



Designing Vocabulary Learning through Psychological and Cognitive Science Principles: A Thirty-Year Retrospective

Robert S. Murphy

University of Kitakyushu, Japan
murphy@kitakyu-u.ac.jp

Abstract

Vocabulary learning appears to be a simple task; learners encounter words, repeat them over time, and recall them when needed. Nevertheless, a paradox persists: in the English as a Foreign Language (EFL) context, vocabulary often proves difficult to retain for an upcoming test and even harder to use in authentic contexts. Despite the promise of the spaced repetition method, forgetting remains common, and, as EFL students will concur, rote lists rarely translate into communicative competence. As a teacher and teacher trainer, I have been curious about this learning phenomenon for over 30 years. During my younger years as a postgraduate researcher, I began tracing the roots of *Mind, Brain, and Education* science, drawing on *Dynamic Skill Theory* (Fischer, 2008) and the early integrative syntheses of neuroscience and pedagogy (Tokuhamu-Espinosa, 2011). These studies laid the foundation for subsequent refinements in the field. As a retrospective monograph, this paper synthesizes decades of research bridging psychology, neuroscience, and pedagogical design to (1) explain why vocabulary learning is harder than it first seems, (2) how it can be made easier and more meaningful for the students, and (3) how assessments can be improved. Drawing on frameworks such as the *Dynamic Area of Total Convergence* (DATC), *NeuroELT*

maxims, and the *3D:CG rubric*, this paper discusses a curated set of wide-ranging pedagogical integrations (diagnostic, dynamic, and convergent) for teacher training, vocabulary lesson design, and course design. These integrations treat vocabulary learning as a developmental coordination problem in which stability emerges only through repeated convergence across linguistic, sociocultural, and affective systems (Fischer & Bidell, 2006; Murphy, 2019).

Keywords: L2 vocabulary, neuroscience, psychology in Language Learning, neuroELT, pedagogical design

Introduction: Why is Vocabulary Learning and Recall so Hard for EFL Students?

This monograph reflects on 30 years of teaching and teacher training, primarily within the Japanese English as a Foreign Language (EFL) context. It offers a retrospective design-theory synthesis, intended for specialists in vocabulary. The analyses and pedagogical designs presented here are grounded in classroom practice at the university level, complemented by work with private institutions and professional development courses for both pre-service and in-service educators. Most of the research and interventions took place in urban universities, with direct engagement of undergraduate students (ages 18–22) as well as future and practicing teachers.

My early postgraduate research in applied linguistics was shaped by the emerging field of *Mind, Brain, and Education* (MBE). Influenced by Fischer's *Dynamic Skill Theory* (2008) and the early integrative syntheses of neuroscience and pedagogy (Tokuhamma-Espinosa, 2011), I began to explore how cognitive development, emotion, and context interact in vocabulary learning. These foundations, later expanded through further refinements in MBE research (with Fischer's mentorship), continue to inform the pedagogical frameworks presented in this retrospective article.

As with many EFL tasks, vocabulary teaching and assessment have been among the most intractable challenges for both teachers and learners. In typical classroom practice, vocabulary instruction often defaults to some form of spaced repetition of lexical items, an approach supported by cognitive science as a useful tool for strengthening recall (Tokuhamma-Espinosa, 2011). While repetition can help learners prepare for immediate tests, it alone rarely secures long-term retention or authentic communication. This limitation is so widespread that it has become a familiar cliché (joke) about the shortcomings of school-based language learning.

Developmental research helps clarify why these problems occur. Fischer's work in *Microdevelopment* identifies three stages of learning-to-output stability, ranging from unstable novice performance to more reliably coordinated output (see Figure 1). When assessments occur during the unstable *Novice* or *Intermediate* stages, as is often the case in everyday classroom practice, the resulting scores capture only momentary performance snapshots with limited statistical significance. Despite this, such scores are routinely recorded and treated as valid indicators of learning. Over time, this practice

will certainly introduce systematic distortions into educational evaluation, significantly obscuring the distinction between short-term recall and sustainably coordinated growth, and, in practice, often confusing teachers, students, and parents alike. From the perspective of Dynamic Skill Theory, this variability comes naturally from learners' ongoing internal coordination and reorganization of understanding (rather than noise or "learner error" *per se*). As a result, learner output and any assessments based on it are inherently sensitive to timing and contextual support and are therefore best interpreted with those conditions explicitly in mind (Fischer & Bidell, 2006; Fischer & Rose, 1998).

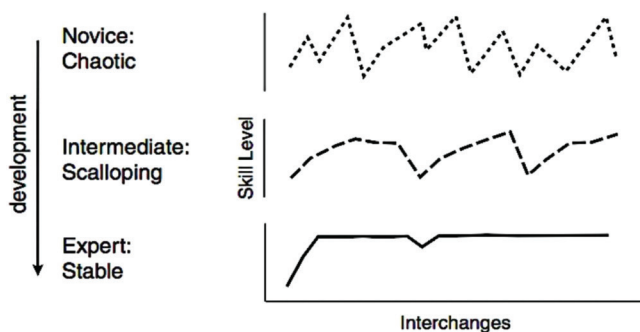


Figure 1 *Learning Outcomes Across Microdevelopment Stages (Fischer, 2008).*

Note. Considering the significant fluctuation, data collected on “bad days” would not be representative of learner’s underlying potential. Within Dynamic Skill Theory, these jagged trajectories are normal and expected (Fischer & Rose, 1998; Murphy, 2019).

A central question, then, is why so many learners remain in transitional stages without progressing toward more stable expert performance. In the case of lexical development, vocabulary often becomes inaccessible even for otherwise high-achieving students unless it is firmly anchored in emotion, culture, and motivational relevance (Murphy, 2019). Microdevelopment studies help clarify why classroom practices centered on rote lists and test preparation tend to have limited long-term impact. Over the years, I have critiqued reliance on short-term memorization as an unreliable solution to a long-term challenge; teaching to the test may produce brief bursts of recall, particularly around assessment periods, yet it rarely builds the stable networks required for fluent communicative use later in life (Segalowitz, 2010).

In many EFL classrooms, instruction proceeds as a ritual of rehearsal and assessment, with little expectation from either teachers or students that genuine proficiency will emerge, particularly in terms of vocabulary memorization. This pattern of misplaced complacency, often overlooked, mirrors long-standing findings in applied linguistics: decontextualized drilling and list learning rarely support durable competence (Lightbown & Spada, 2013). Empirical work on lexical richness confirms this limitation. Laufer and Nation (1995), for example, demonstrated that learners’ *active* vocabulary use is often dramatically smaller and less varied than their *receptive* vocabulary size (which demonstrates the inherent weakness of approaches that emphasize recognition

without nurturing meaningful depth). Thus, one of the most direct answers to the question of why vocabulary learning is so difficult lies not only in pedagogy itself but also in the low expectations surrounding it. When real learning is not anticipated, the conditions for it to occur are not needed (and are therefore not constructed).

Traditional instruction often narrows its focus to decontextualized word lists, a practice long critiqued in vocabulary studies for offering breadth without meaningful depth (Laufer, 1992; Nation, 2013). This distinction aligns with work showing that lexical competence reflects multidimensional lexical sophistication and that frequency or item counts provide an incomplete basis for predicting learning burden and use (Vitta et al., 2023). Vocabulary knowledge involves both quantity and precision, and lexical thresholds for comprehension demand more than mere surface familiarity. Laufer and Nation's Lexical Frequency Profile (1995) demonstrated that learners' productive vocabulary often lags far behind receptive recognition, revealing a persistent gap between rote learning outcomes and communicative capacity.

From a developmental perspective, this narrowing of instructional focus can substantially reduce the *Dynamic Area of Total Convergence* (DATC) (see Figure 2) to the linguistic dimension alone (Murphy & Sin, 2014), a pattern consistent with complex systems accounts of language development (Larsen-Freeman & Cameron, 2008). In such cases, learners may become confined within a limited microcosm of lexical knowledge that lacks sufficient sociocultural and emotional anchoring. As a result, they are often forced to rely on their first language to provide these missing components, leading to a fragile and distorted understanding of the second language. This helps explain why rote learning, even when supported by spaced repetition, so often fails to transfer into flexible, real-life language use.

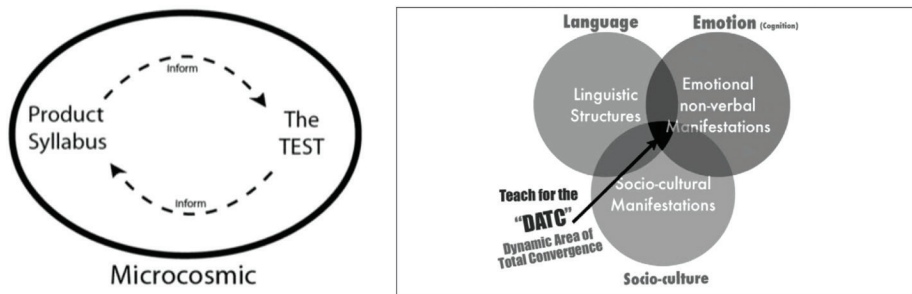


Figure 2 *Microcosm Schematic and DATC Diagram (Murphy, 2009, 2019).*

Note. In the left pane, we see how teaching to the test forms a microcosm. In the right pane, we see the theoretical *Dynamic Area of Total Convergence* (DATC). The more balanced and overlapping the central area is, the more stable the learning is for the student. The DATC diagram can help teachers and students quickly visualize what may be missing in their current learning context (Murphy, 2009, 2019).

The “Teaching for the DATC” approach is intended to make instructional balance visible to the teacher and the student. It is a biology-based metaphor that can be useful for meta-discussions. The DATC framework initially emerged from classroom

observation, where vocabulary development was found to be more stable when learning was distributed over longer periods and embedded in meaningful engagement. The DATC framework offers a practical, quick diagnostic for examining whether instructional designs provide conditions that can help lexical knowledge further develop and reorganize in the long term.

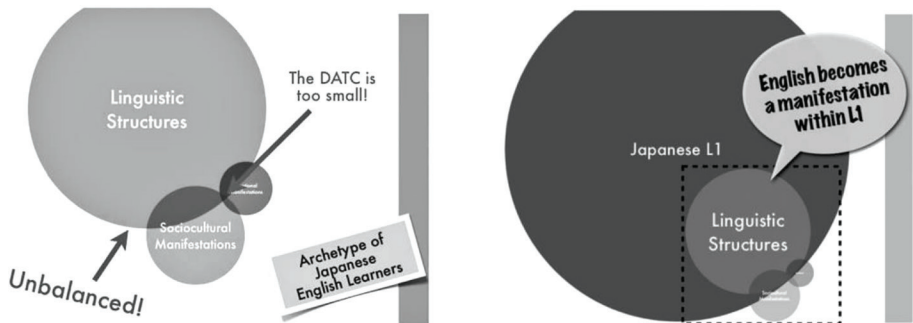


Figure 3 *DATC Outcome Archetypes (Murphy, 2009).*

Note. The left pane shows the theoretical *Dynamic Area of Total Convergence* (DATC) outcome of an archetypical “worst case scenario” of a Japanese learner of English (EFL context); the focus has been on linguistic structure development, with little time spent elsewhere, creating underdeveloped sociocultural and non-verbal (emotional) domains for English, with an almost nonexistent DATC (convergence area). Without their own self-standing “English foundation,” the student must, in theory, rely on their Japanese understanding of the world to make sense of the English they are being taught. “English” becomes a manifestation *within* the student’s Japanese self rather than being its own system within the student’s mind (Murphy, 2009).

In Figure 3, the left pane is a “blow-up” of the right pane’s dotted box. This demonstrates that this student’s fragile English (L2) competency is merely a manifestation of their Japanese (L1) understanding of the world. It need not be this way. Figure 4 shows how the neural networks can develop separately (metaphorically, as separate trees for the most important characteristics of the lexical items, but naturally still sharing a common underlying soil, where the separate trees are rooted). Although these are caricatures of real neuronal network development, they are sufficiently effective in portraying linguistic network development. This depiction aligns with the distinction in Dynamic Skill Theory between functional stability and optimal-level coordination, in which higher-order systems remain accessible but unsustainable without sufficient convergence and support (Fischer & Bidell, 2006; Murphy, 2019).

Returning to the question, “*What makes learning vocabulary hard?*,” Figure 4 sums up much of the answer. Can we possibly foster separate trees for every lexical item that is to be learned? This seems unlikely in most EFL contexts. Moreover, if Figure 3 accurately depicts the reality for many Japanese EFL students—L2 becoming a manifestation within their L1—should we simply accept this as a given? How can we possibly separate the L2 if it has already manifested within the L1 as a subset of their L1? Regardless of the pedagogy we choose, these are core issues regarding why vocabulary is such a challenge in the EFL context.

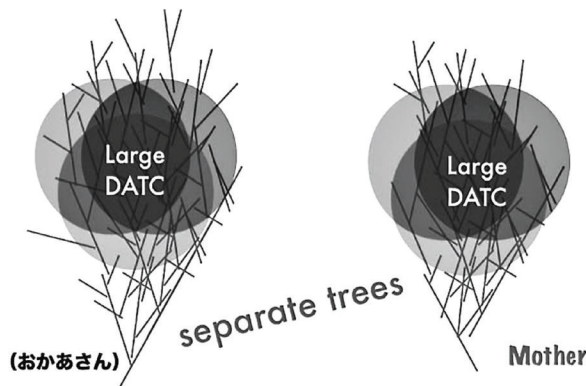


Figure 4 *DATC Development as Separate Networks in Development.*

Note. The left tree (neural network) has developed its understanding of “okaasan” (mother) in Japanese from Japanese contexts. The right tree is depicted as a separate entity to symbolize that the understanding of “mother” has developed independently of the Japanese tree’s development. For a word like “mother,” there should also be plenty of underlying schema overlap (below these trees), not visible here. Compare this outcome to Figure 3 (Murphy, 2009).

While rooted in Japanese EFL, the key principles of this paper (the necessity of deep convergence across linguistic, sociocultural, and affective domains for durable vocabulary learning) and pedagogical innovations such as TRIDENT, CREAM, and 3D:CG should be adaptable for use in EFL contexts, bilingual education, and other international language learning settings. Contextual adaptation may be required for local educational constraints, assessment regimes, and cultural/linguistic backgrounds; however, the underlying principles of the theoretical and pedagogical frameworks are broadly transferable.

What Lies Further Beneath this Difficulty with Vocabulary Teaching?

My first postgraduate dissertation (Murphy, 2009) draws on work from *Mind, Brain, and Education*, a form of Educational Neuroscience that emerged from the Harvard Graduate School of Education. I focused primarily on Dynamic Skill Theory and Fischer’s Self-in-Relationship (SiR) interview research to demonstrate that learner performance varies dramatically across contexts, specifically between high- and low-support contexts. Since I was working with Japanese EFL students in Japan while being mentored by Fischer in Boston, I was positioned to observe how different educational cultures shape classroom engagement and learning behavior (Murphy, 2009). In the Japanese context, collectivist norms and Confucian educational traditions influence how students participate, express uncertainty, and respond to instructional risk (Markus & Kitayama, 1991; Tweed & Lehman, 2002).

Across instructional contexts, emotional engagement is what makes the difference between vocabulary learning that *remains superficial* and learning that *becomes durable and transferable*. Research in affective neuroscience has shown that emotion is integral to learning, shaping valuation, attention, and the consolidation of meaning over time

(Immordino-Yang, 2016; Immordino-Yang & Damasio, 2007). From a developmental standpoint, affect participates in the stabilization of learners' coordinations, influencing whether lexical knowledge continues to reorganize and remain usable beyond the immediate instructional setting (Murphy, 2019). Several recurring issues in vocabulary instruction emerge from this discussion:

- (1) Lexical learning remains fragile when learners do not perceive a reason to value, or to personally connect with, the learning material. Learner recall in such cases may be sufficient for short-term test performance. However, without further support, those lexical items rarely stabilize in ways that support long-term usage, expansion, or transfer.
- (2) This points to the need for instructional designs that support affective engagement in a substantive sense. Here, affective engagement refers to emotionally grounded involvement tied to personal meaning and communicative relevance (cf. Murphy, 2015).
- (3) To be fair, "Teaching to the test" does generate a form of affect, primarily through pressure and consequence. However, because this form of affect is anchored to assessment demands rather than to learners' personal communicative needs, it tends to dissipate along with the dissipation of assessment demands. Why? Because there is little reason for that lexical knowledge to remain active or to continue reorganizing beyond what the test required (Murphy, 2019).
- (4) What is often missing is a strategic pedagogical design that explicitly supports affective engagement strong enough to sustain long-term retention, growth, and transfer potential. When instructional environments provide reasons for lexical material to continue to really matter to learners, vocabulary is more likely to remain active, refined, and usable in the long term.

This list of issues may help clarify the different roles that instructional materials may play. Language textbooks from the 1970s through the 1990s were often more willing to incorporate playful, culturally explicit, or emotionally engaging content. Over the past two decades, mainstream materials have increasingly emphasized predictability, assessment alignment, and institutional safety, particularly with respect to culturally or emotionally charged topics. While these shifts are understandable from commercial and administrative standpoints, from a neuroeducational perspective, the field seems to have taken a few steps back. Although it may be unintentional, several modern publishers are essentially stripping away what gets students' brains "hooked" on the content. This, of course, defeats the purpose of the textbook—unless everyone is simply anticipating test prep, which can be a legitimate goal, albeit one that reinforces the "real-world learning is not expected" issue discussed above.

Is there a viable way forward? Research in *Directed Motivational Currents* suggests that when learning tasks are designed to generate sustained, emotionally rich engagement, learners can enter extended periods of high motivation that support growth far beyond immediate tasks (Dörnyei et al., 2014). From a developmental perspective, I find that these motivational currents are best understood as downstream indicators of emerging stabilization rather than as upstream causes of engagement (Murphy, 2019). In other words, I argue that "motivation" is the end product of successful coordinations

rather than the catalyst that it is typically accepted to be. Thus, this chapter departs from prevailing motivational models by explicitly advancing this reversal as a developmental claim: what is commonly labeled “motivation” reflects the stabilization of underlying coordinations rather than serving as their causal source (Murphy, 2019; Murphy, in press).

This distinction becomes clearer when comparing first- and foreign-language development. In first-language (L1) acquisition, learners typically establish a stable linguistic foundation before formal schooling begins. Classroom instruction then builds upon this foundation, assuming that some level of stability is already in place. However, in foreign language (L2) contexts, learners enter classrooms without an equivalent foundation. Despite this, instructional materials and pedagogical routines often proceed as if stability has already been established, relying on pragmatic shortcuts that prioritize short-term performance over foundational creation and coordination. Over time, this has encouraged instructional practices that produce the appearance of progress while leaving the underlying linguistic system insufficiently stabilized. From a developmental standpoint, the central challenge is therefore not how to accelerate surface performance, but how to support the emergence of genuine L2 stability upon which sustained motivation and communicative competence can reliably grow.

For vocabulary, the brain requires convergence across multiple domains. Classroom practices often fall short of this depth of engagement. Research in extensive reading suggests that lexical competence develops most effectively when learners encounter vocabulary items repeatedly in rich, contextualized input (Coady & Huckin, 1997; Nation, 2013). This aligns with the DATC framework, which emphasizes that emotionally and culturally charged contexts build the stable networks necessary for long-term retention and transfer. In sum, the brain typically requires deeper convergence than can be achieved with classrooms/materials that have been designed to leapfrog over L2 stability development (allowing, or perhaps even expecting, L2 to manifest as a subset of L1). Thus, deep convergence is what is often missing—and what must be included—if a reliable foundation is desirable. Figure 5, my graphic reinterpretation of Ausubel (1977), summarizes this issue from a pedagogical design perspective.

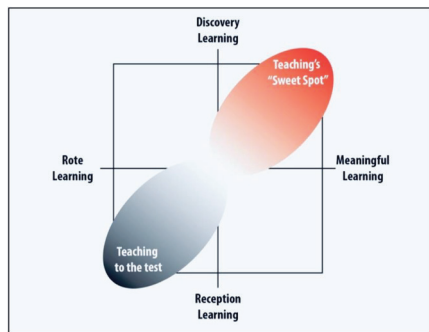


Figure 5 *Possible Types of Learning: Graphic Adaptation of Ausubel (1977).*

Note. If we are not explicitly teaching to the test, the “sweet spot” of EFL teaching often should be the top-right corner. However, we find that much of the EFL teaching in Japanese schools is spent in the bottom-left corner (Murphy, 2019).

Honestly, how much emotional bonding can we expect students to form through teaching that exists in the “Reception-Rote” learning corner? Conversely, the “Discovery-Meaningful” learning corner has an emotional/visceral impact on students, by design. Which corner has a higher efficacy? That naturally depends upon the goal: (a) prioritize lower left if test score improvement is the goal, and (b) prioritize upper right if more long-term usage-oriented learning is the goal. While that may not sound like profound insight, the deeper purpose of this graphic is to clarify how far apart the two goals are and why attempting to achieve both has proven so difficult—they are on opposite poles. I have found this graphical representation useful when training teachers, students, and even when discussing with parents who are curious about why many EFL students may perform well on standardized exams yet exhibit very low real-life communicative fluency. And so, being able to put a finger on the issue is a significant first step. What comes next? *Pedagogy*, leading to deeper convergence for a stable foundation in the brain.

What Creates Stable, Deep Convergence?

As discussed above, the Dynamic Area of Total Convergence (DATC) model visually demonstrates that three regions (linguistic, sociocultural, and emotional) must overlap well enough to ensure stable and sustainable learning development (Murphy & Sin, 2014). This framework is supported by neuroimaging research showing that brain regions engage differently across tasks (Friederici, 2011; Goswami, 2006; Kuhl, 2010). For instance, during vocabulary learning tasks, brain regions respond differently to semantic, phonological, or contextual demands. Additionally, emotion enhances so-called “encoding and recall”, and numerous studies confirm that emotionally salient material is consolidated more robustly in the brain (Immordino-Yang, 2016; LeDoux, 1996; Tyng et al., 2017).

Simply put, the more areas of the brain that are engaged during learning, the more pathways exist to retrieve and make use of that past learning (Sousa, 2017; Tokuhama-Espinosa, 2011). This principle has been consistently emphasized in *Mind, Brain, and Education* research, where integrating multiple modalities creates richer, more durable traces. Therefore, it is logical that, within the DATC framework, optimized stability emerges when linguistic structures, sociocultural manifestations, and emotional resonances overlap significantly (Murphy & Sin, 2014). The DATC framework graphically illustrates how normal vocabulary acquisition (typical L1 learning) is embedded alongside social and emotional experiences; L2 may follow a similar pattern, but this is often not the case—shortcuts seem to be the norm (Ellis, 2008). Traditional grammar-translation tends to underdevelop the convergence area, leading to shallow recall, weak transfer, and quick atrophy. Effective vocabulary teaching requires the deliberate incorporation of culture and emotion into task design (Tokuhama-Espinosa, 2018). Can these principles be meaningfully enacted through classroom pedagogy?

One example of how this can be accomplished is the TRIDENT activity. It can be used in EFL classes for L2-L2 vocabulary mapping (enriching) or for deep translation exercises (see Figure 6). By mapping out synonyms (or translation solutions) along axes such as *literal-figurative* or *unique-universal*, the TRIDENT activity makes visible (“tangible”) how vocabulary’s deeper meaning shifts across cultural and contextual

dimensions (Murphy, 2013). Compared to typical grammar-translation tasks, the uniqueness of this approach (allowing for emotionally evocative and unconventional solutions) makes it significantly more engaging and amusing for students. Group translation tasks—translation of L1 into various L2 (or L2 to various L1) solutions across the ten zones—can heighten cultural awareness and highlight the nuanced interplay of lexical choices, demonstrating that words can exist within webs of figurative and socio-cultural associations rather than just being mundane and disjointed dictionary entries. This can bring lexical items to life in an uncommon yet advantageous way.

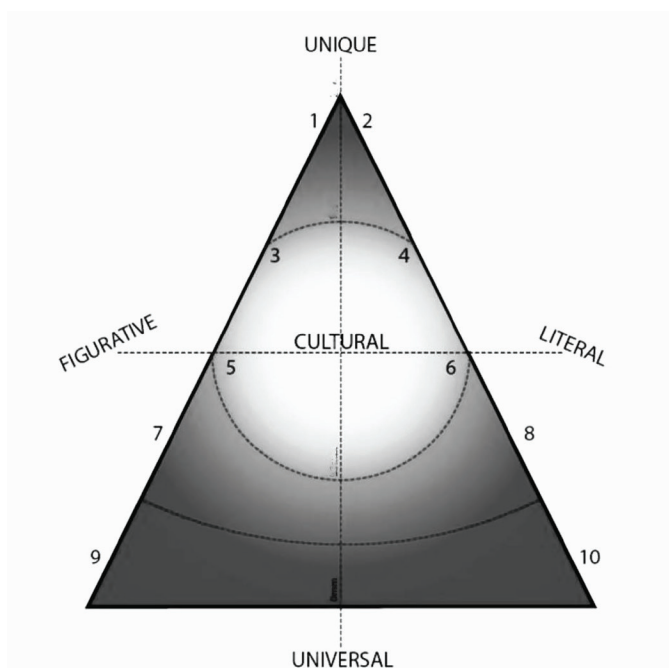


Figure 6 *The TRIDENT Worksheet (Murphy, 2013).*

Note. TRIDENT is an acronym for “TRIangular DENary system for measuring disparity in Translation.” It is a multipurpose tool for the deep exploration of lexical items. I initially developed it for translation task management, but I soon realized how important it could be for L2-L2 deep analyses. How does it work? Provide a lexical item, such as “car,” and then ask students to come up with other words for “car” for each of the 10 zones in the TRIDENT map. Zones 9-10 are the most universally acceptable areas, while the odd numbers (left side of the Y-axis) must be figurative, and the even numbers (right side of the Y-axis) must be literal. Example usage: Zone 10 = “automobile”; 6 = “sedan”; 4 = “Mini Cooper”; 1 = “clunker”; 7 = “my wheels.” This is a naturally engaging and memorable lexical game-like task for students. Because this is time-consuming, it cannot be used for every word that students need to learn, but it can be used to engage deeply with the most important lexical items of the day (Murphy, 2013).

Although I seldom assign traditional “grammar-translation” tasks in my own Japanese EFL courses, TRIDENT grew out of earlier design-based work conducted during my

postgraduate studies, when I was engaged simultaneously with pedagogical theory, second language acquisition, and translation methods. The activity was originally conceived as a classroom translation task that could hold learners' attention more effectively than conventional sentence-level grammar–translation exercises, while still preserving the analytical value that translation can offer as a pedagogical practice.

Through repeated classroom use, TRIDENT gradually revealed a broader pedagogical role. Over time, two recurring patterns have become particularly noticeable. First, many Japanese students encountered, often for the first time, a task structure that openly accepted multiple, viable solutions. This recognition proved liberating, as learners came to appreciate that variation is a normal feature of real-world translations. Second, translation into students' local dialects consistently generated higher levels of engagement, interaction, and enjoyment than translation into standard ("Tokyo dialect") Japanese does. To clarify how these differences played out in classroom practice, the patterns observed through TRIDENT are summarized in Table 1.

Table 1 *Learner Responses to Translation Targets in TRIDENT Activities*

Dimension	Translation into Standard ("Tokyo Dialect") Japanese	Translation into Local Dialect
<i>Perceived authenticity</i>	Experienced as distant from everyday interaction	Experienced as closely aligned with daily communication
<i>Emotional response</i>	Neutral or restrained affect	Positive affect, including humor and enjoyment
<i>Learner engagement</i>	Task completion focused	Sustained engagement with voluntary elaboration
<i>Solution sharing</i>	Limited peer sharing	Spontaneous sharing ("Look at my solution!")
<i>Tolerance for variation</i>	Expectation of a single acceptable answer	Acceptance of multiple valid solutions
<i>Classroom interaction</i>	Individual, quiet work patterns	Peer discussion and collaborative comparison
<i>Teacher training response</i>	Procedural completion of task	Open discussion and reflective comparison of choices
<i>Pedagogical value revealed</i>	Emphasis on form-level accuracy	Emphasis on meaning, stance, and affective alignment
<i>Assessment alignment</i>	Closely aligned with standardized testing norms	Closely aligned with authentic language use

When asked why dialect-based translation felt more engaging, students often explained that standard Tokyo Japanese carried a distant, institutional tone—"like *NHK News*"—that differed from how they spoke with friends or family members in everyday life. Working through their local dialect afforded greater latitude for expressing nuance, humor, and interpersonal positioning, which in turn appeared to support more sustained lexical exploration. Over the years, these observations have emerged

repeatedly across cohorts and instructional contexts, suggesting that they are likely a shared nationwide phenomenon. This aligns well with sociolinguistic research showing that learners' willingness to invest in language use is closely tied to their identity and the perceived legitimacy of their voice within a given linguistic context (Norton, 2013).

Comparable responses also emerged in teacher education settings across Japan, extending this pattern beyond student cohorts. When student teachers were invited to complete TRIDENT-style translation tasks using their own dialects during workshop activities, laughter, spontaneous sharing, and peer comparison became common features of the task. These moments offer a useful view of how linguistic familiarity and affective alignment shape engagement and interaction at both the learner and teacher development levels. From a developmental perspective, these effects reflect reduced coordination strain when learners work within familiar sociolinguistic systems, allowing for a greater allocation of resources to semantic exploration (Murphy, 2019).

In classroom practice, translation activities grounded in learners' lived linguistic repertoires regularly foster higher engagement, tolerance for variation, and deeper "meaning-focused" discussions. In this way, routine vocabulary lists and textbook dialogues involving generic characters can become opportunities for creativity and shared enjoyment when worked through the dialects learners actually use.

It seems that simply adding opportunities for dialect-informed translation was sufficient to shift the character of routine vocabulary work, opening up space for exploration, interaction, and enjoyment. How else, then, might we move toward deeper convergence? In many EFL classrooms, vocabulary instruction settles into a narrow linguistic "microcosm" shaped by testing and coverage demands (see Figure 2). However, real-world vocabulary usage carries specialized sociocultural weight: interpersonal distance, politeness, emotional directness, taboo topics, and power relations are all embedded in word choice and usage. Thus, it is understandable that teachers and students may feel hesitant to engage with these complex dimensions, particularly when instructional time is limited or assessment pressure prevails.

One of the strengths of TRIDENT is that it helps lower these barriers by shifting meaning-making to shared artifacts. Learners explore multiple acceptable interpretations side by side, comparing solutions without being required to immediately reveal personal stances or converge on a single "correct" answer. This creates space for socio-cultural exploration to emerge naturally, often accompanied by laughter, curiosity, and peer discussion.

As discussed earlier, affect plays a central role in pulling these elements together. When tasks evoke emotional resonance alongside linguistic and cultural exploration, deeper connections are formed, and convergence across the three domains of the DATC becomes more attainable (see Figure 7). While this process may initially feel demanding, project book pages like these have been specifically designed to support teachers in implementing "Teaching for the DATC" in everyday classroom practice.

Figure 7 is an excerpt from an older textbook that I wrote (*Optimal Levels! Project Book*; Murphy, 2010), which provides a solid example of how vocabulary can be studied with the DATC in mind, using the tri-elemental model as a project book page. The "Linguistic connections" bubble can house synonyms/antonyms, and other related terms. The "Sociocultural" bubble is for any connections to society/culture that stem from that lexical item. The "Feeling and emotions" bubble is for students

PROJECT A

Tri-elemental Vocab sheet

Use this page to gain a deeper understanding of new words.
Make sentences that connect different elements. Draw images.

Figure 7 *Teaching for the DATC: Example Worksheet (Murphy, 2010).*

to explore their feelings about the word, especially in light of the other two bubbles. Then, students doodle a picture that represents the word—and as a closing, they put it all together in a sentence. Not all bubbles need to be explored in the sentence, but the more they are used, the better (Murphy, 2010).

How Else Can Emotion and Consciousness be Woven into Vocabulary?

The CREAME pedagogy (Consciousness Raising, Emotion Analysis, Manipulation, Expression; Murphy, 2009) invites learners to engage with words through reflective, creative, and emotional tasks. Figure 7 is a good example of this. It emphasizes the cognitive recognition of vocabulary items and the affective and expressive dimensions of language learning. This aligns with other studies on learner training in EFL classrooms, which emphasize the importance of helping students become more personally engaged with language material beyond rote lists (Murphy, 2010). This pedagogical style also aligns with evidence that deeper, meaningful processing supports retention better than simple repetition (Craik & Tulving, 1975; Karpicke & Smith, 2012). Tasks that require learners to link new words with personal narratives, emotional reactions, or cultural associations engage multiple neural networks, facilitating durable recall (Immordino-Yang & Damasio, 2007). Furthermore, manipulation and expression activities, such as adapting words in role plays, debates, or creative writing, encourage flexible use, which enhances both semantic depth and pragmatic competence (Murphy, 2010). Thus, CREAME integrates metacognition with emotion, enabling learners to

transform words from boring/abstract entries into tools for deeper thinking and meaningful communication.

Figure 8 is an example of using CREAMÉ as an advanced form of “mind-mapping.” The root word for this example happened to be “America,” and the lexical items lists were lifted from an actual student’s responses (Murphy, 2009). While the left pane shows a rather typical mind-map with groupings of ideas connected to the theme, it is just the warm-up phase (the “Consciousness Raising” aspect of CREAMÉ). The right pane shows the next steps—how the student put together a web of understanding on paper that mimics the neuronal web co-developing in their brain in real time as they created these new connections on the paper. This forms a feedback loop, cycling between the content on the paper and the developing thoughts in the brain. This is real-time cognitive restructuring (the “Emotion Analysis” and “Manipulation” aspects of CREAMÉ).

What is the task that creates this growing feedback loop? The student was asked to identify cross-connections between the not-directly-related groupings they had come up with earlier in the day. Now they are adding emotional valences while seeing how everything fits together in a bigger picture (both positively and negatively). It is important to note that the teacher does not provide these lexical items. They are discovered by the students who are now self-motivated to create this web of understanding. After this task is completed to their satisfaction, they have a choice: draft a short essay or give a short presentation on this topic (the “Expression” aspect of CREAMÉ).

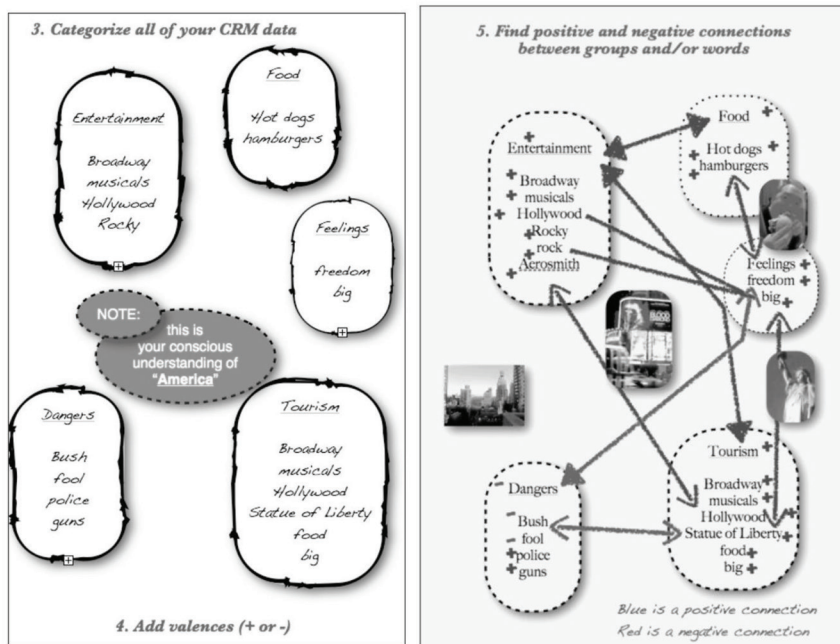


Figure 8 CREAMÉ: Student Output Example (Murphy, 2009).

Note. CREAMÉ in action. The mind map on the left serves as a warm-up to a far more elaborate exploration—the creation of a web of lexical items and rationales on the right.

CREAME offers a structured approach for integrating cultural, affective, and personalized aspects of lexical items. If followed in order, CREAME asks learners to first raise awareness of a word's features, then to analyze its emotional or cultural significance, to manipulate it creatively, and finally to demonstrate/express it in a "performance" of understanding. Such a cycle helps ensure that vocabulary is not merely recognized (then moved on from) but is actively integrated into the learner's communicative repertoire (becoming part of their personal lexicon). CREAME complements all other pedagogical models discussed in this paper (such as 3D:CG assessments discussed below) by offering classroom procedures that make cultural and emotional anchoring explicit (Murphy, 2008, 2019).

Evaluating Pedagogical Design and Vocabulary Outcomes

The pedagogical frameworks discussed thus far point toward deeper, more durable vocabulary learning. A practical question naturally follows: how can such pedagogical design and its outcomes be evaluated in ways that reflect depth, durability, and real-world relevance, rather than short-term recall?

My dissatisfaction with conventional assessment rubrics stemmed from their tendency to flatten learning into additive tallies (simple addition). In response, I developed the *3D Constant Gradient* (3D:CG) evaluation system, which treats learning as a dynamic system rather than a collection of independent parts. The rubric evaluates instructional materials and activities across three intersecting dimensions—*Duration*, *Engagement*, and *Applicability*—represented as the x, y, and z axes of a three-dimensional space (see Figure 9). Instead of simplistically *adding* rubric scores, 3D:CG uses *multiplicative logic*. Also, decimal usage is encouraged to allow finer-grained judgment, as learners are taught to use the rubric for self-evaluation and meta-awareness development. In practice, this has supported learner motivation to understand their own cognitive development (Murphy, 2019).

Rubric-Based Evaluation: Using the 3D:CG in Practice

To apply the 3D:CG rubric, instructors and learners evaluate a task, lesson, or set of materials across three dimensions. *Engagement* captures the degree of active, invested participation elicited by the task. *Duration* reflects the likelihood that learning remains accessible beyond the immediate activity. *Applicability* represents the extent to which the learning transfers to real-world applications (see Figure 9).

For example, an activity that invites learners to select vocabulary to convey stance or emotional nuance in response to a reading passage may rate strongly in *Engagement*, moderately to strongly in *Applicability*, and relatively strongly in *Duration*—especially when contrasted with rote-memorization tasks.

Each dimension of the 3D:CG rubric reflects long-standing insights from neuro-education and applied linguistics. Vocabulary durability requires recurrence across lessons and time frames. Engagement emerges through the active manipulation of lexical items rather than passive exposure. Applicability pressures vocabulary to function in communicative contexts beyond classroom exercises. Multidimensional accounts of lexical competence (Meara, 1996) and systematic approaches to vocabulary growth

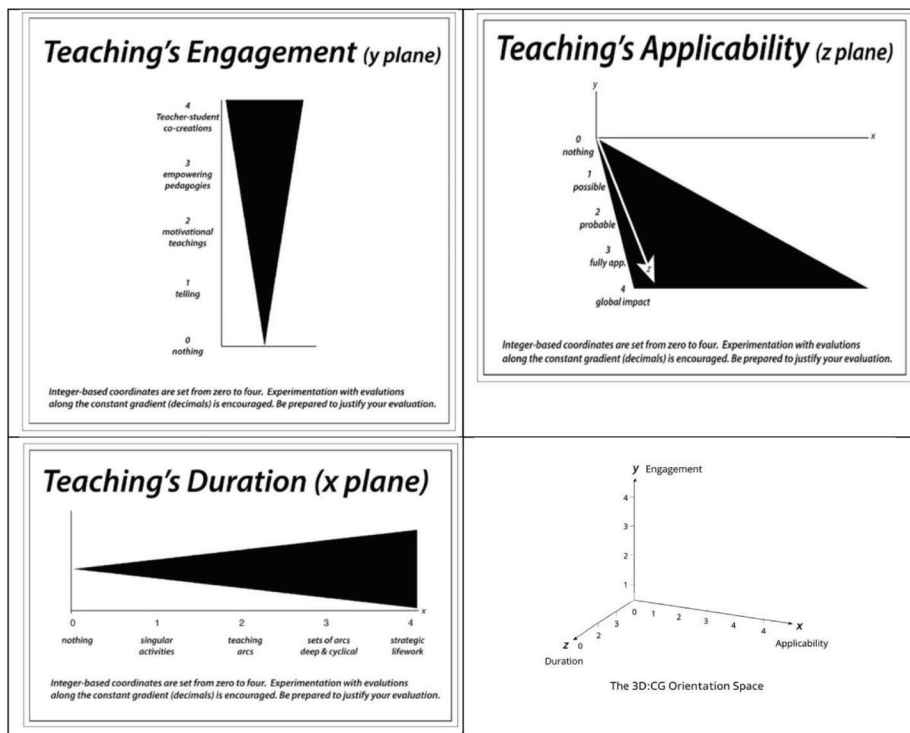


Figure 9 *The Planes of the 3D Constant Gradient (3D:CG) Assessment Framework.*

Note. 3D:CG evaluates instructional design performance across x (Duration), y (Engagement), and z (Applicability) planes using a multiplicative structure that incorporates decimal usage and zero values, hence the term “constant gradient” (Murphy, 2019). The rendered result occupies a three-dimensional space (bottom-right frame).

(Milton, 2009) both emphasize sustained use and meaningful application. The 3D:CG rubric extends these principles into *pedagogical design theory* and also serves as an *evaluation tool* that makes durability and communicative viability visible.

A defining feature of the rubric is its mathematical logic. Because the three dimensions interact multiplicatively, instructional strength appears as volume instead of a simplistically “added up” tally. This distinction matters because a low evaluation, such as a “0.5” in a cell on a standard rubric, is still added to the tally, raising the total rather than lowering it. However, because multiplication is used in 3D:CG calculations, “below 1” values such as “0.5” on any plane significantly compresses the resulting form instead of adding to it. Furthermore, a “zero” evaluation on any plane would collapse the form entirely.

This aligns closely with pedagogical intuition: sustained learning requires durability, engagement, and real-world relevance to co-occur. In contrast, addition-based rubrics can mask weak dimensions behind stronger ones. Thus, the mathematics underlying the 3D:CG rubric makes imbalance issues prominent and proportionate

to their probable pedagogical impact. This non-additive logic mirrors dynamic systems approaches in which learning outcomes emerge from non-linear interactions (Murphy, 2017, 2019).

When I first began using the 3D:CG in program reviews, the most noticeable shift occurred once evaluations were rendered spatially by the users. The discussion rarely centered on explicit numerical outcomes. Instead, attention was on the shape of the result—where volume extended, where it remained compact, and which pedagogical choices may have produced that outcome. Over time, this spatial “idiom” became a shared language, a kind of shorthand, for discussing pedagogical design with staff and students (see Figure 10).

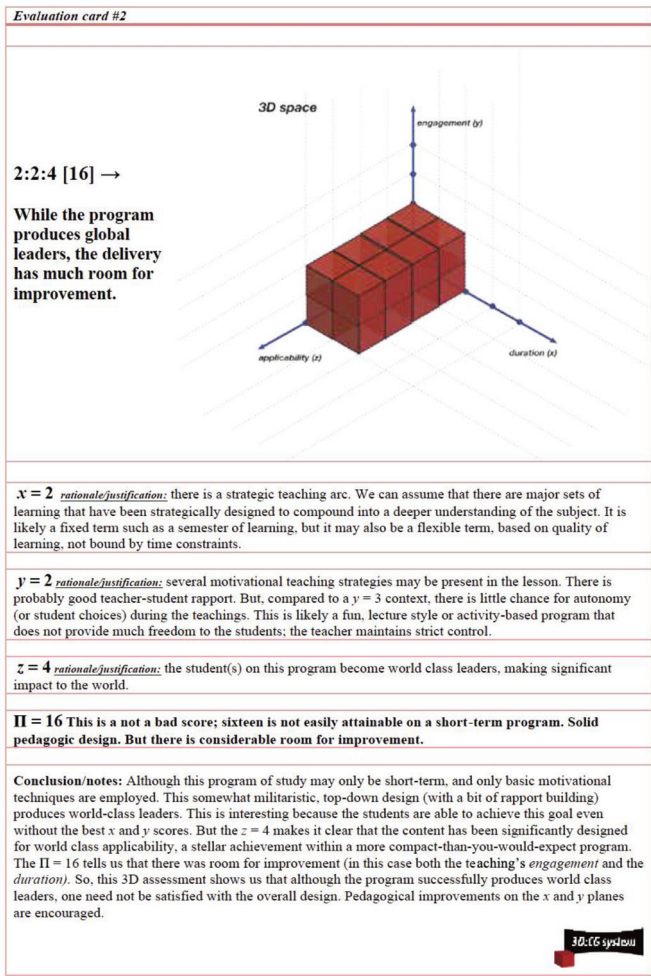


Figure 10 Example 3D:CG Evaluation Card.

Note. An applied instance of the 3D:CG rubric showing a three-dimensional instructional profile. Evaluation rationales and pedagogical implications are recorded beneath the visual representation, supporting reflective discussion and design planning.

Across contexts, these evaluation cards functioned as stable reference points for pedagogical conversations. Multiple 3D:CG profiles are often examined together to compare lessons, units, or programs—an analytical move that is possible only through this three-dimensional, multiplicative structure (see Figure 11).

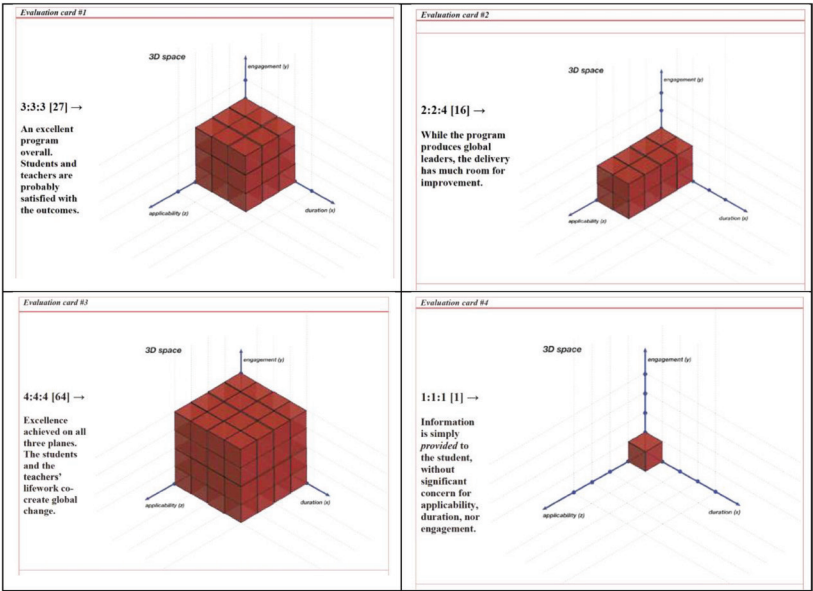


Figure 11 *3D:CG Evaluation Examples: Four Instructional Cases.*
Note. A set of evaluation cards (top half only) illustrating variations in instructional profiles across different pedagogical designs.

In practice, teachers frequently begin by auditing materials using guiding questions aligned with the three dimensions. Do target words recur across weeks or months? Do learners manipulate vocabulary in ways that require choices, stances, or elaboration? Do lexical items serve communicative purposes beyond classroom exercises? Addressing these questions systematically enables the 3D:CG to serve as both an evaluative framework and a design guide (see Figure 12).

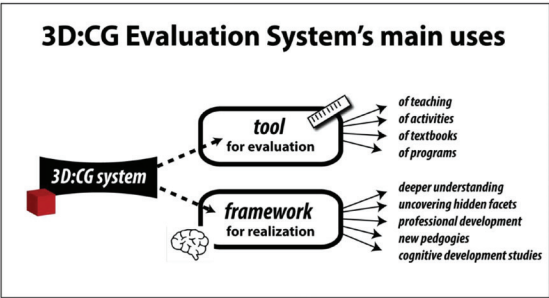


Figure 12 *Uses of the 3D:CG Evaluation System (Murphy, 2019).*

When learners engage in debates, presentations, or collaborative projects, the rubric helps assess whether the learning content has been reinforced across linguistic, cultural, and emotional dimensions. In this way, evaluation becomes directly linked to domain convergence rather than isolated recall. Learner self-evaluation using the 3D:CG has also provided instructors with grounded insight into how students perceive their own learning trajectories, offering a practical window into convergence without specialized instrumentation.

Are There Specific Guidelines That We Can Follow?

This retrospective monograph encompasses a diverse range of content spanning 30 years. A condensed list of “best practice” rules may come in handy. Below, I have a compiled set of rules developed during my doctoral thesis research (2019). I had two independent studies working out “best practices” – one was purely in applied linguistics. It produced a list of excellent “tenets,” while the other, in NeuroELT (Neuroeducation for ELT), produced a list of what we now call “NeuroELT maxims.” Both studies were conducted independently, but seven of the tenets/maxims overlapped in content. This naturally strengthened both studies, and this monograph would not be complete without a display of these results (see Figure 13). This list can serve as a springboard for material design, short-term lesson planning, and long-term course development. The NeuroELT maxims discussed here were later reorganized and further refined into an extended set of 50 maxims for teacher training contexts (Murphy, 2023).

Figure 13 is from my PhD thesis on EFL material design (2019). This is where psychology, neuroscience, education, and applied linguistics converged, and a list of “best practices” for pedagogy and material design emerged. It may seem that if teachers simply followed these rules, they could expect good results in their teaching practices. However, the reality is not as simplistic as that. In my 2019 study, I compared the efficacy of two textbooks (“Old” vs. “New”)—both of which I had designed with the majority of the neuro-education principles discussed in this paper. In many ways, they were similar textbooks. Surprisingly, the “new” textbook showed an 8x faster growth (learning) rate over the “old” textbook of similar design. The focus group discussions I conducted (in Japanese) helped to shed light on the massive discrepancy between the outcomes.

To understand the differences observed between the “old” and “new” textbook users, *Lectical Analysis* (growth analysis system based on Dynamic Skill Theory) was conducted on pre- and post-test writing samples of both groups (Murphy, 2019). We tested their demonstrated growth in written output complexity (demonstration of deeper thinking). As shown in Figure 14, using the *Lectical scale*, undergraduate L1 English users demonstrated a developmental growth rate of 0.05–0.08 points per year (Lectica, benchmark data). Using the upper bound (0.08/year) as the benchmark, results showed that the “old” group ($N = 24$) achieved a mean gain of 0.01 over 15 weeks, equivalent to 0.04/year, roughly half the benchmark rate for L1 learners. In contrast, the “new” group ($N = 30$) achieved a mean gain of 0.08 over the same period, equivalent to 0.32/year, or four times the benchmark rate. When directly compared, the relative growth ratio between OL1 and OL2 was 1:8. To ensure that pre-test level differences between the groups did not distort the results, ANCOVA was applied with

Table 55: The seven neuroELT maxims that are fully compatible with the tenets from this thesis

NeuroELT maxim	Explanation
"Choices" fuel learner motivation	Autonomous choice-making commonly activates neural networks that provide feelings of satisfaction. Thus, acting with autonomy is typically a naturally pleasurable experience.
"Real-time feedback" is at the core of cognitive development	The human love of the prediction-outcome dichotomy stems from the rewarding rush of dopamine received when we <i>perceive</i> our solutions to be correct. Thus, real-time feedback satisfies and even nurtures our craving for the dopamine rushes.
"Aha moments" enhance neural networks	Pleasurable dopamine rushes also result from successful problem-solving. The neuronal networks involved become more <i>efficient</i> through myelination; the most often used networks therefore become enhanced for superior processing (speed and quality actually become enhanced).
"Happy students" learn better	Studies show that happier individuals perform better academically, with enduring motivation. <i>Positive Psychology</i> has found that engagement in <i>perceived</i> meaningful activities may sustain happiness more effectively than the simplistic pursuit of pleasure.
"Performances of Understanding" are essential for good assessment	PoU activities provide learners with the opportunity to make their learning and understanding visible by having them demonstrate their <i>Understanding</i> (real word mastery) of the concepts. PoU-based pedagogies are aimed at synthesizing student learning and advancing it all to the next level in the form of a real-world applicable performance. <i>This is often lacking in traditional top-down, rote-memorization only teaching.</i>
"Assess" in three ways	There are 3 synergistic ways to assess performance (referred to as the <i>Trinity of Assessment</i>): 1. Self-assessment 2. Peer to peer assessment 3. Teacher-student assessment It is the culmination of, and synergy of, <i>all three</i> that is ideal for nurturing students' metacognitive prowess, problem-solving skills, and overall academic growth.
"Encourage" mistakes; celebrate mistakes	People can and do learn from mistakes. However, in contexts where people are <i>afraid</i> of making mistakes—they are in an extreme disadvantage. The negative context may be detrimental to natural learning processes. An environment where mistakes are treated as special teaching/ learning moments that are <i>essential to learning</i> may lead to more neuro-efficient cognitive development.

Figure 13 *The “Best Practices” List Excerpt (Murphy, 2019).*

pre-test scores as covariates. The adjusted means confirmed a significant difference: “old” ($M = 10.09, SE = .04$) vs. “new” ($M = 10.21, SE = .03$); $F(1, 51) = 4.70, p = .035$, partial $\eta^2 = .08$; a medium effect size. In sum, both the descriptive and inferential analyses show that the “new” textbook users displayed a significantly accelerated rate of growth in their L2 output compared to the output of the “old” textbook users. Figure 12 illustrates this difference graphically. The “new” users demonstrated rapid growth over

the 15 weeks, supporting the conclusion that the redesigned materials (fostering metacognitive development) seem to have nurtured growth resulting in approximately 8x the rate of the original materials (Murphy, 2019).

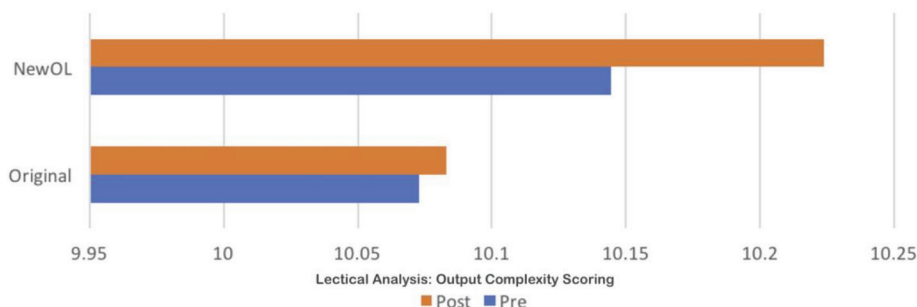


Figure 14 Growth Comparison Using Lectical Analysis (Murphy, 2019).

The “old” textbook focus group discussions became mostly focused on the more tangible aspects of the course. They mostly had discussions of this pattern: “*We did this... and then we did that. I liked this... I did not like that...*” Their feedback was informative, but simplistic. It was neither very deep nor metacognitive.

In comparison, the “new” textbook user focus group consistently jumped to discussions of the *hows* and *whys* of their perceived L2 growth. They did not focus much on *what* we did in class. When I prompted them to discuss certain aspects of the book, they were quick to move on to deeper discussions—they were clearly more interested in discussing the course in more metacognitive self-growth terms. For example: “*I now understand why Section 4.1 was designed this way. It helped us connect with what we had in Section 3.6... This cycle of learning was helpful for us to think more deeply about this topic in English.*”

Figure 15 provides sample feedback from the “new” user group. I requested written feedback from the entire class (not a focus group session activity). I did not specifically request for any diagrams. Interestingly, several students impulsively provided diagrams (such as the one shown in Figure 15). Here, we can see that the student desires to express how the course was designed as a “positive spiral” and has listed what “got better” because of the pedagogical design.

Suppose the textbooks were indeed the decisive factor behind both the accelerated growth rate and the difference in metacognitive maturity; however, the books were similar in design. In that case, the contrast suggests that the “new” book somehow cultivated a deeper self-awareness of L2 development than the “old” book did. Students engaged with their learning at a level that encouraged them to articulate the *hows* and *whys* of their growth, whereas the “old” book offered little support for such reflection.

In retrospect, the “new” book indeed did a better job of nurturing metacognitive development, and my working hypothesis is that this was the critical element—the “secret sauce”—underlying the eight-fold improvement in growth. Students developed

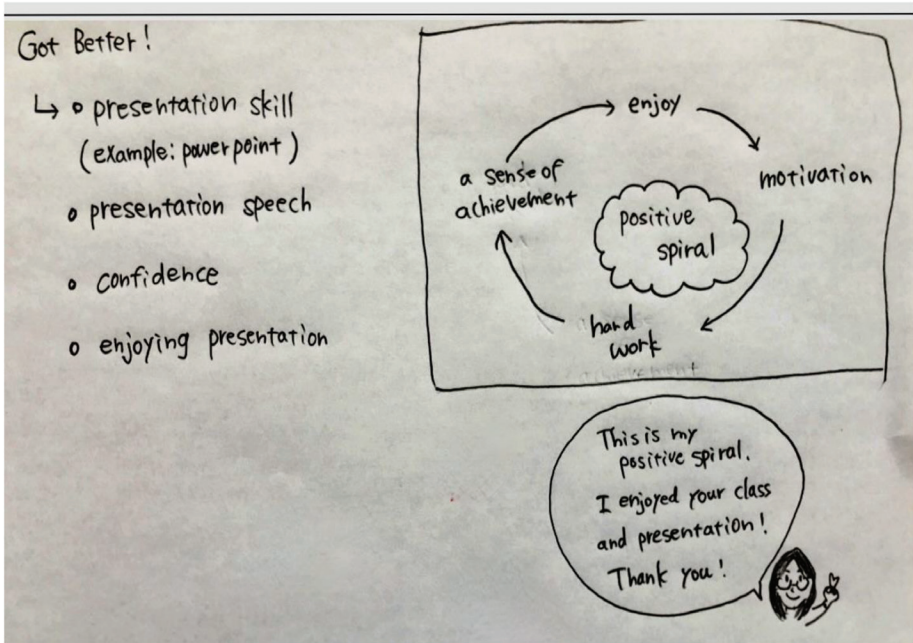


Figure 15 *Example Feedback from a Student in the “New” Textbook Class.*

Note. The feedback is from an EFL course, but it looks more like feedback from a teacher training theory course. Consequently, this “new” textbook class achieved 8x faster average growth (in English output complexity) than the students using the “old” textbook. In retrospect, the “old” textbook did not nurture pedagogical/metacognitive awareness (Murphy, 2019).

an appreciation for the pedagogical design itself, which appears to have amplified motivation and consolidation in a positive spiral, as the student in Figure 13 noted. Although further research is required to confirm this interpretation, the evidence suggests an important implication: when students are involved in the pedagogical process and understand the reasoning behind design choices, their engagement and outcomes may increase considerably.

Beyond the Rules: Transitioning from Rote to Convergent Pedagogy

To make the pedagogical shift discussed in this paper more concrete, Table 2 contrasts two lesson arcs. The first is a common Japanese “orthodox” approach organized around discrete lexical items and assessment cycles. The second is a “Teaching for the DATC” arc that integrates TRIDENT, CREAME, and the 3D:CG assessment framework.

Students using the “Teaching for the DATC” model reported stronger recall, more nuanced understanding, and greater willingness to use new vocabulary in discussions and projects (Murphy, 2019). Teacher feedback likewise pointed to sharper metacognitive insights, particularly increased awareness of how and why words were learned, as well as greater transfer of learning beyond immediate assessment contexts. At the same time, commonly cited drawbacks were that the DATC group’s vocabulary list

Table 2 Comparing an Orthodox Lesson Arc and the “Teaching for the DATC” Arc

Lesson Phase	Orthodox (Rote) Vocabulary Lesson Arc	“Teaching for the DATC” Lesson Arc (TRIDENT, CREAME, 3D:CG)
1. <i>Initial provision</i>	Students receive a fixed list of lexical items, typically selected to align with textbook units, upcoming assessments, or curriculum pacing requirements.	Students explore TRIDENT’s zones as a shared lexical “starter pack” that functions as an entry point into deeper understanding of how words can function in multiple ways.
2. <i>Task framing</i>	Learners prepare to memorize meanings and spellings for a scheduled quiz, with success framed primarily in terms of short-term accuracy and recall under test conditions.	Learners select core words of interest (e.g., <i>journey</i> , <i>achievement</i> , <i>community</i>) and begin trialing them across the various TRIDENT zones.
3. <i>Meaning development</i>	Meaning work centers on repetition, gap-filling, spelling checks, and worksheet-based translation, often emphasizing form–meaning pairings that can be efficiently practiced and checked within limited class time.	Groups use the TRIDENT framework to explore literal and figurative variations, cultural equivalents, and analogous colloquial expressions, mapping discoveries across the TRIDENT axes. New lexical items typically emerge through student-led exploration.
4. <i>Cultural and dialectal input</i>	Vocabulary is treated as standardized and largely context-independent, with priority given to test relevance.	Learners contribute examples from their local dialects and cultural contexts, sharing overlaps and distinctions across groups.
5. <i>Conceptual expansion (CREAME)</i>	Limited time and few structured opportunities for personal association with the lexical items, as instructional focus remains on wide coverage and accuracy.	Learners create mind maps connecting lexical items to personal experiences, emotions, and cultural narratives, working individually or collaboratively.
6. <i>Expressive consolidation</i>	Written practice focuses on controlled sentences or short answers, supporting accuracy and error correction. (Focus on form)	Learners compose micro-narratives, often accompanied by doodling, that portray emotional or societal perspectives related to the lexical items, according to the learner.
7. <i>Knowledge sharing</i>	Knowledge sharing is not typical before or after testing, as instructional time is usually allocated to preparation, review, or assessment cycles.	Learners share discoveries through short presentations or reflective essays, including vocabulary uncovered through exploration.

(Continued)

(Continued)

Lesson Phase	Orthodox (Rote) Vocabulary Lesson Arc	“Teaching for the DATC” Lesson Arc (TRIDENT, CREAME, 3D:CG)
8. <i>Evaluation focus (3D:CG)</i>	Assessment centers on matching or translation accuracy, with performance evaluated against predetermined answer keys or scoring rubrics.	Learners and teachers reflect on the outcomes based on the teachings’ <i>Duration, Engagement, and Applicability</i> .
9. <i>Documentation</i>	Scores are recorded as test results and used for grading, progress tracking, or institutional reporting.	Learners and teachers record evaluations using the 3D:CG rubric and add metacognitive notes to deepen understanding of those outcomes.
10. <i>Metacognitive reflection</i>	Structured reflection on the learning process is typically limited, with feedback delivered through scores, corrections, or brief commentary.	Learners discuss strengths and areas for refinement based on the emerging 3D:CG forms, refine prior notes, and produce a brief summary of their findings.

functioned “merely” as a springboard for discovery and that learning outcomes were not tightly controlled, which limits the usefulness of a “Teaching for the DATC pedagogy” for standardized vocabulary tests (“*We cannot teach to the test with this method...*”). This criticism is, of course, completely valid. As with my depiction of Ausubel’s work earlier (see Figure 5), the two goals are on opposite poles. However, a substantial body of research does show that lexical knowledge developed primarily for short-term assessment tends to fade quickly once testing pressure is removed (Laufer & Nation, 1995; Nation, 2013). From a developmental perspective, lexical knowledge often fails to stabilize into fluent, readily accessible use and remains insufficiently automatized to support long-term retention or transfer beyond the testing context (Segalowitz, 2010). If the vocabulary taught for those tests will be quickly forgotten, we have an interesting conundrum. The larger pedagogical and philosophical questions that surface are: What is the true goal of the course? What is the true goal of education? Do we aim for the bottom-left quadrant of Figure 5 (Ausubel), or the upper-right “sweet spot” quadrant associated with durable, transferable lexical development?

Conclusion: Why is Vocabulary Hard, and How Can it be Resolved?

These questions have shaped my work for more than three decades. Over the years, I have returned to them repeatedly across classrooms, contexts, and research traditions, and the still-developing answers have meandered quite a bit over the years.

Why is vocabulary difficult? My answer today is that vocabulary is difficult because words live at the intersection of multiple developmental systems. Each lexical item carries semantic content, emotional weight, social positioning, and contextual constraints, all of which mature under different conditions and at different rates. In my experience,

learners work to coordinate meaning, form, usage, interpersonal stance, and situational fit simultaneously, often while navigating *uncertainty* regarding how words function across these contexts. Vocabulary growth, therefore, develops unevenly. It shows periods of expansion, stagnation, and episodes of reorganization as support conditions change over time. This is a pattern well documented by Dynamic Skill Theory. Thus, vocabulary learning and use, like any other skill, is always in motion. Vocabulary is commonly presented as having concrete meaning, but this is an illusion. Lexical items reorganize in response to internal states and external contexts, which is what makes them so hard to pin down—the irony of vocabulary is that it defies definition.

How can this irony be resolved? Over the decades, my answer to this question has come to this: the resolution is embedded in pedagogical designs that align with developmental reality. What does that mean? Effective classroom-based vocabulary pedagogy converges across multiple brain domains, engages learners at a visceral level, and sustains motivation for use by actively involving students in the learning process itself. When instruction supports repeated, meaningful use of vocabulary across changing contexts, lexical items gradually stabilize through real experience. Learners develop the capacity to carry functional lexical resources forward, adapt them in real time, and rely on related supporting networks as tools for meaning-making and effective communication in everyday life. I have found that pragmatically supportive networks emerge when learners' developmental insights are paired with purpose-oriented pedagogical design. Frameworks such as DATC, CREAME, the 3D:CG assessment system, and the NeuroELT maxims have provided concrete ways to translate theory into classroom practice, allowing vocabulary teaching to be revisited, evaluated, and refined over the years. These tools have proven useful in supporting growth in both teachers and learners by making visible convergences across cognitive, emotional, and social domains while remaining attentive to changes in contextual support. Thus, for me, the resolution of the irony of vocabulary is to treat words as dynamic resources that exist within dynamic systems and not to be duped by the illusion of concrete definitions. As a teacher, it would therefore be my job to invite learners to take an active part in the legacy and in the growth of our ever-evolving system of languages.

References

- Ausubel, D. P. (1977). The facilitation of meaningful verbal learning in the classroom. *Educational Psychologist*, 12(2), 162–178. <https://doi.org/10.1080/00461527709529171>
- Coady, J., & Huckin, T. (Eds.). (1997). *Second language vocabulary acquisition*. Cambridge University Press.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3), 268–294. <https://doi.org/10.1037/0096-3445.104.3.268>
- Dörnyei, Z., Ibrahim, Z., & Muir, C. (2014). Directed motivational currents: Regulating complex dynamic systems through motivational surges. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 95–105). Multilingual Matters.

- Ellis, N. C. (2008). Usage-based and form-focused language acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 372–405). Routledge.
- Fischer, K. W. (2008). Dynamic skill theory. In R. M. Lerner (Ed.), *Handbook of child psychology* (6th ed., Vol. 1, pp. 313–399). Wiley.
- Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action, thought, and emotion. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (6th ed., Vol. 1, pp. 313–399). Wiley.
- Fischer, K. W., & Rose, S. P. (1998). Growth cycles of brain and mind. *Educational Leadership*, 56(3), 56–60.
- Friederici, A. D. (2011). The brain basis of language processing. *Annual Review of Neuroscience*, 34, 217–243. <https://doi.org/10.1146/annurev-neuro-061010-113830>
- Goswami, U. (2006). Neuroscience and education: From research to practice? *Nature Reviews Neuroscience*, 7, 406–413. <https://doi.org/10.1038/nrn1907>
- Immordino-Yang, M.H. & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education*, 1(1), 3–10. <https://doi.org/10.1111/j.1751-228X.2007.00004.x>
- Immordino-Yang, M. H. (2016). *Emotions, learning, and the brain: Exploring the educational implications of affective neuroscience*. W. W. Norton & Company.
- Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language*, 67(1), 17–29. <https://doi.org/10.1016/j.jml.2012.02.004>
- Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron*, 67(5), 713–727. <https://doi.org/10.1016/j.neuron.2010.08.038>
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). Macmillan.
- Laufer, B., & Nation, I. S. P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–322. <https://doi.org/10.1093/applin/16.3.307>
- Lightbown, P. M., & Spada, N. (2013). *How languages are learned* (4th ed.). Oxford University Press.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253. <https://doi.org/10.1037/0033-295X.98.2.224>
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjær, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35–53). Cambridge University Press.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters.
- Murphy, R.S. (2008). *Models for EFL theory and methodology derived from an SiR-based pilot study on Japanese cognitive development* (Master's thesis, University of Birmingham). https://www.researchgate.net/profile/Robert-Murphy-42/publication/237754434_MODELS_FOR_EFL_THEORY_AND_METHODODOLOGY_

- DERIVED_FROM_AN_SiR-BASED_PILOT_STUDY_ON_JAPANESE_COGNITIVE_DEVELOPMENT/links/542d29360cf27e39fa941d78/MODELS-FOR-EFL-THEORY-AND-METHODOLOGY-DERIVED-FROM-AN-SiR-BASED-PILOT-STUDY-ON-JAPANESE-COGNITIVE-DEVELOPMENT.pdf
- Murphy, R. S. (2010). Learner training: Getting intimate with English. *The Language Teacher*, 34(6), 21–27.
- Murphy, R. S. (2013). TRIDENT: A triangular denary system for translation and lexical analysis. *University of Kitakyushu Fundamental Education Center Bulletin*, 17, 119–162.
- Murphy, R. S. (2015). How can neuroscience inform language teaching? In C. Stillwell (Ed.), *Language teaching insights from other fields* (pp. 99–111). TESOL Press.
- Murphy, R. S. (2017). 3D:CG (Constant Gradient) evaluation system: New rubrics for the evaluation of teaching and professional development programs. *University of Kitakyushu Fundamental Education Center Bulletin*, 28, 77–95.
- Murphy, R. S. (2019). *EFL coursebook innovation: The dynamics of L2 maturity, student engagement, and cognitive development* (Doctoral dissertation, University of Nottingham). <http://eprints.nottingham.ac.uk/59050>
- Murphy, R. S. (2023). Revisiting the maxims of NeuroELT: A new subgrouped order and improved explanations. *University of Kitakyushu Fundamental Education Center Bulletin*, 40, 27–96.
- Murphy, R. S. (in press). Dynamic skill theory and uncertainty. In A. H. Al-Hoorie & T. Gregersen (Eds.), *The motivation of uncertainty*. Applied Linguistics Press.
- Murphy, R. S., & Sin, H. H. Y. (2014). Neuroscientific benefits of including culture and emotion in the second language classroom: Teaching for the DATC. *University of Kitakyushu Fundamental Education Center Bulletin*, 20, 53–69.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Norton, B. (2013). *Identity and language learning* (2nd ed.). Multilingual Matters.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.
- Sousa, D. A. (2017). *How the brain learns: A classroom teacher's guide* (5th ed.). Corwin Press.
- Tokuhamma-Espinosa, T. (2011). *Mind, brain, and education science: A comprehensive guide to the new brain-based teaching*. W. W. Norton & Company.
- Tokuhamma-Espinosa, T. (2018). *The new science of teaching and learning: Using the best of mind, brain, and education science in the classroom*. W. W. Norton & Company.
- Tweed, R. G., & Lehman, D. R. (2002). Learning considered within a cultural context. *American Psychologist*, 57(2), 89–99. <https://doi.org/10.1037/0003-066X.57.2.89>
- Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*, 8, Article 1454. <https://doi.org/10.3389/fpsyg.2017.01454>
- Vitta, J. P., Nicklin, C., & Albright, S. W. (2023). Academic word difficulty and multidimensional lexical sophistication. *The Modern Language Journal*, 107(1), 373–397. <https://doi.org/10.1111/modl.12835>