PUBLISHED 2026-10-20 BY THOMAS C. VEATCH AS A DEFENSIVE PUBLICATION — NOT A PATENT FILING — https://tomveatch.com/tb/

Invention Disclosure: Transliterating Browser and Camera (PUBLISHED 2026-10-20 as a defensive publication)

0. Cover block

Title: Transliterating Browser and Camera: pronunciation-line display that fades with the reader's measured mastery, for reading acquisition

Inventor: Thomas C. Veatch, PhD (linguistics, phonetics), sole inventor. Tom's own opening words: "My invention and responsibility is to create and support ..." the system described here.

Date of this draft: 2026-10-16

Status: PUBLISHED. Decision of the inventor, 2026-10-20: no patent application will be filed. This document, first drafted 2026-10-16 as an invention disclosure for attorney review, is published openly and dated so that it stands as prior art against anyone who would patent these mechanisms. Sections 8-11 are kept as written, as the record of why; section 3.2 gives the reason the decision went this way.

Publication: https://tomveatch.com/tb/disclosure.php (this PDF), https://tomveatch.com/tb/disclosure.php?f=md (source), https://tomveatch.com/tb/ and https://tomveatch.com/tc/ (summary pages), https://tomveatch.com/tb/conception-record.md.php (dated record of the inventor's words). All open to the web and to search engines from 2026-10-20.

Records of conception (dated, kept beside this file as conception-record.md.php):

- 2002, Tamil Nadu: Tom's invention of the phonetic chart labelled with drawings of the mouth instead of the Latin and technical words for place, manner, height and rounding, so that phonetic literacy is open to all and not only to university students of phonetics. This is the alphabet-teaching premise (2) of the present invention and is a separate, earlier invention; the attorney should treat any publication of it as the inventor's own prior art. Tom's dated statement of its origin: 2026-10-16, this session.
- 2026-09-12, Tom's "Reach 2" session: the literacy thesis in his words ("every illiterate person sees every sign as a stop sign"; illiteracy as the hidden driver of poverty, dropout, caste and crime) and a restatement of the 2002 chart invention. Transcript line 1456, 07:17:37.
- 2026-10-16, this session: Tom's statement of the Transliterating Browser and Camera (the quoted statement in section 2.1). An earlier typing of the same statement that day was lost when the session process restarted; the quoted text is Tom's own reconstitution, typed the same day, and is the earliest surviving record of the TB/TC statement. Later messages the same day add the optional user's dialect claim, the two-population point, and the bracket marking.
- 2026-10-16, the page tb/index.php in Tom's own words, with its edit history in the tomveatch.com git repository (commits 967d178e, 38e8db31 and following).
- 2026-10-16, the design discussion recorded in section 3 (elements A-H and the addenda).

Public disclosure: this publication, 2026-10-20, by the inventor's decision; and the inventor's own earlier publications in https://tomveatch.com/tg/ (section 3.2).

Abstract (one paragraph): A browser module and a camera application that take any text a reader meets -- a web page, or printed text seen through a phone camera -- and show, registered under or over each word, a pronunciation line: a phonemic transcription of that word in a phonetic alphabet (IPA or another phonetic alphabet such as tvbet), in a chosen prestige dialect. The plain spelling (S line) and the pronunciation line (R line) are shown together in two lines (HTML ruby, as for Japanese furigana). Three mechanisms distinguish the system from a plain transliterator. First, a confidence rule: each word's pronunciation carries a confidence value; a word with one confident pronunciation gets its pronunciation line; a word with two or more candidates above threshold shows the candidates as options; a word with none is shown in plain spelling only, never a guess, and a visible bracket marking (phonetic, i.e., pronounceable, i.e., trustworthy, a safe space) tells the reader which spans are verified. Second, a mastery rule: a per-reader learner model, keyed by grapheme-phoneme pattern and by word, decides per word whether the pronunciation line is shown, faded, or hidden, so the pronunciation line fades toward plain spelling as the reader's measured mastery grows -- the slope, not the cliff. Third, the learner model is initialized from a baseline read-aloud and updated from the reader's own behavior (taps for audio, dwell) and from read-aloud scoring by automatic speech recognition. Optional user's dialect conversion turns the prestige-dialect transcription into the transcription of the reader's own dialect by systematic correspondence rules keyed on lexical sets. The camera overlay runs under the same checks. An evaluation method measures each learner's time to crossover, the minute at which the learner reads more words aloud correctly per minute from the converted page than from the plain page.

1. Field of the invention

The invention is in computer-assisted reading: software that changes how written text is displayed to a particular reader in order to teach that reader to read. More narrowly it concerns (a) web browsers and browser extensions that modify the rendering of page text, (b) camera applications that detect text in a live image and overlay information on it, (c) grapheme-to-phoneme conversion and pronunciation lexicons, (d) learner models that measure a reader's skill and adapt a display, and (e) automatic speech recognition used to

score reading aloud.

2. Background and the problem

2.1 The invention in the inventor's words

Tom's statement of 2026-10-16 (his own reconstitution, the same day, of a text lost in a session restart; see the cover block):

> "Our job is to create and support a browser or browser module and camera capability, a Transliterating Browser/Camera. All text detected will be transliterated into IPA or tvbet so that semi-illiterates who know the phonetic alphabet can read it as pronounced in the prestige dialect, possibly multi-linear with the original orthography. This will solve the world illiteracy and semi-literacy problem. (1) We can rely on the phonetic/phonological knowledge of native speakers of any dialect or language. (2) We can teach a phonetic alphabet in short order, with ease, with sincere user appreciation and cooperation. (3) Dialect speakers can understand speech in other (perhaps higher prestige) dialects. Given these, any illiterate can see any text on the web or in the environment converted into a form they can read with increasing skill, fluency, comfort. Soon, reading for fun and information. Soon, excellent readership. With multi-linearity (S & R tiers only) regular orthographic reading also follows. Societal consequences are inconceivable under my view that illiteracy is a driving causal factor in most societal ills such as crime, poverty, lack of education, anti-intellectualism, class conflict, caste."

Tom's definition of the invention, later the same day (2026-10-16), separating it from his earlier chart invention:

> "The IPA chart idea was invented in 2002 in Tamil Nadu: drawings instead of educated and latin words for place and manner and height and rounding, make phonetic literacy open to all not just college and graduate students and linguistic scientists in phonetics/phonology classes in linguistics departments. The idea of using a browser/camera transliterator to generate a translation-graph (tiered DAG) with a phonetic alphabetic tier for observed text looked at by or in the browser/camera as a way to provide immediate (phonetic) literacy to illiterates and semi-literates, depending on phonetic-alphabet knowledge by the otherwise semi- or il-literate and on cross-dialectal comprehension from their own to that of the transliteration-target dialect, this is today's idea, 10/16/2026. This is what solves America."

So the thing the browser and camera produce is not a loose line of symbols under each word but a translation graph: the tiered directed acyclic graph (DAG) of Tom's Translation Graph (TG) system, in which a span is identified by (from, to, tier), the nodes carry no content and the edges carry the meaning, and the tiers include the source orthography (S) and the phonetic transcription (R), with further tiers (uncontracted form B, underlying form F, gloss E) attachable later. The two-line display of this disclosure is a rendering of the S and R tiers of that graph.

In the 2026-10-16 rewrite (tb/index.php) Tom made the alphabet explicit: text is "transliterated into IPA or another phonetic alphabet such as tvbet". IPA and tvbet are embodiments; the invention is alphabet-agnostic.

2.2 Why English spelling is the problem

English spelling is a deep orthography: the same letters spell many sounds and the same sound is spelled many ways. Seymour, Aro and Erskine (2003) compared children learning to read in thirteen European orthographies and reported that "the rate of development in English is more than twice as slow as in the shallow orthographies"; at the end of the first school year familiar-word reading accuracy was typically above 90% in most languages but 34% in the English (Scottish Primary 1) sample, and the English group reached the 90% level only at a reading age of about 7 years 4 months (https://eclass-b.uoa.gr/modules/document/file.php/PPP363/Seymour_Aro_Erskine.pdf). (Note for the record: the figure "2.5 years behind" sometimes quoted does not appear in that paper; the paper's own words are "more than twice as slow".)

A pronunciation line in a phonetic alphabet turns English, for the reader who has it, into a shallow orthography.

2.3 How many readers are affected

United States, children: on the 2024 National Assessment of Educational Progress, 40% of fourth-grade students performed below the NAEP Basic level in reading, 60% at or above Basic, and 31% at or above Proficient (https://www.nationsreportcard.gov/reports/reading/2024/g4_8/national-trends/). The National Assessment Governing Board called the below-Basic share "the largest percentage since 2002" (https://www.nagb.gov/content/dam/nagb/en/documents/newsroom/press-releases/2025/naep2024-pressrelease-final.pdf). NAEP levels are cut scores, not grade equivalents; "below Basic" is not the same as "illiterate". Tom's "two thirds" (section 2.4) corresponds to the 69% of fourth graders below NAEP Proficient, the level the Nation's Report Card treats as solid grade-level reading, not to the 40% below Basic.

United States, adults: in the 2023 PIAAC survey, 28% of US adults aged 16-65 scored at literacy Level 1 or below (up from 19% in 2017) and 44% scored at Level 3 or above (down from 48%) (https://ies.ed.gov/learn/press-release/u-s-adults-score-par-international-average-literacy-skills-below-international-average-numeracy-and). The share at Level 2 or below is therefore about 56% by subtraction; the exact distribution should be read from the NCES figure before quoting [CITE].

2.4 The two populations

Tom, 2026-10-14: "The English 3.5th grade mark includes 66% future adult Basic-and-lower semi-literates who by 5th grade are locked on their losing strategy, their known ignorance, and their known non-learnability. And it includes the properly supported sand-dune-climbing better-resourced kids too; that's apples and oranges in one insufficient

statistic."

So there are two groups of readers, and one average hides them:

(i) **Supported readers**, who are climbing the sand dune of English spelling with enough help to get up it. For them, in Tom's words, "a Transliterating Browser or Camera saves time."

(ii) **The locked-strategy semi-literate**, in Tom's words "locked on a losing strategy (guessing words from a first consonant and first vowel and context), with not just known, self-aware ignorance, but also known non-learnability." For them "it is the difference between reading and never reading competently or comfortably."

Evidence that the second group exists and is locked in early: Juel (1988) followed 54 children from first through fourth grade and found that "the probability that a child would remain a poor reader at the end of fourth grade if the child was a poor reader at the end of first grade was .88", and that the poor readers' lack of decoding skill "prevented the poor readers from being able to read as much text -- both in and out of school -- as the good readers" (https://people.uncw.edu/kozloffm/learningtreadandwrite.pdf). Stanovich (1986) named the resulting divergence the Matthew effect, rich get richer and poor get poorer in reading (https://www.readingrockets.org/resources/resource-library/matthew-effects-reading-some-consequences-individual-differences). The guessing strategy itself is documented by Hanford (2019), who quotes Stanovich's 1975 finding that "it was the poorer readers, not the more skilled readers, who were more reliant on context to facilitate word recognition" and describes children who guess from the first letter and the picture and are "stuck" at about third grade when the books lose their pictures (https://features.apmreports.org/files/at-a-loss-for-words-printable.pdf). The "known non-learnability" of the second group is the learned-helplessness pattern reported for poor readers by Butkowsky and Willows (1980) [CITE: full text not opened; wording to be verified] and summarized for learning-disabled students by Arnold (1996/97) (https://www.ldonline.org/ld-topics/self-esteem-stress-management/learned-helplessness-and-attribution-success-and-failure-ld). No primary source was found that names grade 5 specifically as the lock-in point; the best available evidence is Juel's grade 1 to grade 4 persistence and Hanford's "about third grade" [CITE].

2.5 The Initial Teaching Alphabet and its three reasons for dying

Pitman's Initial Teaching Alphabet (ITA; England, from 1961) was a phonemic alphabet for beginning readers. Its director's own report found ITA beginners "significantly in the lead" after five months, with the lead growing, and most children reaching fluency in ITA "during the second year of schooling", after which each child was transferred to standard print (Downing 1964, https://www.gwern.net/doc/psychology/1964-downing.pdf). Downing's 1967 conclusion was that ITA "produces superior results in traditional orthography reading and spelling by the end of third year of school" (as quoted at https://en.wikipedia.org/wiki/John_Downing_(educational_psychologist)). By 1966, 140 of the 158 English education authorities used it in at least one school; it then fell into disuse (Hall 2005, https://www.eyemagazine.com/feature/article/the-initial-teaching-alphabet).

Tom's diagnosis, in his own words: ITA "died for three reasons: no such text existed outside the classroom, and the switch to normal spelling was a cliff, without even the support of [ ]-brackets to indicate trustworthy vs in-the-wild orthographic text. Here every page converts, both forms co-exist with the [ ]-brackets to identify safe space, and the two-line display turns the cliff into a climbable slope, a learning resource. Those three fixes are the whole difference."

The historical record supports the first two reasons directly. Hall (2005) reports that ITA children "could not immediately read the words they saw about them every day" and had "far fewer books available to them", and that the transition to standard spelling was identified "from the outset" as "the major sticking point in the scheme's success" (same URL). Wikipedia's summary adds that some children "could not carry ITA skills over to standard spelling or were confused by two alphabets" (https://en.wikipedia.org/wiki/Initial_Teaching_Alphabet). The third reason -- no visible marker telling the child which text was safe ITA text and which was in-the-wild spelling -- is Tom's observation; no source discussing it was found [CITE].

The invention answers all three: the browser and camera convert every page and every sign (reason one); the mastery rule fades the pronunciation line word by word as the reader masters each pattern, so there is never a single switch-over day (reason two); and bracket marking (phonetic, i.e., pronounceable, i.e., trustworthy, a safe space) shows the reader which spans carry a verified pronunciation line and which are in-the-wild spelling (reason three).

2.6 Two existence proofs: furigana and zhuyin

Japanese children read kanji with furigana, small kana printed beside each character giving its reading; it is "common to use furigana on all kanji characters in works for young children" (https://en.wikipedia.org/wiki/Furigana). A funded Japanese study of 1,699 primary schoolchildren concluded that learning with furigana "has a great effect on mastering not only how to read kanji but also writing kanji" (Tanahashi 2003-2005, https://kaken.nii.ac.jp/grant/KAKENHI-PROJECT-15530589). Taiwan prints zhuyin (bopomofo) beside Chinese characters in children's books for the same purpose [CITE: no source opened by the scouts; well known, attorney may wish a citation]. For a Mandarin speaker, zhuyin beside characters IS this invention's display, done by hand on paper. Both prove that a pronunciation line beside a hard script leads to literacy in the hard script.

2.7 The known limit

Tom: "where the written language diverges incomprehensibly from the spoken one (Arabic, Swiss German, Chinese 'dialects'): there transliteration gives sound but not meaning. For semi-phonetic orthographies like English, French, Thai, Hindi, etc., it gives the whole benefit." (Tom revised the list on 2026-10-20: the targets are the spellings that have drifted far from the sound -- English, French, Thai, and Chinese through zhuyin or pinyin; Hindi, Spanish, Italian and Finnish can be sounded out already and are not targets.) A Cantonese speaker reading Mandarin text with zhuyin gets the Mandarin

sounds, which are not her words. The user's dialect conversion (section 4.5) is the mechanism that moves a case from the limit side to the benefit side wherever the sound correspondence between the prestige dialect and the reader's dialect is systematic.

3. Summary of the invention

A. Two-line rendering. Under or over each word of the plain spelling (S line) the system places a pronunciation line (R line): the word's phonemic transcription in a phonetic alphabet, in a chosen prestige dialect. In a web page this is an HTML ruby element, exactly the mechanism used for Japanese furigana; in the camera it is drawn into the live image at the position of the detected word. In the language of Tom's existing Translation Graph (TG) tiers, "a Transliterating Browser/Camera adds a phonetic/phonemic tier to text in any web page or camera image." More exactly (Tom, 2026-10-16): the browser or camera generates, for the observed text, a translation graph -- the TG tiered DAG, spans identified by (from, to, tier) -- with a phonetic-alphabet tier, and the display reads the S and R tiers of that graph.

B. Coverage and the confidence rule. Every pronunciation the pipeline produces carries a confidence value. The coverage metric is the fraction of words on the page with a confident pronunciation. The confidence check has three outcomes per word: (i) one confident pronunciation, shown; (ii) two or more candidates above threshold (heteronyms such as read, lead, bass; an unresolved Thai tone), shown as options, compactly or on tap, and the reader's pick is learner-model data; (iii) none, plain spelling only, never a guess. Bracket marking (phonetic, i.e., pronounceable, i.e., trustworthy, a safe space) indicates the text is to be interpreted as phonetic or phonemic text; thus, both trustworthy and visibly separate from in-the-wild orthography.

C. The mastery rule. A per-reader learner model stores mastery values keyed by grapheme-phoneme pattern and by word. The pronunciation line for a word is shown in full while the reader's mastery of that word's patterns is below a low threshold, is faded (reduced opacity, size or contrast) between the low and high thresholds, and is hidden above the high threshold, with the plain spelling always present. This is the slope, not the cliff. A never-seen word whose patterns are all mastered already needs no pronunciation line.

D. Prestige-dialect selection. The pronunciation line is the pronunciation of a selected prestige dialect (for English, General American as encoded in Tom's cmudict-derived tvbet lexicon; other choices possible). Optionally, user's dialect conversion rewrites that transcription into the reader's own dialect by systematic correspondence rules keyed on lexical sets, so the reader sees each word as, or more nearly as, they themselves say it.

E. Camera overlay. The camera application detects text in the live image (OCR), passes the words through the same pronunciation pipeline, confidence rule and mastery rule, and draws each word's pronunciation line registered under or over the detected word in the image. In Tom's words, the camera overlay runs "under the same checks" as the browser.

F. Scripts without word spaces. For Thai and similar scripts the pipeline first segments the text into words and derives tone by rule (initial-consonant class x live/dead syllable x vowel length x tone mark) before transliteration; where the rule cannot decide, the confidence rule applies.

G. Learner-model update. The learner model is updated from the reader's behavior while reading (taps for audio, dwell time, picks among candidates) and from read-aloud scoring by automatic speech recognition, at the level of grapheme-phoneme patterns and of words. It is initialized at intake from a baseline read-aloud of a plain page before any alphabet instruction.

H. Evaluation by time to crossover. For each learner, two lines are drawn on one chart: words read aloud correctly per minute from the converted page and from the plain page, against minutes of phonetic-alphabet instruction and practice. The crossover is the minute at which the converted page wins. The result for a population is the fraction of learners who have crossed as a function of minutes; a learner who quits before crossing counts as not yet crossed at the minute they quit. The power of any variant (alphabet, mastery rule on or off, user's dialect conversion on or off, bracket marking on or off) is how much sooner its crossover distribution sits. For adults the crossover must arrive before the hour at which they would have quit; for children the x axis is in school years. The headline number is the crossover of group (ii).

3.1 Relation to the general Translation Graph display (Tom, 2026-10-17)

The Transliterating Browser and Camera are a special case of a more general thing: Translation Graphs (TG) shown in browsers, camera views and media players (for example a YouTube module) for bilingual learning, with translation (T), gloss (E), underlying (F), uncontracted (B) and other tiers. The special case is the easy one because no translation is needed, only lookup, and only the S and R tiers are used; but it is strictly a special case, and claim 21 says that further tiers are attachable to the same word positions.

The general case is not claimed, and the attorney should not spend a claim on it: aligned translation tiers over web pages, video subtitles and camera text are already shipped and published -- Language Reactor (dual subtitles on YouTube and Netflix, word-level glosses, known words hidden), LingQ and Readlang (in-page glosses with known-word marking), Google Lens (camera overlay), interlinear glossed text, and on the data side ELAN tiers and the ISO Linguistic Annotation Framework / GrAF (graph-structured multi-tier annotation, about 2008) [CITE: URLs to be added]. A broad claim to tiered aligned display would be both an abstract idea and anticipated. What is claimed here are the specific mechanisms of the special case, whose effect on decoding is measurable: the confidence rule, mastery per grapheme-phoneme pattern with the fade, live struggle detection, bracket marking, and user's dialect conversion. Mastery per morpheme rather than per whole word is the one mechanism that might carry over to a gloss tier; it is left for a later application.

3.2 The inventor's own prior publications (found 2026-10-19; the attorney must read these first)

The general case is not only unpatentable; it was published by the inventor himself, and those publications are the closest prior art to this disclosure. The directory https://tomveatch.com/tg/ is served to the open web (checked 2026-10-19: every file below answers HTTP 200). What each discloses:

- https://tomveatch.com/tg/tg.apps.txt ("Translation Graph Markup Language: Applications", by Tom Veatch, 2018; the same text appears in https://tomveatch.com/tg/call.for.multilinearity.txt, which carries dates 2000-2017): the tier set including "L2Surface-Script1 (transliterated for readability)" with "A|B is a transliteration system"; a script-teaching application in which each letter gets an arc parallel to "IPA or Roman equivalent, audio form"; "Presentation for learning may be computer controlled based on a model of the reader/learner's knowledge, or manually parameterized"; "A map to an IPA reference and to a pronunciation guide and script description should be a click away"; "All this may be hidden and the image/video/audio media (dis-)played, with the display of all tiers or a parameterized, selected subset, a click away during playback when the audience is puzzled and wants to understand the part they just heard" (the Hunh? button, in substance); "scan them, then build up what you have into parsed, understandable TG documents" (scanned documents into a TG); "Or the TG UI can demand interactivity from the user, to learn what they know. Repeat and get some phonetic training for your accent"; and "confidence-weighted character/word/morph" alternatives in TG construction.
- https://tomveatch.com/tg/tg.introduction.txt (dates 2003, 2008, 2017): the same learner-model-controlled presentation and hidden-until-puzzled passages.
- https://tomveatch.com/tg/tg.definition.txt (2018, 2025): "Citation phonetics and phonology are provided to the learner in integrated but separate presentation, available any time relative to the symbols of the script"; "Fast speech and dialect phonetics are learned by reading phonological or citation-form transcripts associated with segmented audio."
- https://tomveatch.com/tg/BusinessModel.vibe (a design conversation dated 2025-09-30; present in the public directory on the server since 2025-09-30 22:39 PDT by its file date, in the site's git history since 2026-03-22): names the "Hunh? button" ("present for support when Hunh? is indicated or detected"), and a per-user knowledge model with heard/said counts, "struggled_with" flags, "mastery_level" and drills until mastery, feeding the next session.
- https://tomveatch.com/tg/TGonVideo.txt: TG frames over video, timed to segments, bouncing-ball highlighting.

None of these mentions illiterate or semi-literate native speakers, a per-grapheme-phoneme-pattern mastery value, a graded fade of the pronunciation line, a rule that withholds any pronunciation below a confidence threshold, bracket marking of verified spans, conversion of a displayed transcription into the reader's own dialect, a live camera overlay, or a baseline read-aloud intake (searched 2026-10-19 for those words and their variants: none found). But they do disclose the genus -- a transliteration or IPA tier over any document, hidden or shown under a model of the learner's knowledge, with audio a click away -- and an examiner will cite them against claims 1, 2, 7, 8 and 22. The claims therefore rest on the species: the specific rules and their measured effect, not on the tiered display or on the learner model as such.

Grace period consequence: the 2003-2018 publications are far outside any grace period and are prior art everywhere. BusinessModel.vibe, public since 2025-09-30 by its server file date, passed the US one-year grace of 35 USC 102(b)(1)(A) on 2026-09-30, so its content (the Hunh? button; struggle flags and mastery levels in a learner model) is also prior art even in the US. Nothing in it anticipates the species claims above, but claim 22 (live struggle detection) must be distinguished from "present for support when Hunh? is indicated or detected": the distinction is that claim 22 detects the struggle from read-aloud alignment at the word, and answers it by showing that word's pronunciation line in the text, under the mastery rule, on arbitrary page or camera text.

Publication note: nothing further about the general TG-in-media idea should be published before the provisional is filed, and nothing in more detail than the provisional contains; the directory tg/ is already public and should be left as it is (taking it down changes nothing and looks like concealment).

4. Detailed description with embodiments

4.1 Browser module or extension

The browser embodiment is a browser extension (or a built-in browser module) with content-script access to the page's document object model (DOM).

1. **Text collection.** The content script walks the DOM and collects text nodes inside visible elements, skipping script, style, code and editable fields, and skipping elements already inside a ruby element.
2. **Tokenization.** Each text node is split into word tokens and separators. For languages with word spaces, tokens are maximal runs of letters, apostrophes and in-word hyphens; punctuation and spaces are separators and are left in place. For scripts without word spaces (section 4.2, step 2) a segmenter supplies the token boundaries.
3. **Lookup.** Each distinct token (case-folded, with a copy of the original for display) is sent to the pronunciation pipeline (section 4.2), which returns zero, one or several candidate transcriptions, each with a confidence value and the list of grapheme-phoneme patterns it exercises. Lookups are cached per session.
4. **Deciding.** For each token the confidence rule (section 4.2, step 6) and then the mastery rule (section 4.3) decide an outcome: show, show-faded (with a fade level), show-candidates, or plain.
5. **Rendering.** Each token is replaced by a ruby element whose base is the original spelling (S line) and whose ruby text is the pronunciation line (R line), styled

according to the outcome (section 4.9). Separators are untouched, so the page's text content, selection and search behavior are preserved. A plain-outcome token is left as is, outside any bracket marking.

6. **Dynamic pages.** A MutationObserver watches for added or changed text nodes and processes them the same way; nodes the extension itself inserted are tagged and ignored, preventing loops. Infinite-scroll pages and single-page applications are therefore handled as they load.
7. **Controls.** A per-page toggle, a per-site allow list, the prestige dialect selector, the user's dialect selector, the alphabet selector (IPA, tvbet, other), and a mastery-rule override (show all, follow the rule, hide all) are available to the reader. A tap on any word plays its audio (section 4.7) and is recorded as learner-model data (section 4.8).
8. **Where the pipeline runs.** The lexicon and rules can run entirely on the device (preferred for privacy and latency) or on a server; the learner model is stored per user account so that browser and camera share it.

4.2 The pronunciation pipeline

For each word token, in order:

1. **Overrides file.** A curated per-language overrides file is consulted first. An entry here has confidence 1.0 and wins. (This is how Tom's English R tier is already built: cmudict.tvbet.json with an overrides file consulted first.)
2. **Segmentation and tone (scripts without word spaces).** For Thai and similar scripts the text is segmented into words by a dictionary-and-statistics segmenter, and each syllable's tone is derived by the published rule: initial-consonant class (high, mid, low) x syllable type (live or dead) x vowel length x tone mark gives the tone. Tom's existing tools thai-tone-check.py and thai-tone-restore.py implement this rule; applied to 1471 toneless frames, all but 5 were fixed mechanically. The 5 remaining needed linguistic judgment. That experience is why a confidence rule is essential: a word with no confident pronunciation is shown in plain spelling, never a guess.
3. **Lexicon.** The word is looked up in the pronunciation lexicon for the prestige dialect (for English, cmudict.tvbet.json, derived from the CMU Pronouncing Dictionary and re-expressed in tvbet). A single entry gives confidence near 1.0 (reduced slightly for words the lexicon itself marks as uncertain). Multiple entries (heteronyms: read /ɹiːd/ ~ /ɹɛd/, lead /liːd/ ~ /lɛd/, bass /beɪs/ ~ /bæs/) give several candidates; a part-of-speech tagger or context model may raise one candidate above the others, and if it does so with margin the word becomes a single confident pronunciation, otherwise it stays a candidate set.
4. **Rules.** A word not in the lexicon is decomposed by morphology (inflections, common affixes, compounds of lexicon words) and, failing that, by grapheme-to-phoneme rules. Each rule step carries a confidence weight; the product is the word's confidence. Proper names and all-capital tokens are discounted.
5. **Missing-words file.** Every token that reaches the rules stage, with its page count and the rules' best guess, is appended to a missing-words file. A curator reviews that file and promotes entries into the overrides file or the lexicon. Thus the system's coverage rises with use, and no guess is shown to a reader before a curator has approved it.
6. **Confidence rule.** With threshold T (initially 0.9; tunable per language), the outcomes are: exactly one candidate at or above T, show it; two or more candidates at or above T, show the candidate set (section 4.11); none, plain spelling only. The page's coverage metric is the share of tokens in the first two outcomes and is shown to the reader on request.
7. **Output.** For each word: outcome, candidate transcription(s) in the selected alphabet, confidence, and the list of grapheme-phoneme patterns exercised (for example, for "night": n→/n/, igh→/aɪ/, t→/t/), which the mastery rule needs.
8. **Assembly into a translation graph.** The outputs for a sentence (or a camera-detected line) are assembled into a translation graph in the TG form: a tiered DAG whose nodes are word positions and carry no content, whose edges carry the content, and whose spans are identified by (from, to, tier). The S tier holds the original spelling of each word span; the R tier holds the transcription of the same span (or its candidate set, or nothing, according to the confidence rule), with the confidence value and the pattern list attached to the R edge. Because both tiers are keyed to the same word positions, the renderer reads S and R together without any second alignment step, and further tiers (B uncontracted form, F underlying form, E gloss) can be attached to the same positions later by the same tools that build Tom's existing TG bundles. The graph is cached per page and per camera frame set.

The alphabet is a display choice applied last: the internal representation is phonemic; IPA, tvbet (Tom's phonemic alphabet for English, with symbols such as ɔ ɹ ð θ š č iː uː æ ɑ ɚ), or another phonetic alphabet is produced by a symbol table.

4.3 The mastery rule

Learner model. For each reader the system stores a mastery value m in [0, 1] for:

- each grapheme-phoneme pattern of the language (for English, several hundred: single letters, digraphs, vowel teams, silent-e patterns, each paired with the phoneme it spells in context), and
- each word the reader has met, together with counts and timestamps of the evidence behind each value, and a forgetting parameter so that a value decays toward its prior with time since last evidence.

Per-word mastery. The mastery of a word for display purposes is the minimum of the mastery values of the patterns it exercises, combined with the reader's mastery of that word itself (a reader can know a whole word by sight before knowing every pattern in it, and the reverse). Because the patterns are shared across words, a word the reader has never seen whose patterns are all mastered already needs no pronunciation line; this is the property that makes the rule a reading teacher rather than a flashcard deck.

Show / fade / hide schedule. Two thresholds, low L (initially 0.5) and high H (initially 0.9):

- m < L: the pronunciation line is shown in full.
- L <= m < H: the pronunciation line is faded. The fade level f = (m - L) / (H - L) is applied as reduced opacity (for example from 1.0 down to 0.25), optionally reduced size or contrast; the pronunciation line remains legible on a deliberate look but no longer competes with the plain spelling.
- m >= H: the pronunciation line is hidden; a tap on the word restores it for that word and counts as evidence that the reader needed it. The plain spelling is present at every level. This is "the slope, not the cliff": the reader moves from reading the pronunciation line to reading the plain spelling one pattern at a time, on pages they chose to read, with the pronunciation tier aligned and adjacent.

Interaction with the confidence rule. The confidence rule runs first. A word with no confident pronunciation is plain regardless of mastery; a word with a candidate set is shown as candidates regardless of mastery until the reader's picks have settled it.

Thresholds. L, H and the fade curve are settable per reader and per deployment; a teacher's profile may set a steeper or shallower slope. The evaluation method (section 4.10) is the tool for choosing them.

4.4 Prestige-dialect selection

The pronunciation line is the pronunciation of a selected (perhaps, prestige) dialect, because (Tom's premise 3) "dialect speakers can understand speech in other (perhaps higher prestige) dialects", and because a single well-curated lexicon exists for it. For English the default lexicon is General American as encoded in cmudict.tvbet.json. Other prestige dialects (for example Received Pronunciation) are separate lexicons or are derived by rule from a lexical-set-keyed base form (section 4.5). The reader or the deployment chooses the dialect; the choice applies to the whole display and to the audio (section 4.7).

4.5 OPTIONAL: user's dialect conversion

Tom's optional claim (2026-10-16): "with conversion to user's-dialect-transcription". The prestige-dialect transcription is converted, by systematic phonological correspondence rules, into the transcription of the reader's own dialect, so the reader sees each word as, or more nearly as, they themselves say it.

Mechanism. The correspondence rules are keyed on lexical sets (Wells 1982, https://en.wikipedia.org/wiki/Lexical_set): each stressed vowel of the prestige-dialect transcription is tagged with its lexical set (KIT, DRESS, TRAP, LOT, STRUT, FOOT, BATH, CLOTH, NURSE, FLEECE, FACE, PALM, THOUGHT, GOAT, GOOSE, PRICE, CHOICE, MOUTH, NEAR, SQUARE, START, NORTH, FORCE, CURE, plus unstressed happY, lettER, commA), and a rule table for the reader's dialect maps each set, in stated phonological environments, to the reader's vowel. Consonant rules are stated the same way (environment → output). Rules apply to the transcription; the plain spelling is untouched.

Which parts are rule-driven and which need a table. Rule-driven: any correspondence that is regular across a lexical set in an environment (mergers, vowel shifts, consonant-cluster reduction, r-lessness). Table-driven: lexical incidence -- words that belong to a different set in the reader's dialect than in the prestige dialect (for example which words take the BATH vowel), and individual words with their own forms. The table is small relative to the lexicon and is itself a curated file with the same overrides-first logic as section 4.2.

Worked example, General American (prestige) to a Southern US dialect (reader), tvbet/IPA strings. Rules, each keyed on a lexical set:

- PRICE before a voiced consonant or word-finally: /aɪ/ → /aː/. "time" GA /taɪm/ → /taːm/; "my" /maɪ/ → /maː/. (Before voiceless consonants the rule may be withheld for some speakers: "right" /ɹaɪt/ stays /ɹaɪt/; this is a parameter of the reader's rule set.)
- PIN/PEN merger: KIT and DRESS before a nasal → /ɪ/. "pen" GA /pɛn/ → /pɪn/; "ten" /tɛn/ → /tɪn/.
- LOT/THOUGHT kept distinct where GA has merged them for the reader: "caught" /kɔt/ stays /kɔt/, "cot" /kɑt/ stays /kɑt/ (no rule needed if the GA lexicon keeps both; a merger rule is added for readers who merge).
- Final consonant-cluster reduction (optional, a register parameter): "test" /tɛst/ → /tɛs/.

Worked example, General American to African American English. Rules:

- PRICE monophthongization as above: "I" /aɪ/ → /aː/.
- Final cluster reduction: "cold" /koʊld/ → /koʊl/; "hand" /hænd/ → /hæn/.
- Non-prevocalic r-lessness: lettER /ɚ/ → /ə/ ("sister" /sɪstɚ/ → /sɪstə/), START /ɑɹ/ → /ɑː/ ("car" /kɑɹ/ → /kɑː/).
- θ → f and ð → v non-initially: "with" /wɪθ/ → /wɪf/; "brother" /bɹʌðɚ/ → /bɹʌvə/.
- PIN/PEN merger as above. These are illustrations for the disclosure; the shipped rule sets will be written by the inventor (a phonetician) and curated; which rules a given reader gets is a profile setting, chosen at intake (section 4.13) from a short list or inferred from the baseline read-aloud by the speech recognizer.

Why this is not just accent selection. Prior systems let a user pick an accent for a whole display or voice. Here the conversion is a transformation applied to the prestige-dialect transcription of the displayed pronunciation line, under the confidence rule and the mastery rule, with the reader's dialect as the target derived from the reader's profile or read-aloud. Its purpose is literacy: the reader's own knowledge of the sounds of their speech is the one resource the locked-strategy semi-literate (group (ii) of section 2.4) reliably has, and this step puts that resource in contact with the text. In the limit cases of section 2.7, wherever the sound correspondence between the written standard and the reader's spoken variety is systematic, this conversion moves the case from "sound but not

meaning" to "sound that triggers meaning."

4.6 Camera overlay

1. **Capture.** The camera application streams frames from the device camera.
2. **Text detection and OCR.** A text detector returns word-level bounding boxes with orientation; an OCR engine returns the characters in each box with a recognition confidence.
3. **OCR confidence.** A box whose recognition confidence is below a threshold is not passed to the pipeline and receives no pronunciation line (it may receive a faint outline so the reader knows it was seen). OCR confidence and pronunciation confidence are separate thresholds; a word must clear both.
4. **Pipeline and rules.** Recognized words go through the pronunciation pipeline, the confidence rule and the mastery rule exactly as in the browser, using the same user account and learner model ("under the same checks").
5. **Registration.** For each word with a show or show-faded outcome, the pronunciation line is drawn in the live frame along the word's baseline, under (default) or over the detected text, scaled to the word's height, in a color contrasting with the sampled background, and following the word's orientation. The plain text in the image is never covered. Bracket marking is drawn around verified spans.
6. **Tracking and latency.** Word boxes are tracked across frames so that the overlay follows camera motion without re-running OCR on every frame; OCR is re-run when tracking is lost or the scene changes. A freeze button holds the current frame with its overlay so the reader can read at leisure and tap words for audio. Target end-to-end latency for a new scene is under one second on a current phone; the design does not depend on the number.
7. **Behavior data.** Taps on words in the frozen frame, time spent on a frame, and read-aloud of the frozen frame all feed the learner model.

4.7 Audio on tap

A tap (or long press) on any word plays its pronunciation in the selected dialect: a recorded human pronunciation where one exists in the system's recording store (Tom's existing per-language wav stores and recording tools), otherwise synthesized speech from the transcription. In the candidate-set case the tap lists the candidates and plays each. Every tap is recorded with the word, its patterns and the current mastery values as learner-model evidence (section 4.8).

4.8 Learner-model update from behavior and from read-aloud

From reading behavior. Events: a tap for audio on a word (evidence the reader needed help with that word's patterns, weighted by which patterns are currently least mastered); a tap that reveals a hidden pronunciation line (stronger evidence); dwell on a word or line beyond the reader's own typical rate (weak evidence); a candidate pick (evidence about that word and about the reader's attention); smooth reading of a passage containing a pattern with no taps (weak positive evidence for every pattern in it). Each event adjusts the relevant mastery values by a small step toward 0 or 1, with the forgetting parameter applied to elapsed time.

From read-aloud scoring. The reader reads a passage aloud (the existing rasa recorder); an automatic speech recognizer aligned to the known text returns, per word, whether it was read correctly and, where it was not, which phoneme differed. Each word read correctly is positive evidence for every pattern it exercises; a mispronunciation is negative evidence for the pattern that maps to the differing phoneme. This is the strongest evidence the model receives and is weighted accordingly. Read-aloud scoring also supplies the words-correct-per-minute measure of section 4.10, so the same session both teaches and measures.

Live struggle detection (Tom, 2026-10-17). While the reader reads aloud, the recognizer runs in real time, aligned to the displayed text. A pause longer than the reader's own typical inter-word time, a misread, or a restart on a word is a struggle signal for that word. On a struggle signal the system acts at once for that word: it shows the word's pronunciation line if the mastery rule had hidden or faded it, and, if the struggle continues, plays the word's audio (section 4.7); the event is recorded against the word's grapheme-phoneme patterns. This is the moment at which help is worth most, because the reader has just shown exactly what they do not yet know. Amira Learning and Carnegie Mellon's Project LISTEN Reading Tutor (Mostow) detect struggle from read-aloud and prompt in the same way inside their own reading applications on their own texts; here the same signal drives a per-word pronunciation line inside any web page or camera text, and the stored per-pattern mastery sets the default display between struggle events.

From drills. Mastery of the phonetic alphabet itself (symbol to sound) comes from the existing PlayAlong drills, which record per-user mastery scores and drill time. Until a symbol is mastered in the drills, words whose transcription contains it are treated as unmastered regardless of pattern mastery.

4.9 Multi-linear S+R display details

- Element: HTML ruby with the plain spelling as the ruby base and the pronunciation line as the ruby text; rp fallback parentheses for browsers without ruby support.
- Placement: under the base by default for Latin-script pages (ruby-position: under), over for pages where the author's layout leaves room above; the reader can switch. For vertical scripts the pronunciation line sits to the right, as furigana does.
- Font: the pronunciation line uses a font with full IPA coverage and consistent letter shapes for tvbet; size about 55-65% of the base; line height of the page is increased so that the second line does not overlap the next text line.
- Alignment: the pronunciation line is centered under its word; for long words with short transcriptions and the reverse, the base is not letter-spaced, so the plain spelling keeps its normal shape (the reader must learn the normal shape).

- Fade: implemented as opacity on the ruby text element, so the base is unaffected; candidate sets are shown as the first candidate with a small marker, expanding on tap.
- Script direction: right-to-left pages place the pronunciation line right-to-left under each word with the same ruby element; mixed-direction lines follow the bidi order of the base.
- Accessibility: the pronunciation line is marked so that screen readers skip it unless the reader chooses otherwise.

4.10 Evaluation method: time to crossover

Tom's restatement: "the measurement is one curve: minutes elapsed with and without phonetic-alphabet instruction and practice on the x axis, cumulative words read aloud correctly (or per minute) on the y axis, two lines per learner, the converted page against the plain page. The proof is in the crossover, if it occurs. Power is measured by the distribution of the crossover: sooner = more powerful. Practically, for an adult it must arrive before the hour at which they would have quit. For children the X axis is in school years."

As an embodiment:

1. **Intake and classification** (section 4.13): a baseline read-aloud of a plain page classifies the learner as group (i) or group (ii) and sets the starting point of the plain-page line.
2. **Instruction and practice:** the learner does phonetic-alphabet drills (PlayAlong) and reads converted pages in the browser or camera; the system logs minutes.
3. **Measurement:** at intervals the learner reads aloud a passage in the converted display and a matched passage in plain spelling; the speech recognizer (or a person) scores words read correctly; the system records both the per-session rate (words correct per minute) and the cumulative total of words read correctly, for each display.
4. **Crossover:** the learner's time to crossover is the minute of instruction and practice at which the converted-page rate first exceeds and stays above the plain-page rate. A cumulative total crosses later than the rate; both are recorded; the per-session rate is used for the crossover and the cumulative total for total benefit.
5. **Reporting:** for a population, the fraction of learners who have crossed as a function of minutes; a learner who quits before crossing counts as not yet crossed at the minute they quit (a censored observation). The headline number is this curve for group (ii).
6. **Per-variant power:** the same curve is produced for each variant (alphabet; mastery rule on or off; user's dialect conversion on or off; bracket marking on or off); the power of a variant is how much sooner its crossover distribution sits. This per-variant comparison is also the "unexpected results" evidence for section 103, one dependent claim at a time.
7. **Children:** the x axis is school years; the comparison is between classrooms or cohorts.

4.11 The three-way confidence outcome and candidate display

Tom: "A measurable weakness is coverage: the fraction of words for which we have a confident pronunciation; the remainder can be shown as multiple options and/or in plain spelling."

Per word: (i) one confident pronunciation → pronunciation line shown (subject to the mastery rule); (ii) two or more candidates above threshold → the candidates are shown as options, compactly (first candidate plus a marker, or both separated by a slash when short) or on tap (a small chooser listing each candidate with audio); the reader's pick is recorded; when the same reader picks the same candidate for the same word in the same context repeatedly, the pick becomes an override for that reader, and aggregated picks across readers feed the curator; (iii) none → plain spelling only, never a guess. English examples: "read" /ɹiːd/ ~ /ɹɛd/, "lead" /liːd/ ~ /lɛd/, "bass" /beɪs/ ~ /bæs/. Thai example: a syllable whose tone the rule leaves unresolved is shown with its candidate tones.

4.12 Bracket marking (phonetic, i.e., pronounceable, i.e., trustworthy, a safe space)

A visible marker on the page distinguishes spans whose pronunciation line is verified (safe to learn from) from in-the-wild spelling. The default marker is a pair of light brackets [ ] enclosing each maximal run of words in outcome (i) (and, in a distinguishable style, outcome (ii)); an equivalent marker (underline, background tint, a margin bar) may be chosen. Within a bracketed span the reader knows that reading the pronunciation line will give the right word; outside it the reader knows they are on their own and may use whatever strategy they have. Bracket marking persists when the mastery rule has hidden the pronunciation line, so the reader still knows the span is verified and that a tap will reveal it. It is the visible form of the confidence check and answers the third of ITA's three reasons.

4.13 Intake and baseline classification

Tom: "After teaching a phonetic alphabet, our aim is to give the experience of reading a real page correctly, because that experience is what breaks the belief that they cannot learn. Users must therefore be classified by that baseline, so the headline number can be the second group's crossover."

The intake step: before any alphabet instruction, the learner reads aloud a plain page of known text; the speech recognizer (or a person) scores it. The score, together with the pattern of errors (first-consonant-and-vowel guesses, picture-driven substitutions, refusals), classifies the learner as a supported reader (group i) or a locked-strategy semi-literate (group ii), and initializes the learner model: every pattern read correctly starts with positive evidence, every pattern misread starts low, and the reader's dialect profile is set from the recognizer's output where it can be. The sequence is then: alphabet drills (PlayAlong) → the first real page in the browser, chosen so that its coverage is near 100%

and its patterns are the ones the learner has drilled, so the first experience is reading a real page correctly → the mastery rule takes over.

4.14 The zhuyin precedent, the Chinese-"dialects" limit, and how user's dialect conversion moves a case across the line

Taiwan's zhuyin beside characters for children (section 2.6) is this invention's display for a Mandarin speaker: a pronunciation line in a phonetic alphabet beside the hard script, read until the characters are learned. For a Cantonese speaker reading the same text, the zhuyin gives Mandarin sounds, which are not the reader's words: the limit of section 2.7 applies, because the written language diverges from the spoken one beyond any systematic sound correspondence. Between those two poles lie the cases where the correspondence IS systematic -- regional pronunciations of a standard language, dialects of English, the spoken varieties of Thai and French relative to their standards -- and there the user's dialect conversion (section 4.5) rewrites the prestige-dialect pronunciation line into the reader's own and so moves the case from the limit side to the benefit side. The decision of whether a given pair is convertible is a property of the rule set's coverage, which the evaluation method of section 4.10 measures directly.

5. Drawings to be prepared

- **FIG. 1** System overview: browser module and camera application sharing one pronunciation pipeline, one confidence rule, one mastery rule, one learner model and one user account.
- **FIG. 2** Browser flow: DOM text nodes → tokenization → lookup → confidence rule → mastery rule → ruby rendering; MutationObserver loop for dynamic pages.
- **FIG. 3** Pronunciation pipeline: overrides → (segmentation + tone rule for Thai-type scripts) → lexicon → morphology and grapheme-to-phoneme rules → confidence value → missing-words file to curator → overrides.
- **FIG. 4** Three-way confidence outcome: one confident candidate (show), several (candidate display, compact and on-tap forms), none (plain spelling); coverage metric.
- **FIG. 5** Learner-model data structure: per-reader table keyed by grapheme-phoneme pattern and by word, with mastery value, evidence counts, timestamps, forgetting parameter; per-word mastery derived from patterns.
- **FIG. 6** Mastery-rule schedule: mastery value on the x axis, pronunciation-line visibility on the y axis, with thresholds L and H and the fade region; "slope not cliff".
- **FIG. 7** Screen mock-ups of one paragraph at four mastery states: all shown; partly faded; mostly hidden; all hidden with bracket marking still present and one word tapped to reveal.
- **FIG. 8** Bracket marking (phonetic, i.e., pronounceable, i.e., trustworthy, a safe space): a page with bracketed verified spans, candidate-marked words and unbracketed in-the-wild words.
- **FIG. 9** Prestige-dialect selection and user's dialect conversion: transcription tagged by lexical set → rule table for reader's dialect → converted pronunciation line; worked example words from section 4.5.
- **FIG. 10** Camera overlay: frame → text detection boxes → OCR confidence threshold → pipeline and rules → pronunciation line registered under each word with bracket marking; tracking across frames; freeze.
- **FIG. 11** Thai example: unsegmented text → segmentation → tone rule table (consonant class x live/dead x vowel length x tone mark) → transcription; one unresolved syllable shown as candidates.
- **FIG. 12** Learner-model update: event types (audio tap, reveal tap, dwell, candidate pick, read-aloud per-word result) and their weights; forgetting over time.
- **FIG. 13** Intake: baseline read-aloud → classification into group (i) or (ii) → learner model initialized → drills → first real page → mastery rule.
- **FIG. 14** Evaluation: one learner's two lines (converted page, plain page; words correct per minute against minutes) with the crossover marked; population curve of fraction crossed against minutes with censoring; two variants compared.
- **FIG. 15** Translation graph for one sentence: word-position nodes; S-tier and R-tier spans (from, to, tier) over the same positions; confidence and pattern list on each R edge; one candidate-set span; one span with no R edge; the two-line rendering read from the graph.

6. Draft claims

(For attorney review. Written in claim form; numbering and dependencies to be reworked by the attorney.)

1. (Method.) A computer-implemented method of displaying text to a reader, comprising: obtaining text comprising a plurality of words from a source selected from a document rendered by a web browser and an image captured by a camera; for each word of the plurality of words, determining by a pronunciation pipeline one or more candidate phonemic transcriptions of the word in a selected dialect, each candidate having a confidence value, the pronunciation pipeline comprising at least a curated overrides store consulted first, a pronunciation lexicon, and grapheme-to-phoneme rules; applying a confidence rule to each word by comparing the confidence values of its candidates to a confidence threshold, the confidence rule producing one of (i) a single confident transcription, (ii) a set of two or more candidate transcriptions, and (iii) no confident transcription; retrieving, from a learner model stored for the reader, mastery values associated with grapheme-phoneme patterns exercised by each word and with the word, and applying a mastery rule that determines, from the mastery values and at least one mastery threshold, a visibility level for a pronunciation line of the word, the visibility level varying continuously or in a plurality of steps between fully shown and hidden as the mastery values increase; and rendering the text such that each word is displayed in its original spelling as a plain spelling line and, for each word for which the confidence rule

produced a single confident transcription, a pronunciation line comprising the transcription expressed in a phonetic alphabet is displayed registered under or over that word at the visibility level determined by the mastery rule, each word for which the confidence rule produced no confident transcription being displayed in its original spelling without a pronunciation line.

2. (System.) A system comprising one or more processors and memory storing instructions that, when executed, cause the system to: obtain text comprising a plurality of words from a web page document object model or from text detected in a camera image; for each word, determine one or more candidate phonemic transcriptions in a selected dialect, each with a confidence value, by consulting in order a curated overrides store, a pronunciation lexicon, and grapheme-to-phoneme rules; apply a confidence rule producing, per word, a single confident transcription, a candidate set, or no confident transcription; maintain a learner model for a reader comprising mastery values keyed by grapheme-phoneme pattern and by word; apply a mastery rule that sets a visibility level of a pronunciation line per word as a graded function of the mastery values of the grapheme-phoneme patterns exercised by the word; and render each word as a two-line display comprising the original spelling and, where a single confident transcription exists, a pronunciation line in a phonetic alphabet registered to the word at the visibility level, words with no confident transcription being rendered in original spelling only.

3. (Computer-readable medium.) A non-transitory computer-readable medium storing instructions that, when executed by a processor of a device having a display, cause the device to perform the method of claim 1.

4. The method of claim 1, wherein the confidence rule compares each candidate's confidence value to the confidence threshold and (i) when exactly one candidate meets the threshold the word is displayed with that candidate as its pronunciation line, (ii) when two or more candidates meet the threshold the word is displayed with a candidate display presenting the candidates as selectable options, and (iii) when no candidate meets the threshold the word is displayed in original spelling only, such that no transcription below the threshold is ever displayed as a pronunciation line.

5. The method of claim 1, wherein the mastery rule comprises a low threshold and a high threshold, the pronunciation line being displayed at full visibility when a per-word mastery value derived from the mastery values is below the low threshold, at a reduced visibility that decreases monotonically with the per-word mastery value between the low and high thresholds, and hidden when the per-word mastery value is at or above the high threshold, the plain spelling line being displayed at all mastery values.

6. The method of claim 5, wherein the per-word mastery value is derived at least in part from the mastery values of the grapheme-phoneme patterns exercised by the word, such that a word not previously displayed to the reader receives a reduced or hidden pronunciation line when the reader's mastery values for its grapheme-phoneme patterns exceed the thresholds.

7. The method of claim 1, further comprising updating the mastery values of the learner model from reading behavior of the reader comprising at least one of: a tap on a word requesting audio, a tap revealing a hidden pronunciation line, dwell time on a word, and a selection among candidates in a candidate display, each event adjusting the mastery values of the grapheme-phoneme patterns exercised by the word.

8. The method of claim 1, further comprising capturing audio of the reader reading displayed text aloud, aligning the audio to the displayed text by automatic speech recognition to obtain a per-word correctness result and, for incorrectly read words, an identification of a differing phoneme, and updating the mastery values of the grapheme-phoneme patterns exercised by each word according to the correctness result and the differing phoneme.

9. The method of claim 1, wherein the prestige dialect is selected from a plurality of dialects by the reader or by a deployment setting, and the pronunciation lexicon and grapheme-to-phoneme rules are those of the selected prestige dialect.

10. The method of claim 1, further comprising, before rendering, converting the single confident transcription in the prestige dialect into a transcription in a dialect of the reader by applying systematic phonological correspondence rules, each correspondence rule being keyed on a lexical set membership of a vowel of the transcription and on a phonological environment and specifying a replacement in the reader's dialect, the reader's dialect being determined from a reader profile or from automatic speech recognition of the reader's speech, and the pronunciation line displayed being the converted transcription.

11. The method of claim 10, wherein the correspondence rules are supplemented by a lexical-incidence table listing words whose lexical set membership in the reader's dialect differs from that in the prestige dialect, the table being consulted before the correspondence rules.

12. The method of claim 1, wherein the source is an image captured by a camera, the method further comprising detecting text regions in the image with word-level bounding boxes and orientations, recognizing characters in each bounding box with a recognition confidence, discarding bounding boxes whose recognition confidence is below a recognition threshold, and rendering the pronunciation line of each remaining word within the image along the baseline of its bounding box, under or over the detected text and following its orientation, without obscuring the detected text, the bounding boxes being tracked across successive frames so that the pronunciation lines follow the detected text.

13. The method of claim 1, wherein the text is in a script written without spaces between words, the pronunciation pipeline further comprising segmenting the text into words and, for each syllable, deriving a tone from a tone rule combining an initial-consonant class, a live or dead syllable type, a vowel length and a tone mark, a syllable for which the tone rule yields more than one tone being passed to the confidence rule as a candidate set.

14. The method of claim 1, wherein rendering the text in a web browser comprises

replacing each word's text node content with a ruby element whose ruby base is the original spelling and whose ruby text is the pronunciation line, the visibility level being applied as an opacity of the ruby text, separators between words being preserved, and newly added text nodes being processed on detection by a mutation observer.

15. The method of claim 1, further comprising, in response to a tap on a displayed word, playing audio of the word's transcription in the prestige dialect, the audio being a stored human recording when one exists for the word and synthesized speech otherwise, and recording the tap as an update to the learner model.

16. The method of claim 4, wherein the candidate display presents the candidates compactly as a first candidate with a marker and, on selection, as a list of all candidates each with audio playback, and wherein a selection by the reader is recorded in the learner model and, after a repeated selection of the same candidate for the same word, is stored as a reader-specific override.

17. The method of claim 1, further comprising rendering a visible bracket marking enclosing each run of consecutive words for which the confidence rule produced a single confident transcription, the bracket marking distinguishing such verified spans from words displayed in original spelling only, and the bracket marking remaining displayed when the mastery rule has hidden the pronunciation lines of the words within it.

18. The method of claim 1, further comprising, before the reader has received instruction in the phonetic alphabet, capturing audio of the reader reading a plain-spelling passage aloud, scoring the reading per word by automatic speech recognition, classifying the reader into one of at least two reader groups from the score and the pattern of errors, and initializing the mastery values of the learner model from the per-word scores.

19. The method of claim 18, further comprising presenting phonetic-alphabet drills to the reader, recording a per-symbol mastery from the drills, and treating a word whose transcription contains a symbol below a drill-mastery threshold as having a per-word mastery value below the low threshold.

20. (Evaluation method -- include only if the attorney thinks it helps.) A method of evaluating a reading-display system, comprising: for each of a plurality of learners, recording minutes of phonetic-alphabet instruction and practice; at intervals, measuring a first rate of words read aloud correctly per minute from text displayed by the system with a pronunciation line and a second rate from matched text displayed in plain spelling only, the rates being measured by automatic speech recognition aligned to the text; determining for each learner a time to crossover as the minute at which the first rate first exceeds and remains above the second rate; treating a learner who ceases participation before crossover as not yet crossed at the minute of cessation; computing, for the plurality of learners, the fraction of learners crossed as a function of minutes; and comparing the fraction-crossed functions obtained under at least two configurations of the system differing in at least one of a phonetic alphabet, a mastery rule being enabled or disabled, a user's dialect conversion being enabled or disabled, and a bracket marking being enabled or disabled.

21. The method of claim 1, further comprising assembling, for the obtained text, a tiered translation graph comprising nodes representing word positions and edges representing spans, each span being identified by a start position, an end position and a tier, the tiers comprising at least a source-orthography tier whose spans carry the original spelling and a phonetic-transcription tier whose spans carry the candidate transcriptions, confidence values and grapheme-phoneme pattern lists produced by the pronunciation pipeline, the spans of the two tiers being keyed to the same word positions, wherein the rendering step reads the plain spelling line and the pronunciation line of each word from the source-orthography tier and the phonetic-transcription tier of the graph, and wherein further tiers are attachable to the same word positions.

22. The method of claim 8, further comprising, while the reader reads the displayed text aloud, aligning the reader's speech to the displayed text in real time, detecting a struggle signal for a word from at least one of a pause exceeding the reader's own typical inter-word interval, a misread, and a restart, and in response to the struggle signal displaying the pronunciation line of that word at full visibility notwithstanding a reduced or hidden visibility level determined by the mastery rule, and recording the struggle signal as an update to the mastery values of the grapheme-phoneme patterns exercised by the word.

7. Prior art and distinctions

Legend: Y = the reference has the element; N = lacks it; P = partial. Elements: **2L** two-line pronunciation line on arbitrary text; **CR** confidence rule with plain-spelling fallback, never a guess; **MR** per-word mastery rule; **FADE** graded fade (slope); **GPP** mastery per grapheme-phoneme pattern; **BEH** learner model from in-tool reading behavior; **ASR** learner model from read-aloud scoring; **PD** prestige-dialect selection; **UDC** user's dialect conversion by correspondence rules; **CAM** camera overlay registered to text; **CAND** candidate display for ambiguous words; **BRK** bracket marking; **BASE** baseline read-aloud intake; **LIT** aimed at native-speaker literacy.

Reference (URL in scouts' JSON)	2L	CR	MR	FADE	GPP	BEH	ASR	PD
Rakuten US10275445B2 (ruby by learning score, forgetting, show/hide taps)	Y (kana ruby)	N	Y	N	N	Y (show/hide taps)	N	N
Readable English US11544444B2 / US10521511B2 (cues in spelling, accent choice, hard-words-only toggle)	P (cues in-line)	N	N	P (manual global)	N	N	N	Y
Woodward US8131538B2 (pictograms above/below text, dialect key)	P (pictograms)	N	N	N	N	N	N	Y
NTT JPH1139298A (furigana by user rank)	Y	N	P (typed rank)	N	N	P (remembers manual choices)	N	N

Microsoft US10409903B2 (known/unknown word classifier, web page, dwell/gaze)	N (definitions)	N	P	N	N	Y	N	N
Apple US11373544B2 (read-aloud proficiency adapts text difficulty)	N	N	N	N	N	N	Y	N
MetaMetrics US11908488B2 (per-GPC scoring from ASR)	N	N	N	N	Y (assessment only)	N	Y	N
See Word Design US11526654B2 (phonetical font, browser ext., AR mention)	P (glyphs)	N	N	N	N	N	N	N
Fuji Xerox US7386453B2 / US7455522B2 (adaptive aid level from errors)	N	N	P	P	N	Y	Y	N
Cytanovich US5057020A (scanner + augmentation marks)	P	N	N	N	N	N	N	N
Word Lens US20110090253; Google US9239833/US10726212; A9 US9922431; Microsoft US20170255614; Translate Abroad US9275046	N (translation)	N (OCR conf. only)	N	N	N	N	N	N
Furiganai; Wani-Furigani; FuriKani; jpd-breader; Kioku; Satori Reader; Language Reactor (known-item hiding)	Y	N	Y (binary)	N	N	N (external SRS/manual)	N	N
Japanese Furigana Extension; Furigana Maker; Yomu (fixed level filter)	Y	N	P (declared level)	N	N	N	N	N
Transcriber; IPA Replacer; IPAfy (English page to IPA)	N (replace)	N	N	N	N	N	N	N
toPhonetics (paste-in, US/UK, unknown words red, multi-pron. blue)	P (line-by-line)	P (flags, does not suppress)	N	N	N	N	N	Y
PIE Phonetically Intuitive English (2013, Full/Lite/Extra Lite)	P (diacritics)	N	N	P (manual modes)	N	N	N	N
Hanzi Annotator; Pinyinizer; Zhongwen; PinVocab	Y	N	N	N	N	N	N	N
Furigana Lens; HanziTranslator; Furigana Camera; Pleco OCR	Y	N	N	N	N	N	N	P (HanziTranslato mode)
Unisyn / Combilex (lexical-set keyed accent rules, TTS)	N	N	N	N	N	N	N	Y
Henton US6738738B2 (AmE to BrE transcription rules)	N	N	N	N	N	N	N	N
IBM US6343270B1; Adobe US9990916B2; IBM US10930274B2; SoundHound US10319250B2 (user's dialect for speech)	N	N	N	N	N	N	P	P
Microsoft Immersive Reader / Reading Coach	N (syllable dots)	N	N	N	N	N	Y	N
Amira Learning (AI reading tutor: real-time ASR alignment, struggle detection, prompted word playback) [CITE: patents to be pulled]	N (own texts)	N	P (per-word struggle state)	N	P	Y	Y	N
Project LISTEN Reading Tutor, CMU (Mostow, 1990s-2000s: ASR-tracked oral reading, help on hesitation) [CITE]	N (own texts)	N	P	N	N	Y	Y	N
ITA (Downing 1964); Unifon converter	P (paper/paste)	N	N	N	N	N	N	N
INVENTOR'S OWN: tomveatch.com/tg/tg.apps.txt (2018), tg.introduction.txt (2003-2017), tg.definition.txt, call.for.multilinearity.txt -- TG tiers incl. transliteration/IPA tier; learner-model-controlled presentation; tiers hidden until the audience is puzzled; scanned documents into TGs; dialect phonetics transcripts	Y (any tier)	N (confidence-weighted alternatives only)	P (model-controlled presentation, generic)	N	N	P ("learn what they know")	N	P (dialect transcripts)
INVENTOR'S OWN: tomveatch.com/tg/BusinessModel.vibe (2025-09-30): Hunh? button; per-user heard/said/struggled/mastery model; drills to mastery	N	N	P	N	N	Y	P (recordings reviewed)	N

Closest references and the missing combination. Three references come closest, one per axis. (1) Rakuten US10275445B2 already claims a per-word learning score, updated from the reader's show/hide taps with a forgetting curve, that hides ruby readings above a threshold, including in HTML ruby; it is the main obstacle to a bare "hide the pronunciation when known" claim. It lacks a graded fade, mastery per grapheme-phoneme pattern (so it cannot release a never-seen word from its pronunciation line), read-aloud scoring, a phonetic alphabet for an alphabetic language, any confidence rule, dialect handling, and a camera. (2) Readable English US11544444B2/US10521511B2 is the closest in purpose: native-speaker literacy, spelling preserved, regional accent selectable, and a manual global knob that shrinks the cues; it lacks per-word, automatic, mastery-driven fading, a per-word confidence fallback, a separate pronunciation line, camera overlay and dialect conversion. The attorney must read its claims in full [CITE:

claim 1 of US10521511B2 not fully read by the scouts]. (3) Unisyn/Combilex and Henton US6738738B2 already perform rule-based, lexical-set-keyed derivation of accent-specific transcriptions, for speech synthesis; nothing displays the result to a reader under a mastery rule. On the camera side, Word Lens and the Google, A9 and Translate Abroad patents own the overlay mechanism for translation, and Furigana Lens and HanziTranslator ship pronunciation overlays for Japanese and Chinese; a camera claim must therefore depend on the two rules. The shipped furigana show/hide rules (Furiganai, Satori Reader, Language Reactor, jpd-breader) are binary and fed by flashcard state or manual marks. What no reference found has, alone or in combination: a confidence rule that falls back to plain spelling and never shows a guess; a graded fade of the pronunciation line driven by mastery measured per grapheme-phoneme pattern; that mastery learned from the reader's own reading and read-aloud inside the same tool; bracket marking of verified spans; candidate display whose picks feed the learner model; a baseline read-aloud intake; user's dialect conversion of the displayed line; and all of it applied to arbitrary web pages and camera images for illiterate and semi-literate native speakers. The scouts report (angle 2, not_found) that no patent was found claiming the slope, the confidence rule, the user's dialect conversion of a displayed line, or the browser-plus-camera pipeline.

8. Patentability positioning for the attorney

8.1 Section 101 (eligibility)

The risk, stated plainly: an examiner's first reading of "detect text, convert it to a transcription, display it" is Electric Power Group (collect, analyze, display information; https://patentdocs.org/2016/08/01/electric-power-group-llc-v-alstom-sa-fed-cir-2016/), the camera path is Content Extraction (scan, recognize, store; https://patentdocs.org/2015/01/05/content-extraction-and-transmission-llc-vs-wells-fargo-bank-fed-cir-2014/), and the mastery rule and dialect conversion are "tailoring content to the user", which Intellectual Ventures v. Capital One called abstract (https://www.sughrue.com/publications/summary-of-intellectual-ventures-i-llc-v-capital-one-financial-2014-1506). A claim to the transliterated page or to a correspondence table as data is Digitech (https://www.bitlaw.com/source/cases/patent/Digitech.html). A sounding-out step can also be called a mental process.

The answer is to put the "how" in every independent claim, which the draft claims above attempt. The features that map to the favorable line are:

- The mastery rule as a specific display mechanism: a stored per-reader, per-pattern value compared to thresholds, producing a graded visibility of a specifically delimited piece of data (the pronunciation line) in a specific place (ruby text registered to the word). This is the Core Wireless pattern, "precise language delimiting the type of data to be displayed and how to display it" (https://www.knobbe.com/blog/core-wireless-licensing-sarl-v-lg-electronics-inc/; USPTO memo of 2018-04-02, https://uspto.gov/sites/default/files/documents/memo-recent-sme-ctdec-20180402.PDF).
- The confidence rule as a step that enables the browser to do what prior browsers and transliterators could not: show pronunciations without ever showing a false one. This is the Finjan pattern, specific steps that accomplish a result the prior system could not (https://www.knobbe.com/blog/finjan-inc-v-blue-coat-systems-inc/).
- The rules as specific, limited rules: the Thai tone rule (consonant class x live/dead x vowel length x tone mark), the lexical-set-keyed correspondence rules, the show/fade/hide schedule. This is McRO, where the incorporation of specific phoneme-sequence rules, not the computer, improved the process (https://www.bitlaw.com/source/cases/patent/MCRO.html). The claims must recite the rule structure, not "applying rules"; McRO's claims nearly died in the district court by being read as "any rules".
- The framing of the whole as an improvement to a browser rendering engine and reading interface, with the specification stating what prior browsers and prior phonetic-text systems could not do (Enfish, https://en.wikipedia.org/wiki/Enfish,_LLC_v._Microsoft_Corp.). Data Engine warns that within one family the specific interface claim lives and the generic one dies (https://patentdocs.org/2018/10/17/data-engine-technologies-llc-v-google-llc-fed-cir-2018/); every independent claim should therefore carry the rule and registration specifics, not rely on dependents.

Assessment: eligibility is winnable but not assured; the mastery rule and confidence rule are the features most likely to carry it, and a claim without them should not be filed as an independent claim.

8.2 Section 103 (obviousness)

The strongest obviousness combination the examiner will build, under KSR rationale (A), combining known elements by known methods for predictable results (https://www.uspto.gov/web/offices/pac/mpep/s2141.html): Rakuten US10275445B2 (behavior-updated, ruby hidden above a threshold in HTML) + Transcriber or Readable English (English page converted to a phonetic form, with accent choice)

- MetaMetrics US11908488B2 (per-grapheme-phoneme scoring from read-aloud) + Word

Lens or Furigana Lens (camera overlay) + Unisyn (lexical-set rules for the reader's accent).

The answers:

1. **The combination does more than the parts.** Rakuten hides a reading when a word has been seen enough; it cannot fade, and it cannot release a word never seen. Mastery per grapheme-phoneme pattern with a graded fade turns a per-word aid into a per-pattern one that carries over to words the reader has never seen. No reference teaches the fade; the scouts found none in products or patents (angle 1 and 2 not_found lists). The confidence rule is likewise absent: toPhonetics flags unknown

words in red and still shows them; Furigana Maker and jpd-breader admit inaccuracy and still display. The never-a-guess rule, made visible by bracket marking, is a design choice no reference makes, and it is the choice that makes the pronunciation line safe to learn from.

2. **Long-felt need and failure of others.** ITA had positive results (Downing 1964, 1967) and was still abandoned, for the three reasons given in section 2.5, documented by Hall 2005 for the first two. The field's response was to drop phonemic transitional alphabets rather than fix the transition. Whether that amounts to teaching away under MPEP 2145 X.D (a reference must criticize, discredit or discourage the path) depends on finding a reference that does so in words; Hall 2005 and Wikipedia's "confused by two alphabets" are candidates, and Warburton and Southgate 1969 should be pulled [CITE: primary text not opened]. At minimum the ITA history is evidence of a recognized problem (the cliff) that persisted for sixty years and that the mastery rule addresses directly.
3. **Unexpected results, as evidence.** MPEP 2145 requires evidence, not argument, commensurate with claim scope. The evaluation method of section 4.10 produces exactly that evidence, one dependent claim at a time: the crossover distribution with the mastery rule on versus off, with user's dialect conversion on versus off, with bracket marking on versus off. The recommendation is to run at least a small version before the nonprovisional so that data exist, especially for group (ii).
4. **Dialect conversion.** Unisyn and Henton render rule-based accent conversion itself known; the dependent claim is drafted (claim 10) as a step on the displayed pronunciation line under both rules with the reader's dialect as target from profile or read-aloud, which no speech-synthesis reference suggests, because none has a displayed line or a reader.

8.3 Section 112 (written description and enablement)

MPEP 2161.01 is explicit that for software it is not enough that a programmer could write the program; the specification must disclose the algorithm in prose, formula or flowchart (https://www.uspto.gov/web/offices/pac/mpep/s2161.html). Claims that later rely on the provisional date get it only for what the provisional supports (MPEP 2163). The provisional must therefore contain, as drawings and text: the flowcharts of FIGS. 2, 3, 10, 12 and 13; the learner-model data structure of FIG. 5 with the per-word derivation rule; the show/fade/hide schedule of FIG. 6 with numeric starting thresholds; the three-way confidence outcome with the threshold and the candidate-display behavior; the correspondence-rule format and the worked examples of section 4.5 with real IPA/tvbet strings and an example lexical-incidence table; the Thai tone rule table and one worked example; the camera registration steps including OCR-confidence discard and tracking; the bracket-marking rule; the intake classification rule; and screen mock-ups at several mastery states (FIG. 7). Avoid results-only language ("display adapted to the reader"). The claims and description must use one name per thing throughout.

9. Procedure and timing

- **File a provisional application first.** It needs a specification meeting 35 USC 112(a), any drawings needed to understand the invention, a cover sheet identifying it as provisional, and the fee; claims are not required but a claim set should be included so the support map is visible (https://www.uspto.gov/patents/basics/apply/provisional-application). Write it as if it were the nonprovisional specification.
- **Twelve-month window.** A nonprovisional claiming benefit must be filed within 12 months of the provisional's filing date; the same date is the deadline for any foreign or PCT filing claiming priority. The provisional is not examined and lapses at 12 months.
- **Fee.** Provisional filing fee as found on the USPTO fee schedule opened 2026-10-16 (schedule effective 2025-01-19): $325 undiscounted, $130 small entity, $65 micro entity (https://www.uspto.gov/learning-and-resources/fees-and-payment/uspto-fee-schedule). Micro-entity status (37 CFR 1.29) requires small-entity status, not being named on more than four (the USPTO page also says five in one place; attorney to confirm) previously filed applications, prior-year gross income not above three times median household income (USPTO page shows $262,380 effective 2026-09-15), and no obligation to assign to a non-qualifying entity.
- **Grace period versus absolute novelty.** In the US, 35 USC 102(b)(1)(A) excuses the inventor's own disclosures made within one year before the effective filing date (https://www.law.cornell.edu/uscode/text/35/102). Europe and most other jurisdictions have absolute novelty: any disclosure to the public before the filing date, even to one person without a duty of secrecy, destroys novelty, with only narrow exceptions (https://en.wikipedia.org/wiki/Novelty_under_the_European_Patent_Convention) [CITE: EPC Article 55 text not opened]. The grace period is not a plan.
- **No public demo before filing.** No public page, public repository, app-store listing, talk or demonstration of the browser or camera before the provisional is filed. Keep tb/ and tc/ login-restricted. Any pre-filing demonstration under a written non-disclosure agreement.
- **Evaluation data.** If practicable, run a small evaluation (section 4.10) between the provisional and the nonprovisional so that unexpected-results evidence exists.

10. Open questions for the attorney

1. Has any public disclosure occurred? The tb/ and tc/ pages are login-restricted; the "Reach 2" session text of 2026-10-14 is a private record. Please confirm, and advise on whether the AI-assisted drafting sessions themselves constitute disclosure (they are under the provider's confidentiality terms; please check those terms).
2. Readable English US11544444B2 and US10521511B2: please read all claims in full (the scouts read US11544444 claims 1 partial and 2-34, and US10521511 claims 3-18 only). Does any claim read on a separate pronunciation line with a global manual reduction such that our mastery rule claim needs further distinction?
3. Rakuten US10275445B2: is claim 5 (mastery per grapheme-phoneme pattern with

graded fade) sufficiently distinct, or should the independent claims themselves recite the pattern-level model and the fade?

4. Should the confidence rule (claim 4) be moved into the independent claims as drafted (it is there in claim 1) or kept dependent to preserve a broader independent claim? Our view: keep it in the independent claims for section 101.
5. User's dialect conversion (claims 10-11): given Unisyn, Combilex and Henton US6738738B2, is the claim as drafted (conversion of the displayed line under both rules, reader's dialect from profile or read-aloud) defensible, and are Adobe US9990916B2 and IBM US10930274B2 freedom-to-operate concerns for it?
6. Camera claim 12: given Word Lens, the Google and A9 families, Furigana Lens and HanziTranslator, is a camera-only independent claim worth filing, or should the camera remain dependent (our draft)?
7. Evaluation method claim 20: file it, drop it, or hold for a separate application after data exist?
8. Micro-entity status: please confirm eligibility (prior applications count, income threshold) and the "four versus five" wording.
9. Foreign filing: is a PCT application within the 12 months worth the cost given absolute novelty abroad and the markets for a literacy tool (Thailand, India, francophone Africa)?
10. Inventorship record: the design conclusions A-H and the addenda were Tom's decisions in a 2026-10-16 discussion with AI assistance; the records are dated. Is anything more needed to document sole conception?
11. Open-source lexicon licensing: the English lexicon derives from the CMU Pronouncing Dictionary (BSD); Unisyn is non-commercial only. Does any license term bear on the patent or its commercialization?
12. Terminology: the disclosure uses "pronunciation line (R line)", "plain spelling (S line)", "mastery rule", "confidence rule", "prestige dialect", "user's dialect conversion", "camera overlay", "bracket marking (phonetic, i.e., pronounceable, i.e., trustworthy, a safe space)". Please keep one name per thing in the application.
13. Statistics marked [CITE] in section 2: should they be in the application at all, or kept to the background in general terms?
14. Amira Learning and Project LISTEN (added 2026-10-17): please pull Amira's patents and Mostow's publications; is claim 22 (live struggle detection driving the per-word pronunciation line on arbitrary text) defensible over them, and are Amira's patents a freedom-to-operate concern for the read-aloud update of claim 8?
15. The inventor's own publications in https://tomveatch.com/tg/ (section 3.2): the server file dates are tg.introduction.txt 2018-02-06, tg.apps.txt 2023-12-26 (text dated 2018), BusinessModel.vibe 2025-09-30, all in a directory served to the web; please confirm these as publication dates (archive copies if any) and advise which claims need redrafting to stand clear of them. The inventor's view is that they disclose the genus (a tiered display with a transliteration tier under a learner model) and none of the claimed species.

11. Inventor statement placeholders

Statement of inventorship. I, Thomas C. Veatch, declare that I am the sole inventor of the subject matter described in this disclosure; that I conceived it on 2026-10-16 as recorded in the conception record kept beside this disclosure (conception-record.md.php) and in the tb/index.php page of the same date; that the drawing-labelled phonetic chart on which premise (2) rests is my separate invention of 2002; and that the design decisions recorded in section 3 and section 4 are mine.

Note on AI assistance. Artificial-intelligence tools assisted in drafting this document, in background research and in prior-art searching. Conception of the invention is the inventor's. Under the USPTO's revised inventorship guidance of 2025-11-26, which rescinded the February 2024 AI-assisted inventorship guidance, the same conception standard applies to all inventions, only natural persons may be inventors, and AI systems are tools (https://www.uspto.gov/subscription-center/2025/revised-inventorship-guidance-ai-assisted-inventions). The dated conception records are kept separate from AI-drafted material. No AI system is to be named as an inventor.

Inventor signature: _____ Date: _____

Witness (read and understood): _____ Date:

Attorney received: _____ Date: _____