“Teens like science, not science class, study finds”
Clausal evidential parentheticals in Contemporary American English

María José López-Couso
ORCID: 0000-0003-3543-7408
Belén Méndez-Naya
ORCID: 0000-0003-2238-2364

The present paper focuses on a type of evidential parenthetical featuring a noun phrase subject with the noun *study* combined with a verb phrase, as in *Teens like science, not science class, study finds*. These parentheticals are recorded from the early 1940s in American English but become frequent only in the 1990s. Using data from the *Corpus of Contemporary American English* (COCA), we show that *study*-parentheticals are attested in two different patterns, which differ as regards the relative order of subject and verb, the non-inverted pattern being the most frequent in the data. Different variants have been identified for each pattern, depending on the complexity of the subject noun phrase (presence/absence of determiners and modifiers). Our analysis reveals that these structures are closely associated with the written medium, in particular with popular science journalistic texts. We argue that *study*-parentheticals show incipient grammaticalization and that the different parenthetical variants can be classified as constructional theticals (Kaltenböck, Heine, & Kuteva 2011), with one particular sequence, *study finds*, coming

https://doi.org/10.51814/ufy.1041.1460

close to a formulaic thetical which has become a staple in pop-sci headlinese.

Keywords: parentheticals, evidentiality, grammaticalization, journalistic texts, headlinese, American English.

1. Introduction

Evidentiality is a semantic notion concerned with the indication of one's source of information. In languages such as Tibetan or Quechua evidentiality is an obligatory grammatical category (Aikhenvald 2004: 3), while others make use of various types of evidential strategies, including both lexical means and grammaticalized expressions (Aikhenvald 2004).[1]

English lacks a fully grammaticalized evidential system but possesses various strategies to convey epistemic and evidential meanings, ranging from the most grammaticalized ones (e.g., modal verbs) to others showing lower degrees of grammaticalization, such as parentheticals (e.g., *it seems*) and adverbs (e.g., *perhaps*) (cf. Chafe 1986: 261; Mélac 2022: 340). Over the last few years, we have devoted some space to clausal parenthetical constructions in both historical and contemporary English, with a focus on various patterns used with epistemic and/or evidential value. In this chapter we turn our attention to a further evidential parenthetical type, which has become relatively frequent in the last decades. The examples in (1) and (2) below illustrate this recent pattern, which features a third person subject in the form of a phrase that contains a noun denoting an examination or an investigation, such as *study* or *research*, accompanied by a verb. The analysis of these structures shows their close association with a specific genre, namely the press.

1 This research is funded by the Spanish Ministry of Science and Innovation (National Programme for Excellence in Scientific and Technical Research; grant PID2020-114604GB-I00) and the Regional Government of Galicia, *Consellería de Cultura, Educación, Formación Profesional e Universidade*, grant ED431B 2023/03. We also thank an anonymous reviewer and the editors of this volume for their insightful comments. The usual disclaimers apply. The chapter is dedicated to Päivi Pahta, friend and colleague, whose work has always been an inspiration for us.

1 This targeted ridicule, *studies show*, can have lasting effects. (COCA, MAG, 2015, Newsweek)

2 Total or near-total blockage of two or perhaps one major artery may warrant angioplasty. Those noninvasive steps can frequently relieve angina and lengthen life, *research suggests*. (COCA, MAG, 2007, Consumer Reports)

The chapter is organized in the following way. Section 2 sets the scene by providing a characterization of clausal parentheticals; special attention is paid here to their syntactic and prosodic independence, their positional mobility, and their semantico-pragmatic features. Section 3, in turn, is concerned with methodological issues, including the materials used for the analysis (Section 3.1) and the searches run in the *Corpus of Contemporary American English* (COCA; Davies 2008–) (Section 3.2). The analysis of the data is the focus of Section 4, where we describe the parenthetical patterns identified in the corpus, their variants, and their distribution in the short diachrony (Section 4.1), the behaviour of our parentheticals as regards the position they occupy in relation to their host clauses (Section 4.2), the verbal predicates which occur in the different parenthetical variants (Section 4.3), and the distribution of the parentheticals under examination across genres (Section 4.4). Section 5 is devoted to the discussion of how the data described in Section 4 fit in with accounts on the origin and development of clausal parentheticals, special attention being drawn to several formal indicators of grammaticalization. Finally, Section 6 offers a brief conclusion and some suggestions for further research.

2. Setting the scene: Clausal parentheticals

Parentheticals are words, phrases, or clauses which are independent from their host or anchor both syntactically and prosodically, being set off from the rest of the utterance by pauses (and, typically, by commas in writing). The information conveyed by parentheticals is optional (Boye & Harder 2021) and their meaning is non-restrictive (Kaltenböck, Heine, & Kuteva 2011: 857).

Parentheticals may take diverse forms and show different degrees of syntactic complexity. The categories that can occur parenthetically range from clauses (e.g., *I think*, *as you say*, *what is more*) and phrases of various types

(e.g., *between you and me*) to adverbs (e.g., *frankly*), interjections (e.g., *damn*), and discourse markers (e.g., *like*) (Kaltenböck 2007: 29–31; Dehé 2009; Kaltenböck, Heine, & Kuteva 2011: 856, 876). In this chapter we are concerned with a specific type of clausal parentheticals, namely those structures resembling a matrix clause (Quirk *et al.* 1985: 1112) which also have "a non-parenthetical use in which they take a declarative content clause as complement" (Huddleston & Pullum *et al.* 2002: 895). A paradigmatic instance of parentheticals of this kind is *I think*, exemplified in (3a) and (4a), while (3b) and (4b) illustrate the corresponding matrix clause counterparts.

3 a. Doing the show from Philadelphia during the Pennsylvania primaries was, *I think*, hard enough. (COCA, SPOK, 2011, NPR_FreshAir)
b. I think (that) doing the show from Philadelphia during the Pennsylvania primaries was hard enough.

4 a. She would count as a groupie, *I think*. (COCA, MAG, 2015, Esquire)
b. I think (that) she would count as a groupie.

Despite the similarities between the (a) and (b) sequences in (3) and (4), the two constructions show some crucial differences. For example, from a syntactic point of view, *I think* in (3a) and (4a) functions as an adverbial, more specifically, as a disjunct (Quirk *et al.* 1985: 612 ff.), while the rest of the utterance is the main clause. In (3b) and (4b), by contrast, *I think* is the matrix of the subordinate clause that follows, which functions as a complement of the predicate *think*. Another relevant difference between the (a) and the (b) counterparts in (3) and (4) concerns argument structure: while in the sequences in (b) the verb *think* takes a clausal complement, in the (a) examples *think* lacks the object argument, which suggests that parentheticals like *I think* are syntactically defective; hence the terms 'gap-containing parenthetical clauses' (Kaltenböck 2007: 41) or 'reduced parenthetical clauses' (Schneider 2007: 76).

The (a) and (b) sequences in (3) and (4) also differ pragmatically. Thus, in (3a) and (4a) the underlined statements constitute the primary focus of discourse, that is, they convey the most salient information, while *I think* is secondary. For instance, (4a) does not describe an act of thinking; it is rather about her being a groupie. The discourse secondariness of *I think* can be probed by addressability tests (Boye & Harder 2007: 581–585): only discourse-

primary elements can be taken up by question tags, as in (5a), or can be addressed by *really*-queries, as shown in (5b):

5 a. She would count as a groupie, I think, wouldn't she?
b. Really? (i.e., does she really count as a groupie?)

In complementation structures, on the other hand, it is the matrix clause *I think* that is discourse-primary (see Boye & Harder 2007; 2012). This becomes evident in cases where the matrix *I think* is questioned (6a) or is addressed by *really* (6b).

6 a. Well, *what do you think*? *I think* that I'm done looking for answers around here, *that's what I think*. (COCA, 2012, TV, Crazy)
b. Really? (i.e., do you really think that?)

Deriving from their syntactic independence, parenthetical clauses such as *I think* enjoy flexibility when it comes to positional mobility: they may occur before their host, in medial position, or after their host. The literature on parentheticals (see, e.g., Heine & Kaltenböck 2021: 11 for a recent example)[2] has often noted the controversial status of clause-initial clauses of the type shown in (7), which can be indeterminate between a matrix clause reading (taking a zero clause as complement) and a parenthetical interpretation:

7 MATTHEWS: Yeah, I agree with you, *I think* the moment was important. *I think* he was a little angry, little fatigued. (COCA, SPOK, 2009, NBC Matthews)

In most cases, however, this ambiguity cannot be easily solved; it is only when the sequence is followed by a non-declarative complement, as the interrogative clause in (8) below, or when some hesitation sound (e.g., *uh*, *uhm*) or other fillers intervene, as in (9), that the parenthetical analysis seems to be preferrable (see Kaltenböck 2007: 45; Brinton 2008: 12).

---

2 Heine and Kaltenböck (2021) use the label 'thetical' instead of parenthetical. The term thetical was first used in Kaltenböck, Heine, and Kuteva (2011) precisely on account of the fact that not all parentheticals occur in interpolated position (2011: 856).

8 *I mean*, can you think of any other situation, Pop, when a man gets so close to a woman except when he is actually making love to her? (1981 Rendell, *The Best Man to Die* [BNC]; from Brinton 2008: 12)

9 Gloria: (Chuckles) Well, *I think, um*, Candy's assets will work out very nicely with the trust fund set. (SOAP, 2010, Young and Restless)

In terms of their semantics, the most common parenthetical clauses are those expressing the speaker's evaluation of the proposition encoded by the host (e.g., *I think*, *I guess*, etc.). Parentheticals of this kind are known as 'comment clauses' (cf. Quirk *et al.* 1985: 1112–1118; Brinton 2008). Another common type of clausal parenthetical is that featuring predicates entailing some sort of communication. These parentheticals have sometimes been referred to as 'reporting clauses' (e.g., *he said*, *it is said*) (Kaltenböck, Heine, & Kuteva 2011: 856; see also Vandelanotte 2009: chapter 8; Cichosz 2018: 183). An example of a reporting clause is provided in (10).[3]

10 Deep below, *it is rumored,* lies Saddam's main bunker, (COCA, 1996, MAG, Saddam's Inferno)

Among the wide range of clausal parentheticals available in English, the present chapter is concerned with a type of reporting clauses for which we use the umbrella term *study*-parentheticals. As mentioned above, this parenthetical type features a third person inanimate noun phrase as subject, in contrast to paradigmatic parentheticals, such as *I think* or *you know*, which contain a first- or second-person subject pronoun. The nouns occurring in the construction typically denote some kind of investigation or examination. Besides *study* (example (1) above) and *research* (example (2)), representative illustrations are the nouns *report*, *analysis*, *results,* and *survey*, among others. For our purposes in this chapter, however, we focus solely on the noun *study*, not only because it is the most frequent item occurring in the construction in our preliminary searches, but also because it is the most neutral and general noun in the set from a semantic point of view.

---

3 Quirk *et al.* (1985: 1114) would subsume the two groups of clausal parenthetical constructions discussed here under 'comment clauses'.

3. Methodology

3.1. Materials

As stated in Section 1, our study is based on data retrieved from the *Corpus of Contemporary American English* (COCA), which contains over one billion words of text, with 25 million words allotted for each year between 1990 and 2019. Due to its sheer size, COCA is ideal to examine low-frequency constructions, such as *study*-parentheticals, in the short diachrony. Moreover, sample searches in the *Corpus of Historical American English* (COHA; Davies 2010) demonstrated the recency of this parenthetical type: the earliest unambiguous examples of the construction are attested in the early 1940s (see example (11)), but they are highly sporadic in the data. It is only from the 1990s onwards that *study*-parentheticals increase in frequency in the COHA material, that is, precisely in the timespan represented in COCA.

11 High-pressure salesmanship and instalment methods have taken millions from the working class annually, *the studies showed*, with 70 percent of all policies lapsing and another 20 percent terminating by surrender. (COHA, 1942, MAG, NewRepublic)

The current version of COCA contains eight different genres, almost equal in terms of size: spoken texts, fiction, popular magazines, newspapers, academic texts, and (after the March 2020 update of the corpus) TV and movie subtitles, blogs, and other webpages. While all other text categories contain material from each year from 1990 to 2019, web pages and blogs were collected in October 2012 and cannot therefore be used for a diachronic study. Rather, they represent a synchronic picture of the language of those genres in that particular year. For our analysis we consider all eight genres in COCA, except when a diachronic perspective is adopted (cf. Sections 4.1 and 4.2 below); in such cases blogs and other web pages are excluded from the tally.

Whenever it was deemed necessary for the sake of illustration, the material from COCA was complemented with examples taken from observation (mostly from the authors' reading of the press and the internet) and with examples retrieved from random searches in various other sources.

3.2. Searches

As mentioned in Section 2, parentheticals may occur in three different positions in relation to their host clause: initially, medially, or finally. However, given the difficulties of teasing out initial parentheticals and matrix clauses in complementation structures (see Section 2 above), the searches carried out in COCA were intended to yield unambiguous examples of parentheticals, that is, those occurring either in medial or in final position. Instances of the type illustrated in (12) were therefore left aside. Although the default interpretation for *most studies suggest* in this example is that of a matrix clause followed by an object clause introduced by the complementizer zero (*mortality continues to decrease*…), in the absence of clear prosodic clues in examples such as these the parenthetical reading cannot totally be discarded.

12 *Most studies suggest* mortality continues to decrease -- or does not increase -- in the 40 -- 70 ng/mL range, but a few studies have shown an increased mortality within this range (COCA, 2012, WEB, hsph.harvard.edu)

The corpus searches, which were conducted in October–November 2022, are reflected in Table 1, together with the target sequences they sought to retrieve. We made use of the wildcard <*> for optional elements in the search strings and of punctuation marks (commas and periods) as indications of pauses to identify parentheticals in medial or in final position. The concordance lists obtained from the searches in Table 1 had to be further refined manually in order to remove false positives. An example is given in (13). As seen here, the search string (*empirical studies reveal*) appears between commas, but is clearly a matrix clause in a complementation structure, taking a dependent *that*-clause as object:

13 Regrettably*, empirical studies reveal*, however, *that* authoritarians are frequently enemies of freedom, antidemocratic, antiequality, highly prejudiced, mean-spirited, power hungry, Machiavellian and amoral. (COCA, 2016, NEWS, Chicago Sunday Times)

Table 1. Searches run in the corpus.

Search string	Target
, (*) (*) [study] (*) VERB , , (*) (*) [study] (*) VERB .	Examples in medial or in final position, where the nominal head *study* appears in the singular or in the plural and is followed by any simple form of the verb.[4] The nominal head may occur on its own or be preceded by a determiner, by a modifier, or by both a determiner and a modifier. The searches also allow for the presence of one intervening word (e.g., an adverb) between the noun phrase and the verb
, VERB (*) (*) [study] , , VERB (*) (*) [study] .	Same as above, but with inversion of noun phrase and verb

The searches also yielded examples of initial structures which, as discussed in Section 2, could be ambiguous between a matrix and a parenthetical interpretation. Consider in this respect (14) and (15) below. In these examples the existence of punctuation marks after the search strings could be interpreted as an indication of prosody, thus lending support to the analysis of the noun phrase + verb sequence in such instances as an initial parenthetical. However, given that ambiguity of initial strings cannot always be solved, we decided to leave all such cases out of the analysis.

14 And, *studies suggest*, overburdened kids are also at risk of muscle strain, fatigue, and back problems later in life. (COCA, 2005, MAG, Prevention)

4 Since COCA does not allow automatic searches longer than five words unless a premium license is purchased, we decided to leave out complex verb phrases, such as the present perfect in (i) and (ii) below. Due to the restriction to five words (punctuation marks count as a word), it is not possible to run automatic searches with one or two wildcards in front of the noun (e.g., < , * * [study] _vh _v?n .>).

(i) Fungi can be used to treat a violin to make it sound like a rare Stradivarius, *a study has found*. (COCA, 2012, BLOG, bussorah.wordpress.com)

(ii) Women from low socioeconomic backgrounds are 25 per cent more likely to suffer a heart attack than disadvantaged men, *a major new study has found*. (COCA, 2017, MAG, Medical Xpress)

15 If all consumers became as informed as medical experts, *the study concluded*, national headache-remedy brands would see their sales cut in half. (COCA, 2014, MAG, Atlantic)

After filtering the corpus examples in accordance with the criteria discussed in the preceding paragraphs, the number of valid instances amounted to 974.

4. Analysis of the data

This section is devoted to the analysis of the data from COCA, paying attention to the following issues: (i) the patterns in which the evidential parenthetical construction with the noun *study* is attested in the material (Section 4.1), (ii) the position the parentheticals occupy in relation to their host clauses (Section 4.2), (iii) the different predicates occurring in the parenthetical construction (Section 4.3), and (iv) the distribution of the parentheticals under analysis across genres (Section 4.4).

4.1. Parenthetical patterns

Two major patterns are attested in the COCA material, taking into account the relative position of subject and verb: one without inversion and another with subject-verb inversion. For these two patterns a number of variants have been identified in the data, depending on whether the subject noun phrase features determiners and/or modifiers or, on the contrary, the head noun *study* stands alone. Table 2 shows the distribution of patterns and variants of *study*-parentheticals in COCA.

Table 2. Parenthetical patterns with the noun *study* and their variants in COCA.

Pattern	Variants		Nº of tokens (%)	Total
Non-inversion	**I**	[study] VERB	145 (14.9%)	930 (95.5%)
	II	* [study] VERB	565 (58.0%)	
	III	* * [study] VERB	220 (22.6%)	
Inversion	**IV**	VERB * [study]	13 (1.3%)	44 (4.5%)
	V	VERB * *[study]	31 (3.2%)	
Total				974

As the data in Table 2 show, the most recurrent pattern in our American English material is the one without inversion, which accounts for over 95% of the instances. The pattern with inversion, on the other hand, is attested in less than 5% of the cases. In what follows, the variants of these two patterns are illustrated and described in some detail.

Variant I ([study] VERB), which represents 14.9% of all cases, is the least syntactically complex of the five sequences included in Table 2. As shown in examples (16) and (17), it features the noun *study* occurring on its own (with no determiners or modifiers) immediately followed by the verbal predicate, which appears almost invariably in the present tense in our data (141 out of 145 examples; 97.2%). The noun can be inflected for the singular (Variant I.a; 24 exx.), as in (16), or for the plural (Variant I.b; 121 exx.), as in (17).

16 The AJC headline on Monday read, "HPV shots don't make girls promiscuous, *study says*." (COCA, 2012, WEB, clatl.com)

17 Four to eight people grieve intensely for each suicide, *studies show*. (COCA, 2013, NEWS, Pittsburg Post Gazette)

Variant II (* [study] VERB) is the most recurrent sequence in our American English material, accounting for more than half of the instances (58%). Variant II shows a wider range of variability than Variant I, both in the noun phrase and in the verb phrase. Concerning the latter, in contrast to Variant I, the ratio of preterite tense forms to present tense ones is here 2:1 (68.7% vs. 31.3%). As regards the preverbal noun phrase, the noun *study* is preceded by either a determiner (II.a) or a modifier (II.b). Variant II.a is by far the most common, amounting to 540 instances, which represent 95.6% of the total in Variant II. The list of determiners available here is quite large, comprising (in order of decreasing frequency) articles (*a*, *the*; 92.8%), quantifiers (*many*, *most*, *one*, *several*, *some*, *another*, *other*; 4.3%), demonstratives (*that*, *this*, *these*; 1.8%), and possessives (*her*, *his*, *our*; 1.1%). In this set of prenominal elements, the definite article *the* stands out as the preferred choice (475 examples; 88% of the instances belonging to this variant), followed by the indefinite article *a* (26 occurrences; 4.8%); the remaining determiners in the set represent minor options.

Example (18) serves as an illustration of Variant II.a with the definite article *the*, whereas (19)–(22) show the use of other determiners preceding the noun *study*.

18 About a third of all strawberry samples had at least 10 pesticides, *the study found*. (COCA, 2018, NEWS, Chicago Sunday Times)

19 A brief workplace intervention using mindfulness techniques has a positive effect on an individual's sleep quality and sleep, processes that are vital for daily recovery from the demands of work, *a study finds*. (COCA, 2015, ACAD, OccupationalHealth)

20 Not only the quality, but the number of evaluations matter. The more the evaluations, *many studies show*, the more accurately they average out. (COCA, 1994, NEWS, Washington Post)

21 Nor, *this study reveals*, did wartime integration involve the two nations' citizenries as a whole. (COCA, 2004, ACAD, American Review of Canadian Studies)

22 Children are being exposed to unacceptable levels of the chemical, *our study found*. (COCA, 1999, MAG, Consumer Reports)

As mentioned above, in Variant II.b the prenominal slot is occupied by a modifier, rather than by a determiner. Only 25 examples out of 565 of occurrences of Variant II in the data belong to this type (4.4% of the total of instances). These premodifiers add descriptive information on the head noun *study*, restricting or qualifying it. Premodifying items in this pattern include several of the types of premodification identified by Quirk *et al.* (1985: 1321–1337), specifically adjectives, nouns, and genitives. The latter two types are only sporadically attested in our COCA material: 6 (24%) and 4 occurrences (16%) of modification by nouns (both common and proper nouns) and by genitives, respectively. Examples (23)–(24) illustrate the use of nouns as premodifiers of *study*, while (25) shows a genitive in premodifying function.

23 "That's right. With so many people dying from overwork, I thought we should do our part to stay healthy. Having a pet around the office helps you to relax. It improves office efficiency, *government studies show*," Nakamura says. (COCA, 1996, FIC, Audrey Hepburn's Neck)

24 New car quality, in fact, is at an all-time high, but complaints about electronic accessories have stalled automakers' efforts to improve their quality scores, *J. D. Power studies show*. (COCA, 2003, NEWS, USA Today)

25 Personality traits such as being receptive to the unfamiliar may directly explain why life for some centenarians is still as enjoyable and independent as it was in their younger years, *Poon's study finds*. (COCA, 2011, MAG, Psychology Today)

As expected, most of the premodifiers in our data are adjectives (15 out of 25 examples; 60%): *academic* (2), *behavioural*, *independent*, *latest*, *long-term*, *military*, *new* (4), *recent* (3), and *secret*. Note the overall positive semantic prosody resulting from the combination of the noun *study* with positive evaluative adjectives such as those conveying the idea of novelty (e.g., *latest*, *new*, *recent*), as in (26), or of prestige (e.g., *academic* in (27)).[5]

26 Those warmer air temperatures are significantly boosting soil temperatures in many regions, *new studies show*. (COCA, 2007, MAG, Science News)

27 But the majority of mergers, *academic studies show*, fail to deliver for shareholders (COCA, 2005, SPOK, CNN_Next)

This also applies to Variant III (* * [study] VERB; 22.6% of the total), where the preverbal noun phrase contains both a determiner and a modifier. The most frequent sequence in this variant is that which combines the indefinite article *a* and the adjectival premodifier *new* (i.e., *a new study*), as in (28), which is found in COCA in 132 instances (60% of the total of occurrences of Variant III in the data). In addition to *new*, the premodifiers attested here mostly convey the ideas of novelty and recency (e.g., *recent*, *first-of-its-kind*) or of an authoritative study (e.g., *Harvard*, *MIT*, *Oxford*, *NASA*). On other occasions, they indicate the source (e.g., *federal*, *Canadian*, *Japanese*) or the date (e.g.,

5 Positive evaluative adjectives modifying the noun *study* can be seen as 'epistemic extensions' (Aikhenvald 2004: 6), with the speaker underlining the validity of their statement by referring to a reliable or trustworthy source. We thank an anonymous reviewer for this insightful suggestion.

2012, *2013*) of the study being referred to in the parenthetical. See (29)–(31) below as illustrations.

28 The action of a gene called *ATDC* is required for the development of pancreatic cancer, *a new study finds.* (COCA, 2019, MAG, Medical Xpress)
29 Fair-haired folks and redheads have a higher risk of Parkinson's than those with dark hair, *a Harvard study found.* (COCA, 2013, MAG, Men's Health)
30 It also provides motivational messages -- which, *a British study says*, may increase your chances of being smoke-free at the 6-month mark. (COCA, 2012, MAG, Men's Health)
31 Over forty percent of them, *a 2005 study showed*, descend from just four "founding mothers" having Middle-Eastern-profile mitochondrial DNA. (COCA, 2008, ACAD, Commentary)

In contrast to the three parenthetical variants just discussed, Variant IV (VERB * [study]; 1.3% of the total) and Variant V (VERB * * [study]; 3.2%) illustrate so-called 'quotative inversion' (see, among others, Collins & Branigan 1997; Cichosz 2018), a type of inversion which is particularly frequent in written registers (Joseph 2000: 312). In Variant IV (shown in (32)) the noun phrase contains only a determiner (mostly the definite article *the*, which occurs in 10 of the 13 instances of this type; 76.9%) or the modifier *recent* preceding the noun *study* (33). By contrast, in Variant V the noun phrase contains both a determiner and a modifier as prenominal elements. In the latter type, the most recurrent combination is that of the indefinite article *a* with an adjective, mostly *new*. Example (34) illustrates the sequence *a new study* following the verb, which occurs 12 times in the COCA material (38.7% of the examples belonging to Variant V), while (35) provides an example of the combination *a*/*one*/*two recent*, which is attested in the data on 7 occasions.

32 The patient purchased the medicine over the internet, *said the study*. (COCA, 2018, NEWS, USAToday)
33 Diners using big forks ate less of a * large serving than those using small ones, *says recent study*. (COCA, 2015, MAG, Saturday Evening Post)

34 After a hectic day - when the lure of the drive-thru is most magnetic - your metabolism tanks, *says a new study*. (COCA, 2014, MAG, Prevention)

35 "What are the scientists saying? Gatorade works," one print ad says. (Sports drinks don't improve athletic performance, *concludes a recent study*, but do replenish needed minerals). (COCA, 1990, NEWS, USA Today)

Interestingly, the vast majority of examples with inversion in COCA (36 out of 44) are attested in formal written genres (academic texts, newspapers, and, most notably, popular magazines). In addition to the potential correlation between Variants IV and V and a higher degree of formality, inversion can also be explained by the principle of end-weight. By way of illustration, consider (36) and (37) below, which show a relative clause (Quirk *et al.* 1985: 1244 ff.) and an *-ed* participle clause (Quirk *et al.* 1985: 1264–1265), respectively, as postmodifiers of the noun *study*.

36 Almost 11 million of these uninsured Americans are children age 18 or younger, *said the study*, which was conducted by the department's Agency for Health Care Policy and Research. (COCA, 1997, MAG, San Francisco Chronicle)

37 The drop-off in ordering screenings could be linked to "decision fatigue," the mental burn-out that makes it harder for people to think through big decisions the farther they get into their day, *said the study*, published Friday in JAMA Network Open. (COCA, 2019, MAG, MarketWatch)

In order to provide a snapshot of the diachronic development of the five parenthetical variants described in the preceding paragraphs, Table 3 shows their distribution across time from the 1990s to the year 2019. For practical purposes, we have subdivided COCA into three subperiods, each covering a decade. Since, as mentioned in Section 3.1 above, the material for blogs and other webpages in COCA was collected for just one year (2012), examples belonging to these two genres have been excluded from the count.

Table 3. Distribution of parentheticals with the noun *study* across time in COCA (excluding blogs and other webpages).

	Variant I	Variant II	Variant III	Variant IV	Variant V	Total
1990–1999	36 (26.3%)	114 (27.5%)	28 (17.9%)	3 (27.3%)	5 (20%)	186 (25%)
2000–2009	52 (37.9%)	107 (25.8%)	28 (17.9%)	2 (18.2%)	10 (40%)	199 (26.7%)
2010–2019	49 (35.8%)	194 (46.7%)	100 (64.2%)	6 (54.5%)	10 (40%)	359 (48.3%)
Total	137	415	156	11	25	744

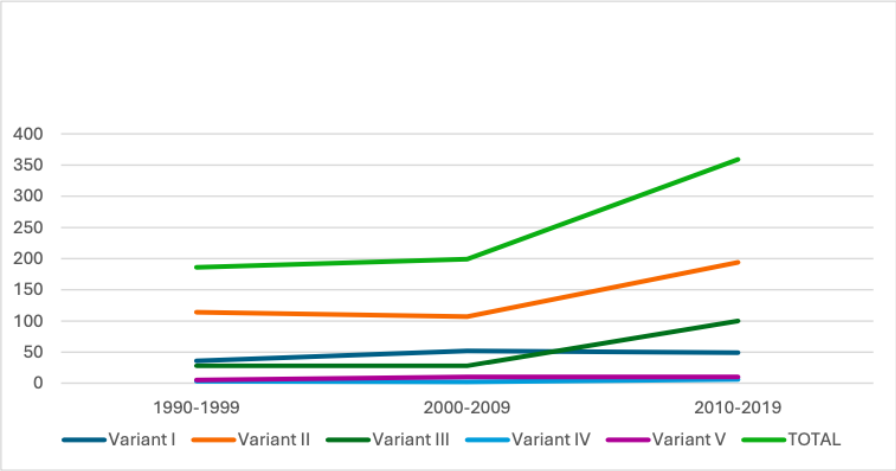


Figure 1. Parenthetical variants with the noun *study* across time in COCA (excluding blogs and other webpages).

As seen in Table 3, all five variants are attested in the three subperiods in COCA. The table also shows that parentheticals with the noun *study* have been on the rise over the last 30 years in American English. Note that the number of the constructions under analysis almost doubles from subperiod 1 (186; 25%) to subperiod 3 (359; 48.3%). Figure 1 shows that such an increase is especially noticeable from the 2000s to the 2010s (light green line) and particularly marked for Variants II and III (orange and dark grey lines).

4.2. Position of *study*-parentheticals

The distribution of our corpus instances (excluding material from the web and blogs) according to position is provided in Table 4.[6] As shown here, medial position occurs in just one sixth of the instances (125 exx., 16.8%).[7] On the other hand, final position is overall the preferred choice for the parenthetical constructions under analysis (619 exx., representing 83.2% of the total of examples),[8] in line with Biber *et al.*'s (1991: 923) findings for reporting clauses in fiction and news. Final-position parentheticals also predominate over those in medial position in all three individual subperiods in COCA. Note the slight decrease of medial position from subperiod 1 (22.6%) to subperiod 3 (13.4%).

Table 4. Position of parentheticals with the noun *study* across time in COCA (excluding blogs and other webpages).

	Medial position	Final position	Total
Subperiod 1: 1990–1999	42 (22.6%)	144 (77.4%)	186
Subperiod 2: 2000–2009	35 (17.6%)	164 (82.4%)	199
Subperiod 3: 2010–2019	48 (13.4%)	311 (86.6%)	359
Total	125 (16.8%)	619 (83.2%)	744

Final position is attested for all the parenthetical patterns described in Section 4.1. Examples (16), (22), (29), and (34) above show Variants I-IV, while (38) below illustrates Variant V.

38 How to make a food instantly higher in fiber, lower in fat, and less caloric? Slap an "organic" sticker on it, *says a Cornell study*. (COCA, 2011, MAG, Good Housekeeping)

---

6 As discussed in Section 3.2, initial sequences have been excluded from the analysis given their ambiguous status between matrix and parenthetical clauses.

7 In this respect, the behaviour of *study*-parentheticals differs from that of standard reporting clauses, for which medial position is not so marginal (Quirk *et al.* 1985: 1022).

8 Final position is also the unmarked position for *study*-parentheticals occurring in the web and in blogs in COCA: 114 exx. (89.8) and 93 exx. (90.3%), respectively.

In turn, medial position is restricted in our data to examples without inversion (Variants I-III). Our parentheticals in medial position occur mostly between the subject and the verb (80 exx., representing 64% of the relevant instances), as in examples (30) and (31) above. This is precisely one of the "weak spots" identified by Kaltenböck (2007: 43) in which parentheticals can easily be inserted. The second most common pattern is that illustrated in (39), where the parenthetical intervenes between an adjunct and the subject. Although parentheticals most frequently occur at the boundaries of phrases, they can sporadically be found in positions in which they "cut across phrasal constituents" (Kaltenböck, Heine, & Kuteva 2011: 869). Such disfavoured positions (cf. Kaltenböck 2007: 42 ff.; Brinton 2008: 8 and references therein) include, for instance, between an adjective and its complement, as in (40).

39 During pregnancy, *our studies suggest*, the trophoblast functions as a guest conductor. (COCA, 2007, MAG, Natural History)

40 They're also more likely, *studies show*, to drive up an exit ramp and into a head-on collision. (COCA, 2007, NEWS, Houston)

4.3. Predicates across parenthetical variants

A look at the verbal predicates occurring in the parentheticals under analysis also affords interesting insights into the nature of the construction. Table 5 provides the list of predicates attested in each of the five variants discussed in Section 4.1 above. As can be gathered from this list, all the verbs identified in the COCA material convey, in different ways, the source of information; in other words, they express evidential meaning.[9] Moreover, all these verbs belong to particular predicate types in Noonan's (1985) and Rudanko's (1989) well-known typologies of complement-taking predicates:

9 *Study*-parentheticals, with their evidential function, are connected with one of the types of reporting clauses studied by Vandelanotte (2009), namely his 'Subjectified Distancing Indirect Speech or Thought' (S-DIST), which are characteristic of news reports (Vandelanotte 2009: 325).

(i) Utterance predicates (Noonan 1985: 110–113), i.e., those which denote a transfer of information, describing the manner of transfer; e.g., *say*, *suggest*, *state*, *report*, etc.

(ii) Predicates of knowledge and acquisition of knowledge (Noonan 1985: 118–119), which describe the state or the manner of acquisition; e.g., *find*, *read*, *conclude*.

(iii) Demonstration predicates (Rudanko 1989: 77); e.g., *show*, *reveal*.

Table 5. Predicates attested in COCA in the five parenthetical variants with the noun *study*.

Variant I	**Variant II**		**Variant III**	**Variant IV**	**Variant V**
find (17)	*add* (1)	*note* (14)	*conclude* (8)	*conclude* (1)	*conclude* (1)
indicate (4)	*argue* (4)	*observe* (1)	*contend* (1)	*praise* (1)	*find* (7)
reveal (1)	*assert* (2)	*predict* (3)	*declare* (1)	*report* (2)	*note* (1)
say (12)	*calculate* (1)	*read* (4)	*find* (115)	*say* (8)	*report* (4)
show (94)	*claim* (2)	*recommend* (1)	*indicate* (7)	*suggest* (1)	*say* (14)
suggest (17)	*conclude* (40)	*report* (19)	*predict* (1)		*show* (1)
	confirm (1)	*reveal* (13)	*report* (3)		*suggest* (3)
	contend (1)	*say* (176)	*reveal* (1)		
	estimate (2)	*show* (45)	*say* (33)		
	explain (1)	*state* (2)	*show* (9)		
	find (201)	*suggest* (20)	*state* (1)		
	indicate (6)	*warn* (3)	*suggest* (40)		
	maintain (2)				

Interestingly, six of these predicates appear repeatedly in the non-inverted parenthetical pattern (Variants I, II, and III), namely *find*, *indicate*, *reveal*, *say*, *show*, and *suggest*.[10] Most importantly, these six predicates represent 86.8% of all occurrences of the parentheticals in the data, with *find* (340 exx.; 34.9%), *say* (243 exx.; 24.9%), and *show* (149 exx.; 15.3%) taking the lead. These three predicates are, semantically speaking, the most general in their respective categories. Thus, for example, *say* is the hyperonym of utterance verbs and *show* is the hyperonym of predicates of demonstration. Note also that the six predicates are the only ones attested in our Variant I, which

10 Some of these predicates (*find*, *say*, *show*, and *suggest*) also occur sporadically in the pattern with inversion (Variants IV and V).

contains the lowest number of types. By contrast, Variants II and III show a much wider range of types, including predicates with a more specific kind of meaning. For example, within the group of utterance predicates, together with the hyperonym *say*, we find *read*, *claim*, *explain*, *warn*, etc., i.e., predicates which also convey the manner of information. The predicate *say* is also the most frequent in the inversion pattern (8 and 14 exx. in Variants IV and V, respectively). The dominance of *say* here suggests the existence of a link with quotative inversion (cf. Section 4.1 above).

4.4. Distribution of *study*-parentheticals across genres

As mentioned in Section 3, COCA contains material belonging to eight different genres, nearly evenly divided in size. Table 6 shows that *study*-parentheticals are attested in all text categories represented in the corpus, though with marked differences among them.

Table 6. Distribution of *study*-parentheticals across genres in COCA.

Genre	Nº of tokens (%)
ACAD	45 (4.6%)
BLOGS	103 (10.6%)
FIC	3 (0.3%)
MAG	331 (34%)
MOV	1 (0.1%)
NEWS	355 (36.4%)
SPOK	9 (0.9%)
WEB	127 (13.1%)

Perhaps the most salient feature of the distribution of the parentheticals at issue here is their strong correlation with the written language, in particular with the press category, which comprises popular magazines (MAG) and newspapers (NEWS). Note that over 70% of the relevant corpus instances belong to these two genres (34% and 36.4%, respectively). Academic texts (ACAD) show a lower proportion of the parenthetical constructions under analysis (4.6%). On the other hand, *study*-parentheticals are far less common in those genres with a higher degree of speech-likeness, such as blogs (10.6%), and are practically non-existent in those categories which can be taken to

represent speech, such as spoken (0.9%), fiction (0.3%), and TV and movie subtitles (0.1%). The strong connection of our parentheticals with the press category and, therefore, with the written language brings them close to Kaltenböck, Heine, and Kuteva's (2011: 856) 'reporting parentheticals' and Vandelanotte's (2009) S-DIST clauses, in contrast with other types of parentheticals, which are considered a typical feature of conversation and of written texts which resemble speech (Thompson & Mulac 1991; Biber *et al.* 1999: 197, 862–865, 983; López-Couso & Méndez-Naya 2014).

Interestingly, within the press category in COCA, our examples typically occur in texts which can be described as popular science (pop-science or popsci), a written genre concerned with the interpretation of science intended for a general audience which has experienced a boost over the last few years. Pop-science, which can be produced by either scientists or non-scientists, represents a bridge between scientific literature and journalism and aims at capturing the methods and accuracy of science, while making it more accessible to the general public. By contrast, academic writing is written by scientists and addressed to their peers. Even though popular science shares some linguistic features with academic writing (e.g., the use of epistemic modals indicating different degrees of likelihood; see Biber & Conrad 2009), in pop-science reference to the authorial sources can be vaguer than in academic texts, where precise mention of the sources is mandatory. As Chafe (1986: 268) puts it, "[i]n academic writing knowledge obtained through language is indicated with the formal device of citing a reference or personal communication". The difference between these two ways of conveying evidential meaning in scientific literature and in pop-science is illustrated in examples (41) and (42). While in (41) the evidential source (*Hilpert 2013: 133*) is expressed in its "most precise and deliberate form" (Chafe 1986: 269), in (42) the reference is undoubtedly more imprecise, though the use of the premodifier *Harvard*, a very prestigious academic institution, in the noun phrase *a Harvard study* serves to lend credibility to the statement made in the host clause, providing the source of evidence with a high degree of authority (see Section 4.1 above).

41 *Hilpert (2013: 133)* suggests that one of the Baayen's hapax-based measures of productivity, namely expanding productivity, could be used. (Hilpert 2017: 58)

42 A man with a good shot at getting a girlfriend may live three months longer than men who outnumber local women, *a Harvard study finds*. (COCA, 2011, MAG, PsychToday)

The growing importance of popular science among the general public and the strong connection of our *study*-parentheticals with pop-science become manifest in the existence of a popsci webpage precisely entitled *StudyFinds* (https://studyfinds.org/), of *The New York Times Magazine* popsci column *Studies show. What medical research really says* (https://www.nytimes.com/column/studies-show), and of the Wiley podcast called *This Study Shows* (https://audioboom.com/channels/5003898-this-study-shows). The latter is presented in the following way: "The only way for research to change the world is if the world knows and cares about it. In This Study Shows, a podcast from Wiley, we explore how to connect research with emotions and experiences and transform the way science is shared".

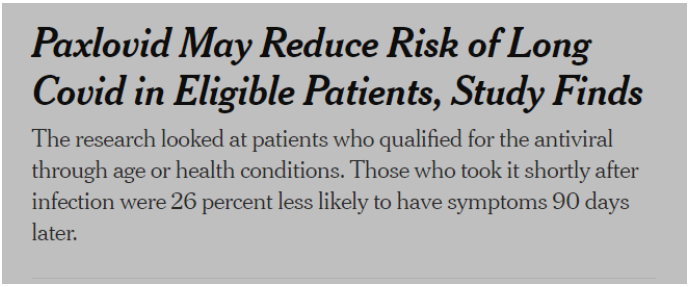
Another interesting feature of the parentheticals under examination is that they commonly occur in headlines, as shown in (43) and (44), where we find an explicit reference to a headline (see also example (16) above).

43 Greenland Ice Melt Reaching a Tipping Point, *Study Says*. (COCA, 2019, NEWS, Minneapolis Star Tribune)

44 But the Times apparently wasn't interested in the headline, "*Bartenders Get Flu and Recover, Study Says*." Such a headline would also not go far in justifying a smoking ban. (COCA, 2001, MAG, ConsumResrch)

Outside COCA, examples of this kind are not difficult to find, as shown by (45) and (46), taken from *The New York Times* (7 November 2022) and from *The Guardian* (20 September 2023), respectively.

45



Paxlovid May Reduce Risk of Long Covid in Eligible Patients, Study Finds

The research looked at patients who qualified for the antiviral through age or health conditions. Those who took it shortly after infection were 26 percent less likely to have symptoms 90 days later.

46

Fathers have 'unique effect' on children's educational outcomes, study finds

Research claims children whose fathers read and play with them see a 'small but significant' increase in their educational attainment

Interestingly, it is in headlines where Variant I is normally found. Headlines serve as clickbaits, to spur the reader's curiosity which is later quenched in the main body of the text. As illustrated in the examples below, when the topic advanced in the headline is taken up in the initial part of the article itself, the more complex Variant II (example (47) below) or Variant III sequences (examples (48) and (49)) are typically attested.

47 'Gigantic' power of meat industry blocking green alternatives, *study finds*
[…] The "gigantic" power of the meat and dairy industries in the EU and US is blocking the development of the greener alternatives needed to tackle the climate crisis, *a study has found*. (*The Guardian*, 18 August 2023) [headline: Variant I.a; body of the article: Variant II.a]

48 Gulf Stream could collapse as early as 2025, *study suggests*
[…] The Gulf Stream system could collapse as soon as 2025, *a new study suggests*. (*The Guardian*, 25 July 2023) [headline: Variant I.a; body of the article: Variant III]

49 People who took the antiviral drug Paxlovid within a few days after being infected with the coronavirus were less likely to be experiencing long Covid several months later, *a large new study found*. (*The New York Times*, 7 November 2022) [corresponding to the headline in (45) above]

Note that in the body of the article itself the source of information can also be conveyed by means of an adjunct (50) or by a main clause (51).

50 Erythritol, an ingredient in stevia, linked to heart attack and stroke, *study finds*
A sugar replacement called erythritol — used to add bulk or sweeten stevia, monkfruit and keto reduced-sugar products — has been linked to blood clotting, stroke, heart attack and death, according to a study. (https://edition.cnn.com/2023/02/27/health/zero-calorie-sweetener-heart-attack-stroke-wellness/index.html)

51 Humans 'may need more sleep in winter', *study finds*. Research shows people get more deep REM sleep than in summer, and may need to adjust habits to season. (*The Guardian*, 17 February 2023)

As is well known, headlinese is characteristically condensed and can be regarded as "stylistically 'deviant'" (Molek-Kozakowska 2017: 897).[11] One of its most salient linguistic features is its fragmentary nature, as evinced by, for instance, the lack of determiners in cases where they would normally be required. As Moncomble (2018) puts it, absence of determiners is "part and parcel of the production of headlines". In fact, all our examples of Variant I.a (that featuring the singular noun *study* on its own) occur in headlines. Absence of the determiner is also attested in the inverted pattern, specifically in

11 On the pragmatics of headlines, see the 2024 special issue of *Journal of Pragmatics*, edited by Rita Finkbeiner.

Variant IV (VERB * [study]), but only in one example in the data (see example (33); *says recent study*), which is also found in a headline.

Another feature of headline language is the predominant use of the simple present to refer to past events, where a present perfect or even a preterite would be preferred in ordinary language. This is clearly seen in example (47) above (headline: *finds* vs. text: *has found*). Note that, as mentioned in Section 4.1, Variant I almost invariably shows a present tense verb phrase.

Despite the close association of Variant I with headlines, the structure may also occur in the body of the text, as illustrated in example (52). Such examples, however, feature the noun *study* in the plural (i.e., they correspond to Variant I.b). By contrast, all instances of singular *study* (Variant I.a) occur in headlines.

52 # For those born in that first wave and now entering adulthood, it's a tough, uncertain future. Some, like Abenes, will go to college or find jobs and eventually move out on their own. # But most will not, *studies show*. Most will continue to live at home and will, at best, find part-time, minimum-wage work -- or no work at all. (COCA, BLOG, 2012, specialedpost.com)

5. Discussion

As may have become apparent from the examples provided in the preceding section, *study*-parentheticals have a clear evidential function, that is, they convey the source of information the speaker/writer has for their statement. Of the three major categories of evidentials, namely direct perception, inference, and hearsay (Mélac 2022: 333), *study*-parentheticals present the information in the proposition as hearsay, so that the speaker's commitment is not really at stake. Hearsay evidentials of this kind fit in well with pop-sci texts, which portray the evidential source as serious and reliable, though not in such a precise manner as in academic writing. This is specially the case in examples containing modifiers referring to prestigious institutions (e.g., *Oxford*, *NASA*) or indicating that the study is up-to-date (e.g., *new*, *recent*) (see Section 4.4 above).

As far as their development is concerned, *study*-parentheticals seem to illustrate a similar pathway to that described in the literature for other clausal

parentheticals. The emergence of parenthetical clauses has been explained by invoking several processes of change, among them grammaticalization, pragmaticalization, lexicalization, and co-optation (see Heine & Kaltenböck 2021: 4–7 for a good summary of the different proposals). Of these processes, grammaticalization stands out as the most widespread account and is the one we will adopt in the discussion that follows.

In a highly influential paper, Thompson and Mulac (1991) proposed the so-called matrix clause hypothesis, according to which epistemic parentheticals of the type *I think* originate in matrix clauses taking *that*/zero-clauses as complements. The increase in frequency of the complementizer zero is regarded as the trigger for the syntactic reanalysis of the matrix as a parenthetical, which, once reanalysed, can also occur in medial and in final position. A related grammaticalization account is the one posited by Boye and Harder (2007; 2012), who consider not only the changes in syntactic function, but also in discourse prominence, arguing that the development from a matrix clause to a parenthetical clause requires a previous usage reanalysis in which the dependent clause acquires primary discourse status (see Section 2). In fact, in line with Boye and Harder (2007), Kearns' (2007) study of *I think* suggests that semantico-pragmatic changes are precisely the catalyst for structural changes.

Indeed, the development of clausal parentheticals illustrates semantic, pragmatic, and structural indicators of grammaticalization. Among the former, clausal parentheticals show semantic bleaching and non-addressability (Boye & Harder 2007), together with the acquisition of (inter)subjective functions. As far as structural concomitants of grammaticalization are concerned, they typically show decategorialization and fixation (see López-Couso & Méndez-Naya 2014 and 2021 on *like*-parentheticals and on *chances are*- and *odds are*-parentheticals, respectively).

In an account compatible with grammaticalization, Heine and Kaltenböck (2021: 5) suggest that the development of (paren)theticals primarily involves a process of co-optation, whereby "main clauses without their complements […] are taken out of sentence grammar and transferred to the metatextual level of discourse processing to give rise to comment clauses and discourse markers in general". Co-opted elements can conventionalize and grammaticalize over time, giving rise to different types of (paren)theticals (Heine & Kaltenböck 2021: 7–8): (i) instantaneous theticals, which can be created spontaneously,

maybe not to be uttered again (cf. also Quirk *et al.* 1985: 1114 on "freely constructed" comment clauses); (ii) constructional theticals, which have "a schematic format and function", while involving a certain amount of variability (comment clauses are examples of this parenthetical type); and (iii) formulaic theticals, which are invariable (e.g., discourse markers, interjections, politeness formulae, etc.).

In our view, the data discussed in Section 4 suggest that *study*-parentheticals are moving along a grammaticalization pathway, as is also the case with other reporting clauses found in news reports (see Vandelanotte 2009: 325, 329). The three major parenthetical variants (that is, those illustrating the non-inverted pattern) can be placed along a cline of complexity which represents increasing grammaticalization, in line with the general pathway attested for other parentheticals. As shown in Figure 2, Variant III, which enjoys the highest degree of variability, would be placed at the leftmost end of the cline, while Variant I.a, the least syntactically complex, would be located at the right end of the continuum. Variant II.a is taken to be more constrained than Variant II.b, since the prenominal position is taken by a grammatical word (a determiner, mostly the definite article), while in Variant II.b various types of open-class modifiers (adjectives, nouns, etc.) can occupy the prenominal slot.

Higher degree of complexity **Lower degree of fixation**				**Lower degree of complexity** **Higher degree of fixation**
Variant III *a new/recent study finds*	> Variant II.b *recent studies find*	> Variant II.a *the study finds*	> Variant I.b *studies find*	> Variant I.a *study finds*

Figure 2. Complexity/fixation cline of *study*-parentheticals.

The growing grammaticalization of Variant I ([study] VERB) is supported by a number of features:

(i) The restricted range of predicates available in this variant (*find*, *indicate*, *reveal*, *say*, *show*, and *suggest*), all of which are fairly general from a semantic point of view, with *find*, *say*, and *show* (the hyperonyms of their respective semantic classes, namely knowledge and acquisition of knowledge, utterance, and demonstration) standing out in terms of frequency. Variants II and III, by contrast, feature a larger number of types,

including verbs with more specific meanings conveying the manner of transfer of information, of knowledge acquisition, or of demonstration (see Section 4.3). Moreover, the range of predicates occurring in Variant I.a is even more constrained than in Variant I.b, thus pointing at specialization (Hopper 1991: 22). Only four predicates occur in Variant I.a (*find*, *say*, *show*, and *suggest*), with *find* present in over half of the tokens (13/24 exx.).

(ii) The almost complete restriction of the verb phrase to the present tense form (see Section 4.1), thus suggesting morphosyntactic fixation. The extensive literature on parentheticals has proved that the more grammaticalized a given parenthetical type, the more likely for it to show a verb phrase inflected for the present tense (Brinton 2008: 2; van Bogaert 2010: 401 and the references therein).

(iii) The disallowance of insertion of intervening material (e.g., an adverb), while this is perfectly possible in Variants II and III, as shown in (53), an example of Variant II. This points at an increasing bonding between the noun phrase and the verb phrase in Variant I. Increasing bonding is a well-known feature of grammaticalization, namely fusion (Brinton & Traugott 2005: 27).

53 According to the CDC study, the most common wound after head injuries involved upper limb fracture, followed by lower limb fractures. About 14 rides out of every 100,000 lead to injuries, *the study reportedly found*. (COCA, 2019, MAG, Gizmodo)

(iv) The absence of a determiner in Variant I.a, with the noun *study* in the singular, could be suggestive of decategorialization (Hopper 1991: 22, 30 ff.), in as much as the parenthetical type *study finds* lacks a determiner which would otherwise be mandatory, making the noun phrase somewhat deviant and defective. However, the omission of the determiner here could also be interpreted as genre-licensed (headlinese) (see Section 4.4 above).

In view of the above, it can safely be concluded that the parenthetical *study finds* is on its way towards becoming a formulaic phrase, with a growing restriction to just one predicate (*find*) (Hopper 1991: 22, 25). In Kaltenböck, Heine, and Kuteva's (2011) classification of thetical elements, Variants II and III would qualify as constructional theticals, while Variant I.a is moving towards becoming a formulaic thetical in a particular type of discourse.

6. Concluding remarks

This chapter has examined the development and use of an evidential strategy in the recent history of American English, whereby clausal parenthetical structures containing a noun phrase with *study* as head in combination with a verb phrase are employed to indicate hearsay, thus illustrating a type of reporting clause. The data from COCA (1990–2019) and complementary material have allowed us to identify a number of interesting trends:

(i) While s*tudy*-parentheticals occur sporadically from the 1940s, they only gain momentum from the 1990s, experiencing a marked increase in frequency over the last 30 years.

(ii) Although final position is by far the preferred choice for *study*-parentheticals, the construction is also attested medially, mostly at the boundaries of phrases (e.g., between the subject and the verb), but also, though only occasionally, across phrasal constituents (e.g., between an adjective and its complement).

(iii) In contrast to other parenthetical types, the *study*-parentheticals examined here are closely associated with the written mode of expression, in particular with journalistic discourse (the text-categories NEWS and MAG in COCA). More specifically, these parentheticals frequently occur in texts which can be characterized as pop-science. In texts of this kind *study*-parentheticals present the hearsay evidence as coming from a reliable, authoritative source, even though in a less precise manner than that required in academic writing.

(iv) *Study*-parentheticals show a certain degree of variability, appearing in the COCA material in two patterns and five different variants. The pattern featuring quotative inversion (e.g., *says a* (*new*) *study*) is only sporadically found. In the non-inverted pattern, the most frequent in the data is Variant II (* [study] VERB), in particular that showing a determiner in prenominal position (Variant II.a), followed by Variant III (* * [study] VERB) and Variant I ([study] VERB), in order of decreasing frequency.

(v) The major three variants in which *study*-parentheticals are attested in COCA can be placed along a cline of complexity/fixation, with Variant I.a, with the singular noun *study*, on its way to becoming a fossilized expression: (a) the predicates in Variant I.a are limited to just four verbs, namely *find*, *say*, *show*, and *suggest*, with *find* occurring in over 50% of the relevant cases (specialization); (b) the verb is almost invariably marked for the present tense (fixation); and (c) intervening material is disallowed between the noun and the verb (fusion). Finally, the lack of a determiner in Variant I.a, exclusively found in headlines in our data, would be a clear sign of decategorialization if this variant were ever attested outside headlines.

(vi) In Kaltenböck, Heine, and Kuteva's (2011) taxonomy, Variants I, II, and III would qualify as constructional theticals, with the parenthetical *study finds* coming close to a formulaic thetical which has become a staple in pop-sci headlinese, as testified by the existence of a website called *StudyFinds* (see Section 4.4 above).

The evidence presented in this chapter suggests that *study*-parentheticals represent an appealing but understudied type of reporting parenthetical construction, which clearly deserves investigation. One of the possible avenues for further research implies extending the analysis to other semantically related nouns, including *research*, *report*, or *survey*, as well as to animate nouns such as *expert*, *scientist*, or *researcher* in order to check whether animacy plays a role in the behaviour of these evidential parentheticals.

Bibliography

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
Biber, Douglas & Susan Conrad 2009. *Register, Genre and Style*. Cambridge: Cambridge University Press.
Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, & Edward Finegan 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
Boye, Kasper & Peter Harder 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in Language* 31 (3): 569–606. https://doi.org/10.1075/sl.31.3.03boy
Boye, Kasper & Peter Harder 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88 (1): 1–44. https://doi.org/10.1353/lan.2012.0020
Boye, Kasper & Peter Harder 2021. Complement-taking predicates, parentheticals and grammaticalization. *Language Sciences* 88: 101416. https://doi.org/10.1016/j.langsci.2021.101416
Brinton, Laurel J. 2008. *The Comment Clause in English. Syntactic Origins and Pragmatic Development*. Cambridge: Cambridge University Press.
Brinton, Laurel J. & Elizabeth C. Traugott 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
Chafe, Wallace 1986. Evidentiality in English conversation and academic writing. *Evidentiality: The Linguistic Coding of Epistemology*, ed. by Wallace Chafe & Johanna Nichols. 261–272. Norwood, NJ: Ablex.
Cichosz, Anna 2018. Parenthetical reporting clauses in the history of English. The development of quotative inversion. *English Language and Linguistics* 23 (1): 183–214. https://doi.org/10.1017/s1360674317000594
Collins, Chris & Phil Branigan 1997. Quotative inversion. *Natural Language and Linguistic Theory* 15 (1): 1–41. https://doi.org/10.1023/A:1005722729974
Davies, Mark 2008–. *COCA: Corpus of Contemporary American English*. Available online at https://www.english-corpora.org/coca/.
Davies, Mark 2010–. *The Corpus of Historical American English (COHA): 400 million words, 1810–2009*. Available online at https://www.english-corpora.org/coha/.
Davies, Mark. 2011–. SOAP: *Corpus of American Soap Operas*. Available online at https://www.english-corpora.org/soap/.
Dehé, Nicole 2009. Clausal parentheticals, intonational phrasing, and prosodic theory. *Journal of Linguistics* 45 (3): 569–615. https://doi.org/10.1017/s002222670999003x
Finkbeiner, Rita (ed.) 2024. *The Pragmatics of Headlines*. Special issue of *Journal of Pragmatics*.
Heine, Bernd & Gunther Kaltenböck 2021. From clause to discourse marker: On the development of comment clauses. *Language Sciences* 87: 101400. https://doi.org/10.1016/j.langsci.2021.101400
Hilpert, Martin 2017. Frequencies in diachronic corpora and knowledge of language. *The Changing English Language. Psycholinguistic Perspectives*, ed. by Marianne Hundt, Sandra Mollin, & Simone E. Pfenninger. 49–68. Cambridge: Cambridge University Press.
Hopper, Paul J. 1991. On some principles of grammaticalization. *Approaches to Grammaticalization*. Vol. I, ed. by Elizabeth C. Traugott & Bernd Heine. 17–35. Amsterdam: John Benjamins.

Huddleston, Rodney & Geoffrey Pullum *et al.* 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
Joseph, Brian D. 2000. Textual authenticity: Evidence from medieval Greek. *Textual Parameters in Older Languages*, ed. by Susan C. Herring, Pieter van Reenen, & Lene Schøsler. 309–330. Amsterdam: John Benjamins.
Kaltenböck, Gunther 2007. Spoken parentheticals in English: A taxonomy. *Parentheticals*, ed. by Nicole Dehé & Yordanka Kavalova. 25–52. Amsterdam: John Benjamins.
Kaltenböck, Gunther, Bernd Heine, & Tania Kuteva 2011. On thetical grammar. *Studies in Language* 35 (4): 848–893. https://doi.org/10.1075/sl.35.4.03kal
Kearns, Kate 2007. Epistemic verbs and zero complementizer. *English Language and Linguistics* 11 (3): 475–505. https://doi.org/10.1017/s1360674307002353
López-Couso, María José & Belén Méndez-Naya 2014. From clause to pragmatic marker: A study of the development of *like* parentheticals in American English. *Journal of Historical Pragmatics* 15 (1): 66–91. https://doi.org/10.1075/jhp.15.1.03lop
López-Couso, María José & Belén Méndez-Naya 2021. From complementizing to modifying status: On the grammaticalization of the complement-taking-predicate-clauses *chances are* and *odds are*. *Language Sciences* 88: 101422. Special Issue *Complement and Parenthetical Constructions: Theory and Description*, ed. by An Van linden, Lieselotte Brems, & Lieven Vandelanotte. https://doi.org/10.1016/j.langsci.2021.101422
Mélac, Eric 2022. The grammaticalization of evidentiality in English. *English Language and Linguistics* 26 (2): 307–329. https://doi.org/10.1017/s1360674321000101
Molek-Kozakowska, Katarzyna 2017. Stylistic analysis of headlines in science journalism: A case study of *New Scientist*. *Public Understanding of Science* 26 (8): 894–907. https://doi.org/10.1177/0963662516637321
Moncomble, Florent 2018. The deviant syntax of headlinese and its role in the pragmatics of headlines. *e-REA* 15 (2). https://doi.org/10.4000/erea.6124
Noonan, Michael 1985. Complementation. *Language Typology and Syntactic Description*. Vol. 2, ed. by Timothy Shopen. 42–140. Cambridge: Cambridge University Press.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, & Jan Svartvik 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
Rudanko, Juhani 1989. *Complementation and Case Grammar*. State University of New York Press: Albany.
Schneider, Stefan 2007. *Reduced Parenthetical Clauses as Mitigators. A Corpus Study of Spoken French, Italian and Spanish.* Amsterdam: John Benjamins.
Thompson, Sandra & Anthony Mulac 1991. A quantitative perspective on the grammaticalization of epistemic parentheticals in English. *Approaches to Grammaticalization*. Vol. II, ed. by Elizabeth C. Traugott & Bernd Heine. 313–339. Amsterdam: John Benjamins.
Van Bogaert, Julie 2010. A constructional taxonomy of *I think* and related expressions: Accounting for the variability of complement-taking mental predicates. *English Language and Linguistics* 14 (3): 399–427. https://doi.org/10.1017/S1360674310000134
Vandelanotte, Lieven 2009. *Speech and Thought Representation in English: A Cognitive-Functional Approach*. Berlin and New York: Mouton de Gruyter.