Applied Linguistics, 2024, XX, 1–20
https://doi.org/10.1093/applin/amae084
Advance access publication 13 December 2024
Article

Linking adverbials in children's writing: Exploring variation across year groups, genres, and disciplines

Philip Durrant[1,*], Erdem Akbaş[2], Elif Barbaros[3], and Arwa Aldawood[1,4]

[1]School of Education, University of Exeter, Exeter, United Kingdom
[2]Department of English Language Teaching, Faculty of Education, Erciyes University, Kayseri, Türkiye
[3]School of Foreign Languages, Erciyes University, Kayseri, Türkiye
[4]Department of English, Prince Sattam Bin Abdulaziz University, Al-Kharj, Kingdom of Saudi Arabia

[*]Corresponding author: Philip Durrant, School of Education, University of Exeter, St. Luke's Campus, Heavitree Road, Exeter, EX2 2LU, UK. E-mail: p.l.durrant@exeter.ac.uk

Linking adverbials is a crucial element in successful academic writing that is particularly challenging for both first and second-language learners to master. Drawing on a corpus of writing by mainstream students in UK schools, the current article explores the under-researched issues of how these forms develop across levels of study in an Anglophone context and how their use and development vary across text genres and academic disciplines. We demonstrate that, excluding a small number of high-frequency *pathbreaking* items (such as *and, but* and *so*), linking adverbials are markedly more frequent in children's non-literary than literary writing and that the former, but not the latter, shows an increase in use of linkers as children mature. Linkers are equally prevalent across academic disciplines. However, the specific linkers used are strongly dependent on both text genre and academic discipline, reflecting functional differences between these. The analysis further demonstrates how students move from using characteristically spoken-style linkers towards more written-style linkers as they progress through school.

Introduction

A core skill in academic writing is that of showing the relationships between stretches of text. Writers need to show where one claim is an inference from another; where one idea contrasts with, or supports, another; where items are cited to illustrate a point that has been made; and where they form a list or temporal sequence. In Anglophone academic contexts, writers are expected to make such relationships explicit, often through the use of *linking adverbials* (LAs): words and phrases such as *therefore, furthermore,* and *on the other hand*.

Skilful use of LAs is central to helping readers interpret a text and is a key marker of successful school writing (Halliday and Hasan 1976; Liu 2008; Yin 2016). However, mastering these items is far from straightforward. Conceptualizing a link between two stretches of discourse, and identifying a connector that effectively expresses that link, is a cognitively complex task (Larsen-Freeman and Celce-Murcia 2016); and is made especially challenging by the diversity and subtlety of meanings that can be expressed by LAs.

Appropriate use of LAs also requires a high level of audience awareness. At a macro-level, this means taking account of the expectations of different registers (Yin 2016) and academic disciplines (Peacock 2010). At a micro-level, it means keeping track of how readers' understandings and expectations are likely to be developing as they progress through a text (Hyland and Tse 2004). A further challenge is the fact that the LAs used in academic writing often differ from those that students may be used to from more informal settings (Biber et al. 2021).

LAs are, therefore, a crucial site of both cognitive and social development for young writers. However, as we will see below, relatively little is known about their use and development in school writing. The present study aims to increase our understanding by tracing LA use in children from Years 6 to 11 in schools in England. It will investigate the extent to which LAs are used overall, which individual items are distinctive of different school genres and disciplines, and how their use varies across year groups. By exploring the evolution of these forms, we aim to gain insights into students' progression as mature writers and to inform pedagogical decisions about what forms may need special attention at what points in the learning process.

What are linking adverbials?

The category of LA encompasses several grammatical forms (Biber et al. 2021): simple adverbs (e.g. *anyway; furthermore*); adverb phrases (*even so; first and foremost*); prepositional phrases (*in other words; for example*); finite clauses (*that is; that is to say*) and non-finite clauses (*added to that; to conclude*). They are functionally related to both adverbial subordinators (*although, because*) and coordinating conjunctions (*and, but, or*). However, whereas these forms link clauses within a sentence, LAs can connect a wide range of structures, including sentences and paragraphs.

Several taxonomies have been suggested to organize LAs into semantic groups. We can identify two main sets of taxonomy: those based on Halliday and Hasan's (1976) *Cohesion in English* (Liu 2008; Larsen-Freeman and Celce-Murcia 2016); and those built around the categories in Quirk, Greenbaum, Leech and Svartvik's (1985) *Comprehensive Grammar of the English Language* (Carter and McCarthy 2006; Yin 2016; Biber et al. 2021). Huddleston and Pullum (2002) are distinct from both groups, but are closer to the latter than the former.

The Halliday and Hasan (1976) framework distinguishes between:

- Additive conjunctions: *and*/*or*-type relationships.
- Adversative conjunctions: *and but*-type relationships.
- Causal conjunctions: *so*-type relationships.
- Temporal conjunctions: *then*-type relationships.

It also distinguishes what Halliday and Hasan call *external* and *internal* relations. The former refers to relationships in the world outside of the text (as in Example 1), whereas the latter refers to relations created by the speaker/writer within their text (Example 2).

1. Next he inserted the key into the lock.
2. *Next, he was incapable of inserting the key into the lock.*

Example 1 illustrates an external relationship, with *next* indicating the sequence in which events took place. Example 2 illustrates an internal relationship, with *next* indicating that the speaker/writer is adding an additional point to an argument they are constructing. Using these primary distinctions, Halliday and Hasan unpack their categories into 57 detailed sub-types.

Quirk et al. (1985)'s main categories are:

- Listing: giving a structure or orientation to a list.
- Summative: introducing an item that embraces the preceding items.
- Appositive: expressing preceding items in other terms.

- Resultive: a category that they do not define, but which is well captured in Carter and McCarthy (2006)'s formulation that one segment expresses a result or outcome of the events or states described in the other.
- Inferential: indicating a logical conclusion.
- Contrastive: presenting something that contrasts what has preceded.
- Transitional: shifting attention to another topic or an event related in time.

It is important not to essentialize these categories. Halliday and Hasan emphasize that their system should not be seen as rigid and that alternatives are possible, with little basis for choosing one typology over another (1976: 239). Similarly, Huddleston and Pullum (2002: 778) caution that the 'few broad subcategories' they set out should not be seen as 'either sharply distinctive or exhaustive'.

These caveats are well founded. Classifying LAs requires us to group items into categories with a meaning that is common to all but that distinguishes them from other items. Both aspects of this task—determining what LAs have in common, and determining what distinguishes members of one group from those of another—pose problems. Regarding the first, a characterization that captures several different LAs must, by definition, be vaguer than one tailored to an individual item. For example, Quirk et al. (1985: 639) characterize items in their *concessive* category as showing that 'one unit is seen as unexpected in the light of the other'. This includes items such as *anyway, besides, else, however, nevertheless*, and *still*. Clearly, these items are not interchangeable; they show differences in meaning, in preference for sentence position, and in register. As Liu (2008) notes, something is always lost in categorization, and adequate category definitions are difficult to provide.

This difficulty is reflected in the diverse ways that what appears to be the same category is defined in different grammars. To continue with Quirk et al'.s *concessive* group, both Biber et al. (2021) and Carter and McCarthy (2006) include categories based on this, having both a similar name and including a similar set of words. However, the characterizations offered by the three grammars are not closely aligned. We have seen that Quirk et al. (1985: 639) describe concessive LAs as where 'one unit is seen as unexpected in the light of the other'. Carter and McCarthy in contrast, focus on acknowledgement of conflicting perspectives:

> the speaker or writer is prepared to accept part of an argument or proposition (that is expressed in the second text segment) which typically in some way contrasts with what has already been stated (Carter and McCarthy 2006: 259)

Biber et al. (2021) describe concessive LAs as throwing doubt on a previous idea:

> the subsequent discourse expresses some reservation about the idea in the preceding clause (Biber et al. 2021: 870).

The understanding of contrast highlighted here is clearly different in each case. For Quirk et al., it is the unexpectedness of one statement, given another; for Carter and McCarthy, it is acknowledging aspects of truth in an opponent's argument; for Biber et al., it is casting doubt on a previous statement. These definitions are not incompatible—all three might be true of the same example—but they represent strikingly different interpretations of what LAs have in common.

We have noted that linguists need to define not only what LAs in a group have in common, but also how they are different from other groups. However, sharp boundaries often cannot be drawn. Biber et al. (2021), for example, note that concessive meaning is often combined with a meaning of *contrast*, which they define as showing an incompatibility. Similarly, Halliday and Hasan (1976) note how their category of *dissimilar comparison*, a sub-category within the broad *additive* group (which includes LAs such as *on the other hand, by contrast*, and *conversely*), approaches an *adversative* meaning.

Given these points, it is unsurprising that taxonomies vary widely. This is seen even at the broadest levels of categorization. Halliday and Hasan's *additive* category, for example, include subgroups that Quirk et al. place under the separate headings of *listing, transitional, appositive,* and *contrastive*. Similarly, Halliday and Hasan's *temporal* category includes groups that Quirk et al would describe as *listing*, *transitional,* and *summative*. Thus, groups that Halliday and Hasan place together are spread across multiple groups in Quirk et al, while groups that Halliday and Hasan distinguish are included within a single group in Quirk et al.

In some cases, broad groups do match. Halliday and Hasan's *adversative* group, for example, corresponds roughly to Quirk et al'.s *contrastive* group. At a more fine-grained level, however, striking differences remain, with Quirk et al. conflating many contrastive subcategories which are kept distinct by Halliday and Hasan. Moreover, each taxonomy mentions sets of LAs that appear to have been omitted by the other: for example, Halliday and Hasan's *purpose* category, within the *causal* group, and Quirk et al'.s *replacive* category, within the *contrastive* group.

This discussion has several implications for studying the development of LAs in children's writing. First, semantic categories should be seen as interpretive lenses, rather than objective things in the world. Second, a specific use of an LA may legitimately be interpreted under different semantic categories. Finally, because they abstract away from the meanings, syntactic associations, and register preferences of individual LAs, categorizations may obscure some of the complexity involved in the learning of those LAs.

Given these points, we argue that it is not productive to include semantic category as an a priori variable in learner language analysis. This is for four main reasons: First, because semantic categories are not objective types, there is no reason to believe that reliable generalizations can be built up on them. Any apparent patterns may be due to extraneous factors. Second, different taxonomies may yield different results, and there is no principled way of choosing between them. Third, the different interpretations given to LAs by different analysts, and the ways that different meanings bleed into each other, suggest that individual uses cannot be reliably categorized. Fourth, semantic categories may hide important distinctions between different LAs. While categories of meaning may be useful in an interpretive heuristic, therefore, grouping LAs into categories from the outset may produce misleading results.

Linking adverbials in school writing

Research into LAs in Anglophone school contexts has been surprisingly sparse. However, it does allow for some tentative conclusions and points towards priorities for further research. First, both Crowhurst (1987) and Barbaros and Akbaş (2022) find that a small group of individual items dominate students' use of LAs. In particular, Halliday and Hasan's (1976) four semantic categories are dominated by *and* (additive), *but* (adversative), *so* (causal), and *then* (temporal). Not coincidentally, these are precisely the words that Halliday and Hasan (1976) use to explain the basic meaning of each category—in their terms, the *logical notions* on which they are based (p. 234). These items can be seen as akin to the *pathbreaking* items of usage-based learning models (Ellis and Ferreira-Junior 2009): high-frequency words whose meaning is prototypical of their category and which learners pick up early and use extensively before broadening their repertoire to other category members. In line with this model, use of *and, but,* and *so* is reported to be highest in the early years and to decrease dramatically as learner mature (Crowhurst 1987; Barbaros and Akbaş 2022).

Second, regarding the central question of whether overall frequency of LA use changes over time, studies conflict. While Nippold, Ward-Lonergan and Fanning (2005) found a substantial increase from ages 11 to 17, similar patterns were not found in other studies. Crowhurst (1987) found no change across a similar age span for additives and adversatives, and significant decreases for causal and temporal items. Barbaros and Akbaş (2022) found a complex pattern of significant decrease from Years 2 to 6, a significant increase from Years 6 to 9, and a slight decrease from Years 9 to 11.

A possible explanation for these contradictory findings is that Nippold et al. (2005) did not include what they called *early developing* conjuncts. Though they do not say how these items were

identified, they are likely to have included the high-frequency items that dominated the findings in the other two studies (indeed Nippold et al. explicitly mention that *and, so*, and *then* were excluded). These items weighed heavily in the overall token counts for the other studies; *but*, for example, accounted for 62 per cent of contrastives in Barbaros and Akbaş (2022) and 74 per cent in Crowhurst (1987). As we have discussed, items of this sort also tended to decrease in frequency across year groups. We might hypothesize, therefore, that their inclusion in the overall counts has obscured increases in other items.

Third, some findings suggest that genre is an important variable in LA use. In Crowhurst (1987), the shift away from the *pathbreaking* additive and adversative conjunctions (*and*; *but*) was seen in argumentative, but not narrative texts. Conversely, there was a decrease in temporal conjunctions in narrative, but not argumentative texts. Further evidence of genre-specificity is seen in Barbaros and Akbaş (2022), who found that literary texts used more contrastive LAs than non-literary texts.

While these findings are suggestive, it remains difficult to draw strong conclusions. Crowhurst's finding that different genres have different developmental tendencies is intriguing, but the strong influence of high-frequency items noted above obscures the behaviour of the broader range of LAs. Barbaros and Akbaş's (2022) finding of an overall higher use of LAs in literary texts is also potentially interesting. However, their research looked only at contrastive markers, again limiting its generalizability to a broader set of LAs. It also remains unclear why a similar effect was not found by Crowhurst (1987).

An area that has not yet been explored is that of disciplinary variation. It is well known that school and university vocabulary differs widely between subject areas (Hyland and Tse 2007; Durrant 2014). We would therefore expect corresponding differences in the LAs. Such variation is important as it may offer insight into how the use and development of LAs are influenced by the communicative needs of particular topic areas and so into the learning challenges and opportunities presented at particular points in the curriculum. However, we are not aware of any research addressing this issue.

Building on the literature reviewed above, the current research aims to study both the overall prevalence of LAs across year groups and genres and the developing roles of prominent individual items across year groups, genres, and academic disciplines. In particular, we will address the following research questions:

1. How does overall frequency of linking adverbial use vary across
   a. school years?
   b. text genres?
   c. academic disciplines?
2. To what extent is variation across school years genre- or discipline-dependent?
3. Which individual linking adverbials increase or decrease in use across school years?
4. To what extent are individual linking adverbials associated with specific genres or disciplines?
5. To what extent are increases/decreases in the use of individual linking adverbials associated with specific genres or disciplines?

Because the frequent, early-developing items *and, but, so*, and *then* appear to have heavily skewed findings in previous studies, possibly obscuring developmental patterns, these items will be excluded from our analysis to allow patterns for a wider set of LAs to emerge.

Methodology

Corpus

Our data come from the Growth in Grammar[1] corpus, which comprises 2,898 texts and was collected between 2015 and 2017 from 24 schools in England. Texts were produced by approximately

1,000 children across Years 2 (when most children are 6–7 years old), 4 (8–9 years old), 6 (10–11 years old), 9 (13–14 years old), and 11 (15–16 years old) for classes in English, Science, and the Humanities (History, Geography, Religious Studies). All texts were written as part of students' regular school work, rather than being elicited by the researchers. Writing for English was divided into two main types: *literary* and *non-literary*. The former refers to texts, such as *creative fictions* and *literary imitations*, which can be evaluated without reference to their factual accuracy or the persuasiveness of their arguments. The latter refers to texts such as *autobiographies, essays, literary criticism*, and *experimental reports*, which aim to describe, evaluate, or argue for something and which are typically evaluated based on the accuracy of the information they contain or the persuasiveness of the argument they present (for further discussion of these categories, see Durrant 2022).

For the present study, we excluded Year 2 because the use of LAs was strongly influenced by some heavily scaffolded tasks where students have clearly been asked to follow a set format with LAs supplied by the teacher (see Durrant 2022 for a discussion). We also excluded Year 4 (of which there are very few texts) and a small number that were rewritten versions of another text. We removed extra-textual features (i.e. headings, tables, symbols, graphics, annotator comments), and quotations from source texts. This resulted in a corpus of 2,189 texts (76 per cent of the full corpus. See Table 1 for a summary of corpus contents).

Identifying LAs

LAs were defined as words or word sequences that function as an adverbial and that form a semantic link with another stretch of text. To identify LAs in the corpus, we first compiled a list of items attested in one of the following sources: Halliday and Hasan (1976); Quirk et al. (1985); Huddleston and Pullum (2002); Carter and McCarthy (2006); Liu (2008); Larsen-Freeman and Celce-Murcia (2016); Yin (2016); Biber et al. (2021). This yielded a list of 275 candidate items. As discussed in the *Linking adverbials in school writing* section, the high-frequency, early-developing, items *and, but, so*, and *then* may obscure developmentally important patterns so were excluded from our analysis. Of the remaining items, 168 were found to occur in our corpus and were retained for further study.

Many items were polysemous and many can play multiple grammatical functions. For example, while *rather* can function as an LA, as in example 1,

1. *Eddie says (or **rather** accuses) Catherine of 'walking wavy'.*

it can also function as a modifier within an adjective phrase (as in 2), or within an adverb phrase (3), or as an element within a verb phrase (4).

2. *As an audience, we would be **rather** anxious watching.*
3. *He decided, **rather** unjustly, to dismiss her…*
4. *However Lady Macbeth commits suicide as she can't cope with what she has done and she would **rather** die than live with a guilty conscience.*

Table 1. Summary of corpus contents.

	N. texts (mean words per text)		
	Year 6	Year 9	Year 11
English (literary)	274 (362)	220 (395)	63 (406)
English (non-literary)	296 (177)	305 (299)	368 (355)
Humanities	106 (222)	113 (263)	49 (291)
Science	171 (193)	166 (146)	58 (363)

Moreover, items functioning as adverbials do not always have a linking function. This is seen in 5, where *at this moment* refers exophorically to the time of reading, rather than to another segment of the text.

5. *Parents think carefully about what your children are doing* ***at this moment***.

Identifying instances in which a word or phrase was used as an LA required manual analysis. To this end, we retrieved all instances of the candidate LAs from the corpus using AntConc (Anthony 2022). For each LA, the concordance was exported as a spreadsheet and saved. During coding, the status of each instance as an LA or non-LA was recorded in this spreadsheet. Coders also had access to the original AntConc concordance and to the full text of the corpus, which they were able to consult when a decision could not be made based on the concordance line alone.

A coding document was created, with an initial operational definition of LAs and examples of items that should and should not be treated as LAs. In the first round of coding, 10 LAs were then coded independently by all four authors of this paper. After coding, we compared our codes and discussed cases of disagreement. Based on this, we agreed on final codes and further clarifications and examples were added to the guidance document.

The coding task for the remaining 158 LAs was distributed such that each author checked 118 or 119 LAs and each LA was coded by three people. Coding took place in five rounds of 38–40 LAs. After each round, codes were compared and the majority decision for each concordance line was provisionally accepted as correct. After each round, we each received a set of spreadsheets showing how our codes compared to the majority decision. In a follow-up meeting, we then had the chance to appeal any cases in which we believed the majority to be incorrect and to discuss any issues or questions that had arisen. Final codes were agreed based on this discussion. The coding guidance document was updated after each meeting as required (see Supplementary Materials A for the final document).

Our list of candidate LAs included several embedded items. For example, *after* was listed, but so too were *after all* and *after that*. To avoid double counting, cases of longer adverbials were not counted as cases of the shorter. Thus, instances of *after all* and *after that* were not coded as instances of *after*. During the coding process, we also encountered several cases in which an item was used as part of a longer adverbial that was not listed. For example, the item *first* was used multiple times as part of the longer adverbials *first of all* and *at first*, which did not appear in our original list. If a longer form appeared in the corpus more than once, it was treated as a potential LA in its own right and added to our list. These additional items were coded in the same way as the others in a sixth round of coding, once the original list had been completed.

In one case, a group of items that were closely related in form and function were combined into a single item. Specifically, items that gave a quantity, then a time duration, then the word *later* (e.g. *a few days later; two months later; a second later*) were combined into a single item which we have labelled *N TIME later*.

If a longer adverbial appeared only once, it was counted as an instance of the listed term only if (1) the listed LA was the head of the adverbial phrase AND (2) the rest of the phrase could be deleted without substantially changing the meaning. This is seen, for example, in 6, where the listed term *to conclude* is embedded in the longer *to conclude with*, which is found only once in the corpus.

6. ***To conclude with****, these magnificent birds are terrifically adapted to the inhospitable ice country of Antarctica*.

Readers will note that 6 was accepted as an LA, despite its being, arguably, infelicitous. Because we aimed to capture development, we decided to include all obvious attempts to use an item, even if grammatically, semantically, or pragmatically imperfect. If a use was so infelicitous that the intended meaning was unclear, it was not counted as an LA.

Inter-rater reliability was calculated separately for each LA as the percentage of cases on which all three raters agreed. The average percentage agreement is shown for each round in Table 2. Because values were not normally distributed across LAs, median agreement it used to show the central tendency.

All attested LAs are listed, along with their frequency distributions, in Supplementary Information, B. The final coded spreadsheets are available on the IRIS Database (https://iris-database.org/details/CUspA-cAUBt).

Analysis

In quantifying the overall frequency of LA use, we take individual texts as our unit of analysis. This enables us to account for the variation in use between texts within a particular year group or genre. To this end, we first calculate the number of LA tokens in each text. These counts are divided by the total number of words in each text to normalize frequencies across texts of different lengths. These normalized frequencies are then multiplied by 100 to give an easily interpretable figure (LA tokens per 100 words).

Comparisons across year groups are made following the descriptive methods developed by Staples, Gray, Biber and Egbert (2023), Larsson, Sardinha, Gray and Biber (2023). Developmental trajectories are evaluated by calculating the mean normed frequency for texts in each year/genre along with 95 per cent confidence intervals surrounding the mean. As in Staples et al. (2023) and

Table 2. Median percentage agreement on codes by round.

Round	1	2	3	4	5
Median per cent agreement	67.12	96.43	87.19	100	95

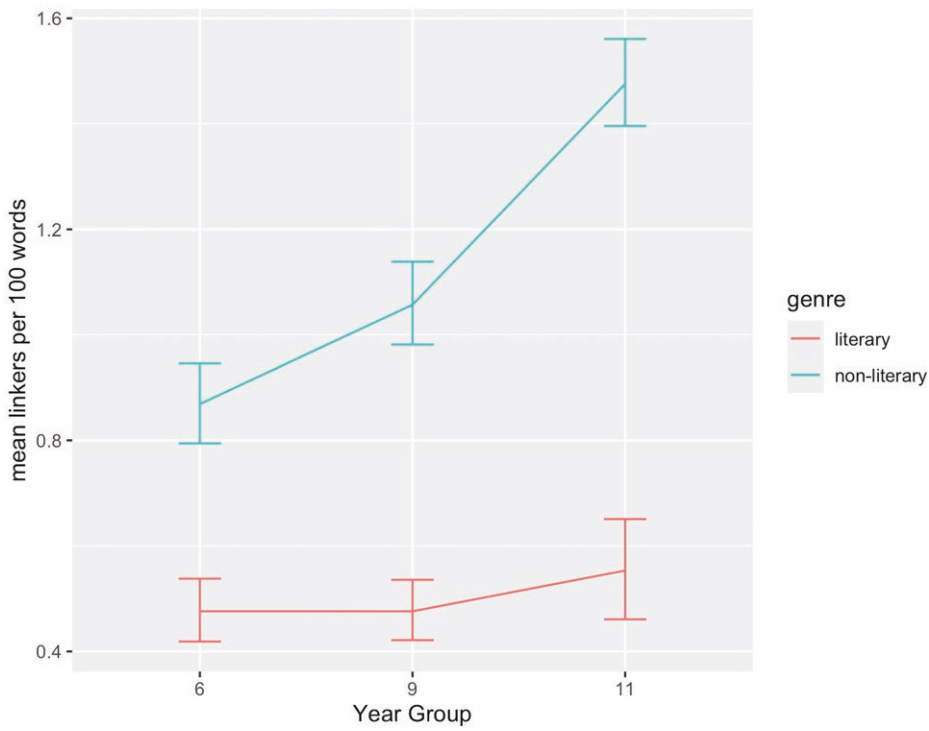

Figure 1. linker tokens per 100 words across genres and year groups.

Larsson et al. (2023), development is considered to have occurred if the 95 per cent confidence intervals around the means for the two groups do not overlap. Unlike those studies, we do not specify a minimum percentage difference necessary for change to have occurred. As we have argued elsewhere (Durrant, Dirdal and Tveitan 2024), the meaning of this number is variable, depending on the linguistic feature studied. Thus, rather than percentage change, we quantify the size of difference between year groups using a standardized effect size, *Cohen's D*.

Comparison of the use of individual LA types across year groups and genres does not allow us to use texts as the unit of analysis because most LAs are not found in most texts. For this part of the analysis, therefore, we use the overall normalized frequency of each item (i.e. occurrences per 10,000 words) within the subcorpus for each year*genre group. Because this approach obscures possible differences between texts, such that an item may attain high frequency within a group even if it is used in a small number of texts, these figures are supplemented with the percentage of texts in which each item is found. This enables us to show both how frequent an item is overall, and how widespread it is within a group.

Findings

Linking adverbial token frequencies across year groups, genres, and disciplines

Figure 1 and Table 3 show the mean frequencies of LA tokens across year groups and text genres along with their 95 per cent confidence intervals. These suggest: (1) a clear distinction between non-literary and literary writing, with the former using far more LA tokens than the latter; (2) a developmental trend by which the use of LAs in non-literary writing increases dramatically across year groups, while that in literary writing remains stable. The increase from Years 9 to 11 is especially strong (Cohen's $D = 0.44$), suggesting that LAs become especially prevalent in the later stages of secondary school.

Figure 2 and Table 4 show the mean frequencies of LA tokens across year groups and non-literary writing in different disciplines, along with their 95 per cent confidence intervals. Unsurprisingly given the smaller sample sizes, the increases across year groups are less robust than those for non-literary writing as a whole, but there are still non-overlapping confidence intervals across all year groups for English, from Years 6 to 9 for Humanities, and from Years 9 to 11 for Science. The confidence intervals for different disciplines overlap strongly in each year group. Thus, the overall frequency of use of LAs does not appear to vary across disciplines.

Comparing specific linking adverbials across genres and disciplines

To establish which LAs are substantially more frequent in one genre than the other, we compared the frequencies of each LA in the two genres. In comparisons of this sort, it is important to guard against giving undue prominence to very rare items. For example, the LA *soon enough*

Table 3. Mean linker tokens per 100 words across genres and year groups.

	Literary		Non-literary	
	Mean(sd)	**95 per cent CI**	**Mean(sd)**	**95 per cent CI**
Year 6	0.48(0.48)	0.42-0.53	0.87(0.89)	0.80-0.94
Year 9	0.48(0.44)	0.42-0.53	1.06(0.96)	0.98–1.14
Cohen's *D* Year 6–9	0.00			0.20
Year 11	0.55(0.39)	0.45-0.65	1.48(0.97)	1.39–1.56
Cohen's *D* Year 9–11	.19			0.44

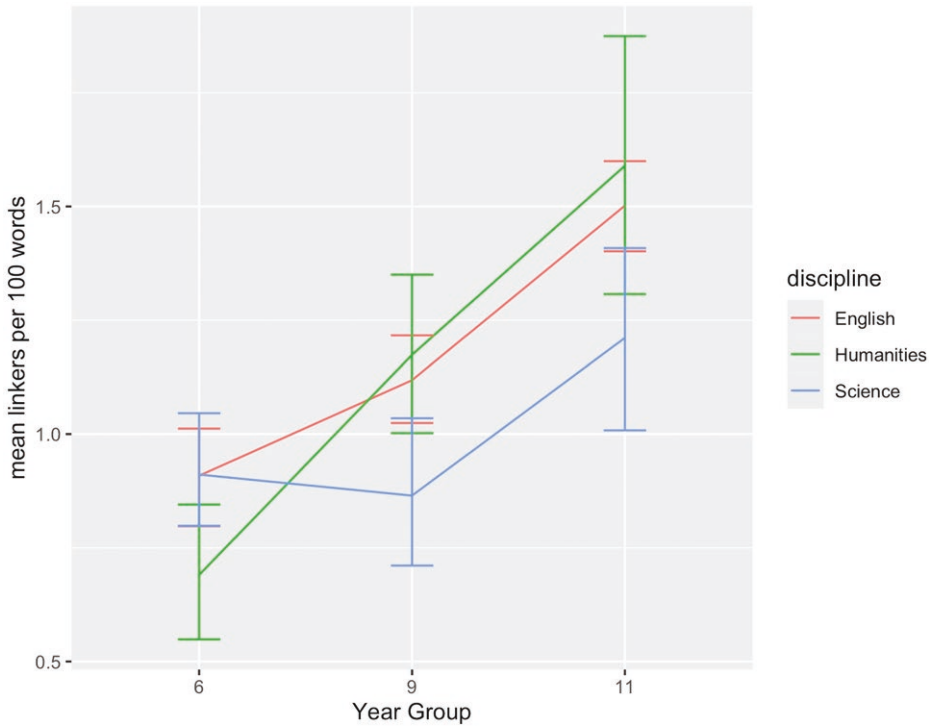

Figure 2. linker tokens per 100 words in non-literary writing across disciplines and year groups.

Table 4. Mean linker tokens per 100 words in non-literary writing across disciplines and year groups.

	English		Humanities		Science	
	Mean(sd)	95 per cent CI	Mean(sd)	95 per cent CI	Mean(sd)	95 per cent CI
Year 6	.91(0.95)	0.80–1.02	0.69(0.76)	0.54-0.84	0.91(0.93)	0.79–1.04
Year 9	1.12(0.86)	1.02–1.22	1.17(0.95)	1.00–1.35	0.86(1.12)	0.69–1.04
Cohen's *D* Year 6–9	.23		.56		-0.05	
Year 11	1.51(0.98)	1.41–1.61	1.59(1.00)	1.30–1.88	1.21(0.78)	1.01–1.42
Cohen's *D* Year 9–11	.42		.43		.36	

occurred three times in the literary sub-corpus and was absent in the non-literary corpus. It is therefore infinitely more frequent in the former than the latter. However, its very low frequency (approximately one appearance every 70,000 words in literary writing) means that we cannot say that it is distinctively literary. With this in mind, LAs were only considered potentially distinctive if they were reasonably frequent within a genre. There is, of course, no magic number beyond which a LA can be defined as 'reasonably frequent'. The balance needs to be maintained between reliability (by not setting the threshold too low) and including a substantive set of items (by not setting the threshold too high). Pilot analyses suggested that a mean frequency of at least one occurrence per 10,000 words yielded a set of LAs that achieved this balance. LAs meeting this criterion in either genre were considered distinctive if their frequency was at least double that in the other genre. Distinctive LAs are shown in Table 5.

Table 5. Linkers distinctive of each genre.

	Linker	Frequency/10k			Range (per cent)		
		Lit	Non-lit	Lit/non-lit	Lit	Non-lit	Lit/non-lit
Distinctive to literary	this time	1.12	0.00	Inf	3.30	0.00	Inf
	N TIME later*	2.33	0.71	3.27	7.82	1.30	6.03
	soon	1.30	0.49	2.64	3.71	0.94	3.95
	now	2.23	0.93	2.40	7.56	2.09	3.62
	eventually	1.93	0.86	2.25	6.60	1.99	3.31
	once again	1.14	0.57	2.00	4.17	1.61	2.59
Distinctive to non-literary	furthermore	2.67	0.00	Inf	6.38	0.00	Inf
	further	1.86	0.00	Inf	4.32	0.00	Inf
	here	1.83	0.00	Inf	3.83	0.00	Inf
	in conclusion	1.72	0.00	Inf	5.08	0.00	Inf
	overall	1.67	0.00	Inf	4.37	0.00	Inf
	firstly	1.34	0.00	Inf	3.30	0.00	Inf
	for example	4.21	0.13	32.28	8.68	0.53	16.38
	also	38.09	2.65	14.38	48.74	9.80	4.97
	therefore	5.35	0.51	10.42	10.98	1.82	6.03
	later on	1.08	0.17	6.26	2.62	0.64	4.12
	as well	2.34	0.59	3.95	5.05	2.33	2.17
	however	14.87	4.86	3.06	27.78	13.35	2.08
	next	1.47	0.50	2.93	2.61	2.01	1.30
	first	2.50	0.93	2.68	5.13	3.65	1.41

A large majority of LAs were distinctive. Of 27 items meeting the minimum frequency criterion in at least one genre, 20 (74 per cent) are distinctive, including 7 (26 per cent) that are not found in the other genre at all. This suggests that genres differ strongly, not only in the extent to which LAs are used but also in the specific LAs that are chosen.

Table 5 also suggests a broad functional divergence between genres. All distinctive literary LAs organize discourse in relation to time. In contrast, non-literary LAs commonly relate to organizing points in an atemporal list (*furthermore; further; firstly; also; as well; next; first*). This likely reflects the focus of literary writing on temporally sequenced narratives and of non-literary writing on non-temporal sequences of facts or arguments.

We have seen that non-literary texts use more LAs overall, and the list of distinctive LAs is, accordingly, longer and more diverse. In addition to listing items, these include LAs used to draw conclusions (*in conclusion; overall*); show logical links (*therefore; however*); and refer to points in an external text that have been cited (*here; later on*). This suggests that non-literary school writing employs a wider range of linking functions than literary writing.

We can get a more fine-grained view of LAs in non-literary writing by considering their distribution across disciplines. Tables 6–8 list items that were distinctive of each discipline, in the sense that they occurred at least once per 10,000 words and at least twice as frequently in that discipline than in each of the other two.

Table 6. Linkers distinctive of English non-literary.

Linker	frequency (per cent range)			frequency ratio	
	English	Humanities	Science	English / Humanities	English / Science
Additionally	1.25 (2.08)	.00 (0.00)	.30 (0.77)	Inf	4.20
Again	3.18 (7.46)	.79 (2.06)	1.13 (2.53)	4.05	2.81
Here	2.27 (4.97)	.28 (0.63)	.00 (0.00)	8.01	Inf
Later on	1.58 (3.78)	.37 (0.90)	.00 (0.00)	4.31	Inf
Well	2.44 (4.37)	.71 (0.63)	.40 (0.58)	3.45	6.03
Yet	2.83 (6.07)	.62 (1.51)	.56 (1.54)	4.58	5.10

Table 7. Linkers distinctive of Humanities non-literary.

Linker	Frequency (per cent range)			Frequency ratio	
	Humanities	English	Science	Humanities/English	Humanities/Science
Because of this	1.62 (4.58)	0.37 (0.83)	0.76 (1.93)	4.42	2.13
Consequently	2.28 (3.35)	0.33 (0.84)	0.00 (0.00)	6.84	Inf
In conclusion	4.17 (11.97)	1.52 (4.77)	0.92 (1.58)	2.75	4.54
Overall	3.49 (6.37)	1.51 (2.73)	1.17 (2.73)	2.30	2.99

Table 8. Linkers distinctive of Science non-literary.

Linker	Frequency (per cent range)			Frequency ratio	
	Science	English	Humanities	Science/ English	Science/ Humanities
After that	1.39 (3.30)	0.55 (1.00)	0.54 (0.92)	2.52	2.60
First	5.67 (10.36)	1.41 (3.13)	0.80 (1.94)	4.02	7.10
Lastly	1.43 (3.45)	0.40 (1.00)	0.47 (1.36)	3.57	3.05
Next	4.19 (6.46)	0.43 (0.83)	0.80 (1.94)	9.73	5.25
Thus	1.48 (1.57)	0.68 (1.65)	0.48 (1.20)	2.19	3.11

Tables 6–8 suggest substantial discipline-specificity, though this is less striking than the differences between genres. Of 32 LAs meeting the minimum frequency criterion, 15 (47 per cent) were distinctive. No LAs were exclusive to a single discipline. Distinctive LAs in Humanities and Science fall into clear functional groups. Humanities LAs relate to showing reasons (*because of this; consequently*) and drawing general conclusions (*in conclusion; overall*), both functions reflecting the explanatory text types frequent in History writing, especially at secondary school (Coffin 2006). Science LAs relate to sequences (*after that; first; lastly; next*) and making logical links (*thus*). The former are used primarily in descriptions of experimental processes, as in Example 6.

6. ***First*** *I labelled both cups one saying yeast and water and the other saying yeast, water and sugar. < gram > Then I added four desserts of flour in each cup.</gram >* ***Next*** *I added a solution of yeast to cup A until it had the consistency of a thick, milkshake.* ***After th****at, I added enough yeast water and sugar to the solution.*

Thus was mostly used to describe cause-effect relationships (Example 7).

7. *Cloning could help solve this problem by increasing the population of that species and* ***thus*** *bringing it back into the animal kingdom without danger of re-extinction in the foreseeable future*

Its use appears to be strongly dependent on personal preference; all seven appearances in Year 9 Science come from a single author, as do both uses at Year 11.

English LAs are more diverse: *additionally* and *again* are both used primarily to add points to an explanation or argument. *Here* and *later on* refer to points in a literary text under discussion. *Yet* is a contrast marker. *Well* is a characteristically spoken-style LA (as the usage notes in the Cambridge Dictionary attest[2]), which tends to look inappropriately informal in these contexts (see examples 8-9). At Years 9 and 11, English is the only discipline where this is used in non-literary writing, suggesting that this subject area may be more likely to elicit such a spoken style.

8. *Snow can cause accidents and can damage the economy if it snows enough. How?* ***Well*** *first of all snow means ice, which can cause someone to slip and fall.* (English Year 11 non-literary, 13_762d).
9. *And I disagree with that festivals are bad,* ***well*** *it a festival what do expect.* (English Year 11 non-literary, 23_1112a).

Specific linking adverbials across year groups and genres

We turn now to the question of which LAs increase or decrease in use across year groups and the extent to which these patterns are genre-specific. Because of the smaller number of texts per year group in each discipline, we will not attempt a full analysis of development within disciplines, but reference will be made to disciplinary factors where a strong pattern is evident.

As before, we apply a mean frequency threshold of one occurrence per 10,000 words. Because the present analysis focuses on development across year groups, items need to meet this criterion within an individual year*genre combination. For example, *indeed* was not considered as a developing LA within the non-literary genre because it did not occur more than once per 10,000 words for any year group within this genre. In contrast, *further* was included because, although it did not meet the criterion in Years 6 and 9, it exceeded it in Year 11.

Change in frequency from one year group to the next was quantified as a percentage increase or decrease. Of course, no LAs maintained identical frequencies across year groups. We therefore needed to define a level of change that could be considered developmentally interesting. As with the minimum frequency threshold, trial-and-error identified a cut-off that yielded a good number of items while maintaining a robust difference between year groups. Items were considered developmental if they showed a consistent frequency increase or decrease of at least 10 per cent both from Year 6 to Year 9 and from Year 9 to 11. LAs whose direction of change reversed (e.g. increasing from Years 6 to 9 but decreasing from Years 9 to 11) were not included. Table 9 shows the developmental LAs in each genre.

A small number of LAs show consistent developmental patterns regardless of genre: *however, yet, instead*, and *in fact* increase in frequency in both genres; *well* decreases in both. Of the four shared increasing LAs, three mark relations of contrast, suggesting this is an important area of development, regardless of genre. *In fact* is much less frequent than these. Though increasing in both genres, its characteristic use differs in each. In literary writing, it is primarily used to reinforce a previous statement (example 10). In non-literary, it is more often contrastive, (example 11).

10. *The city of Karachi has never been a safe place.* ***In fact****, the moment you step off the plane, there are armed policemen along the gangway, and the menacing looks of all the officials in the airport strike fear into my heart.* (Year 11 English literary, 22_1063f).

11. This shows how Mr Birling acts as though he knows what is going on in the world when he **in fact** has no idea. (Year 11 English non-literary, 11_608).

As we noted above, the decreasing LA *well* is markedly spoken in style and looks out of place in non-literary writing. Its decrease therefore suggests development towards the more formal writerly style expected at school.

The remaining LAs in Table 9 show genre-specific developmental patterns. Five of the LAs that increase only in non-literary writing have the overlapping function of adding points to a sequence (*also; again; furthermore; further*, and *lastly*). This function therefore seems to increase in prominence as students progress in this genre.

A second group of items increasing only in non-literary writing are three LAs used to draw conclusions or summarize previous points: *in conclusion* and *overall,* and *ultimately*. As we have seen, *in conclusion* and *overall* are particularly characteristic of explanatory texts in History. *Ultimately* is usually used in a similar way, either drawing together previous points into a conclusion or discussing the outcome of events in a text that is being discussed. While occurring with similar frequencies in English and Humanities writing, it is absent from Science.

Three further items increase in non-literary writing: *for example, here* and *thus. For example* is, overall, the fourth most frequent LA in non-literary writing. By Year 11, it is found in 14 per cent of all non-literary texts, suggesting the increasing importance of this function.

We have already seen that *here* is strongly associated with writing for English. Indeed, at Years 9 and 11, it is found only in writing for English classes. Its characteristic use is to refer to a part of a text that is being studied (as in Examples 12 and 13). While rather infrequent at Year 6 its increase to Year 11 (per cent) is the second highest of all LAs (behind only *further*).

12. *When Lady Macbeth is alone she opens with 'Nought's had, all's spent'. **Here** she is saying that she feels she has got nothing from Duncan's death.* (Year 11 English non-literary, 23_1088b).
13. *Sissy's thought process is clever as she thinks outside the box when deciding if it's a prosperous nation. **Here** she shows true wisdom* (Year 11 English non-literary, 22_1060e).

As we noted above, *thus* is strongly associated with Science writing and its use appears to be dependent on personal preference. At Year 11, it is also rather frequent in English non-literary writing (1.8/10,000 words). As in Science, its use here appears to be highly personal: of 24 uses, 13 come from just two writers.

The functionally related LAs *first* and *next* are the only items to increase exclusively in literary writing. Interestingly, their use actually decreases in non-literary writing. Thus, while in Year 6 both LAs are strongly distinctive of non-literary writing (*first* being 8 times, and *next* 35 times, more frequent than in literary writing), at Year 11 this relationship has reversed (*first* being 1.6 times, and *next* 4 times, more frequent in literary than non-literary writing).

In literary writing, they are used to order points in a narrative; their increased use may therefore be due to increased temporal complexity. In non-literary writing, both LAs function to structure sequences. We saw above that they occur most frequently in Science writing to explain the steps in an experimental procedure. In other disciplines, they are used to order points in a historical narrative or in instructions. At Years 9 and 11, historical narratives and instructions become rare, which partly accounts for their decrease. While experimental reports are common at higher year groups, writers either do not use LAs (often preferring numbered lists) or use a wider range of LAs, such that *next* becomes less prominent. This suggests that, even when task is held relatively constant, use of this LA decreases. So, in non-literary writing, this appears to be an early-developing item whose use decreases as students' repertoire of linking techniques increases (e.g. through increased use of the sequencing LAs discussed above). This may also partly explain the decreasing frequency of other sequencing LAs in non-literary writing (*after that; N TIME later; now*).

Two items decrease only in literary writing: *finally* and *anyway*. The use of *finally* in this genre is primarily not in its 'listing' function, but rather in the sense of 'later than wanted', and tends to

Table 9. Developmental linkers across year groups in each genre.

	Literary							Non-literary						
		Frequency (per 10K words)			Range (per cent)				Frequency (per 10K words)			Range (per cent)		
		Year			Year				Year			Year		
	Linker	6	9	11	6	9	11	Linker	6	9	11	6	9	11
Increasing frequency	however	2.623	4.146	7.822	6.93	10.91	22.22	also	87.3	115.286	153.345	70.16	75.34	90.32
	yet	1.211	2.879	6.258	4.38	8.64	19.05	however	26.96	38.773	48.523	33.16	45.89	67.16
	instead	0.605	0.806	3.129	2.19	3.18	12.7	for example	5.41	16.666	22.544	8.73	28.94	45.68
	first	0.706	0.921	1.173	2.55	3.64	4.76	furthermore	1.284	5.372	5.967	2.27	10.1	13.68
	in fact	0.101	0.461	1.173	0.36	1.82	4.76	again	0.55	1.928	5.546	0.87	3.94	14.32
	next	0.101	0.23	1.173	0.36	0.91	4.76	yet	0.917	2.41	4.581	1.75	4.97	11.37
								further	0.459	2.204	4.099	0.87	4.45	8.42
								here	0.183	0.826	4.581	0.35	1.88	10.74
								in conclusion	0.275	0.826	4.4	0.52	1.71	9.26
								overall	0.734	1.584	2.833	1.4	3.94	9.89
								instead	0.275	2.135	2.592	0.52	4.79	7.79
								thus	0.459	0.758	1.808	0.87	1.88	5.47
								in fact	0.367	0.689	1.567	0.7	1.03	3.79
								ultimately	0.367	0.62	1.145	0.7	1.54	3.37
								lastly	0	0.138	1.386	0	0.34	3.16

Table 9. Continued

	Literary							Non-literary						
		Frequency (per 10K words)			Range (per cent)				Frequency (per 10K words)			Range (per cent)		
		Year			Year				Year			Year		
	Linker	**6**	**9**	**11**	**6**	**9**	**11**	**Linker**	**6**	**9**	**11**	**6**	**9**	**11**
Decreasing frequency	well	9.52	3.43	2.649	1.173	12.04	7.73	first	5.686	1.102	0.723	10.3	2.57	2.53
	too	8.8	3.531	1.612	1.173	11.68	6.36	as well	2.934	2.204	1.869	4.71	4.97	5.47
	finally	5.57	2.421	0.921	0.782	8.03	3.18	too	2.843	1.446	1.206	4.89	3.25	4
	anyway	3.77	1.513	0.806	0	5.11	3.18	well	3.21	1.24	0.301	4.89	2.23	0.84
								next	3.485	0.62	0.301	5.58	1.2	1.05
								now	1.1	0.964	0.723	1.75	2.4	2.11
								N TIME later	1.467	0.482	0.181	2.27	1.2	0.42
								after that	1.376	0.482	0.181	2.27	1.2	0.63
								soon	1.009	0.413	0.06	1.75	0.86	0.21

imply a sense of impatience or discomfort with a process (as in example 14). The sharp decrease in its use by Year 11 may indicate that this emotion is more characteristic of younger writing.

14. ***Finally***, *he woke up from his nightmare* (Year 6, English literary, 2_72b)

Anyway is very spoken in style and is often, though not always, used in dialogue (see Example 15). It seems to be strongly dispreferred in more mature writing, its use decreasing to zero at Year 11. As with the decrease in *well*, this suggests a move away from a spoken style.

15. *He hugged me in relief. '****Anyway****, time for dinner'.* (Year 6, English literary, 15_840b)

Discussion and conclusions

Our analysis has shown that LAs were used much more in non-literary than literary writing. This gap widened over time, with use increasing across year groups in the former but not the latter. Greater use of LAs thus appears to be an important marker of maturity for non-literary, but not for literary, texts.

The majority of individual LAs were also strongly genre-specific. Many of these reflect a functional difference between a literary focus on temporal events and a non-literary focus on facts and arguments. Non-literary texts also used LAs for a range of functions that are not common in literary writing; most notably, drawing conclusions and showing logical links. Our analysis of *in fact* further showed how individual LAs can have different functions in the two genres. Such fine-grained genre differences could be usefully explored by future research.

Together, these patterns imply that, as Rose and Martin (2012) have argued with regard to writing skills more generally, story writing is unlikely to adequately prepare students to use LAs in the non-literary genres that are central to academic success. It also highlights how development in writing can vary depending on the types of writing in which learners are engaging, emphasizing the importance of including a detailed account of text types at the heart of any developmental model (Author; Britton et al. 1975; Coffin 2006). While the present research has employed a broad division into two categories, there is reason to believe that a more elaborated model of text types will provide both more adequate developmental models and more specific pedagogical guidance (see Durrant 2022 for discussion of the challenges and possibilities for such development).

The finding that linkers were more frequent in non-literary writing contrasts strikingly with Barbaros and Akbaş (2022), who, using the same corpus, came to the opposite conclusion. One possible explanation is the range of items analysed. Unlike the present study, Barbaros and Akbaş (2022) included both LAs and subordinating conjunctions, and focused only on contrastive items. This suggests that contrast markers and/or subordinating conjunctions may be more frequent in literary than non-literary texts. This would be an interesting finding, further complexifying the story we have told above. However, further research is needed to confirm this systematically[3].

Another potential explanation is methodological. Whereas the current study used individual texts as the unit of analysis, Barbaros and Akbaş (2022)used normalized counts across each genre sub-corpus, not accounting for variation between texts. Moreover, their corpus sample was skewed towards younger writers. Unlike the present study, it included Year 2 texts and more texts were included from the lower than the higher year groups. For example, it included over 110K words of Year 6 literary writing but less than 27K words for Year 11. The overall comparison between genres was therefore strongly weighted towards Year 6 and away from Year 11. This is important because Year 11 was the one group in Barbaros and Akbaş (2022)found contrastives to be more common in non-literary writing.

Unlike some studies of mature academic writing (Peacock 2010; Gao 2016; Alanazi 2024), we did not find a quantitative difference in the total numbers of linkers used across disciplines. However, we did find variation in the individual linkers preferred across disciplines. This reflected

functional differences, with English texts making greater use of LAs that refer to points in another text under discussion (*here; later on*); Humanities texts drawing conclusions (*in conclusion; overall*) and showing logical links (*because of this; consequently*); and Science texts sequencing events in an experiment (*after that; first; lastly; next*). This illustrates how different subject areas present different linguistic challenges and different opportunities for learning key aspects of academic writing. This underlines the importance of teachers across all areas being aware of and paying explicit attention to such features.

Turning to variation across year groups, the increase found for non-literary writing supports our hypothesis that *pathbreaking* (Ellis and Ferreira-Junior 2009) items (*and, but, so, then*), which decrease in frequency as writers mature (Crowhurst 1987; Barbaros and Akbaş 2022), have obscured the overall increase in LA use in previous studies. We can therefore suggest a developmental model whereby writers start from a small repertoire of pathbreaking items that are common in English and which young writers repeat frequently. As they mature, writers come to use these items less and, simultaneously, adopt a wider repertoire of items. These conflicting tendencies may make the overall prevalence of LAs relatively stable across year groups. Thus, it is not the overall need for linking that changes as learners mature, but rather the repertoire of items that are used to perform that function.

This study has differed from much previous research by not categorizing LAs semantically from the outset. We have argued that semantic categories are not objective types; that there is no principled way of choosing between the diverse, and often contradictory, taxonomies available; that categories have low interrater reliability; and that predefined categories may conceal important semantic distinctions. We chose, instead, to employ functions in an interpretive way to aid the discussion of key patterns of use that emerged from our data. This analysis has shown some semantic types to be important to understanding variation in LA use. Key examples are the distinction between temporal (prominent in literary writing) and nontemporal sequences (prominent in non-literary writing); the prevalence of logical and concluding links in non-literary but not in literary writing; and the overall increase in contrastive markers across year groups. While these categories are common in taxonomies, others are not. The use of *here* to refer to quotations from cited texts, which was strongly distinctive of non-literary writing in English is not widely used (though it could be seen as an example of the *here and now* relationship referred to by Halliday and Hasan [1976]). The difference in the use of temporal LAs between Science writing and literary writing can be understood only by linking the former to the specific context of sequenced experimental procedures, a shade of meaning not captured by taxonomies. Similarly unattested but important for understanding our data was the use of *finally* to express impatience, and of *in fact* to reinforce a statement. Such examples underline the utility of a contextually sensitive and interpretive approach to LA functions.

Future frameworks for LA writing instruction benefit from the genre and discipline-based distinctions we have identified, ensuring that teaching supports contextual approaches to teaching LAs. Methodologically speaking, our study also demonstrates the importance of adopting a flexible, context- and data-driven approach to analysis that is sensitive both to genre and discipline-specific language and that enables targeted pedagogical strategies and a more nuanced understandings of writing development across academic settings.

Supplementary data

Supplementary data is available at *Applied Linguistics* online.

Funding

The Growth in Grammar corpus was created as part of a project supported by a grant from the Economic and Social Research Council (ES/M00967X/1).

Notes on contributors

Philip Durrant is Professor of Applied Linguistics and Education at the University of Exeter. He has been a language teacher and researcher for over twenty-five years, working in both the UK and Türkiye. He has published widely on corpus linguistics, vocabulary learning, and the language of school and university writing.

Erdem AKBAŞ holds his PhD from the University of York, UK (2015) and now works as an Associate Professor at Erciyes University, Türkiye. Presenting his research extensively at various reputable international conferences, he published articles in national and international journals/books. Dr. Erdem Akbas co-organized the first-of-its-kind conference on Metadiscourse in 2017, 2019, and 2021; published an edited book and two edited special issues on Metadiscourse.

Elif Barbaros is a lecturer of General English at the School of Foreign Languages at Erciyes University, Türkiye. She has an MA in vocabulary learning strategies, vocabulary size, and lexical threshold. She is currently a PhD candidate, working on corpus linguistics. She has published a co-authored article on children's writing.

Arwa Aldawood is a lecturer of Applied Linguistics at the University of Prince Sattam bin Abdulaziz. She has been a language teacher for over five years, working in Saudi Arabia. She is now a PhD candidate at Exeter University. She has published some articles on corpus linguistics, language learning, and learner writing.

END NOTES

[1] https://phildurrant.net/creating-the-growth-in-grammar-corpus/
[2] https://dictionary.cambridge.org/grammar/british-grammar/well
[3] A further difference is that Barbaros and Akbaş's (2022) analysis included the high-frequency item *but*. However, reanalysis of Barbaros and Akbaş's (2022) data show that this cannot explain the conflicting results, as this item bucked the overall trend, being more common in non-literary than literary writing.

References

Alanazi, Z. (2024) 'An Exploratory Analysis of Linking Adverbials in Research Articles Across Different Disciplines.', *Theory and Practice in Language Studies*, 14: 1181–92. https://doi.org/10.17507/tpls.1404.26

Anthony, L. (2022). *AntConc (Version 4.1.4): Waseda University*. Retrieved from https://www.laurenceanthony.net/software. Accessed 27 October 2022.

Barbaros, E., and Akbaş, E. (2022) 'The Developmental Trajectory of Contrast Markers in Children's Writing Across Four Grade Levels', *Lingua*, 279: 103427. https://doi.org/10.1016/j.lingua.2022.103427

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (2021). *Grammar of Spoken and Written English*. Amsterdam: John Benjamins Publishing Company.

Britton, J., Burgess, T., Martin, N., and Rosen, H. (1975). *The Development of Writing Abilities*, pp. 11–8. London and Basingstoke: Macmillan.

Carter, R., and McCarthy, M. (2006). *Cambridge Grammar of English*. Cambridge: Cambridge University Press.

Coffin, C. (2006). *Historical Discourse: The Language of Time, Cause and Evaluation*. London: Continuum.

Crowhurst, M. (1987) 'Cohesion in Argument and Narration at Three Grade Levels.', *Research in the Teaching of English*, 21: 185–201. https://doi.org/10.58680/rte198715585

Durrant, P. (2014) 'Discipline and Level Specificity in University Students' Written Vocabulary.', *Applied Linguistics*, 35: 328–56. https://doi.org/10.1093/applin/amt016

Durrant, P. (2022) 'Studying Children's Writing Development With a Corpus.', *Applied Corpus Linguistics*, 2: 100026. https://doi.org/10.1016/j.acorp.2022.100026

Durrant, P., Dirdal, H., and Tveitan, V. D. (2024) 'Vocabulary Sophistication in Children's L2 School Writing.', *International Journal of Learner Corpus Research*. https://doi.org/10.1075/ijlcr.23025.dur

Ellis, N. C., and Ferreira-Junior, F. (2009) 'Construction Learning as a Function of Frequency, Frequency Distribution, and Function', *The Modern Language Journal*, 93: 370–85.

Gao, X. (2016) 'A Cross-Disciplinary Corpus-Based Study on English and Chinese Native Speakers' Use of Linking Adverbials in Academic Writing', *Journal of English for Academic Purposes*, 24: 14–28. https://doi.org/10.1016/j.jeap.2016.08.002

Halliday, M. A. K., and Hasan, R. (1976). *Cohesion in English*. Harlow: Longman.

Huddleston, R., and Pullum, G. K. (2002). *Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hyland, K., and Tse, P. (2007) 'Is there an 'Academic Vocabulary?'', *TESOL Quarterly*, 41: 235–53. https://doi.org/10.1002/j.1545-7249.2007.tb00058.x

Hyland, K., and Tse, P. (2004) 'Metadiscourse in Academic Writing: A Reappraisal.', *Applied Linguistics*, 25: 156–77.

Larsen-Freeman, D., and Celce-Murcia, M. (2016). *The Grammar Book: Form, Meaning, and Use for English Language Teachers*. Boston: National Geographic Learning.

Larsson, T., et al. (2023) 'Exploring Early L2 Writing Development Through the Lens of Grammatical Complexity.', *Applied Corpus Linguitics*, 3: 100077. https://doi.org/10.1016/j.acorp.2023.100077

Liu, D. (2008) 'Linking Adverbials: An Across-Register Corpus Study and Its Implications.', *International Journal of Corpus Linguistics*, 13: 491–518. https://doi.org/10.1075/ijcl.13.4.05liu

Nippold, M. A., Ward-Lonergan, J. M., and Fanning, J. L. (2005) 'Persuasive Writing in Children, Adolescents, and Adults: A Study of Syntactic, Semantic, and Pragmatic Development.', *Language, Speech, and Hearing Services in Schools*, 36: 125–38. https://doi.org/10.1044/0161-1461(2005/012)

Peacock, M. (2010) 'Linking Adverbials in Research Articles Across Eight Disciplines.', *Iberica*, 20: 9–34.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. *London and* NY: Longman.

Rose, D., and Martin, J. R. (2012). *Learning to Write, Reading to Learn: Genre, Knowledge and Pedagogy in the Sydney School*. Sheffield: Equinox.

Staples, S., et al. (2023) 'Writing Trajectories of Grammatical Complexity at the University: Comparing L1 and L2 English Writers in BAWE.', *Applied Linguistics*, 44: 46–71. https://doi.org/10.1093/applin/amac047

Yin, Z. (2016) 'Register-Specific Meaning Categorization of Linking Adverbials in English', *Journal of English for Academic Purposes*, 22: 1–18. https://doi.org/10.1016/j.jeap.2016.01.004