languages

MDPI

Article

Derivational Morphology in L2 English: Investigating the Role of Affixal Neutrality Through the Lens of Linguistic Theory

Xingcheng Wang and Helen Zhao *

School of Languages and Linguistics, The University of Melbourne, Melbourne 3010, Australia; alex.wang2@unimelb.edu.au
* Correspondence: helen.zhao@unimelb.edu.au

Abstract

This study investigates how second language (L2) learners acquire morphologically complex English words, focusing on affixal neutrality—whether suffixes preserve the phonological form and semantic transparency of the base (e.g., *-ness* in *happiness*) or trigger phonological/orthographic changes (e.g., *-ity* in *activity*). Drawing on linguistic theories of morphological decomposition and lexical representation, we examine how this property influences different dimensions of derivational knowledge. Fifty-four Mandarin-speaking secondary school EFL learners completed three receptive tasks targeting relational knowledge (morphological relatedness), syntactic knowledge (category awareness), and distributional knowledge (contextual appropriateness). Lexical items varied in affixal neutrality, and participants' accuracy and response times were analysed across three L2 proficiency levels. Affixal neutrality significantly affected performance in the relational knowledge task, with neutral suffixes facilitating accuracy and faster responses. Effects were attenuated in syntactic and distributional tasks, suggesting domain-specific sensitivity to neutrality. L2 Proficiency was associated with higher accuracy across all three domains but did not substantially affect processing speed. These findings highlight the selective role of a theoretically motivated morphological property in L2 lexical acquisition and show how linguistic concepts such as affixal neutrality can form the basis of targeted hypotheses, bridging theoretical linguistics and empirical research in second language learning.

Keywords: morphological awareness; derivational knowledge; affixal neutrality; second language acquisition

1. Introduction

Acquiring a new morphological system is challenging yet cognitively efficient: language users can understand and productively use morphologically complex words through decomposing a limited number of stems and affixes. In first language (L1) acquisition of English, derivational morphology has been shown to function as a potent metacognitive resource for vocabulary growth and reading development (e.g., Carlisle & Fleming, 2003; Jiang & Kuo, 2019; Kieffer & Lesaux, 2008, 2012; Nagy et al., 2006). Research in linguistics and psycholinguistics has also examined how morphologically complex words are processed and represented in the L2 mental lexicon. However, much less is known about the development of L2 derivational knowledge itself, and particularly about how different aspects of this knowledge interact with specific morphological properties during L2 processing. Existing studies have documented difficulties L2 learners face in acquiring derivational knowledge (e.g., Schmitt & Meara, 1997; Schmitt & Zimmerman, 2002), yet the

Academic Editor: Alessandro Benati

Received: 5 November 2025
Revised: 4 February 2026
Accepted: 24 February 2026
Published: 5 March 2026

complexity of this knowledge is often underrepresented, as it is frequently subsumed under broader constructs such as morphological awareness (Apel, 2014) or examined primarily in relation to reading and literacy development (e.g., Zhang & Koda, 2013).

One key theoretical concept with a longstanding influence in morphological theory and L1 morpheme acquisition research, but rarely applied in the L2 context, is *affixal neutrality*. Affixal neutrality refers to whether suffixation preserves the base's surface form (Chomsky & Halle, 1968). In this study, neutral suffixation is defined as suffixation that preserves the base's segmental form and does not induce stress shift or major orthographic alternation (e.g., happy → happiness), whereas non-neutral suffixation is defined as suffixation that alters the base's surface form via phonological change, stress shift, and/or orthographic modification (e.g., active → activity).

Incorporating this concept into L2 research offers a theoretically grounded means of examining how inherent morphological properties interact with learners' proficiency to shape their relational, syntactic, and distributional knowledge of derivational morphology (Tyler & Nagy, 1989). In this light, the present study undertakes a refined analysis of affixal neutrality in L2 English derivational morphology, exploring its role in the acquisition and processing of morphologically complex words.

2. Literature Review

2.1. Derivational Affixes: Neutral vs. Non-Neutral

English derivational affixes can be organized into two broad classes[1] according to their morphological and phonological behaviour: neutral and non-neutral (Chomsky & Halle, 1968; Halle & Mohanan, 1985; Spencer, 1991). Neutral affixes, predominantly of Germanic origin (e.g., *-ness*, *-ful*, *-less*, *-ly*), typically attach to independent words. For example, when *-less* is removed from *colourless*, the remaining base (*colour*) is itself a free morpheme. Moreover, neutral affixation generally does not induce significant orthographic or phonological alternations in the base (e.g., *home-homeless*, *serious-seriousness*).

By contrast, non-neutral affixes, commonly of Latinate origin (such as *-ity*, *-ify*, *-ion*, *-ic*), typically attach to bound morphemes. Removing *-ify* from *qualify* or *quantify*, for instance, does not yield a free morpheme (**qual*, **quant*). Crucially, non-neutral affixation is often accompanied by phonological and/or orthographic alterations. These include stress shifting (*situate* /ˈsɪtʃʊæɪt/ → *situation* /sɪtʃʊˈæɪʃən/), vowel shifting (i.e., from *divide* /dɪˈvɑed/ to *division* /dɪˈvɪʒən/, second vowel /ɑe/ moves toward first vowel /ɪ/), and consonantal change (the final consonant /d/ → /ʒ/ in *divide-division*). Additional processes such as resyllabification (i.e., the coda in /dɪ.ˈvɑed/ becomes an onset in /dɪ.ˈvɪ.ʒən/, '.' marks the boundary of syllables) and orthographic modification (*consume* → *consumption*) are also common.

Neutral and non-neutral affixes further differ in their degree of productivity (Katamba & Stonham, 2006; Lieber, 2021; Plag & Baayen, 2009; Tyler & Nagy, 1989). Neutral suffixes are generally highly productive, as they can attach to a wide range of stems. For instance, *-er* can readily combine with numerous verbs to form agentive nouns (e.g., *teach* → *teacher*, *run* → *runner*). In contrast, non-neutral affixes such as *-tion* and *-ence* are more restricted, typically combining only with specific bound bases (e.g., *perceive-perception*). Consequently, neutral affixes tend to be more productive in forming new complex words (Booij, 2016). Empirical evidence also indicates that English-speaking children acquire neutral affixes earlier and demonstrate greater metalinguistic awareness of them than of non-neutral affixes (Carlisle, 2000).

In the literature, several partially overlapping terms are used to characterise derived words: affixal neutrality (neutral vs. non-neutral), morphological/semantic transparency (transparent vs. opaque), and the presence of base-form alternations (phonological and/or

orthographic change, often including stress shift). These terms are related but not interchangeable. Affixal neutrality is the theoretical construct adopted here, operationalised as a property of the derivational process. We use "transparent/opaque" only as a descriptive label for semantic transparency, which often correlates with neutrality but is not equivalent to it: neutral derivatives are often more semantically transparent, yet transparency varies item-by-item. Likewise, "phonological/orthographic change" refers to the surface diagnostic used to classify items as neutral vs. non-neutral in our stimulus set (with non-neutral items further subtyped as orthographic change, phonological/stress change, or both). For clarity, in the remainder of the paper we primarily use the terms neutral vs. non-neutral when referring to the manipulated factor, and reserve "transparency" for semantic interpretation when relevant.

In the following section, we review different aspects of derivational knowledge and examine how affixal neutrality has been addressed in each. Drawing on Tyler and Nagy's (1989) influential framework, derivational knowledge is operationalised in terms of three dimensions: *relational*, *syntactic*, and *distributional* knowledge.

2.2. Relational, Syntactic, and Distributional Knowledge

Relational knowledge refers to the ability to recognize the morphological and semantic relationships among words that share a common lexical base (e.g., *employ-employer* as opposed to *corn-corner*) (Apel, 2014; Tyler & Nagy, 1989). Among the three dimensions of derivational knowledge, relational knowledge has received more empirical attention. One approach to measuring it involves asking learners to judge whether a base word and a derived form are morphologically related. Research on L1 acquisition has shown that native English-speaking children begin to demonstrate emerging awareness of morphological relations, specifically, the ability to identify stems within derived words, by around the fourth Grade (e.g., Carlisle, 2000; Mahony et al., 2000; McBride-Chang et al., 2005; Nagy et al., 2006; Sparks & Deacon, 2015). Tyler and Nagy (1989) observed that children in primary-school Grades 4, 6, and 8 exhibited progressively greater sensitivity to morphological relatedness, particularly for words containing neutral suffixes compared to non-neutral ones. However, they argued that this difference might reflect broader development in reading ability, vocabulary knowledge, and test-taking strategies rather than genuine growth in knowledge of affixal neutrality. More generally, the ease with which learners perceive morphological relatedness is partly influenced by affixal neutrality: words formed with neutral affixes tend to be more semantically transparent than those with non-neutral affixes (Diependaele et al., 2009; J. Li & Taft, 2020).

Syntactic knowledge refers to the knowledge of the syntactic category that a derivational suffix assigns to its base (Lieber, 2021; Tyler & Nagy, 1989). For example, a learner can infer the part of speech of a derived form (e.g., *reflection*, *reflective*) without necessarily knowing the grammatical category or meaning of its base (e.g., *reflect*), since the suffixes *-ion* and *-ive* signal nominal and adjectival functions, respectively. Empirical findings concerning the role of affixal neutrality in syntactic knowledge are mixed. Some studies have reported significant effects of neutrality (Carlisle, 2000; Singson et al., 2000), whereas others have not (Tyler & Nagy, 1989). For instance, Singson et al. (2000) employed sentence completion and grammaticality judgment tasks to assess children's syntactic knowledge of derived words, including both real words and pseudowords, among learners from Grades 3 to 6. Their findings revealed age-related improvement and consistently higher accuracy for real words than for pseudowords. Notably, pseudowords containing neutral suffixes yielded higher accuracy than those with non-neutral suffixes, suggesting that the morphological transparency of neutral affixes facilitates the access to syntactic category information.

Distributional knowledge refers to an individual's awareness of the restrictions governing how stems and suffixes can concatenate (Tyler & Nagy, 1989). A given derivational suffix can attach only to stems that meet specific morphological, semantic, phonological, orthographic, or syntactic requirements (Dixon, 2014; Katamba & Stonham, 2006; Plag & Baayen, 2009). For instance, *un-* attaches only to free words rather than bound morphemes (a morphological constraint) and typically denotes opposition between gradable adjectives (e.g., *happy-unhappy*), but not between contradictory adjectives (e.g., *artificial-genuine*) (a semantic constraint). The suffix *-en* forms verbs from adjectives ending in obstruents (e.g., *blacken*, *loosen*), reflecting a phonological constraint, while *-ness* attaches to adjectives (e.g., *kind*) but not to verbs (e.g., *fight*), reflecting a syntactic constraint. Similarly, *-ic* adds a final consonant to *metal* to form *metallic*, demonstrating both orthographic and phonological constraints. Although derivational suffixes generally follow systematic and predictable patterns, their productivity varies considerably due to the influence of these constraints.

Using a 2 × 2 factorial design crossing neutrality (neutral vs. non-neutral) and formedness (well-formed vs. pseudo-formed), Tyler and Nagy (1989) examined children's distributional knowledge of derivatives. Well-formed derivatives were defined as forms that conform to conventional distributional constraints (e.g., *dreamful*), whereas pseudo-formed derivatives violate these concatenation rules (e.g., **harshful*, which breaches a syntactic constraint). Distributional knowledge was assessed using an untimed lexical decision task in which participants indicated whether they knew the meaning of each stimulus. The results showed that children could differentiate conventionally formed derivatives from pseudo-formed ones, and that items containing neutral suffixes were judged as more acceptable than those with non-neutral suffixes. Importantly, an interaction between formedness and neutrality emerged, suggesting that affixal neutrality modulates learners' sensitivity to morphological well-formedness.

Among the three dimensions of derivational knowledge, distributional knowledge appears to be the most complex, as it requires sensitivity to multiple properties of both the stem and suffixes and to the intricate constraints governing their combination. Consequently, this dimension tends to emerge later in L1 acquisition than relational and syntactic knowledge (Tyler & Nagy, 1989).

2.3. Derivational Knowledge in L2 Learning

Traditionally, L2 studies on derivational knowledge have examined it as a component of L2 Morphological Awareness (MA) and explored its relationship with L2 reading ability. Studies have consistently shown that L2 MA is moderately correlated with intralingual measures of word decoding (see Ke et al., 2021, for a review) and that it makes a unique contribution to L2 literacy development (Kieffer & Lesaux, 2008, 2012) and reading comprehension (Wu & Juffs, 2022). For example, Kieffer and Lesaux (2008) reported substantial improvement in Spanish-speaking English learners' performance from Grade 4 to Grade 5 on a derivative decomposition task (Carlisle, 2000). Their results further demonstrated that derivational knowledge significantly predicted reading comprehension, even after controlling for word reading skill, vocabulary knowledge, and phonological awareness. Similarly, Zhang (2017) found that Chinese ESL learners showed considerable growth in English derivational awareness over the course of one academic year, and that L2 MA significantly contributed to their reading development.

In contrast to the extensive literature linking L2 MA to reading outcomes, relatively few studies have directly examined the componential nature of L2 derivational knowledge, specifically its relational, syntactic, and distributional dimensions. Studies that have addressed these dimensions (e.g., Ramirez et al., 2010; Schmitt & Meara, 1997; Schmitt

& Zimmerman, 2002; Zhang, 2017; Zhang & Koda, 2012, 2013) generally report positive associations between L2 learners' relational and syntactic knowledge of derivational morphology. Zhang and Koda (2012), for example, developed a base morpheme identification task to assess Chinese EFL learners' knowledge of morphological relatedness (e.g., identify *forest* in *reforestation*), alongside a lexical inferencing task designed to measure syntactic knowledge of derivatives (i.e., inferring meaning based on knowledge of the suffixal syntactic categories). They found a moderate correlation between the two constructs (r = 0.44), although a subsequent study with younger Chinese EFL learners yielded a weaker correlation (r = 0.29) (Zhang & Koda, 2013). Hayashi and Murphy (2011), however, reported counterevidence, finding that Japanese ESL learners performed considerably better on a receptive test of morphological relatedness than on a productive test of syntactic knowledge, with no significant correlation between the two dimensions. Despite these insights, the distributional aspect of L2 derivational knowledge remains largely unexplored, leaving an important gap in understanding how learners acquire and internalize the constraints governing stem–affix concatenation.

Beyond these componential distinctions, there is even a great dearth of empirical research examining how specific morphological properties, such as affixal neutrality, shape different aspects of L2 derivational knowledge. Existing findings on how affixal neutrality influences learners' derivational processing remain mixed. To et al. (2016), using derivation and decomposition tasks adapted from Carlisle (2000), found that L2 learners showed significantly lower accuracy and slower response for derived forms involving phonological change (e.g., *admit-admission*) compared to those without such change (e.g., *enjoy-enjoyment*). In contrast, Jiang and Kuo (2019) found no significant effect of phonological or orthographic change in a multiple-choice word association task (Kuo et al., 2017), where stimuli varied in lexical frequency and morphological complexity, although potential alternation effects were not examined in their syntactic knowledge component. These mixed results suggest that the impact of surface form variation may depend on the type of task and the aspect of derivational knowledge being assessed.

Crucially, insights from L2 lexical decision research further suggest that morphologically well-formed pseudowords can pose particular challenges for learners, especially when they involve productive and salient suffixes. Studies have shown that pseudowords composed of real stems and real affixes are often harder to reject and elicit longer response times than less morphologically structured nonwords, reflecting increased word-likeness and decision uncertainty. For example, Amenta et al. (2025) demonstrated that adult L2 learners exploit morphological information productively when processing unfamiliar strings, and that the presence of a real suffix substantially increases processing cost in lexical decision. This difficulty has been attributed to competition between decompositional analysis and lexical rejection processes rather than to simple orthographic familiarity. Such findings underscore the importance of examining distributional knowledge and provide a strong rationale for investigating how affixal neutrality, which is closely tied to productivity and morphological salience, modulates L2 learners' judgments of derived word well-formedness across different task contexts.

3. The Current Study

While previous research has provided valuable insights into the contribution of L2 learners' derivational knowledge to reading, very few studies have systematically examined all three core aspects of derivational knowledge (relational, syntactic, and distributional) within the same empirical design (Zhang, 2017). Existing L2 studies have tended to focus on relational and syntactic knowledge, leaving the distributional aspect largely unexplored, despite its relevance to more advanced forms of morphological awareness. Moreover,

the theoretical linguistic concept of affixal neutrality, as a long-standing construct in morphological theory and in L1 morpheme acquisition research, has rarely been applied to understanding L2 morphological acquisition. In theoretical morphology, affixal neutrality is central to accounts of morphological decomposition, lexical storage, and productivity; however, its potential role in shaping L2 morphological knowledge remains underexplored. The present study addresses this gap by using affixal neutrality as the basis for the formulation of specific research questions and hypotheses, thereby aligning with the Special Issue's aim of grounding L2 research in linguistic theory.

Methodologically, this study also addresses limitations of prior L2 research by employing both well-formed and pseudo-formed derivatives (e.g., Ramirez et al., 2010). This allows us to examine derivational knowledge unconfounded by lexical familiarity and frequency, while capturing how authenticity of lexical contexts interacts with morphological knowledge across proficiency levels. We assess relational, syntactic, and distributional knowledge through three receptive tasks, analysing both accuracy and response times (RTs) to capture differences in knowledge depth and processing efficiency. Based on the literature reviewed, we formulate the following hypotheses (H):

H1 (affixal neutrality effect by task)**.** *Affixal neutrality will facilitate L2 learners' performance across all three derivational knowledge tasks, such that derivatives with neutral suffixes will generally elicit higher accuracy and/or faster RTs than those with non-neutral suffixes. However, the magnitude of this effect is expected to vary by task, being strongest in the relational knowledge task, where surface transparency directly supports base–derivative mapping, and weaker in the syntactic and distributional knowledge tasks, which rely more heavily on grammatical cues and constraint-based or lexical decision processes.*

H2 (formedness effect)**.** *Well-formed derivatives will yield higher accuracy and faster RTs than pseudo-formed derivatives in all tasks, reflecting the influence of conventional morphological relations and lexical well-formedness.*

H3 (neutrality × formedness interaction)**.** *The advantage of neutral suffixes will be particularly pronounced for pseudo-formed items, as greater transparency may help learners evaluate unfamiliar or ill-formed derivatives more effectively than non-neutral suffixes.*

H4 (proficiency effect)**.** *Higher-proficiency learners will outperform lower-proficiency learners across all tasks, with the greatest gains in the more complex distributional knowledge task.*

4. Methodology

4.1. Participants

Fifty-eight Chinese-L1 learners of English participated in the study; however, four participants were excluded due to incomplete task data. The participants were recruited from three private secondary schools in China and comprised 28 males and 26 females, with a mean age of 16 years. Eighteen participants were enrolled in the Grade 10, twenty-eight in Grade 11, and eight in Grade 12. Language proficiency was indexed by participants' IELTS Academic scores. The mean overall score was 5.7 (SD = 0.58), with mean reading and writing scores of 5.8 (SD = 0.77) and 5.4 (SD = 0.49), respectively. These correspond to CEFR B1-B2.

Information regarding participants' language background was collected using a simplified Chinese version of the Language History Questionnaire 3.0 (P. Li et al., 2020). On average, participants began learning English at age 6 (SD = 1.46), with the earliest start at age 3 and the latest at age 9. They reported using Mandarin (L1) for over 75% of their daily

communication (SD = 0.12), while English (L2) was used primarily during English classes or when completing English homework (approximately 25%, SD = 0.15).

4.2. Research Design

Affixal neutrality, derivational formedness, and L2 proficiency were factored in to investigate learners' multifaceted derivational knowledge. Neutral suffixes were operationalised as those that produce minimal orthographic or phonological changes to the stem with which they combine, whereas non-neutral suffixes were defined as those that alter the consonantal or vocalic segments, induce stress shift in the stem, and/or trigger orthographic modifications (Katamba & Stonham, 2006). Non-neutral suffixed items were further subcategorised into three subtypes: *orthographic change*, *phonological change*, and *combined orthographic–phonological change*.

Following Tyler and Nagy's (1989) framework, derivational knowledge was operationalized in terms of three dimensions: relational, syntactic, and distributional knowledge. The distinction between well-formed and pseudo-formed derivational relations was determined by whether the base and its derivative constituted a conventionalized and linguistically valid pairing under each dimension of derivational knowledge.

Stimuli

Suffixes—The current study examined derivational suffixes only. The selection of suffixes drew on Bauer and Nation's (1993) graded list of English affixes, which classifies affixes into seven hierarchical levels. These levels were established according to several factors, including frequency, productivity, predictability of derived meaning, regularity of the base's written and spoken forms, regularity of the affix's spelling and pronunciation, and consistency of function. The current study focused on frequent derivational suffixes at Level 3 (the most frequent and regular derivational suffixes), Level 4 (frequent, orthographically regular suffixes), and Level 6 (frequent but irregular suffixes). Level 1, 2, 5, and 7 were excluded because Level 1 treats each derived form as a separate word (e.g., *active* vs. *activity*), whereas Levels 2, 5, and 7 include inflectional suffixes, infrequent affixes, and classical roots and affixes, respectively.

The study targeted eight derivational suffixes in total, comprising four neutral suffixes and four non-neutral suffixes (Table 1). To examine the syntactic and distributional knowledge of suffixes, suffixes carrying different parts of speech were recruited, including nominal suffixes (e.g., *-ness*, *-ity*) and adjectival suffixes (e.g., *-ful*, *-ic*) in both neutral and non-neutral categories. Verbal suffixes (e.g., *-ify*, *-ize*) were excluded to maintain comparability across tasks and to keep syntactic and distributional judgments within parallel nominal and adjectival contexts. This decision also avoids additional semantic and structural complexity associated with verbal derivation (e.g., organ → organize), thereby reducing cognitive load and isolating the effects of affixal neutrality.

Table 1. Target neutral and non-neutral suffixes.

Neutrality	Parts of Speech	
	Nominal	**Adjectival**
Neutral	*-ness*, *-er*	*-ful*, *-less*
Non-neutral	*-ion*, *-ity*	*-ic*, *-al*

Real words and pseudowords—all real-word stimuli were drawn from the British National Corpus (BNC) word family list. To align with participants' proficiency levels, only items within the first and second 1000 word-frequency bands were used to construct target words

(both base and derived words) and corresponding sentence contexts for the derivational knowledge tests.

To ensure that observed differences between neutral and non-neutral affix conditions were not confounded by unrelated lexical variables, variation in basic lexical properties, such as word length (number of letters and syllables) and word-form frequency, was minimized. Stems and their derived counterparts across the two conditions were selected to be broadly comparable on these properties. In addition, log-transformed word-form frequency (based on the BNC) was included as a covariate in all mixed-effects models for both accuracy and RT analyses, thereby statistically controlling for any residual variation in frequency. The following additional criteria guided the selection of real-word stimuli:

- Each stem and its derived forms belong to different grammatical categories, enabling the assessment of participants' syntactic knowledge and knowledge of distributional constraints.
- Each selected stem had at least three derivative forms, ensuring sufficient answer choices for the syntactic knowledge task.
- Suffixation involved only minor alterations to the base (e.g., *commerce* → *commercial*) to minimize additional cognitive load (Tyler & Nagy, 1989).
- Suffix stacking was avoided. Complex derivatives containing multiple concatenated suffixes (e.g., *commerc-ial-ize*, *commerc-ial-iz-ation*) were excluded to avoid morphological layering that could confound analysis.

Pseudowords were generated using *Wuggy* (Keuleers & Brysbaert, 2010), a pseudoword generator producing phonotactically and orthographically plausible nonwords. The anchor words used for generation were drawn from the first and second 1000 word-frequency bands of the BNC word list (e.g., *slor* comes from *slow*). All pseudowords conformed to English orthographical and phonological rules, and all derived pseudowords adhered to English phonotactic constraints (e.g., *slor-ity*, *pank-ful*, *lind-less*). These words were manually screened to ensure that they are not existing English words and not obvious near-neighbours of existing derivatives, and not accidental plausible derivatives of real bases. See Supplementary Material B for detailed descriptive statistics on the word length, syllable count, and log frequency by neutrality and formedness across all three tasks.

4.3. Instruments

Three instruments were employed to measure learners' receptive relational, syntactic, and distributional knowledge. The dependent variables were L2 learners' accuracy and RTs on the judgment tasks corresponding to these three dimensions. All tasks were programmed and administered online via Qualtrics (2005, Version 12.2020) using a laptop/desktop. Test items were balanced across suffix types (nominal vs. adjectival), neutrality conditions (neutral vs. non-neutral), and formedness (genuine vs. pseudo relation or real words vs. pseudowords). Four practice items with feedback were provided for each task. They were also randomised and presented one at a time on the screen, and Qualtrics automatically recorded both response accuracy and RTs. The complete set of lexical stimuli used in the study is provided in Supplementary Material A.

4.3.1. Relational Knowledge Judgement Task

This task was modelled on the morphological relatedness task developed by Nagy et al. (2006). Participants were asked to judge whether the second word in a pair (e.g., *fighter*) was derived from the first word (e.g., *fight*). Eight pairs represented well-formed derivational relations (e.g., neutral: *aware-awareness* vs. non-neutral: *minor-minority*), whereas the remaining eight pairs represented pseudo-formed derivational relation (e.g., neutral: *cent-center* vs. non-neutral: *polite-politic*), in which the second word was not morphologically

derived from the first. Each formedness category included four pairs with neutral suffixes and four with non-neutral suffixes. The Cronbach's α was 0.72.

4.3.2. Syntactic Knowledge Task

This task was designed as a sentential cloze task, adapted from Nagy et al. (2006). Participants read sentences each containing an empty slot representing a target word, which could be a real (well-formed) or pseudo (pseudo-formed) derivative. For each item, participants selected the correct completion from three derivative options that shared a common base. The sentence context provided only enough information to support a syntactic judgement of parts-of-speech appropriateness. For example, in the pseudo-formed item '*He shared a ___ experience with us*', participants chose among *mounic* (the correct choice due to the adjectival suffix), *mounity* and *mounize*. Only one option fit the sentence context syntactically. Because the target suffixes in this study were nominal and adjectival, the syntactic choices for target items involved only these two parts of speech. To prevent this from becoming predictable, additional filler items containing other suffix types were included to diversify part-of-speech contexts. The task comprised eight real-word target items (four neutral and four non-neutral items) (e.g., neutral: *darkness* vs. non-neutral: *selection*) and eight pseudoword items (also four neutral and four non-neutral) (e.g., neutral: *pankful* vs. non-neutral: *slority*). The task demonstrated good internal consistency, with a Cronbach's $\alpha = 0.82$.

4.3.3. Distributional Knowledge Task

The distributional knowledge task was implemented as a timed lexical decision task (Meyer & Schvaneveldt, 1971). Participants were instructed to decide, as quickly and accurately as possible, whether a visually presented stimulus was a real word (i.e., a well-formed derivative) or a pseudoword (i.e., a pseudo-formed derivative). The task included 16 real-word items (8 neutral and 8 non-neutral) (e.g., neutral: *informer* vs. non-neutral: *equality*) and 16 pseudoword items (also 8 neutral and 8 non-neutral) (e.g., neutral: *metalful* vs. non-neutral: *engagity*). All stimuli were either selected or generated to have a word length of 8 to 10 letters, with an average of approximately 9 letters for both neutral and non-neutral items. The Cronbach's α of the task was 0.83.

4.3.4. Construct Validity, Task Demands, and Item Structure

Although no single task provides a process-pure measure of derivational knowledge, each task in the present study was selected because it places its primary decision demand on one targeted dimension of derivational knowledge (Tyler & Nagy, 1989), while other task demands are acknowledged.

The relational judgement task primarily targets relational knowledge, as participants judge whether two real words stand in a genuine derivational relation (e.g., *aware-awareness*) or only share surface similarity (e.g., *cent-center*). This task necessarily involves orthographic and lexical processing, but successful performance depends on recognising morphological relatedness. The syntactic cloze task primarily targets syntactic knowledge, because all response options share a common base and differ only in suffix, making suffixal category information the critical cue for grammaticality in context. At the same time, the task also involves sentence reading and option comparison. Therefore, response times are interpreted as reflecting sentence-level processing in addition to suffix processing. The distributional knowledge task, implemented as a lexical decision task, targets learners' sensitivity to derivational well-formedness by contrasting attested derivatives with pseudowords formed through plausible but non-conventional stem–suffix combinations. As lexical decision is known to reflect orthographic familiarity and word-likeness, we interpret this task as indexing distributional/lexical well-formedness sensitivity.

All three tasks instantiated a 2 × 2 design crossing Affixal Neutrality (neutral vs. non-neutral) with a task-specific notion of formedness. The operationalisation of formedness necessarily differs by task because each task targets a different aspect of derivational knowledge. In the relational task, the contrast is between genuine derivational relations (e.g., base and derivative) and pseudo-related real-word pairs that share surface overlap but lack a derivational relation. In the syntactic cloze and lexical decision tasks, the contrast is between attested well-formed derivatives and pseudowords created to be orthographically/phonotactically plausible but not legitimate English derivatives.

4.4. Data Analysis

We analysed accuracy (binomial) using logit mixed-effects models[2] (Breslow & Clayton, 1993) and response times (RTs, continuous, log-transformed) using linear mixed-effects models[3]. Separate models were built for each aspect of derivational knowledge (relational, syntactic, and distributional) because the tasks differed in format (relational judgement, multiple-choice selection, lexical decision).

Each model included fixed effects for affixal neutrality, formedness of derivational relation, L2 proficiency, and their interactions. L2 proficiency was *z*-transformed before entry. Participants and items were included as crossed random effects to account for repeated measures and variation across both learners and stimuli (Barr et al., 2013; Linck & Cunnings, 2015).

We also compared baseline mixed-effects models with extended models including log word frequency as an item-level covariate for both accuracy and RT across all three tasks. For accuracy, likelihood-ratio tests indicated that adding log frequency did not significantly improve model fit for the relational knowledge task, $\chi^2(1) = 2.94$, $p = 0.086$, the syntactic task, $\chi^2(1) = 0.43$, $p = 0.513$, or the distributional task, $\chi^2(1) = 0.02$, $p = 0.884$. A comparable pattern was observed for RTs. In the relational knowledge task, the likelihood-ratio test did not reach significance, $\chi^2(1) = 3.74$, $p = 0.053$. Similarly, for the syntactic knowledge task, the inclusion of log frequency did not improve model fit, $\chi^2(1) = 0.01$, $p = 0.932$, and for the distributional knowledge task, the frequency-augmented model again failed to improve fit, $\chi^2(1) = 0.54$, $p = 0.461$. Taken together, these results indicate that lexical frequency does not make a reliable contribution to either accuracy or reaction times once morphological factors and proficiency are considered.

To reduce the influence of extreme values, RTs were log-transformed and trimmed using the 1.5 × interquartile range (IQR) criterion, excluding observations outside Q1 − 1.5 × IQR and Q3 + 1.5 × IQR. RT analyses were restricted to correct responses, and 121 missing data points (1.6%) were excluded. Models were fitted in R 4.0.3 (R Core Team, 2020) using the *glmmTMB* (Brooks et al., 2017) and *lme4* (Bates et al., 2018) packages. We assessed model assumptions (e.g., multicollinearity, residual normality) with *performance* (Lüdecke et al., 2021). Pairwise comparisons were conducted with *emmeans* (Lenth et al., 2023) and adjusted using the Bonferroni method (Haynes, 2013). Model reporting follows Baayen et al. (2008) and Linck and Cunnings (2015). The data are available on the Open Science Framework: https://osf.io/63gu5/?view_only=cc6cbff93f5f43fa9e83b26d46f13fee (accessed on 21 January 2026).

We ran sensitivity checks for our mixed-effects models. Although the number of suffixes we examined was limited, our observation-level number was substantial due to the crossed participants × items design. We leveraged crossed random effects for Participants and Items, which improves generalisation over both sources of variability and typically increases power relative to by-subject or by-item analyses alone, as effects are estimated from thousands of observations rather than from aggregated means. Across tasks, a power analysis (Green & MacLeod, 2016) indicated high power (>0.80) to detect small-to-moderate

fixed effects (odds ratios ≈ 1.4–1.6 for accuracy; ≈ |0.10–0.15| in log-RT units) under our observed random-effect structure.

5. Results

5.1. Descriptive Statistics

As shown in Table 2, the relational knowledge task was the easiest (0.84), followed by the syntactic knowledge task (0.66) and the distributional knowledge task (0.61). Participants achieved higher mean scores on items with neutral than those with non-neutral suffixes in both relational (0.90 vs. 0.79) and syntactic (0.70 vs. 0.63) knowledge tasks. However, their performances were comparable for neutral (0.60) and non-neutral (0.62) items in the distributional knowledge task. Regarding formedness, participants responded more accurately to well-formed than to pseudo-formed stimuli across all three tasks. Table 2 also demonstrates the mean RTs for each condition. In the relational task, participants spent substantially less time processing neutral-suffixed derivatives than non-neutral ones (1525 ms vs. 2760 ms). In the syntactic task, RTs were similar for the two suffix types (4333 ms vs. 4513 ms). In contrast, participants took longer to respond to neutral items than non-neutral items in the distributional knowledge task (2116 ms vs. 1285 ms). Across all tasks, pseudo-formed stimuli elicited slightly longer RTs than well-formed stimuli, indicating a modest processing cost associated with non-conventional morphological relation.

Table 2. Mean accuracy and mean response times (RTs) (in milliseconds) (N = 54).

	Relational	**Syntactic**	**Distributional**
	Mean accuracy (SD)	Mean accuracy (SD)	Mean accuracy (SD)
Neutral	0.90 (0.14)	0.70 (0.20)	0.60 (0.15)
Non-neutral	0.79 (0.16)	0.63 (0.26)	0.62 (0.13)
Well-formed	0.89 (0.12)	0.70 (0.23)	0.69 (0.14)
Pseudo-formed	0.86 (0.20)	0.64 (0.22)	0.53 (0.19)
Total	0.84 (0.13)	0.66 (0.22)	0.61 (0.11)
	Mean RT (SD)	Mean RT (SD)	Mean RT (SD)
Neutral	1524 (223)	4333 (830)	2116 (660)
Non-neutral	2760 (459)	4513 (1246)	1285 (177)
Well-formed	1805 (509)	4875 (1639)	1862 (543)
Pseudo-formed	1856 (506)	5000 (1821)	2063 (629)

5.2. Modelling Receptive L2 Derivational Knowledge: Accuracy Analysis

Three logit mixed-effects models were built to analyse participants' accuracy scores across the relational, syntactic, and distributional knowledge tasks. Supplementary Material C1 provides a detailed description of the accuracy models along with tables of estimates, standard errors, test statistics, p-values, and confidence intervals.

Results showed that the relational knowledge data yielded significant main effects of affixal neutrality ($\chi^2(1) = 4.89$, $p = 0.03$), formedness ($\chi^2(1) = 5.42$, $p = 0.02$), and L2 proficiency ($\chi^2(1) = 25.22$, $p < 0.001$). Post hoc comparisons using the Bonferroni adjustment demonstrated that participants performed better on items targeting neutral than non-neutral suffixes ($p = 0.003$, Cohen's d = 0.95) and better on well-formed than pseudo-formed items ($p = 0.002$, Cohen's d = 1.0). These main effects are illustrated in Figure 1.

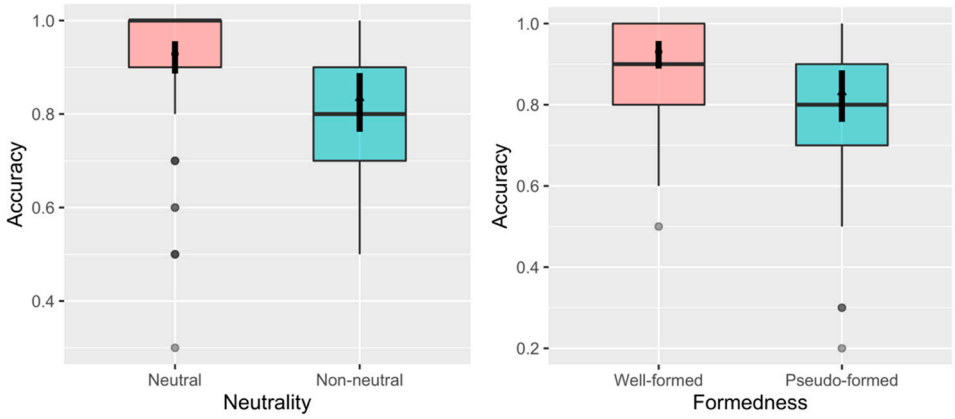

Figure 1. Effects of affixal neutrality and formedness of derivational relation on accuracy in the relational knowledge test.

In terms of accuracy in the syntactic knowledge task, affixal neutrality and formedness of derivational relation showed no significant main effects, but there was a significant interaction between neutrality and formedness. The ANOVA analysis demonstrated a significant neutrality and formedness interaction (see Figure 2): $\chi^2(1) = 4.93$, $p = 0.03$. The post hoc comparison revealed that participants showed no accuracy difference on well-formed items (i.e., real words) with neutral or non-neutral suffixes, but they performed better on neutral than non-neutral items that were pseudo-formed ($p = 0.01$, Cohen's d = 1.22). There was also a significant L2 proficiency effect ($\chi^2(1) = 36.83$, $p < 0.001$).

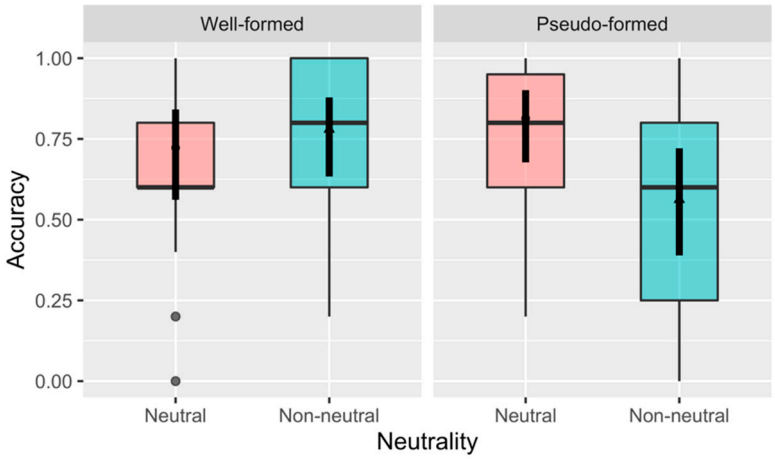

Figure 2. *Neutrality* × *formedness* interaction effect on accuracy in the syntactic knowledge task.

Regarding the distributional knowledge, the ANOVA analysis revealed a significant main effect of formedness ($\chi^2(1) = 17.53$, $p < 0.001$). Post hoc comparisons with Bonferroni adjustment showed that participants performed significantly better on well-formed items (i.e., real words) than on pseudo-formed items (i.e., pseudowords) ($p < 0.001$, Cohen's d = 1.45) (Figure 3). The effect of L2 proficiency was also significant ($\chi^2(1) = 4.47$, $p < 0.05$). As illustrated in Figure 4, participants with higher proficiency levels demonstrated greater accuracy across all three derivational knowledge tasks.

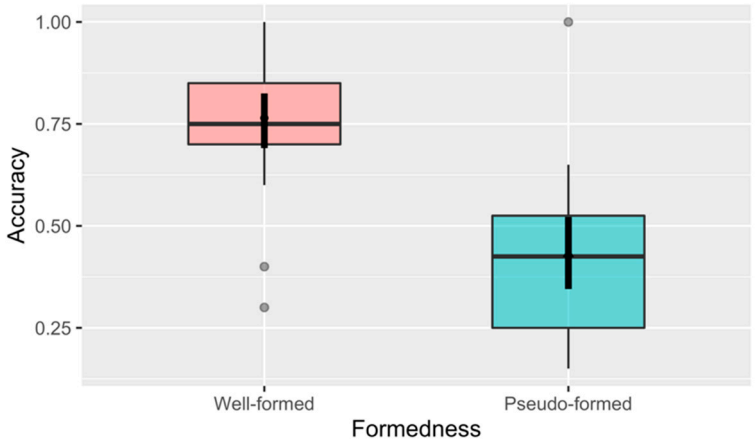

Figure 3. Effect of formedness of derivational relation on accuracy in the distributional knowledge test.

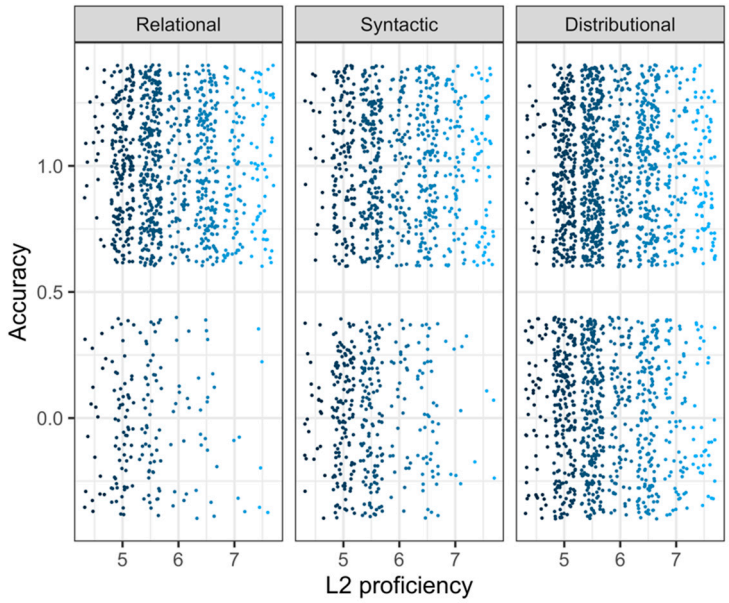

Figure 4. Effects of L2 proficiency (based on IELTS reading scores) on accuracy in the relational, syntactic, and distributional knowledge tasks.

5.3. Modelling Receptive L2 Derivational Knowledge: RTs Analyses

Linear mixed models were constructed to analyse participants' RTs in the receptive relational, syntactic, and distributional knowledge tasks. Supplementary Material C2 provides a detailed description of the RT models along with tables of estimates, standard errors, test statistics, *p*-values, and confidence intervals.

In terms of relational knowledge, the ANOVA analysis showed only a significant main effect of affixal neutrality ($F(1, 19) = 5.03$, $p = 0.04$). Participants spent significantly less time on neutral-suffixed than non-neutral-suffixed items in the relational knowledge test (Cohen's d = 0.43). This effect is illustrated in Figure 5. For syntactic knowledge, no main effects of neutrality and formedness or any interaction effects were found. For distributional knowledge, an ANOVA revealed a significant interaction between neutrality and formedness, $F(1, 42) = 6.11$, $p < 0.05$ (see Figure 6). L2 participants spent a comparable amount of time judging well-formed and pseudo-formed items with neutral suffixes, but they spent significantly longer on pseudo-formed than on well-formed items with non-neutral suffixes ($p = 0.03$, Cohen's d = 0.41).

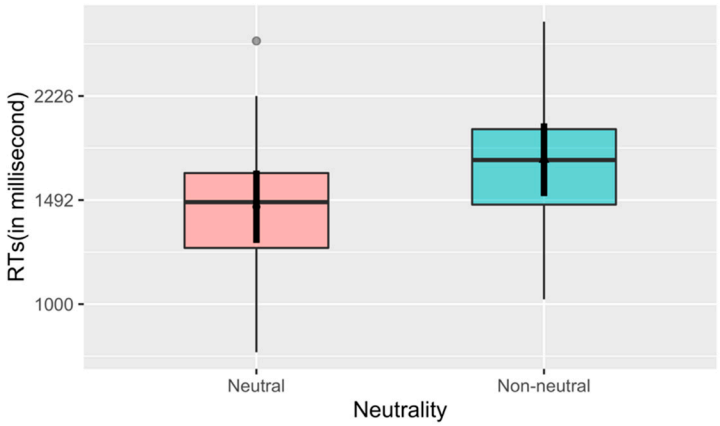

Figure 5. Effect of affixal neutrality on RTs in the relational knowledge test.

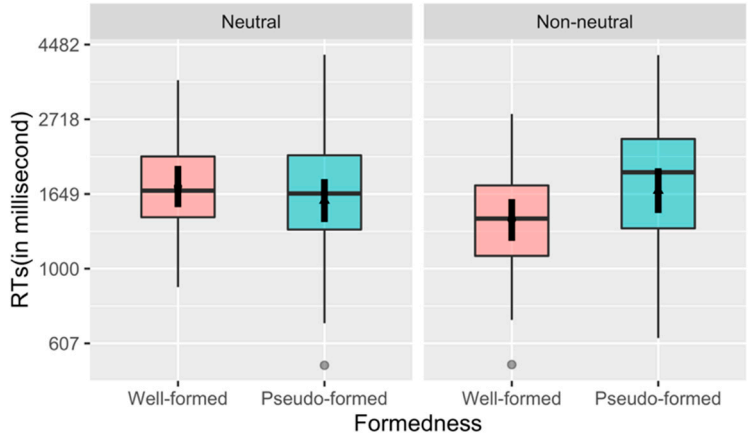

Figure 6. *Neutrality* × *formedness* interaction effect on RTs in the distributional knowledge test.

6. Discussion

Overall, the findings provided partial but theoretically meaningful support for our hypotheses and offer new insights into the linguistic concept of affixal neutrality in L2 derivational morphology. As predicted, neutral suffixes facilitated both accuracy and speed in the relational knowledge task, aligning with claims from theoretical morphology that such affixes are more semantically transparent and morphologically productive. In contrast, neutrality exerted no main effect in the syntactic or distributional tasks, supporting the hypothesis that its influence would be weaker in domains involving more abstract grammatical or constraint-based knowledge (H1). Consistent with our formedness hypothesis (H2), well-formed derivatives yielded higher accuracy in relational and distributional tasks. The predicted neutrality × formedness interaction (H3) was confirmed in syntactic and distributional knowledge, with the advantage of neutral suffixes most evident for pseudo-formed items, suggesting that transparency helps learners detect morphological violations. Higher proficiency was associated with overall accuracy gains (H4) but did not meaningfully alter RT patterns. By examining a long-standing theoretical construct—affixal neutrality—across multiple dimensions of derivational knowledge, the study demonstrates how questions grounded in linguistic theory can be addressed through L2 data, offering insight into the interaction between morphological properties and learner knowledge.

6.1. Neutrality Effect on L2 Derivational Knowledge

Our results demonstrated that participants had a higher probability of responding correctly and spent significantly less time on items with neutral than non-neutral suffixes in

the relational knowledge task, with medium to large effect sizes. The main effect of neutrality on relational knowledge ties well with Carlisle (2000), who found L1 learners performed better on transparent (neutral) than phonologically shifted (non-neutral) derivatives in a word decomposition task. However, this finding contrasts with some L1 studies (e.g., Mahony et al., 2000; Tyler & Nagy, 1989) that reported no facilitative effect of neutrality on recognizing morphologically related word pairs. This divergence may indicate that neutrality plays a more crucial role in the development of L2 than L1 relational knowledge of derivational morphology. One possible explanation is that L2 learners rely more heavily on sublexical information to determine morphological relatedness between words (Silva & Clahsen, 2008). Moreover, as Hayashi and Murphy (2011) suggest, L2 learners' metalinguistic awareness and recognition of morphological relatedness may benefit from explicit classroom instruction and systematic exposure to derivational morphology, conditions that differ markedly from those experienced by native-speaking children. Thus, EFL learners in classroom contexts may rely on the greater transparency of neutral suffixes to identify morphological relatedness more rapidly and accurately.

However, no significant main effect of affixal neutrality was found in the syntactic and distributional knowledge tasks. The absence of a neutrality effect on accuracy in the syntactic knowledge task is consistent with Tyler and Nagy (1989), who reported neutrality did not facilitate L1 learners' recognition of syntactic categories of derived words from Grades 4 to 8. As Tyler and Nagy (1989) noted, "Neutral and non-neutral suffixes do not differ in their regularity in terms of marking the part of speech" (p. 665). This suggests that the syntactic function of a derivational suffix is not determined by whether its attachment induces orthographic and/or phonological changes in the base, but rather by the suffix's inherent grammatical role. In this respect, the relative independence between affixal neutrality and syntactic category marking appears to hold both L1 and L2 acquisition, which may account for the null effect of neutrality on syntactic knowledge in the current study. Furthermore, the absence of a neutrality effect on RTs in the syntactic knowledge task may reflect the limited sensitivity of RT measures in processing sentence-level stimuli. Because participants were required to parse an entire sentence and evaluate multiple answer options, the time devoted to syntactic integration and sentence comprehension likely overshadowed the processing time associated with individual derivatives, thereby obscuring potential differences attributable to affixal neutrality.

Likewise, no significant effects of affixal neutrality were observed for either accuracy or RTs in the distributional knowledge task. This seems contradictory to Tyler and Nagy's (1989) finding that participants accepted items with neutral suffixes significantly more often than that with non-neutral suffixes. However, the apparent discrepancy likely arises from task differences: whereas Tyler and Nagy's task measured acceptability judgements, the present timed lexical decision task indexed accuracy. It is possible that L2 participants' greater acceptability of neutral-suffixed items led them to erroneously accept neutral pseudowords, thereby offsetting their higher accuracy on neutral real words and reducing their overall accuracy in the task (approximately 60%). For instance, L2 learners may have a strong tendency to accept plausible pseudowords such as *metalful* as real words, since both components of the pseudo-derivative (*metal* and *-ful*) appear morphologically legitimate. Friedline (2011), although not specifically examining affixal neutrality, similarly found that non-native speakers were less likely than native speakers to reject deviant derivatives containing orthographically and phonologically regular suffixes. Consistent with psycholinguistic evidence (e.g., Bowden et al., 2010; Silva & Clahsen, 2008), this suggests that L2 learners may have reduced access to rule-based knowledge and procedural memory limiting their ability to accurately apply affixation rules or detect violations of compositional structure during real-time processing. It is important to acknowledge that

the relatively small sample size and restricted item pool may have limited the precision of the estimates and the ability to detect subtle effects of affixal neutrality across knowledge domains. Future studies can expand the range of test items, which would help strengthen the robustness of the results.

6.2. Effects of Formedness and Neutrality × Formedness on L2 Derivational Knowledge

L2 participants' accuracy scores were also influenced by the formedness of the target stimuli (well-formed vs. pseudo-formed) in both relational and distributional knowledge tasks. Across tasks, participants consistently performed more accurately on well-formed than on pseudo-formed items. In the relational knowledge task, pseudo-formed items were real-word pairs lacking genuine morphological relation (e.g., *cent-center*), whereas in the distributional knowledge task, pseudo-formed items were pseudowords (e.g., *informness*). Because all lexical items used in the relational knowledge task were drawn from high-frequency word levels, participants were expected to know their meanings. If responses had been based solely on semantic similarity, accuracy should have been comparable between real and pseudo-related items. Thus, the elevated error rates for pseudo-related items suggest that L2 learners were influenced by lexical surface forms, particularly orthographic similarity, during online lexical processing (Heyer & Clahsen, 2015).

Similarly, L2 participants showed significantly lower accuracy in judging pseudowords than real words in the distributional knowledge task. This pattern may reflect a variant of the word superiority effect (McClelland & Rumelhart, 1981), where speakers recognise a morpheme (e.g., *-ness*) more efficiently when it appears in a real word (e.g., *weakness*) than in a nonword (e.g., *informness*), indicating that morphemic processing is modulated by higher-level lexical processing. Meanwhile, participants' incorrect acceptance of pseudowords may signal underdeveloped knowledge of derivational concatenation rules. Specifically, while L2 learners may be able to decompose a complex form into a stem and a suffix (e.g., *informness* = *inform* + *ness*), they may not recognize that *-ness* typically attaches to adjectival, rather than verbal or nominal, bases. Taken together, the main effect of formedness suggests that morpho-orthographic activation plays a significant role in L2 learners' online processing of morphologically complex words.

The syntactic function of a derivative is marked by its attached suffix, and L2 participants could rely exclusively on this information for syntactic parsing, regardless of whether an item was well- or pseudo-formed (Tyler & Nagy, 1989). Therefore, formedness alone did not yield a significant main effect on accuracy in the syntactic knowledge task. Interestingly, however, a significant interaction emerged between neutrality and formedness, such that L2 learners scored higher on pseudo-formed items with neutral than non-neutral suffixes, with a large effect size. One possible explanation is that learners may not need to engage in morphological analysis to identify the syntactic category of well-formed (real words) items, as they could rely on lexical memory to retrieve syntactic information (Wu & Juffs, 2022). Thus, neutrality, as a morphological property, does not influence their retrieval of syntactic information of real words. In contrast, for pseudowords (e.g., *informness*), morphological decomposition becomes essential to access the syntactic information encoded by the suffixes. Because neutral suffixes are generally easier to isolate and identify than non-neutral suffixes, learners could more readily exploit their grammatical cues for syntactic identification in pseudo-formed contexts.

No significant main effect of formedness was found for RTs across any aspects of derivational knowledge. Thus, while orthographic properties influenced the accuracy of certain aspects of derivational knowledge, they did not affect the speed of knowledge application. Although RTs are generally considered a more sensitive measure of morphological

processing, the distributional knowledge task yielded the most time-sensitive data among the three tasks.

This distinction likely reflects task-specific processing demands. The relational knowledge task involved exclusively high-frequency real words (both genuine- or pseudo-morphologically related), and RTs capture the total judgment time for word pairs. In contrast, RTs in the syntactic knowledge test reflected the time required to parse the entire sentence, process answer choices, and make decisions. Consequently, these two tasks were less effective in isolating the online processing of individual words than the distributional knowledge task, which directly measured participants' lexical decision speed on single-word stimuli.

Thus, the distributional knowledge task was the only task in which response times revealed a significant neutrality × formedness interaction. Participants spent comparable amounts of time judging well-formed and pseudo-formed items with neutral suffixes (2.03 s vs. 1.91 s), whereas they responded substantially more slowly to pseudo-formed than well-formed items in the non-neutral condition (1.91 s vs. 1.57 s). This RT pattern can be interpreted alongside accuracy. Neutral pseudowords were accepted as real words more frequently than non-neutral pseudowords (47.5% vs. 42.1%), indicating that they were perceived as more word-like. Together, these findings suggest that affixal neutrality modulates lexical decision difficulty in a manner that depends on lexical status. In lexical decision, response times are known to reflect graded familiarity and decision uncertainty rather than categorical judgments of well-formedness (Balota & Chumbley, 1984; Baayen et al., 2011), and highly plausible pseudowords are known to attenuate RT contrasts between real and nonword items.

One plausible explanation is that neutral suffixes are highly productive and morphologically salient, strongly inviting morphological decomposition. While this transparency facilitates parsing, it may also increase processing time by activating multiple competing analyses. For neutral pseudowords in particular, both the base and the suffix are independently well-formed and familiar, yielding combinations that are morphologically plausible but lexically unattested, thereby heightening competition between "word" and "nonword" responses. By contrast, non-neutral real-word derivatives are less easily segmented and can benefit from entrenched whole-word representations, allowing faster responses, whereas non-neutral pseudowords lack such support and require additional evaluation of stem–suffix compatibility. This pattern is consistent with previous findings on L2 morphological processing and real-word advantages (see also Diependaele et al., 2009; Silva & Clahsen, 2008). Thus, the slower RTs observed for neutral items should not be interpreted as greater processing difficulty per se, but rather as reflecting increased analytic engagement and competition triggered by highly segmentable and productive morphology (Amenta et al., 2025).

Situated within existing research on L2 derivational knowledge, the present findings contribute to theory by clarifying how linguistically defined morphological properties constrain different components of L2 morphological acquisition. Much prior work has examined derivational knowledge primarily as part of global morphological awareness and in relation to reading development (e.g., Kieffer & Lesaux, 2008, 2012; Zhang, 2017), often treating morphology as a relatively unitary construct. By contrast, the current results support a multidimensional view of L2 derivational knowledge, showing that affixal neutrality exerts domain-specific effects. These findings extend earlier L2 work distinguishing relational and syntactic knowledge (e.g., Zhang & Koda, 2012, 2013; Hayashi & Murphy, 2011). Moreover, by incorporating distributional knowledge—an aspect rarely examined in L2 research—the study shows that learners' evaluation of derivational well-formedness is shaped by interactions between morphological transparency and lexical status, helping reconcile mixed findings on phonological and orthographic change effects (To et al.,

2016; Jiang & Kuo, 2019). More broadly, the pattern of results supports views in which systematic structural properties of morphology (e.g., affixal neutrality) play a central role in L2 morphological acquisition, consistent with structure-sensitive and constraint-informed perspectives in morphology and SLA that emphasise how linguistic regularities shape learners' developing representations (Baayen et al., 2011; Clahsen et al., 2010).

6.3. L2 Proficiency Effect on L2 Derivational Knowledge

L2 proficiency was shown to have significantly facilitative effects on accuracy but not on RTs in all aspects of derivational knowledge. The proficiency effect was strongest for syntactic knowledge (odds ratio = 2.48), followed by tasks assessing morphological relatedness (odds ratio = 1.95) and distributional constraints (odds ratio = 1.22). As indicated by Pienemann's (1998) Processability Theory, L2 learners progressively develop their processing abilities to handle increasingly complex linguistic operations. In the framework of Tyler and Nagy (1989), relational knowledge represents the earliest acquired aspect of derivational knowledge. Consistent with this view, learners' extensive exposure to L2 input, combined with explicit classroom instruction on frequent derivational morphemes, likely facilitates the ability to recognise morphological relations between derivatives and their components. Once learners became proficient in rule-based decomposition, however, the development of relational knowledge tends to stabilize, leading to a smaller proficiency effect in this domain.

By contrast, syntactic knowledge poses greater challenges for learners with limited proficiency. Differentiating and memorising the syntactic functions of derivational suffixes are cognitively demanding processes that require sustained time and effort. However, identifying the grammatical role that a derivative fulfills in a sentence is a common exercise in EFL instruction. As learners gained proficiency, they continue to develop competence in this domain, yet full mastery typically emerges only at advanced proficiency levels. This developmental pattern explains why syntactic knowledge exhibited the strongest proficiency effect among the three aspects of derivational knowledge.

Finally, distributional knowledge, the most complex and demanding type of derivational knowledge, showed delayed development. By nature, distributional constraints on affixation depend on multilayered metacognitive awareness, including morphological, semantic, syntactic, and phonological dimensions. Although these constraints are, in principle, predictable, they are less systematically organised, making them challenging for both lower- and higher-proficiency learners to internalize and manipulate. This difficulty was reflected in the only above-chance accuracy rates observed in the distributional knowledge task.

7. Conclusions

This study set out to bring the theoretical construct of affixal neutrality into the domain of second language morphology. While affixal neutrality has traditionally been examined to explain patterns of morphological transparency, productivity, and acquisition in native speakers, it has rarely been operationalised in L2 research. By embedding this concept within tasks targeting distinct aspects of derivational knowledge, the present study demonstrates how a linguistic construct can generate fine-grained research questions in L2 acquisition and shape hypotheses about learners' morphological awareness.

The findings suggest that L2 learners are indeed sensitive to the transparency and productivity associated with neutral affixes, yet this sensitivity manifests differently across domains of derivational knowledge. These results highlight the value of integrating theoretically grounded morphological distinctions into L2 research designs, as doing so can reveal domain-specific profiles of learners' morphological awareness that broader measures

may obscure. In turn, second language data can offer linguistic theory an expanded testing ground. L2 learners' differential responsiveness to affixal neutrality across knowledge domains raises questions about the extent to which theoretical distinctions that were formulated primarily from native-speaker data can apply universally across language users. Such evidence invites further theoretical refinement, particularly regarding how morphological transparency interacts with factors like proficiency, lexical exposure, and processing constraints in shaping morphological knowledge.

Pedagogically, affixal neutrality offers a principled basis for sequencing derivational instruction in L2 vocabulary learning. Transparent and highly productive neutral affixes (e.g., *-ness*, *-less*, *-ful*) can be introduced early to help learners establish stable form–meaning mappings and develop awareness of derivational structure. Instruction can then shift toward non-neutral affixes that involve phonological or orthographic alternations, with explicit attention to stem changes, suffix-specific constraints, and lexicalized word families (e.g., *divide–division*; *active–activity*). In addition, the observed difficulty in distributional judgments suggests that learners may benefit from instructional activities targeting usage-based derivational knowledge, such as evaluating the appropriateness of derived forms in short contexts or comparing minimally contrasting derivatives. Combined with increased exposure through reading and input-rich tasks, such practices may support learners in extending their derivational repertoire to less frequent and structurally complex affixes.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/languages11030046/s1. Supplementary material A: Lexical stimuli in receptive derivational knowledge tasks; Supplementary material B: Descriptive characteristics of lexical stimuli, and Supplementary material C: Statistical reporting, including Table S1: Word length, syllable count, and log frequency by neutrality and formedness across tasks, Table S2: Mixed logit model outputs of accuracy of performance in the relational, syntactic, and distributional knowledge tests, and Table S3: Linear mixed-effects model outputs of RTs in the relational, syntactic, and distributional knowledge tests.

Author Contributions: Conceptualization, X.W. and H.Z.; methodology, X.W. and H.Z.; software, X.W.; validation, X.W. and H.Z.; formal analysis, X.W.; investigation, X.W.; resources, X.W.; data curation, X.W.; writing—original draft preparation, X.W.; writing—review and editing, X.W. and H.Z.; visualization, X.W.; supervision, H.Z.; project administration, X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of The University of Melbourne (protocol code 2057495.1 and 12 August 2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data available in a publicly accessible repository. The data presented in this study are openly available in [Open Science Framework] at [https://osf.io/63gu5/?view_only=cc6cbff93f5f43fa9e83b26d46f13fee, accessed on 21 January 2026].

Conflicts of Interest: The authors declare no conflict of interest.

Notes

[1] The classification of neutral versus non-neutral affixes is not absolute, as dual membership (i.e., an affix belongs to both classes) exists (Giegerich, 1999; Plag & Baayen, 2009). However, the appearance of the same affix in both neutral and non-neutral strata can be attributed to "an ongoing language change" (Katamba & Stonham, 2006). The authors also remarked that "The affix morpheme may be splitting into two. So, what we have are two closely related homophonous morphemes, with different

characteristics, occurring at different strata, much as there are two *-ing* morphemes, one indicating the verbal participle and the other a nominalisation of the verb." (p. 135).

[2] Formula: Accuracy(relational/syntactic/distributional) ~ Neutrality * Formedness + Proficiency + (1 | Participant) + (1 | Item).

[3] Formula: log(RTs)(relational/syntactic/distributional) ~ Neutrality * Formedness + Proficiency + (1 | Participant) + (1 | Item).

References

Amenta, S., Foppolo, F., & Badan, L. (2025). The role of morphological information in processing pseudo-words in Italian L2 learners: It's a matter of experience. *Journal of Cognition*, *8*(1), 14. [CrossRef]

Apel, K. (2014). A comprehensive definition of morphological awareness: Implications for assessment. *Topics in Language Disorders*, *34*(3), 197. [CrossRef]

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. [CrossRef]

Baayen, R. H., Milin, P., Filipović Đurđević, D., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*(3), 438–481. [CrossRef]

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340–357. [CrossRef]

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. [CrossRef]

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models. *arXiv*. [CrossRef]

Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, *6*(4), 253–279. [CrossRef]

Booij, G. (2016). Construction morphology. In A. Hippisley, & G. Stump (Eds.), *The Cambridge handbook of morphology* (pp. 424–448). Cambridge University Press.

Bowden, H. W., Gelfand, M. P., Sanz, C., & Ullman, M. T. (2010). Verbal inflectional morphology in L1 and L2 Spanish: A frequency effects study examining storage versus composition. *Language Learning*, *60*(1), 44–87. [CrossRef] [PubMed]

Breslow, N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, *88*(421), 9–25. [CrossRef]

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). GlmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*(2), 378–400. [CrossRef]

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing*, *12*(3), 169–190. [CrossRef]

Carlisle, J. F., & Fleming, J. (2003). Lexical processing of morphologically complex words in the elementary years. *Scientific Studies of Reading*, *7*(3), 239–253. [CrossRef]

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Harper Row.

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and non-native language processing. *Language Learning*, *60*(1), 21–43. [CrossRef]

Diependaele, K., Sandra, D., & Grainger, J. (2009). Semantic transparency and masked morphological priming: The case of prefixed words. *Memory and Cognition*, *37*(6), 895–908. [CrossRef]

Dixon, R. M. W. (2014). *Making new words: Morphological derivation in English*. Oxford University Press.

Friedline, B. E. (2011). *Challenges in the second language acquisition of derivational morphology: From theory to practice* [Ph.D. thesis, University of Pittsburgh].

Giegerich, H. J. (1999). *Lexical strata in English: Morphological causes, phonological effects*. Cambridge University Press.

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. [CrossRef]

Halle, M., & Mohanan, K. P. (1985). Segmental phonology of modern English. *Linguistic Inquiry*, *16*(1), 57–116.

Hayashi, Y., & Murphy, V. (2011). An investigation of morphological awareness in Japanese learners of English. *Language Learning Journal*, *39*(1), 105–120. [CrossRef]

Haynes, W. (2013). Bonferroni correction. In W. Dubitzky, O. Wolkenhauer, K.-H. Cho, & H. Yokota (Eds.), *Encyclopedia of systems biology* (p. 154). Springer. [CrossRef]

Heyer, V., & Clahsen, H. (2015). Late bilinguals see a scan in scanner and in scandal: Dissecting formal overlap from morphological priming in the processing of derived words. *Bilingualism: Language and Cognition*, *18*(3), 543–550. [CrossRef]

Jiang, Y. L. B., & Kuo, L. J. (2019). The development of vocabulary and morphological awareness: A longitudinal study with college EFL students. *Applied Psycholinguistics*, *40*(4), 877–903. [CrossRef]

Katamba, F., & Stonham, J. (2006). *Morphology*. Macmillan Education.

Ke, S., Miller, R. T., Zhang, D., & Koda, K. (2021). Crosslinguistic sharing of morphological awareness in biliteracy development: A systematic review and meta-analysis of correlation coefficients. *Language Learning*, *71*(1), 8–54. [CrossRef]

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627–633. [CrossRef]

Kieffer, M. J., & Lesaux, N. K. (2008). The role of derivational morphology in the reading comprehension of Spanish-speaking English language learners. *Reading and Writing*, *21*(8), 783–804. [CrossRef]

Kieffer, M. J., & Lesaux, N. K. (2012). Development of morphological awareness and vocabulary knowledge in Spanish-speaking language minority learners: A parallel process latent growth curve model. *Applied Psycholinguistics*, *33*(1), 23–54. [CrossRef]

Kuo, L.-J., Ramirez, G., de Marin, S., Kim, T.-J., & Unal-Gezer, M. (2017). Bilingualism and morphological awareness: A study with children from general education and Spanish-English dual language programs. *Educational Psychology*, *37*(2), 94–111. [CrossRef]

Lenth, R. V., Bolker, B., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Riebl, H., & Singmann, H. (2023). *Emmeans: Estimated marginal means, aka least-squares means* (1.8.7) [Computer software]. Russell V. Lenth.

Li, J., & Taft, M. (2020). The processing of English prefixed words by Chinese-English bilinguals. *Studies in Second Language Acquisition*, *42*(1), 239–249. [CrossRef]

Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, *23*(5), 938–944. [CrossRef]

Lieber, R. (2021). *Introducing morphology*. Cambridge University Press.

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second Language research. *Language Learning*, *65*(S1), 185–207. [CrossRef]

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *PsyArXiv*. [CrossRef]

Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and Writing*, *12*(3), 191–218. [CrossRef]

McBride-Chang, C., Wagner, R. K., Muse, A., Chow, B. W.-Y., & Shu, H. (2005). The role of morphological awareness in children's vocabulary acquisition in English. *Applied Psycholinguistics*, *26*(3), 415–435. [CrossRef]

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407. [CrossRef]

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234. [CrossRef]

Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology*, *98*(1), 134–147. [CrossRef]

Pienemann, M. (1998). Developmental dynamics in L1 and L2 acquisition: Processability Theory and generative entrenchment. *Bilingualism: Language and Cognition*, *1*(1), 1–20. [CrossRef]

Plag, I., & Baayen, H. (2009). Suffix ordering and morphological processing. *Language*, *85*(1), 109–152. [CrossRef]

Qualtrics. (2005). *Version 12.2020. Provo, UT, USA*. Available online: https://www.qualtrics.com (accessed on 13 February 2020).

Ramirez, G., Chen, X., Geva, E., & Kiefer, H. (2010). Morphological awareness in Spanish-speaking English language learners: Within and cross-language effects on word reading. *Reading and Writing*, *23*(3), 337–358. [CrossRef]

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 1 July 2020).

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, *19*(1), 17–36. [CrossRef]

Schmitt, N., & Zimmerman, C. B. (2002). Derivative word Forms: What do learners know? *TESOL Quarterly*, *36*(2), 145–171. [CrossRef]

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, *11*(2), 245–260. [CrossRef]

Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing*, *12*, 219–252. [CrossRef]

Sparks, E., & Deacon, S. H. (2015). Morphological awareness and vocabulary acquisition: A longitudinal examination of their relationship in English-speaking children. *Applied Psycholinguistics*, *36*(2), 299–321. [CrossRef]

Spencer, A. (1991). *Morphological theory: An introduction to word structure in generative grammar*. Wiley.

To, N. L., Tighe, E. L., & Binder, K. S. (2016). Investigating morphological awareness and the processing of transparent and opaque words in adults with low literacy skills and in skilled readers. *Journal of Research in Reading*, *39*(2), 171–188. [CrossRef]

Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language*, *28*(6), 649–667. [CrossRef]

Wu, Z., & Juffs, A. (2022). Effects of L1 morphological type on L2 morphological awareness. *Second Language Research*, *38*(4), 787–812. [CrossRef]

Zhang, D. (2017). Derivational morphology in reading comprehension of Chinese-speaking learners of English: A longitudinal structural equation modeling study. *Applied Linguistics*, *38*(6), 871–895. [CrossRef]

Zhang, D., & Koda, K. (2012). Contribution of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge and reading comprehension among advanced EFL learners: Testing direct and indirect effects. *Reading and Writing*, *25*, 1195–1216. [CrossRef]

Zhang, D., & Koda, K. (2013). Morphological awareness and reading comprehension in a foreign language: A study of young Chinese EFL learners. *System*, *41*(4), 901–913. [CrossRef]