



CJKI Comprehensive Hanzi Pinyin Database Special Features

The CJKI Comprehensive Hanzi Pinyin Database for Simplified Chinese is a comprehensive lexical resource covering over 600,000 entries of general vocabulary, technical terms, and proper nouns. In addition to its extensive coverage and high level of accuracy, the following special features help distinguish this resource as one of the most complete and accurate pinyin databases available for licensing today.

Type and POS Codes

The database contains explicit codes to indicate whether the headword is a technical term, company name, personal name, general vocabulary, etc., as well as POS (part-of-speech) codes.

Polyphone Coverage

The database properly covers all polyphones, e.g., words written with identical pinyin but having different tones and different meanings.

东西	dong1-xi1	east-west
东西	dong1-xi0	thing

Classical Pinyin Coverage

The database covers classical pinyin, and distinguishes pinyin based on meaning whether the word is read in the classical pronunciation as opposed to the modern one.

	Modern	Classical		
见	jian4	xian4	图穷匕见	tu2-qiong2-bi3-xian4
景	jing3	ying3	蹶景追飞	nie4-ying3-zhui1-fei1
渐	jian4	qian2	沉渐刚克	chen2-qian2-gang1-ke4

Alternative Pinyin

The database provides correct pinyin for the neutral tone, along with alternatives, in accordance with the 6th edition of XianDai Hanyu Cidian.

机器	ji1-qi0	ji1-qi4
因为	yin1-wei0	yin1-wei4

Complements

Careful attention has been given to pinyin variation when used in complements, as opposed to ordinary ways.

起来	qi3-lai0
看起来	kan4-qi0-lai0
说得来	shuo1-de0-lai2
起不来	qi3-bu0-lai2
拿得起来	na2-de0-qi3-lai2
跑得起步来	pao3-de0-qi3-bu4-lai0
干得了	gan4-de0-liao3
搞不懂	gao3-bu0-dong3
办得不好	ban4-de0-bu4-hao3
哭个不停	ku1-ge0-bu4-ting2

Attention to Neutral Tone

Close attention has been paid to the accuracy of neutral tones based on the actual usage within compound words.

学生	xue1-sheng0
大学生	da4-xue2-sheng1
精神文明	jing1-shen2-wen2-ming2
精神旺盛	jing1-shen0-wang4-sheng4
正经	zheng4-jing0
正经八百	zheng4-jing1-ba1-bai3