

Monophthongal Transcription of Pinyin <ao>

It is overwhelmingly taken as fact that the Standard Chinese (SC) sound inventory contains the diphthong /ao/, represented in Pinyin as <ao> (Duanmu 2007). Its classification as a diphthong is essential to SC phonemic analysis (Duanmu 2011). I present evidence from acoustic analysis of every monosyllabic word containing <ao> that it is a monophthong. Specifically, it is the low-mid back rounded vowel /ɔ/. This analysis explains why <ao> functions differently from other SC diphthongs in that it does not rhyme with the vowel that makes up the latter part of the diphthong, /u/. It also raises yet another question regarding missing forms in SC. If <ao> is a monophthong, it is supposedly eligible to be followed by a nasal coda, but such words are nonexistent.

Based on generalizations about vowel qualities and their correlations with formant frequencies given in Ladefoged & Johnson (2011) and Yavaş (2020), I predicted that the mean F1 and F2-F1 values of the target vowels would show slight bidirectional shifts between halves. I also predicted that the average F1 frequency would remain just above 300 Hz and the average F2-F1 difference would remain under 250 Hz, indicating that the target vowel is a low-mid back monophthong.

To test these hypotheses, I sent a list of every monosyllabic word containing the target phoneme and emailed it to two native speakers, who recorded themselves reading the list aloud. This analysis excludes the words with glides in their onsets. I used Praat 6.4.47 (Boersma & Weenink 2025) to visualize and annotate the spectrograms and retrieve the mean F1 and F2 frequency from each half of each vowel duration. I then compared, between the two halves, the F1 and F2-F1 values.

The results of the acoustic analysis strongly support the hypothesis that <ao> is a monophthong. RM ANOVAs fitted using the *rstatix* package (Kassambara 2023) in R (version 2025.09.2+418) confirmed that the vowel consistently became higher and less back, although minimally. F2-F1 was highly varied and bidirectional across speakers and items. Some of this variation is due to issues of recording quality causing outliers in the F2 frequency data.

The ranges of the mean F1 and F2-F1 values of each word indicate that <ao> is a mid-low and back or near-back vowel. Since its roundedness can be confirmed by seeing a person articulate the vowel, I propose that it should be transcribed as /ɔ/. This answers one question about SC rhyming sets, but raises questions about current theories of maximal syllable structure and its power to explain some missing forms.

References

- Boersma, Paul & Weenink, David (2025). Praat: doing phonetics by computer [Computer program]. Version 6.4.48, retrieved 13 December 2025 from <https://praat.org>
- Duanmu, S. (2007). *The phonology of Standard Chinese* (2nd ed.). Oxford University Press.
- Duanmu, S. (2011). Chinese syllable structure. In Marc Van Oostendorp, Colin Ewen, Elizabeth Hume, & Karen Rice, eds., *The Blackwell Companion to Phonology*, (vol. 5, pp. 2754-2777). Wiley-Blackwell.
- Kassambara A (2023). *rstatix*: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.2, <https://rpkgs.datanovia.com/rstatix/>.
- Ladefoged, P. & Johnson, K. (2011). *A course in phonetics* (6th ed.). Wadsworth, Cengage Learning.
- Yavaş, M. (2020). *Applied English phonology* (4th ed.). Wiley-Blackwell.