

국립국어원 국어 역사 말뭉치 (버전 1.0)

- **자료명:** 국립국어원 국어 역사 말뭉치
- **공개일**
 - (버전 1.0) 2024. 6. 28.
- **자료 유형:** 텍스트
- **관련 사업:** 21세기 세종계획(1998~2007)
역사자료 종합 정비(2014, 2015, 2016, 2018)
- **자료 설명**
 - **내용**
 - 15세기 한글 창제 이후부터 20세기 초기까지 한글로 기록된 문헌자료 1,153종의 원시 말뭉치
 - **분량**
 - 총 1,153건
 - **파일 형식:** XML(UTF-8 인코딩)
 - **파일 수 및 크기:** 파일 1,153개, 총 149MB
- **인용:**

(국문) 국립국어원(2023). 국립국어원 국어 역사 말뭉치(버전 1.0). URL: <https://kli.korean.go.kr/corpus>
 (영문) National Institute of Korean Language(2023). NIKL Korean History Corpus(v1.0).
 URL: <https://kli.korean.go.kr/corpus>

· 파일 명명 규칙

자리	1	2	3	4	5	6	7	8	9	10	11	12	13	14
속성	매체	장르	주석 단계		구축 연도		일련번호(8자리)							
정의	H (역사 자료)	X (무작위 구성)	RW (원시 말뭉치)		23 (2023년)		20000001 ~ 20001153							
※ 예시: HXRW2320000001.xml - 2023년에 최종 정비한 국어 역사 말뭉치의 첫 번째 파일														

· 예시

```
<doc>
<teiHeader>
<titleStmt>
<title lang="kor">훈민정음</title>
<title lang="chi">訓民正音</title>
<volume n=""/>
<author></author>
<date>1459</date>
</titleStmt>
<notesStmt>
<bibl>
<title>언해본 훈민정음(서강대 월인석보 권두본)</title><!-- 국가기록유산, 디지털한글박물관 등-->
<pubPlace>서강대</pubPlace>
<publisher></publisher>
<date></date>
</bibl>
</notesStmt>
<revisionDesc>
<date></date>
<respStmt>
<resp>입력자</resp><name>김송희</name>
</respStmt>
<respStmt>
<resp>교정자</resp><name></name>
</respStmt>
</revisionDesc>
</teiHeader>
<sent type="title" lang="chi" page="1a" n="1">世宗御製訓民正音【製는 글 지술씨니 御製는 님금 지스산 그리라】
【訓은 ㄱㄹ칠씨오 民은 百姓이오 音은 소리니 訓民正音은 百姓 ㄱㄹ치시논 正은 소리라】</sent>
<sent type="main" lang="chi" page="1a" n="2">國之語音이【國은 나라히라 之는 입겨지라 語는 말쑤미라】
</sent>
<sent type="main" lang="kor" page="1a" n="2">나랏 말쑤미</sent>
<sent type="main" lang="chi" page="1b" n="1">異乎中國하야【異는 다를씨라】【乎는 아모그에 ㅎ는 거체 쓰는 ㄱ
라】【中國은 皇帝 겨신 나라히니 우리 나랏 常談에 江南이라 ㅎ는 나라】</sent>
<sent type="main" lang="kor" page="1b" n="1">중등國語에 달아</sent>
<sent type="main" lang="chi" page="1b" n="2">與文字로 不相流通홀씨【與는 이와 더와 ㅎ는 거체 쓰는 ㄱ라】
【文은 글와리라】【不은 아니 ㅎ는 ㅍ디라】【相은 서르 ㅎ는 ㅍ디라】【流通은 흘러 스믈출씨라】</sent>
<sent type="main" lang="kor" page="1b" n="2">文字와로 서르 스믈디 아니홀씨</sent>
<sent type="main" lang="chi" page="2a" n="1">故로 愚民이 有所欲言하야도【故는 전치라】【愚는 어릴씨라】【有는
이실씨라】【所는 배라】【欲은 ㅎ고져 홀씨라】【言은 니를씨라】</sent>
<sent type="main" lang="kor" page="2a" n="1">이런 전초로 어린 百백姓성이 니르고져 홀 배 이셔도</sent>
<sent type="main" lang="chi" page="2b" n="1">而終不得伸其情者 多矣라【而는 입겨지라】【終은 ㅁ츄미라】【得은
시를씨라】【伸은 펼씨라】【其는 제라】【情정은 ㅍ디라】【者는 노미라】【多는 할씨라】【矣는 말 ㅁ는 입겨지라】</sent>
<sent type="main" lang="kor" page="2b" n="1">ㅁ츄내 제 ㅍ들 시러 퍼디 ㅁ홀 노미 하니라</sent>
<sent type="main" lang="chi" page="2b" n="2">予 爲此憫然하야【予는 내 ㅎ습시논 ㅍ디시니라】【此는 이라】【憫
然은 어엿비 너기실씨라】</sent>
<sent type="main" lang="kor" page="2b" n="2">내 이를 爲하야 어엿비 너겨</sent>
<sent type="main" lang="chi" page="3a" n="1">新制二十八字하노니【新은 새라】【制는 ㅁ ㄱ실씨라】【二十은
스믈여들비라】</sent>
<sent type="main" lang="kor" page="3a" n="1">새로 스믈여들 ㅁ를 ㅁ ㄱ노니</sent>
<sent type="main" lang="chi" page="3b" n="1">欲使人人으로 易習하야 便於日用티니라【使는 히여 ㅎ는 마리라】
【人은 사르미라】【易는 쉬블씨라】【習은 니길씨라】【便은 便安홀씨라】【於는 아모그에 ㅎ는 거체 쓰는 ㄱ라】【日은 나
리라】【用은 ㅍ씨라】【耳는 ㅍ르미라 ㅎ는 ㅍ디라】</sent>
<sent type="main" lang="kor" page="3b" n="1">사름마다 히여 수비 니겨 날로 ㅍ메 便安키 ㅎ고져 홀 ㅍ르미니
라</sent>
```

<sent type="main" lang="chi" page="4a" n="1">ㄱ 는牙音이니如君^ㄱ字初發聲하니竝書하면如虬^ㄱ字初發聲하니라【牙는 어미라】【如是 그틀씨라】【初發聲은 처섬 퍼아 나눈 소리라】【並書는 글바 쓸씨라】</sent>

<sent type="main" lang="kor" page="4a" n="1">ㄴ 은엄소리니君^ㄴ字처섬퍼아나눈소리ㄴ특니글바쓰면虬^ㄴ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="4a" n="2">ㅋ 는牙音이니如快^ㄱ字初發聲하니라</sent>

<sent type="main" lang="kor" page="4a" n="2">ㅋ 은엄소리니快^ㄱ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="4b" n="1">ㅇ 는牙音이니如業^ㄱ字初發聲하니라</sent>

<sent type="main" lang="kor" page="4b" n="1">ㅇ 은엄소리니業^ㄱ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="5a" n="1">ㄷ 는舌音이니如斗^ㅌ字初發聲하니竝書하면如覃^ㄷ字初發聲하니라【舌은 혀라】</sent>

<sent type="main" lang="kor" page="5a" n="1">ㄷ 는혀소리니斗^ㅌ字처섬퍼아나눈소리ㄴ특니글바쓰면覃^ㅌ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="5a" n="2">ㅌ 는舌音이니如呑^ㄷ字初發聲하니라</sent>

<sent type="main" lang="kor" page="5a" n="2">ㅌ 는혀소리니呑^ㄷ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="5b" n="1">ㄴ 는舌音이니如那^ㄴ字初發聲하니라</sent>

<sent type="main" lang="kor" page="5b" n="1">ㄴ 는혀소리니那^ㄴ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="6a" n="1">ㅍ 는脣音이니如警^ㅍ字初發聲하니竝書하면如步^ㅍ字初發聲하니라【脣은 입시우리라】</sent>

<sent type="main" lang="kor" page="6a" n="1">ㅍ 는입시울소리니警^ㅍ字처섬퍼아나눈소리ㄴ특니글바쓰면步^ㅍ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="6a" n="2">ㅂ 는脣音이니如漂^ㅂ字初發聲하니라</sent>

<sent type="main" lang="kor" page="6a" n="2">ㅂ 는입시울소리니漂^ㅂ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="6b" n="1">ㅁ 는脣音이니如彌^ㅁ字初發聲하니라</sent>

<sent type="main" lang="kor" page="6b" n="1">ㅁ 는입시울소리니彌^ㅁ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="7a" n="1">ㄸ 는齒音이니如卽^ㄸ字初發聲하니竝書하면如慈^ㄸ字初發聲하니라【齒는 나라】</sent>

<sent type="main" lang="kor" page="7a" n="1">ㄸ 는니소리니卽^ㄸ字처섬퍼아나눈소리ㄴ특니글바쓰면慈^ㄸ字처섬퍼아나눈소리ㄴ특니라</sent>

<sent type="main" lang="chi" page="7a" n="2">ㅊ 는齒音이니如侵^ㅊ字初發聲하니라</sent>

<sent type="main" lang="kor" page="7a" n="2">ㅊ 는니소리니侵^ㅊ字처섬퍼아나눈소리ㄴ특니라</sent>