization_rules`에 남긴다. GA /æ/는 원문을 보존하고 /ə/로 분석한 이력을 기록한다. /ɜ/의 미결 기저 모음을 같은 방식으로 임의 보충하지 않는다.

토큰화는 /tʃ/와 /dʒ/ 같은 등록 파찰음을 하나의 음운 단위로 인식한다. 음절분석은 입력 경계·강세 표지와 현재의 추정 규칙을 사용한다. 여러 음절의 자동 분석에는 검토 필요 상태를 남긴다.

4.2 음소와 자모 요소의 대응

`canonicalModel`은 `phonemeTokens`, `mappingUnits`, `blocks`를 가진다. 음소를 곧바로 한 글자와 동일시하지 않는다. /u:/ 하나는 두 JV-U 요소로, 종성의 /t/와 /s/ 두 토큰은 JV-TS 하나로 연결된다. 각 블록은 초성군·음절핵·종성군의 요소 참조를 순서대로 가진다.

boost의 단위	소비한 음소	출력 자모 요소
m1	/b/	JV-B
m2	/u:/	JV-U, JV-U
m3	/s/	JV-S
m4	/t/	JV-T

boost의 블록은 onset에 m1의 첫 요소, nucleus에 m2의 두 요소, coda에 m3·m4의 첫 요소를 둔다. 같은 입력 토큰을 중복 소비하거나 같은 출력 요소를 두 자리에 다시 두면 구조 검사에서 거부한다.

4.3 직렬화와 반환

모든 위치형이 배당되고 분석 문제가 없으면 L+ V+ T* 자모 연쇄를 만든다. 미배당이나 미지원 입력이 있으면 단어 전체 `serialization.text`를 null로 두고 원문과 구조, 사유를 반환한다. 출력의 `status`·`필드명`은 제3본 문서 예시와 일부 다르므로 §10의 대응표를 함께 읽는다.

출처: 제3본 6-13-17쪽; `js/engine.js`의 `tokenize-syllabify-validateModel-convertWord`.

5. 출력층과 화면 표시

현재 화면은 linear-trial-v8-vowel-carrier 표시 프로파일을 쓴다. 엔진 레코드의 display는 conjoining-jamo-v1이며 화면에서 요소별 역할과 크기를 계산한다. 화면용 core-nucleus-tail-loose-consonant 역할과 크기는 이 저장 필드에 추가하지 않는다.

실제 필드	담는 정보	용도
source, normalization_rules	원철자·원문 IPA·출처·정규화 이력	원자료 확인
canonicalModel	음소·대응 단위·음절·강세	구조 보존
serialization	프로파일·상태·자모열·코드포인트	복사와 부호 검사
display	엔진 표시 입력 프로파일, 블록별 onset·nucleus·coda, 강세·IPA	화면 표시의 입력과 대체 정보
analysis, issues, warnings	분석 방식·미결·경고	검토와 오류 추적

5.1 한 블록과 여러 글리프

주강세·부강세 핵은 100%, 무강세 핵은 75%, 뒤 모음은 50%, 독립 자음은 75%다. 부강세가 있는 단어만 주강세를 125%로 키운다. 뒤 모음에 읽기용 무음 ㅇ을 붙인다. 부우스트은 부(100)·우(50)·스트(75)이며 새 음절을 더하지 않는다.

종성 /s/와 /t+s/는 한국어 받침으로 합쳐 숨기지 않고 독립 표시 요소로 남긴다. /t+s/의 정본 NFD는 두 음소를 복합값 JV-TS로 대응하되 /tʃ/와 구별한다. 표시 프로파일을 바꾸어도 canonicalModel과 NFD 직렬화는 그대로 유지한다.

5.2 크기와 보존 정보

지원되지 않는 합성이나 미배당 자모는 다른 음가로 대체하지 않는다. 블록 경계, 선형 요소, IPA와 상태 설명을 제공해 사용자가 미결 또는 대체 표시임을 알 수 있게 한다.

현재 구현은 boost·pass·cats·play·car·go 등 핵심 예시에서 표시 요소와 복사 가능한 정본 NFD를 분리한다. 실제 브라우저·보조기술별 가독성은 계속 검증해야 한다.

근거: js/app.js, js/display.js, js/optical.js와 표시 시험. 2026-09-17 저자 결정: 무음 초성 ㅇ은 표시용이며 canonicalModel과 정본 NFD는 유지한다.

6. RP·GA와 Broad·Narrow 제어

6.1 모듈을 바꾸는 것은 출발 발음을 바꾸는 일

RP·GA는 서로 다른 참조모듈이다. 영어 조희에서 모듈을 바꾸면 그 모듈의 문서 예시·강제 교정 자료 또는 사전 IPA를 다시 읽는다. 화면 자모만 바꾸는 기능이 아니다. 직접 입력에서는 사용자가 넣은 IPA가 선택한 모듈과 맞는지 확인한다.

예시	RP 문서 입력	GA 문서 입력
go	/gəʊ/	/goʊ/
car	/kɑ:/	/kɑɪ/
pass	/pɑ:s/	/pæs/

이 표는 수록된 문서 예시의 차이다. 모든 영국·미국 화자가 이 한 가지 발음을 사용한다는 주장이 아니다. 사전의 발음 변이 전체를 선택하는 화면도 현재는 없다.

6.2 새 앱의 Broad 범위

새 앱은 Broad 변환을 제공한다. []로 감싼 Narrow 입력, 미등록 발음기호나 지원하지 않는 발음 부호를 완성된 Broad 출력으로 바꾸지 않는다. /j, w/는 기저값 명시형을 제공하되 모음과의 최종 결합형은 미결 경고를 남긴다.

Narrow는 기저 음소를 지우는 모드가 아니다. 제2·3본에서는 Broad를 보존하고 표면 이음·환경·규칙을 별도 층에 추가하도록 요구한다. GA 탄설음, dark L, 성문파열과 음절자음 등을 같은 완성도와 지원 상태로 묶어 말하지 않는다.

6.3 이전 앱을 비교할 때

Narrow·연음·조음 안내를 살필 때는 ‘이전 앱 비교’로 이동한다. 출력 차이가 보이면 우선 그 화면이 legacy인지 확인한다. 특히 구판의 /l/ 중복, /ʊ/와 /u:/의 처리, /t+s/와 /tʃ/의 구별을 최신 결과와 섞지 않는다.

새 앱에서 water의 /t/가 자동으로 [ɾ]로 바뀐다고 설명하지 않는다. 현재의 공개는 Narrow 정합화 완료를 뜻하지 않으며, 관련 후속 작업은 §10에 남긴다.

출처: 제2본 31–37쪽, 제3본 7·10·37쪽; js/registry.js-engine.js, legacy/.

7. 핵심 시험어로 읽는 구현

7.1 음절·모음·자음군

pick /pɪk/에서는 KIT의 U+119D를 확인한다. 이를 현대 ㅏ의 U+1175로 바꾸면 다른 값이 된다.

bank /bæŋk/와 friend /frɛnd/에서는 같은 블록 안에 자음이 여러 개 있음을 확인한다. 자모가 길게 펼쳐진다는 이유로 소리에 없는 모음을 더하지 않는다.

play /pleɪ/의 초성군에는 /p/ 다음 /l/이 한 번 있다. /eɪ/는 같은 핵의 두 모음 요소다. 2026-09-17 저자 결정에 따라 뒤 모음에 표시용 무음 ɔ를 붙인다. blue /blu:/는 [블루우]이며, /l/ 한 번과 하나의 긴 핵을 보존한다.

7.2 한 토큰과 두 요소, 두 토큰과 한 요소

boost /buːst/는 /b uː s t/ 네 음운 토큰으로 분석되며 다섯 자모 요소를 가진다. 길이 부호를 독립 자음처럼 읽지 않는다. cats /kæts/는 /k æ t s/ 네 토큰을 세 자모 요소로 적는다. 이 차이를 'IPA · 자모 대응 · 코드포인트'에서 확인한다.

비교	살펴볼 차이
/ʊ/와 /uː/	FOOT 하나와 U U 연쇄는 다른 값
cats와 catch	/t+s/ 두 음소와 /tʃ/ 한 음소
pass와 pat (GA)	종성 /s/와 /t/의 대립
/pɪk/와 /'pɪk/	자모열은 같아도 강세 기록은 다름

7.3 미결을 읽는 예

‘catch 미결 예’ 버튼은 IPA 직접 입력으로 전환하여 /kætʃ/를 넣는다. 이때 원문과 /tʃ/ 토큰, 종성 자모값은 구조에 남지만 종성 위치형 부호가 미배당이므로 평문을 제공하지 않는다. cats의 성공 결과를 가져와 catch의 대체 표기로 삼지 않는다.

GA /nɜːs/와 음절자음 입력도 별도의 미결 이유가 있다. 같은 ‘미결’ 표시라도 부호의 부재, 분석의 미결, 지원 범위 밖을 issues에서 구별해 읽는다.

출처: 제3본 14-17.39-41쪽; tests/book3-vectors.json, tests/engine.test.mjs.

계속 · 교사를 위한 사용해례

아래는 이번 구현본에서 제안하는 수업 절차다. 교육 효과를 검증한 실험 결과나 필수 교수 순서로 제시하는 것은 아니다. 학습자가 IPA와 훈민바름을 대조하고 한 음절의 범위를 읽게 하는 짧은 활동으로 사용할 수 있다.

7.4 활동 하나 · 자음군을 음절로 늘리지 않기

pick, bank, friend, boost를 차례로 제시한다. 먼저 교사가 실제 발음을 들려주거나 출처가 확인된 발음 자료를 준비한다. 이 앱에는 영어 낭독 기능이 없으므로 표기만 보고 발음 모델을 새로 정하지 않는다.

학습자에게 각 단어의 괄호가 몇 쌍인지, 음절핵이 어디인지, 앞뒤 자음이 어느 순서로 놓이는지 물어본다. bank의 끝에서 /ŋ/ 뒤 /k/가 이어져도 새로운 모음핵이 생기지 않는다는 점을 확인한다. 학습자의 발음은 별도로 듣고 피드백한다.

7.5 활동 둘 · 같은 철자, 다른 참조 발음

go, car, pass를 RP와 GA로 각각 조회한다. 먼저 source IPA의 차이를 찾고 그다음 자모열의 차이를 확인한다. 어느 쪽이 더 한국어처럼 읽히는지보다 어떤 원어 소리 정보를 보존하는지 설명하게 한다.

결과를 기록할 때는 단어만 적지 않고 모듈과 원문 IPA를 함께 적는다. 사전 출처가 다른 자료와 비교할 때는 발음 차이가 구현 오류인지, 출처의 다른 변이인지 먼저 판단한다.

7.6 활동 셋 · 저장할 정보 선택하기

/ə.'bʌʊt/를 변환한 뒤 자모열 복사와 전체 기록 저장의 차이를 설명한다. 강세가 어느 출력에 남는지, source IPA와 출처를 나중에 확인하려면 무엇을 보관해야 하는지 학습자가 답하게 한다.

마지막으로 cats와 catch를 비교한다. ‘미결’이 단어를 발음할 수 없다는 뜻이 아니라 현재 프로파일이 완성 평문을 내보낼 조건을 갖추지 못했다는 뜻을 확인한다. 결과를 비판적으로 읽는 능력도 활동의 목표다.

교사의 준비 기록: 목표 단어·모듈·확인한 IPA·음성자료 출처·앱 커밋·기대 결과를 함께 남긴다. 이 절은 새로 제안한 활용안이며 기존 문서의 확정 규칙과 구분한다.

8. 글꼴·세이핑과 대체 표시

8.1 부호·문자열·화면을 각각 확인한다

글꼴에 코드포인트가 있다는 사실만으로 임의의 복합 음절이 읽기 좋은 한 자형으로 표시된다고 판단하지 않는다. 글꼴 로드, 세이핑, 표시 문자열과 복사되는 NFD 기준형을 각각 검사한다.

앱은 fonts/의 두 옛한글 WOFF를 사용한다. linear-trial-v8-vowel-carrier는 합쳐야 할 핵심 요소와 독립 식별할 요소를 나눠 보여 주며, Unicode 정준 합성 결과를 정본으로 덮어쓰지 않는다.

8.2 표시 검증의 기준

조건	확인할 내용
단순 음절	핵심 요소 100%와 KIT-FOOT 등 모음값 식별
복합 음절핵	뒤 모음 요소 50%, 같은 음절 경계와 순서 유지
종성 /s/	받침으로 합치지 않고 75% 독립 요소로 표시
/t+s/ 복합값	/tʃ/와 구별하고 75% 독립 요소로 표시
좁은 화면	입력·결과·대응표의 잘림과 가로 넘침
글꼴 실패	누락 문자 인지와 정본 NFD-IPA 대체 정보

8.3 2026-09-16 확인 범위

당시 270개 자동시험은 /s/ 독립 표시, /t+s/와 /tʃ/ 구별, 뒤 모음 요소 50%, 독립 자음 75%, 정본 NFD 보존을 포함한다. 정적 검사와 배포 빌드도 통과했다.

8.4 남은 환경 검증

라이브 Pages에서 핵심 예시의 화면 표시를 확인했지만, 모든 운영체제·모바일·화면낭독기·클립보드 조합을 전수 검증한 것은 아니다. 글꼴 실패와 누락 글리프 자동 감지도 후속 과제다.

근거: tests/display.test.mjs, npm test 270/270, npm run check, npm run build, GitHub Pages 배포 기록. 확인일: 2026-09-16.

9. 자동시험·전수감사·재현성

9.1 2026-09-16 실행에서 확인한 결과

2026-09-16 기준판의 검사·시험·빌드를 다시 실행했다. 엔진·사전·표시 자동시험 270건이 모두 통과했다. 제3본 예시, 오류 처리, /s/와 /t+s/ 독립 표시, 크기 규칙과 NFD 보존을 포함하며, 정적 검사는 53개 로컬 링크·자산과 입력 표지를 확인한다.

검사 묶음	당시 결과	의미
엔진·사전·표시 자동시험	270/270 통과	당시 수록 시험 조건의 통과
제3본 코드포인트 대조	통과	문서 예시와 registry 대조
구문·경로·라벨 검사	통과	53개 로컬 링크·자산과 접근성 표지
정적 배포 파일 생성	통과	dist 생성 완료
GitHub Actions build-deploy	성공	실행 35074663177
라이브 핵심 예시 표시	확인	boost·pass·car·stop 등 선택 예시
사전 군집 전수 집계	재현	tools/cluster_census.mjs 결과 일치

9.2 통과와 분모

270건을 270개 단어나 사전 전체의 발음 정확도라고 소개하지 않는다. 정상·부정·미결 입력, 미배당 종성, /ʒ/, 음절자음, 토큰 중복, 표시 규칙과 오류 처리 시험이 함께 포함된다. ‘미결이어야 하는 입력이 미결로 반환됨’도 올바른 통과 조건이다.

구판 제4본의 70/70·183/183·515,370회와 원 앱 보고의 115건은 서로 다른 판본·범위의 수치다. 당시의 270건에 합산하지 않는다. 사전 군집 집계 재현과 전체 발음 정확도 평가는 자동시험 통과와도 구분한다.

9.3 전수감사에 필요한 다음 기록

원 사전 해시와 항목 수, 모듈별 전처리, 엔진·등록부 버전, 제외·미결 기준, 성공·실패·미결의 전체 분모를 남긴다. 구조가 맞는 비율과 발음이 맞는 비율은 별도의 검사다. 사전의 IPA가 옳은지 판단하려면 출처와 화자·전문가 검토가 필요하다.

출처: tests/*.test.mjs, tools/check.mjs, tools/build.mjs, tools/cluster_census.mjs의 직접 실행 및 GitHub Actions 35074663177. 확인 일: 2026-09-16.

계속 · 개발자가 실행하는 재현 절차

9.4 실행 환경과 기본 명령

저장소 루트에서 Node.js 22 이상으로 다음을 실행한다. 앱의 검사·시험·빌드에는 npm 패키지 추가 설치가 필요하지 않다. 문서 편집·렌더링 도구는 이 앱 실행 조건과 별개다.

```
node tools/check.mjs
node --test tests/*.test.mjs
node tools/build.mjs
python3 -m http.server 8000 --directory dist
```

브라우저에서 <http://localhost:8000>을 연다. HTML을 파일로 직접 여는 방식은 ES 모듈과 JSON 요청의 정상 동작 조건과 다를 수 있으므로 위의 HTTP 실행을 사용한다. 개발자가 이 명령을 실행하는 컴퓨터와 브라우저가 접근하는 컴퓨터가 같은지도 확인한다.

9.5 엔진 API를 직접 확인하기

```
import { convertWord } from './js/engine.js';
const result = convertWord('bu:st', {
  accent: 'rp', spelling: 'boost'
});
console.log(result.serialization.codepoints);
console.log(result.canonicalModel.mappingUnits);
```

예시는 저장소 루트의 ES 모듈 파일에서 실행한다. 기대 부호열은 U+1107 U+116E U+116E U+11BA U+11C0이다. 이 코드는 직접 호출 예시이며 별도 provenance를 전달하지 않았으므로 사전 검수 출처를 뜻하지 않는다. 운영 데이터에서는 확인한 provenance를 함께 전달한다.

9.6 무엇을 보존할 것인가

시험 결과와 함께 실행 명령, Node 버전, 커밋, 입력 벡터, 사전·글꼴 해시를 남긴다. 배포 빌드는 index.html·css·js·fonts·data·docs와 legacy의 화면 자산을 dist로 복사한다. 소스 시험과 생성 도구는 웹사이트 배포본에 포함하지 않는다.

빌드 성공과 원격 공개는 구분한다. 2026-09-16 판은 GitHub Actions 35074663177의 deploy 성공과 공개 URL을 확인했다. 최신 기록은 부록 D를 참조한다. 이후 변경에서도 workflow 결과와 서비스 주소를 다시 확인한다.

출처: README, package.json, tools/build.mjs; 예시 부호열은 제3본 14쪽과 현재 엔진 직접 실행.

10. 현 구현과 전산 규격의 정합화 과제

현재 앱은 제3본의 핵심 Broad 직렬화 예제를 실행하지만 문서 전체 스키마·입출력·접근성 요구를 모두 구현한 것은 아니다. 실제 필드 이름이나 값이 다른 항목은 교환 어댑터 없이 동일 형식으로 취급하지 않는다.

항목	현재 구현	남은 작업
직렬화 성공 상태	available	제3본 예시 complete와 대응 명시
직렬화 profile	hmbr-broad-nfd-draft-2026-09-08	문서 시범 식별자와 버전 대응
codepoints	U+로 시작하는 문자열 배열	문서의 정수 배열과 변환 정의
장모음	토큰과 두 요소의 대응; 뒤 모음 50% 및 읽기용 무음 ㅇ 표시 구현.	문서의 length 속성과 실제 데이터 필드의 대응 명시.
표시 데이터	display.blocks	요소 참조.render_status 확장
입력 검증	내부 생성 모델의 참조·소비 검사	외부 JSON 전체 스키마 검사
자모열 읽기	parseCanonical 내부 함수	원입력 보존:NFD 수신 어댑터
역변환	사용자용 역변환기 없음	후보·모듈·출처 조건 구현
Narrow	legacy 비교 기능	새 모델의 별도 표면층 구현
글꼴·접근성	코드포인트·설명 제공	누락 감지·실환경 시험

10.1 해결한 핵심 차이

// 한 번, 장모음·복모음의 같은 핵 보존, FOOT와 U U 구별, /t+s/와 /tʃ/ 구별, 미배당 종성의 unavailable 처리를 새 엔진에 반영했다. 새로운 필드가 원 저자의 최종 승인을 받은 표준 스키마라는 뜻은 아니다.

10.2 후속 개정의 원칙

하나의 수정이 입력 분석, 토큰 소비, 직렬화, 표시 가운데 무엇을 바꾸는지 기록한다. 특히 완전한 역변환·외부 레코드 가져오기·Narrow 구현은 대응 시험 없이 기존 함수 이름만으로 완료를 선언하지 않는다. §9의 부분 시험 통과를 전체 규격의 적합성 등급으로 확대하지 않는다.

출처: 제3본 7-8·16-17·37·41쪽; js/engine.js-registry.js·app.js의 실제 필드 대조.

11. 공개 배포와 버전 확인

11.1 표시 프로파일과 배포 기록

현재 표시 프로파일은 linear-trial-v8-vowel-carrier다. 저자의 2026-09-17 결정으로 뒤 모음에 무음 ㅇ을 붙이며 boost를 부우스트로 표시한다. 아래 커밋은 9월 16일 구현 검증의 기준판이다.

2026-09-16 확인 코드판: f049e24da7c34614a2129946071c4094c5f0964

저장소: <https://github.com/ivanpark/hmbr>

웹앱: <https://ivanpark.github.io/hmbr/>

2026-09-16 확인한 GitHub Actions 실행 35074663177의 build와 deploy가 success로 완료됐다. 배포 성공은 자동시험·정적 빌드와 공개 URL을 확인한 것이며, 모든 플랫폼의 조판·접근성 적합성을 보증하지 않는다.

11.2 설정과 재배포

Pages의 게시 원본은 GitHub Actions다. main 변경 시 npm test, npm run check, npm run build를 실행하고 산출물을 배포한다. 실패하면 해당 실행의 로그와 커밋을 함께 기록한다.

문서와 앱이 함께 바뀌면 문서판·커밋·시험 수·배포 실행을 한 묶음으로 식별한다. README와 docs/INDEX.md의 링크도 같은 파일명으로 갱신한다.

11.3 이용조건과 출처

앱 소스와 문서의 이용조건은 저장소 LICENSE 및 docs/LICENSE.md를 따른다. 사전 데이터와 웹 글꼴은 원출처·권리자·재배포 조건을 별도로 확인한다. 공개 저장소라는 사실만으로 모든 포함 자료의 무제한 재이용을 뜻하지 않는다.

확인일: 2026-09-16. 근거: 원격 main f049e24da7c34614a2129946071c4094c5f0964, Actions 35074663177, 공개 Pages와 저장소 파일 목록.

계속 · 오류 제보와 문제 해결

11.4 재현 가능한 오류 제보

결과 문자열이나 화면 사진 하나만으로는 입력과 분석의 차이를 찾기 어렵다. 다음 항목을 함께 기록하면 오류가 발생한 층을 좁힐 수 있다. 공개 제보에는 재현에 필요한 단어·기호와 환경 정보만 포함한다.

항목	기록할 내용
입력	영어 또는 IPA 원입력
모듈·화면	RP/GA, 새 앱/legacy, 선택 기능
버전	문서판과 코드 커밋
기대	해당 제2.3본 조항 또는 시험 벡터
실제	JSON의 상태·issues·코드포인트
표시	화면 사진, 글꼴, 글자 크기
환경	운영체제·브라우저·버전·입력 방식
재현	같은 입력으로 반복되는 순서

11.5 증상별로 먼저 볼 것

‘사전에 없는 단어’이면 철자와 문장부호를 확인한다. 사전이 없다는 이유로 발음을 추측하지 않으며, 확인한 IPA가 있을 때 직접 입력한다. ‘사전 파일을 불러오지 못했습니다’이면 HTTP 실행과 파일 경로를 확인한 뒤 다시 시도한다. 실패한 사전 요청은 캐시에서 제거해 재요청할 수 있도록 구현돼 있다.

‘직렬화 미결’이면 issues에서 미배당 위치형과 분석 미결을 구별한다. /θ ð ʒ tʃ dʒ/의 종성이 해당되면 현 등록부의 제한일 수 있다. 익숙한 현대 한글로 대체해 성공 처리하지 않는다.

자모가 깨져 보이면 원문 IPA와 코드포인트부터 확인한다. 부호열이 맞다면 글꼴 다운로드와 표시 경로를 살핀다. 복사한 내용에서 강세가 보이지 않는 것은 평문 복사의 범위일 수 있으며, 전체 기록의 블록 stress를 확인한다.

11.6 수정 제안

구현 오류는 최소 재현 입력과 기대값을 시험으로 남긴다. 음가·조자법 변경 제안은 먼저 제2본의 공개 검토 항목으로 다루며, 문서와 코드의 판정을 함께 갱신한다.

출처: js/dict.js-app.js-engine.js; 제3본 17:37-41쪽. 제보 양식은 이 구현본의 운영 제안.

12. 향후 다언어 구현 원칙

12.1 영어를 공통 규칙으로 강제하지 않기

영어 참조구현은 다른 언어의 정답을 미리 정하지 않는다. 새 언어는 원문 문자, 언어·지역 식별, source IPA, 자체 음운 단위, 음절구조, 강세·성조와 실제 화자 검토를 갖춘 독립 모듈로 추가한다.

공통으로 재사용할 부분은 원문 보존, 자모값 등록 절차, 음운 토큰과 자모 요소의 분리, 구조와 표시의 분리, 버전·시험 기록이다. 영어의 RP·GA 모음값이나 음절 추정 규칙을 다른 언어에 그대로 적용하지 않는다.

12.2 새 모듈을 위한 최소 자료

출처가 확인된 입력 집합과 발음 변이, 자모값 대응표, 위치형의 지원 상태, 음절분석 근거, 양성·부정·미결 시험을 마련한다. 기존 자모값을 재사용할 때는 음가와 역규칙이 맞는지 확인한다. 모양이나 Unicode 이름의 유사성만으로 동치를 선언하지 않는다.

신규 모듈이 기존 자료의 해석을 바꾸면 새 모듈·등록부 버전으로 구별한다. 과거 레코드를 새 규칙으로 다시 계산할 때는 source와 이전 버전을 보존하며, 원자료가 없는 출력만으로 원래 발음을 확정하지 않는다.

12.3 바르미의 다른 구성과 연결하기

MetaVoc 선택 화면, Seesaurus와의 자료 연결, 종합 학습 과정은 현재 새 웹앱에 포함돼 있지 않다. 향후 어휘 자료를 연결할 때도 표제어만 넘기지 않고 IPA·모듈·출처·검수 상태를 함께 전달해야 한다. 교육용 화면을 확장하는 일과 훈민바름의 음가 규칙을 바꾸는 일은 별도로 관리한다.

맺음말 · 구현은 보존을 보여 준다

좋은 참조구현은 가능한 결과를 보여 주고, 그 결과를 다시 검사할 근거를 남긴다. 원어의 소리를 정본에 담는 원리, 데이터를 잃지 않는 형식, 사람이 읽을 수 있는 화면이 서로를 검증할 때 연구는 다음 단계로 나아간다. 이 구현본은 그중 오늘 실제로 작동한 부분과 다음에 검증할 부분을 함께 기록한다.

출처: 제1본 2쪽, 제3본 37–38쪽. 다언어·바르미 연계 절차는 후속 설계를 위한 제안이며 현재 제공 기능이 아니다.

부록 A. 적합성 체크리스트

이 표는 2026-09-16 구현의 점검표이며 후속 검증은 부록 D를 따른다. ‘확인’은 아래 범위에서 확인했다는 뜻이며 제3본 전체 적합성 인증이 아니다. 미실시·미구현 항목을 전체 통과 수에 넣지 않는다.

점검 항목	상태	확인 범위
원문 IPA·선택 모듈·출처 보존	확인	엔진·사전 시험
문서 예제 코드포인트	확인	43개 모듈별 대조
// 중복 소비 방지	확인	play·blue·모음 사이 //
/u:/ 한 토큰·두 요소	확인	boost
/t+s/ 두 토큰·한 요소	확인	cats 및 경계 분리
FOOT/U U, AW/O O 구별	확인	엔진 대립 시험
미배당 종성 구조 보존	확인	다섯 종성의 부정 시험
/ʒ/·음절자음 임의 보충 금지	확인	미결 반환 시험
원문 강세·정규화 이력	확인	엔진·레코드 검사
filler·호환자모·순서 오류 거부	확인	기준형 부정 시험
토큰·요소 참조와 소비	확인	내부 모델 검사
HTTP 배포 파일 생성	확인	로컬·원격 build
원격 파일 완전성	확인	46개 Git blob 해시
브라우저 복사·JSON 다운로드	코드 확인	코드 확인; 환경별 복사·다운로드 완료 검증 필요
반응형 화면·키보드·낭독기	부분 확인	데스크톱 화면·코드 표지 확인; 모바일·낭독기 검증 필요
사전 군집 전수 집계	집계 재현	군집 통계 재현; 전체 발음 정확도 판정은 별도
외부 JSON 가져오기·역변환	미구현	사용자용 기능 없음
새 모델의 Narrow 층	미구현	legacy와 구분
공개 Pages 서비스	확인	Actions 35074663177 build-deploy success

새로 시험한 환경이나 수정된 판본은 결과·분모·버전을 별도 기록한다. 한 브라우저의 성공을 모든 환경의 성공으로 바꾸지 않는다.

출처: §9의 직접 실행 기록, 원격 Git 트리·Actions 기록, 앱 코드와 제3본 요구사항 대조.

부록 B. 핵심 시험 벡터

아래 부호열의 공백은 읽기용이다. 실제 단어 자모열에는 포함되지 않는다. 이 표는 핵심 회귀 사례의 발체이며, 저장소의 43개 모듈별 대조 전체를 대체하지 않는다. 모든 정상 행은 한 음절블록이다.

입력·모듈	기대 부호열 (각 항 U+)	확인할 것
pick /pɪk/, 공통	1111 119D 11BF	KIT
pat /pæt/, 공통	1111 1162 11C0	종성 /t/
bank /bæŋk/, 공통	1107 1162 11BC 11BF	두 종성 순서
friend /frɛnd/, 공통	114B A976 1166 11AB 11AE	자음군 보존
play /pleɪ/, 공통	1111 1105 1166 119D	/l/ 한 번
blue /blu:/, 공통	1107 1105 116E 116E	긴 핵
boost /bu:st/, 공통	1107 116E 116E 11BA 11C0	/s/ 뒤 /t/
cats /kæts/, 공통	110F 1162 11BE	JV-TS
pass /pæs/, GA	1111 1162 11BA	/s/ 독립 식별
car /kɑ:/, GA	110F 1161 D7DB	권설 자음
go /gəʊ/, RP	1100 119E 118D	RP 핵
go /goʊ/, GA	1100 1169 118D	GA 핵

부정·미결 벡터

입력·조건	기대 판정
/kætʃ/	E-POSITION, text=null, 종성 구조 보존
/nɜːs/, GA	E-VOWEL, normalized_ipa=null
음절자음 ɲ	E-SYLLABIC, 임의 모음 삽입 없음
[pɪk]	E-NARROW, 완성 Broad 출력 없음
filler 포함 기준형	E-CHAR
같은 음소 중복 소비	E-COVERAGE
존재하지 않는 요소 참조	E-REFERENCE

출처: 제3본 14-17·41쪽 및 tests/engine.test.mjs. 오류명은 실제 앱의 식별자다.

부록 C. SOURCE AUDIT와 작성 근거

C.1 직접 확인한 원자료

2026-09-08의 제1-4본, 2026-09-14 PDF 교정본과 교정 작업파일을 대조했다. 9월 14일 판은 편집 원본을 새로 내보낸 것이 아니라 기존 PDF 위의 선택 영역을 교체한 교정판임을 확인했다. 제3·4본은 DOCX 원본을 직접 고쳐 2026-09-16 판으로 다시 만들었다.

2026-09-16 교정 당시 제1본 36쪽, 제2본 59쪽, 제3본 46쪽, 제4본 23쪽의 페이지 구성과 교정 범위를 확인했다. 제1·2본의 최신 편집 원본은 확보되지 않아 9월 8일 PDF를 기준으로 패치를 재구성했다.

제3본의 사전 수치는 당시 저장소의 `tools/cluster_census.mjs`로 재현했다. 제4본은 당시 코드·시험·배포 상태를 반영했다. 문서 검토와 음성학적 타당성·교육 효과·법적 효력 검증은 서로 구분한다.

원본·교정본·코드의 판본 날짜와 역할을 섞지 않는다. PDF에서 확인한 수정은 편집 가능한 원본에 반영한 뒤 다시 렌더링·대조하고, 원본이 없는 책은 그 제한을 공개 기록에 남긴다.

C.2 구현 자료

2026-09-16 기준판의 `js/`, `data/`, `fonts/`, `tests/`, `tools/`, `docs/`와 README를 확인했다. 앱 표시 규칙, 270개 자동시험, 사전 전수조사 도구, 빌드 및 문서 링크를 같은 코드판에서 대조했다.

당시 `npm test`는 270/270, `npm run check`와 `npm run build`는 통과했다. 라이브 Pages 배포와 문서 링크를 확인했다. 다만 모든 브라우저·모바일·화면낭독기·클립보드 조합을 전수 시험했다고 주장하지 않는다.

C.3 근거의 구분

문서의 정본 규칙, 실제 코드·시험, 공개 배포 기록과 이번 편집 판단을 각 절의 출처 줄에 구분한다. 사전 집계 재현은 현재 파일과 실행 코드에 대한 결과이며, 사전 원출처의 판·라이선스 확인이나 영어 전체의 음운 필요량에 대한 영구적 결론은 아니다.

부록 C 계속 · 판본 차이와 후속 갱신

C.4 판정과 구조의 정합화

이전 판본의 상태	2026-09-16 교정의 처리
전체 블록 NFC로 /s/가 받침에 합쳐짐	linear-trial-v2에서 독립 요소로 표시
장모음·복합핵의 뒤 요소 크기 미구현	nucleus-tail 50% 적용
독립 자음의 구별 크기 미구현	loose-consonant 75% 적용
문서와 앱 상태의 시점 차이	당시 main 커밋·시험·배포 실행으로 갱신
사전 집계 재검증 대기	cluster_census.mjs로 수치 재현
반복 장 번호	계속 제목으로 바꾸어 위계 명확화
PDF 위 선택 패치	제3·4본 편집 원본에 직접 반영

C.5 2026-09-16 교정의 핵심 차이

linear-trial-v2는 종성 /s/와 /t+s/를 독립 표시 요소로 남기고, 핵심 요소 100%·뒤 모음 요소 50%·독립 자음 75%를 적용한다. 이 표시 규칙은 canonicalModel과 NFD 정본을 바꾸지 않는다.

제3본의 사전 전수조사는 저장소의 코드로 재현됐고 종성 소계의 1종 차이는 compound_ts로 설명됐다. 제4본은 자동시험 270/270, 빌드 성공, 원격 커밋과 Pages 실행을 당시 값으로 갱신했다.

C.6 재현 기준과 후속 갱신

2026-09-16 기록의 코드판은 f049e24da7c34614a2129946071c4094c5f0964이고, Pages 실행은 35074663177이다. 문서·코드·사전이 바뀌면 새 커밋에서 시험·집계·배포와 링크를 다시 확인한다. 2026-10-09 재확인과 후속 앱 변경은 부록 D를 참조한다.

제4본 DOCX와 Markdown은 편집 원고, PDF는 고정 쪽수의 읽기용 판본이다. 세 형식의 판본 날짜와 핵심 문언을 함께 갱신하고, README·INDEX·체크섬이 같은 파일을 가리키는지 확인한다.

구현은 정본의 권위를 대신하지 않는다. 구현본은 실제로 확인한 작동과 남은 과제를 다음 검토자가 이어받을 수 있게 기록한다.

부록 D. 2026-10-09 구현 상태 갱신

D.1 판본과 정정

네 본을 2026-10-09 교정판으로 정리했다. 제3본 37쪽의 표시 설명과 이 책 16쪽의 장모음 행을 바로잡고 전권의 문언·참조·판본 설명을 교정했다. 제4본의 조회 순서와 저장 display 필드를 실제 코드에 맞췄다. 뒤 모음의 50% 크기와 읽기용 무음 ㅇ 표시는 구현되어 있으며 저장 V 배열과 정본 NFD에는 표시용 ㅇ을 더하지 않는다.

D.2 코드와 검증 기록

2026-10-09 앞선 앱 검증의 기준판은 9f750e2이며 자동시험 283/283, 로컬 링크·자산 70개 검사와 정적 빌드가 통과했다. Actions 37870009195의 배포도 성공했다. 이번 전권 교정의 코드 대조 기준판은 592fe8c이다. 본문의 270건·53개 링크·f049e24·실행 35074663177은 9월 16일 당시 기록이다.

D.3 사용자 흐름의 보완

사전 요청은 20초 후 시간 초과로 처리하며 실패한 요청을 다시 시도할 수 있다. 상세 대응표는 펼칠 때 생성한다. 입력 한도는 100단어·5,000자이며 Ctrl/⌘ + Enter로 변환할 수 있다. 전체 기록 저장은 다운로드 요청을 알리고 브라우저의 다운로드 목록을 확인하도록 안내한다. 네 본의 판본 날짜, 표시와 저장의 차이, 사전 검수 상태와 미지원 기능을 앱에 명시했다.

D.4 남아 있는 검증

데스크톱 브라우저에서 boost 표시, pass의 RP·GA 비교와 catch의 미결 처리를 확인했다. 모바일·화면낭독기·환경별 클립보드 호환성과 JSON 다운로드 완료 검증은 남아 있다. 사전 전체의 발음·출처·라이선스 검수, Narrow 입력, 외부 JSON 가져오기와 사용자용 역변환은 완료되지 않았다. 자동시험 통과는 이 범위의 완료나 학습 효과를 뜻하지 않는다.

연결 발음과 GA t 변이 사용 안내

2026년 10월 12일 개정 부록. 본문 §6-§10과 부록 D의 미구현 표시는 당시 구현 상태다. 이번 갱신에서 연결 발음과 일부 GA t 표면형을 별도 연습층으로 제공한다. 기본 Broad 엔진에 임의의 Narrow 입력을 넣는 기능은 여전히 제공하지 않는다.

실행 순서

1. 웹앱에서 발음 기준을 GA로 고르고 구문장을 입력한 뒤 별도의 연음화 버튼을 누른다. 일반 훈민바름으로 보기 버튼은 원형만 변환한다. button, winter, get me 예제 버튼은 GA로 원형을 보여 주며, 그 뒤 연음화를 눌러 후보를 연다. GA 단어 하나라도 t 변이가 가능하면 같은 버튼으로 연습할 수 있다.
2. 처음에는 모든 후보가 OFF다. 추천 배지를 확인한 뒤 후보별 스위치를 켜거나 추천대로 적용을 누른다. 속도 선택과 단어별 강조·쉽으로 추천을 바꾼다. 모두 해제로 원형을 되찾는다. 자동 추천과 자동 적용을 혼동하지 않는다.
3. 선택한 흐름의 표면 IPA와 훈민바름을 비교한다. [ʔ][ŋ]처럼 자형이 미결인 소리는 해당 단어의 훈민바름 아래에 원형 참조가 표시된다. 카드의 적용 조건과 조음 연습 설명을 읽는다.
4. get me의 성문음화와 미파열을 번갈아 선택한다. 같은 t의 앞 선택이 해제되는지 확인한다. get 뒤에 쉽을 넣으면 경계는 끊어진다. 쉽 앞 t 변이는 수동으로 선택할 수 있지만 기본 추천에는 넣지 않는다.

예문	설정과 관찰
button	보통 GA에서 [ʔŋ] 추천. 원형 훈민바름 참조 표시
winter	빠른 GA에서 t 탈락 추천. 파생 훈민바름 비교
get me	[ʔ]과 [t] 중 한 선택. 혀와 성문 동작 대조
next day	빠른 회화에서 자음군 t 탈락. 강조·쉽 보호

5. 전체 기록 저장으로 원형과 연습형을 함께 내보낸다. 원형 자모열 복사에는 연습층의 IPA나 연결 부호가 포함되지 않는다. 입력을 새로 변환하거나 RP/GA를 바꾸면 이전 선택을 이어 쓰지 않는다.

지원 현황과 다음 음성 단계

기능	2026년 10월 12일 상태
기본 연결 발음	자음-모음 연결, RP linking r, 일부 약형·동화·탈락
GA t 확장	n 뒤 t 탈락, 성문음화, 어말 미파열의 선택형
조음법 설명	기존 표기 해례와 연결 후보별 기관 이동 설명
글꼴	한글 표기 나눔명조 옛한글, UI 나눔바른고딕 옛한글, IPA Doulos SIL
미구현	자동 flapping, 완전한 Narrow 변환, 문장 억양, TTS

본 갱신의 기준은 연결 발음 1차 공개판 66398de와 그 위에 더한 t-variants 변경이다. npm test에서 307개 시험이 통과했다. 새 시험 6개는 button의 분리 주석, winter의 강세 조건, 동일 t의 충돌, 쉽 앞 선택, 여러 단어 조합과 원형 보존을 검사한다. 시험 통과는 음성학적 보편성이나 학습 효과의 입증이지 않는다.

GA winter의 기존 사전 IPA에서 빠진 r을 Collins 미국식 /ˈwɪntər/와 대조하고 앱 관례에 따라 /ɹ/로 정규화해 보완했다. button 원형도 대조했다. 일부 단어의 교정이 전체 RP-GA 사전 검수나 권리 확인의 완료를 뜻하지 않는다. 미결 단어에는 추측한 연결 후보를 만들지 않는다.

확인할 동작

개별 토글과 모두 해제, 추천대로 적용, 속도 변경, 강조·쉽, 키보드 스위치 조작, IPA와 훈민바름의 참조 관계, JSON의 원형 보존을 확인한다. 구현은 js/connected.js, js/t-variants.js와 js/connected-ui.js로 나뉜다. 실제 배포판과 검사 결과는 저장소의 날짜별 상태 검토와 체크섬 목록을 기준으로 확인한다.

TTS를 연결할 때의 기준

다음 단계는 재생·중지·반복과 원형/선택형 비교다. 단순한 철자 읽기를 선택된 표면형의 재생이라고 표시하지 않는다. 음성 엔진이 개별 [ʔ], [ɾ], Ø를 제어할 수 있는지 먼저 확인하고 제어하지 못하는 차이를 화면에서 명시한다. 이 판에는 TTS 재생 버튼이나 API 연동을 포함하지 않는다.

네 본의 PDF는 기존 본문 쪽수를 보존하고 이 개정 부록을 끝에 덧붙였다. 기존 구현 설명 옆에는 갱신 위치를 표시했다. DOCX도 같은 갱신 내용을 담으며, 편집 환경에 따른 줄바꿈 차이는 허용하되 쪽수 인용의 기준은 PDF로 한다.