

The North Wind and the Sun in Lingua-U

An experimental transcription — working note, not part of the whitepaper

The passage below is the standard specimen text of the International Phonetic Association, chosen because it is the most widely transcribed paragraph in phonetics and so invites comparison with existing IPA renderings.

Method and conventions

Every glyph here was generated arithmetically from the tables of the whitepaper rather than selected by hand.

- **Consonants** are tetragrams, by the formula of Chapter 14: `codepoint = U+1D306 + address3`. The block runs U+1D306 to U+1D356 and is exactly eighty-one wide, which was checked against Unicode before anything was transcribed.
- **Cardinal vowels** take monograms and **secondary vowels** digrams, from the lookup table of Table 14.1. These cannot be derived arithmetically, since four of them were unified with existing Yijing symbols.
- **Diphthongs** are written as sequences of two vowel symbols — /eɪ/ as 11·21 and so on. At glyph level the sequence separator becomes unnecessary: each phoneme is exactly one character, so adjacency is unambiguous without it.
- **Rhotacized vowels** /ɜ̃/ and /ɔ̃/ are written 10^r — the mid-central digram bearing a rhotic modifier, per Chapter 8, not the Center followed by /ɹ/.
- **The schwa** has no glyph, because it has no address. It is shown as ☉ below. **This placeholder is an ad-hoc reading aid invented for this experiment and forms no part of the system.** Without it the Center would be invisible, which turns out to be the most interesting thing in the transcription — see the last section.

The reference variety is General American. Two segmentations are contestable: *traveler* is taken as two syllables, and every /ɔ/ here belongs to a speaker who has not merged *cot* and *caught*. A merged speaker would read ☰ as — throughout — six substitutions in the passage.

Word by word

Word	IPA	Addresses	Lingua-U
The	/ðə/	1011 —	☰☉

Word	IPA	Addresses	Lingua-U
North	/nɔɹθ/	2211 12 2121 1010	☰--☶☰
Wind	/wɪnd/	2001 21 2211 0111	☳--☱☳
and	/ənd/	– 2211 0111	○☱☳
the	/ðə/	1011 –	☳○
Sun	/sʌn/	1110 10 2211	☳--☶
were	/wə/	2001 10 ^r	☳-- ^r
disputing	/dɪspʊtɪŋ/	0111 21 1110 0100 2021 2 0110 21 2221	☳--☳☳☶--☳--☶
which	/wɪtʃ/	2001 21 0220	☳--☳
was	/wəz/	2001 – 1111	☳○☳
the	/ðə/	1011 –	☳○
stronger,	/stɹɔŋgə/	1110 0110 2121 12 2221 0121 10 ^r	☳☳☶--☱☶-- ^r
when	/wɛn/	2001 11 2211	☳--☶
a	/ə/	–	○
traveler	/tɹævlə/	0110 2121 01 1001 2111 10 ^r	☳☶--☳☳-- ^r
came	/keɪm/	0120 11·21 2201	☳----☶
along	/əlɔŋ/	– 2111 12 2221	○☳--☶
wrapped	/ɹæpt/	2121 01 0100 0110	☶--☳☳
in	/ɪn/	21 2211	--☶
a	/ə/	–	○
warm	/wɔɹm/	2001 12 2121 2201	☳--☶☳
cloak.	/kloʊk/	0120 2111 12·22 0120	☳☳----☳
They	/ðeɪ/	1011 11·21	☳----

Word	IPA	Addresses	Lingua-U
agreed	/əɡɹɪd/	– 0121 2121 1 0111	○ ☰ ☰ _ _ ☰
that	/ðæt/	1011 01 0110	☰ _ ☰
the	/ðə/	1011 –	☰ ○
one	/wʌn/	2001 10 2211	☰ _ ☰
who	/hu/	0020 2	☰ _
first	/fɜːst/	1000 10 ^r 1110 0110	☰ _ ^r ☰ ☰
succeeded	/səkʰsɪdɪd/	1110 – 0120 1110 1 0111 21 0111	☰ ○ ☰ ☰ _ _ ☰ _ ☰
in	/ɪn/	21 2211	_ ☰
making	/meɪkɪŋ/	2201 11·21 0120 21 2221	☰ _ _ ☰ _ ☰
the	/ðə/	1011 –	☰ ○
traveler	/tɹævlə/	0110 2121 01 1001 2111 10 ^r	☰ ☰ _ ☰ ☰ _ ^r
take	/teɪk/	0110 11·21 0120	☰ _ _ ☰
his	/hɪz/	0020 21 1111	☰ _ ☰
cloak	/kloʊk/	0120 2111 12·22 0120	☰ ☰ _ _ ☰
off	/ɔf/	12 1000	_ ☰
should	/ʃʊd/	1120 22 0111	☰ _ ☰
be	/bi/	0101 1	☰ _
considered	/kənsɪdəd/	0120 – 2211 1110 21 0111 10 ^r 0111	☰ ○ ☰ ☰ _ ☰ _ ^r ☰
stronger	/stɹʌŋgə/	1110 0110 2121 12 2221 0121 10 ^r	☰ ☰ ☰ _ ☰ ☰ _ ^r
than	/ðæn/	1011 01 2211	☰ _ ☰
the	/ði/	1011 1	☰ _
other.	/ʌðə/	10 1011 10 ^r	_ ☰ _ ^r

The passage

Orthography

The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other.

Phonemic

/ðə nɔ:θ wɪnd ænd ðə sʌn wə dɪspju:tɪŋ wɪtʃ wəz ðə strɒŋgə wɛn ə trævələ keɪm əlɒŋ
 ɹæpt ɪn ə wɔ:m klɒk ðeɪ əɡɹɪd ðæt ðə wʌn hu fɜ:st sɛksɪdɪd ɪn meɪkɪŋ ðə trævəl-
 teɪk hɪz klɒk ɒf ʃʊd bi kɛnsɪdəd strɒŋgə ðæn ði ʌðə/

Lingua-U

NO H=HE H=HN ONN NO U=H H=r
 H=HEH--E=H H=E HON HON HEH=HEH=r
 H=H O EH=HN=r E=H ON=H H=EH =H
 O H=HE EH=E H= OEH..H H=EH HON
 H=H E-- H=rEE UOEH..H=H =H H=EH=H
 HON EH=HN=r E=EH E=H EH=E EH=H
 H.. EOH=H=r H EH=HEH=r H=H H.. =H=r

What the transcription shows

One hundred and sixty segments, thirty-five distinct. The passage is built from a small alphabet used at very unequal rates.

The most frequent sound in the passage has no glyph. The schwa occurs thirteen times — more than any other segment, ahead of /ð/ and /n/ at ten each. The system’s busiest sound is the one it declines to address, and a transcription of ordinary English prose is therefore punctuated with holes. This is a consequence of the Chapter 8

decision rather than a defect in it, but it is much more visible in running text than it is in a table, and it is worth deciding whether that is acceptable before the notation is used for anything longer.

Spelling and sound come apart immediately. *Wrapped* opens on 𐍂𐍅, the /w/ tetragram, with nothing corresponding to its written *w*. *Who* opens on 𐍂𐍃, the /h/ tetragram, likewise. *One* opens on 𐍂𐍄, /w/, which its spelling does not contain at all. Three words, three failures of the written form to predict the spoken one, in a single sentence.

Identical sounds produce identical strings. *Cloak* appears twice as  and *stronger* twice as ^r. Less obviously, *were*, *first*, *stronger*, *traveler*, *considered*, and *other* all end in ^r — a rhotacized ending that English spelling distributes across *-ere*, *-ir*, *-er*, and *-ered*.

The tetragram names remain arbitrary, as Chapter 14 warns. /ð/, the second most frequent segment here, falls on TETRAGRAM FOR LEGION; /m/ lands on CLOSURE, which is apt by chance; /g/ on HOLDING BACK, which is not. Nothing in the transcription should be read through those names.

A practical caution

These characters lie in the Supplementary Multilingual Plane and are absent from most common typefaces. The tetragrams have reasonable coverage; several of the digrams do not, and will render as boxes in many environments. Anyone reproducing this will need to specify a font that covers both U+2680-U+268F and U+1D300-U+1D356, and should expect the vowels to break before the consonants do.