Mode of production and (referential) cohesion: An L2 English corpus-based study of syntactic coordination

Ana Díaz-Negrillo [a,*], Mª Carmen Espinola Rosillo [b]

[a] *University of Granada, Spain*
[b] *University of Jaén, Spain*

A B S T R A C T

This study investigates mode of production effects on narrative discourse configuration and referring expression (RE) selection in contexts of reference maintenance in syntactic coordination. It looks at 3rd person singular grammatical subjects across intermediate and advanced L1 Spanish-L2 English, in contrast with L1 English. The data are written and spoken narratives from the COREFL corpus produced by the same participant and under the same task conditions. The study reveals an effect of the mode of production in L2 English (referential) cohesion, while no such an effect has been found in L1 English performance. In general, in L2 English deficits in spoken production seem to persist while those in written production disappear by advanced level. In particular, referential cohesion seems more challenging for L2 English learners than the discourse-syntactic configuration of their narrative texts. By advanced level, the learner's discourse configuration is comparable to that of native speakers' both in their written and in their spoken narratives. However, the spoken productions by advanced learners still show a higher amount of fuller REs, in comparison to the spoken productions by the native participants.

1. Cohesion and referring expressions (REs) in English

The tendency of human discourse is to maintain the reference to the same participant over stretches of language (DuBois, 1987, p. 829; Givón, 1983, p. 8; Givón, 2001, p. 423). This is typical of narrative discourse, where there is maintenance of reference to the narrative participants who are presented in relation to events and in a spatio-temporal frame (von Stutterheim & Klein, 1989, p. 51). As a result, in cohesive texts reference maintenance discursive configurations are expected to abound over other configurations where a new, or a different but not new, participant is activated.

Cohesive devices, and among them referential cohesive devices, also contribute to discourse cohesion. The REs for subject position speakers can choose from in English can be displayed in a scale of fuller to lighter forms:

indefinite NP - definite NP - (unstressed, stressed) pronoun - zero

$\longleftrightarrow$

+ formal code	− formal code
− mental activation	+ mental activation
+ discourse disruption	− discourse disruption

---

* Corresponding author.
 E-mail addresses: anadiaznegrillo@ugr.es (A. Díaz-Negrillo), espinola@ujaen.es (M.C. Espinola Rosillo).

In cognitive terms, the scale matches lower and higher degrees of mental activation of the referent denoted and hence, of cognitive accessibility to the referent in the addressee's mental representation of the discourse. A fuller form (+ formal code) is expected if the referent is new in the discourse or has a low degree of activation in working memory, and it is therefore associated with lower accessibility (see example (1), where "a woman" is new in the discourse[1]). Conversely, a lighter form (– formal code) is expected if the referent's mental activation is high and so is its degree of accessibility (see example (2) where the previously mentioned referent is now referred to using the pronoun "she" and zero). This activation status has been referred to as *accessibility* (Ariel, 1990), *cognitive status* (Gundel et al., 1993) or *attentional activation* (Givón, 2017). More generally, the encoding of reference into REs has been explained in terms of the iconicity principle, whereby the more disruptive, discontinuous, surprising, less predictable or hard to process the information, the more coding material must be assigned to it (Givón, 1983, p. 18).

(1) ***A woman_i*** *arrived at the station.*
(2) ***She_i*** *entered the building and* ***Ø_i*** *looked at the departure board.*

RE selection is also regulated by language-specific factors. In Spanish, a null-subject language, an anaphoric zero is grammatically licensed in subject position, being the default choice to mark maximal reference continuity. In English, in contrast, a non-null subject language, zero is restricted to very few specific discourse-syntactic contexts and, instead, the default selection to mark maximal reference continuity is an anaphoric pronoun (see Quesada, 2015, p. 32 for syntactic features of null and non-null subject languages like Spanish and English, respectively). Givón (2001, pp. 418ff) explains that an anaphoric zero is licensed in English in equi-topic,[2] equi-subject contexts and intra-sententially, typically in coordinate and participle clauses, as in (3) and (4) respectively[3] (see also Quirk et al., 1972, p. 555 for the grammaticality of zeros in coordinate clauses):

(3) ***Mary_i*** *came into the room and* ***Ø_i/she_i*** *looked around.*
(4) ***Ø_i*** *Walking away,* ***she_i*** *saw someone approaching.*

An anaphoric pronoun also establishes subject continuity in coordinate clauses, as shown in (3) above. Still, it has been argued that L1 English speakers often resort to coordinate clauses with zero anaphors as a device to enhance discourse cohesion (Leclercq & Lenart, 2013, p. 13–14; see also Williams, 1988, p. 356). The use of zero anaphors in English has also been described as a device to maximize the connectedness/sequentiality of successive of events in oral narratives, while the use of overt REs often creates a sense of the discreteness instead (Oh, 2006, p. 831–832). Similarly, the use of participle clauses and to-infinitive clauses, where zero anaphors are often used, has been described as a typical subject/topic continuity device (Givón, 1983, p. 23, 24). It seems then that, even if English is a non-null subject language, zero anaphors, and hence the constructions where they are licensed, are resorted to in order to create maximal discourse cohesion. This paper focuses on one of the discourse-syntactic contexts where maximal reference continuity may be established and where the least full of forms, zeros anaphors, are licensed in English, namely reference maintenance in contexts of syntactic coordination.

Grammatical factors readily interact with a number of discourse factors in RE selection, as shown above in the connection between syntactic role, syntactic structure and topicality. In addition to topicality, RE selection has been explored in relation to other factors like the occurrence of competing referents (Ariel, 1990; Arnold & Griffin, 2007; Fedorova, 2014; Kibrik, 2011), distance (Ariel, 1990; Chafe, 1976; Clark & Sengul, 1979; Givón, 1983), protagonisthood (Anderson et al., 1983; Karmiloff-Smith, 1985; Morrow, 1985), unity (Ariel, 1990; Chafe, 1976; Givón, 1983; Van Dijk, 1981; Vonk et al., 1992), etc. Among them, the distance between the RE and the co-referent antecedent stands as a crucial determining factor in RE selection. Givón (2001, p. 419) explains that the co-referent antecedent of zero anaphors and unstressed pronouns typically occurs in the immediately preceding clause, while NP co-referents tend to stand much further away in the discourse. This paper argues that in reference maintenance coordination distance may be explored in terms of the presence (or lack) of intervening subordination, and also the number of intervening subordinate clauses occurring between the equi-topic subjects in the parallel coordinate clauses in question.[4] One important aspect in these configurations where there is intervening subordination is whether the subjects/topics in the intervening subordinate clause(s) actually switch the reference when they do not co-refer with the subjects in the parallel coordinate clauses. Compare, for example (5), where the subjects in the subordinate clauses co-refer with those in the parallel main clauses, and (6), where the intervening subject does not co-refer with the subjects in the main clauses.[5] This paper follows the claims that subjects in subordinate clauses are syntactically less salient and discursively less topical than those in main clauses (Gundel et al., 1993, p. 279; see also Lozano, 2016, pp. 258–259), and therefore do not necessarily switch the reference chain in the main parallel coordinate clauses. The later would possibly explain why zero is still

---

[1] The RE under discussion and its co-referent antecedent are marked in bold.

[2] Topics are discourse participants. They are closely related to the syntactic roles they take in the clause structure. Not all topics are equally topical. Those participants in subject position outrank those in object position and the latter rank any others in the clause structure (Givón, 2017, p. 34).

[3] Examples (3) and (4) are from Givón (2001, p. 418).

[4] In fact, distance has usually being measured in terms of the number of clauses standing between the anaphoric form and the co-referent antecedent (see, for example, Givón, 1983).

[5] The examples from now onwards, including (5) and (6) are from the COREFL corpus. The corpus file codes are given at the end of each example.

selected in (6), despite the same gender but non-coreferential intervening subject ("he").

(5) *He$_i$ finally sat down Ø$_i$ to contemplate the other options that he$_i$ had and Ø$_i$ realized there was a note in the clothes of the baby* (EN_WR_20_14_AB).

(6) *The police man$_j$ sees what he$_k$ is doing and Ø$_j$ keeps a close eye on him* (EN_WR_20_14_CP).

In line with the above, this study explores whether a discourse-syntactic configuration contributing to text cohesion (reference maintenance coordination) and the REs used in this configuration are comparable in written and spoken performance across intermediate and advanced L1 Spanish-L2 English and in L1 English. For a wider view of reference maintenance coordination contexts, the paper also explores distance and RE selection, by additionally looking at configurations with intervening subordination. The next two sections review, in this order, RE selection in L2 English, and written vs. spoken production in L2 performance.

2. Reference maintenance and RE selection in L2 English

Reference maintenance has proven particularly problematic for L2 English learners when it comes to the selection of REs. This seems to be the case irrespective of the learners L1s (see, e.g., Crosthwaite, 2011 for L1 Korean; Quesada & Lozano, 2020 for L1 Spanish).[6] Regardless of whether research specifies the discourse context or not (for example, reference maintenance vs. switch reference), L2 English learners typically select fuller forms than those by native speakers, even at advanced levels (Crosthwaite, 2011; Hendriks, 2003; Kang, 2004; Leclercq & Lenart, 2013; Quesada & Lozano, 2020; Ryan, 2015).[7]

L2 English RE acquisition studies are particularly scarce. Experimental studies often explore L2 English learners whose L1 (vs. English) is a null subject language, i.e, in their L1 zero subjects are licensed in independent clauses, in order to verify whether L1 transfer occurs (see, e.g., Ballester, 2013 for L1 Spanish; Mitkovska & Bužarovska, 2018 for L1 Macedonian; Prentza, 2014 for L1 Greek). Experimental studies have also looked at REs in relation to syntactic factors. They have tested the Position of Antecedent Strategy (PAS; Carminati, 2002) to explore the learner's interpretation of subject pronouns in relation to two potential (subject and non-subject) antecedents (see, e.g., Cunnings et al., 2017 for L1 Greek learners; Santoro, 2020 for L1 Chinese learners). They have also investigated REs in relation to the Interface Hypothesis (IH), so as to test whether RE selection is fully acquirable, which the IH denies (see Quesada & Lozano, 2020 for a more detailed review of experimental studies on L2 English RE acquisition).

Corpus-based studies, like the present study, have often looked at RE selection along with aspects of (narrative) discourse, like the influence of the discourse configuration, competing antecedents, distance, character status, etc. Kang (2004) examined the oral production of intermediate L1 Korean-L2 English and Korean native speakers using a picture-based story retell task. The study found a character effect on RE selection in the learners' production –fuller forms were used for secondary characters– and deficits in their use of pronouns and zero anaphors. This was explained as redundancy and absence of L1 transfer, since learners used more pronouns in English than in their L1 Korean where zeros are possible. The study, however, did not consider the discourse configuration (reference maintenance/switch reference) or the REs' syntactic conditions (coordination, subordination, etc.), which largely determines the choice of REs. Additionally, the tasks included animate and inanimate characters, while animate characters have been reported to pose more difficulties than inanimate characters to L2ers when it comes to RE selection (cf. Lozano, 2009).

Crosthwaite (2011) looked at RE selection in upper-beginner L1 Korean-L2 English and in L1 English oral production using oral picture-based tasks. This study aimed to examine the effect of the interlocutor's intervention during the learners' oral production, which the results supported. In reference maintenance contexts, L1 English speakers produced a similarly high rate of NPs and pronouns, while they produced low rates of zeros. In contrast, learners produced higher rates of NPs than native speakers, lower rates of pronouns than native speakers and a very low rate of zeros. The results are in line with Kang's (2004) results since learners used a lower amount of pronouns and zeros than native speakers, hence being overexplicit. The study did not consider the REs' syntactic contexts, which limits the understanding of zero anaphor use.

Ryan (2015) analysed upper-intermediate L1 Mandarin Chinese-L2 English[8] and L1 English from an oral fill-retell task with a view to tapping into the reasons for redundancy in RE selection. The study analysed the presence of various types of discourse contexts both in the learners and native speakers' productions to investigate whether the possibilities to use lighter and fuller forms were comparable across both populations. The study confirmed redundancy in RE selection by learners and a character effect since learners were overexplicit when referring to the story main characters. The study did not find fewer opportunities to use lighter forms in the learners' texts. Rather, Ryan explains learners' redundancy in RE selection in pragmatic terms. For him, learners choose to be clear over being economical (cf. also Williams, 1988, p. 364, after 1987 for the concept of "hyperclarity"; and Lozano's 2016 Pragmatic Principles Violation Hypothesis PPVH). Again, Ryan did not consider the REs' syntactic contexts and therefore did not account for the related use of zeros in his study.

---

[6] RE selection in reference maintenance contexts has also been reported to be problematic in other L2s (e.g., Chini, 2005, as to L2 Italian; Lozano, 2016, 2018, as to L2 Spanish).

[7] Interestingly, it has been observed that L2 learners seem to accompany their speech with gestures to localize referents in the space of the interaction, which may be interpreted as an added degree of redundancy (Gullberg, 2006).

[8] Ryan (2015: 834) describes his L2 English participants' proficiency level as having recently passed the IELTS exam (or equivalent) and achieving a score "of at least" 6 in each band. This score is associated with B2 level according to the CEFR (see https://ielts.org/organisations/ielts-for-organisations/compare-ielts-ielts-and-the-cefr).

Contemori and Dussias (2016) looked at very advanced L1 Spanish-L2 English vs. L1 English from an oral story-retell task to investigate whether RE selection depends exclusively on discursive factors or, rather, also on the consideration of the listener's perspective, which their results finally supported. The study reported, first, that reference maintenance contexts are particularly problematic for L2 English speakers; and, second, interestingly, that Spanish L2 English learners produced more pronouns where natives selected NPs. For them, the high number of pronouns is a learners' default selection to deal with a high cognitive overload and less automatic processing in contexts they have more problems with.

The focus of the present study is the selection of REs in reference maintenance coordination which, as discussed above, favours the selection of the least marked of forms, zero anaphors. While the use of zero anaphor has been mentioned in some studies, its use in relation to the syntactic context which licenses it has not been closely looked at. Leclercq and Lenart (2013) looked at RE selection in oral film-retell narratives of various learner and native groups, including intermediate and advanced L1 French-L2 English and a native English control group. They reported a very low overall use of zero anaphors (6%–7%) in L1 French-L2 English learners, regardless of their proficiency level and, in turn, a higher proportion of NPs. In contrast, their native control group used a higher amount of zeros (20%–22.7%) and thus were more coherent than the learners. The authors explain that L2 English speakers' texts are even more coherent than L1 French texts given the higher proportion of zeros which, in turn, they explain in a "presumably" higher presence of coordination. Still, the syntactic context was not controlled for in their investigation and the latter explanation remains a hypothesis.

Quesada and Lozano (2020) examined the written story-retell production of L1 Spanish-L2 English learners across beginner to advanced levels against an L1 English control group. The authors looked at the discourse and syntactic configuration of the contexts, which made it possible to examine the participant's RE choices in reference maintenance coordination. They reported a high use of zeros (about 87%) in L1 English in reference maintenance coordination, while the rates are (much) lower in L2 English narratives in this discourse-syntactic context (beginners 19.2%, intermediate 20% and advanced 60%). The latter was explained again as L2 overexplicitness. The study showed that the syntactic context and RE choice that most favour discourse coherence in English is also an area of incomplete acquisition, even at advanced levels of competence. And this is the case of L1 Spanish-L2 English learners whose L1 typically allows zeros in the subject position. The latter, Quesada and Lozano (2020) explain, rejects the existence of L1 transfer of zeros as a mechanism at play in L1 Spanish-L2 English RE selection. Still, the study did not control for the incidence of reference maintenance coordination contexts in the narratives, so whether the learners and the native participants create similar opportunities for zero anaphors was unexplored. The latter is important because the presence of reference maintenance coordination determines the use of zero anaphors and, eventually, the degree of discourse cohesion.

Leclercq and Lenart's (2013) and Quesada and Lozano's (2020) touched upon, among other things, RE selection and coordination. The studies looked at either spoken performance (Leclercq & Lenart, 2013) or written performance (Quesada & Lozano, 2020). The present study aims to look closely at this syntactic context in reference maintenance configurations, both in oral vs. written performance.

3. Written vs. spoken L2 performance

According to L2 literature, the written and oral modes of production differ in various ways (Grabowski, 2007, pp. 168–170; Kuiken & Vedder, 2011, p. 92; Kuiken & Vedder, 2012, pp. 365–366; Vasylets et al., 2017, pp. 396–399; Williams, 2012, p. 322). Two fundamental aspects set them apart: time and permanence of record (Williams, 2012, p. 322). Speaking takes place online and there are usually time constraints at play at the stage of message formulation, involving planning and encoding. Time pressure may also be higher for the speaker at this stage because in oral communication the addressee may somehow be waiting to receive information from the addresser. At the stage of message execution, the articulation of an oral message is also considerably faster than writing.[9] Additionally, the offline time conditions of writing also allow monitoring and editing of the written message, while this is not favoured in spoken language. As to permanence of record, spoken language is not permanent while writing is. This has cognitive implications at the level of memory. Speaking imposes a higher cognitive strain on memory, as previous discourse information must be kept in working memory as speaking continues. In contrast, in writing the previous discourse can be read again anytime, demanding as a result reduced working memory resources. Interestingly, however, the very permanent nature of writing is often associated with demands for higher precision in language, while speaking is felt to be more tolerant to lower degrees of formal accuracy (Schoonen et al., 2009, pp. 79–80; see also Williams, 2012, p. 326).

The above means, first, that writing may benefit from greater opportunities to use explicit knowledge during message planning, linguistic encoding, monitoring and editing, because time conditions allow so. At the same time, writing is also felt to demand higher degrees of accuracy because of its permanence of record. Second, at a cognitive level, writing may benefit from lower cognitive demands compared to speaking, while higher demands on working memory, added to higher time pressure in speaking, may increase the cognitive effort in speakers. In contrast, writing benefits from access to long-term memory. In line with this, Williams (2012, p. 328) concludes that the cognitive window in writing is open somewhat wider and for a longer period of time so learners can test their hypotheses, while this is not always possible in speaking.

Accordingly, greater deficits in L2 RE selection may be expected in the spoken mode. L2 research looking closely at the effect of the mode of production on a variety of linguistic aspects has sometimes shown an effect on language performance and acquisition. Weissberg (2000) found that the written mode was first to show the emergence and development of language accuracy in L1

---

[9] Kuiken and Vedder (2012, p. 366) and Williams (2012, p. 325) report that writing is 5–8 times slower than speaking.

Spanish-L2 English. Vasylets et al. (2017) also concluded that L1 Spanish-L2 English writers obtained better results at syntactic complexity and lexical diversity, and at creating more complex, informationally dense idea units. Martínez-Flor (2006) obtained a higher frequency of linguistic forms in written production in a study on speech acts in L1 Spanish-L2 English. As to RE selection, studies contrasting written vs. spoken L2 performance do not seem to be available. In the specific case of corpus-based studies on L2 English RE selection, what is gathered is that similar findings are revealed whether the data explored is either written or oral (see Quesada & Lozano, 2020, for written data; see Contemori & Dussias, 2016; Crosthwaite, 2011; Kang, 2004; Leclercq & Lenart, 2013, Ryan, 2015, for spoken data). In contrast to L2 RE selection, there are studies on the effect of the mode of production on L1 RE selection yielding varying results depending on the language. No effect was found for L1 Vietnamese (e.g., Ngo et al., 2019), for Catalan L1 acquisition (Bel et al., 2010) and for L1 Spanish (Perales & Portillo, 2007). The opposite, however, was found for L1 Chinese (Christensen, 2000). No studies on L1 English RE selection seem to be available.

This study will explore the effect of mode of production (written vs. spoken) on text cohesion in and across intermediate and advanced L1 Spanish - L2 English and L1 English. In particular, it will explore possible mode effects on the occurrence of reference maintenance coordination contexts, and on RE selection in this discourse-syntactic configuration. By looking at within group performance in the two types of performance data, the paper aims to find out whether the learners' choices are affected by the mode of production. By looking at written and spoken performance across proficiency levels, the paper will explore whether deficits persist (or not) in one of the two modes in the population under study. Ultimately, the results may be taken as an indication of whether mode of production constrains evidence of L2 acquisition.

4. Research questions

RQ1 looks at possible mode effects on the discourse-syntactic configuration under study, i.e. reference maintenance coordination. The aim is to explore whether the opportunities created for the use of zero anaphors are the same in both modes and across proficiency levels. This research question has two parts:

RQ1a: Is there a mode effect on the discursive configuration of the narratives across intermediate and advanced L2 English vs. L1 English discourse?

H1a: The tendency of human discourse is to maintain the reference to the same participant over stretches of language. Ryan (2015) also found that the opportunities of minimal RE use were comparable in oral upper-intermediate L1 Mandarin Chinese-L2 English and in L1 English narratives. Hence, no mode effects are expected in L1 English and L2 English.

RQ1b: Is there a mode effect on the incidence of reference maintenance coordination across intermediate and advanced L2 English vs. L1 English discourse?

H1b: Coordination is one of the few constructions licensing a zero anaphor in subject position in English. This is the least marked of the subject REs in English, and its use contributes to discourse cohesion. This discourse-syntactic configuration is also present in L1 Spanish with the same cohesive function. According to this, and also in line with Ryan's (2015) results mentioned above, both in L1 and L2 English coordination is predicted to be used in combination with reference maintenance, and regardless of the mode of production, as a discourse-syntactic device fostering discourse cohesion.

RQ2 looks at a possible mode effect on REs selection in reference maintenance coordination. The aim is to explore whether the use of zero anaphors is comparable in both modes and across proficiency levels. This research question has two parts:

RQ2a. Is there a mode effect on the selection of REs in reference maintenance coordination across intermediate and advanced L2 English vs. L1 English discourse?

H2a: L1 literature on RE selection has often found no effect of the mode of production, while in L2 literature deficits in RE selection have been found in studies exploring either written production or spoken production. In L2 English the actual characteristics of the spoken mode may impose additional cognitive demands on L2 learners. No mode effects are expected in L1 English, while mode effects are predicted in L2 English with deficits surfacing more markedly in the spoken mode of production.

RQ2b: Is there a mode effect on the selection of REs in reference maintenance coordination across intermediate and advanced L2 English vs. L1 English discourse in contexts of distant coreference?

H2b: This is an exploratory RQ which looks at mode effects in contexts of reference maintenance coordination of distant coreference, specifically, when there is intervening subordination between the RE and the co-referent antecedent in the parallel coordinate clause. It has been shown that the REs used in contexts of distant coreference are typically fuller, but mode effects in these contexts are unexplored. No mode effects are expected in L1 English, while mode effects are predicted in L2 English with deficits surfacing again more markedly in the spoken mode of production.

5. Method

The study uses corpus data from the COREFL (*The Corpus of English as a Foreign Language*), a corpus of beginner to advanced L2

Table 1
Participants.

Intermediate participants	Test score	Sex	Advanced participants	Test score	Sex	Native participants	Sex
B1_19_13_14_JMR	34	M	C1_18_13_14_RLR	50	F	19_14_SC	F
B1_22_14_14_GG	34	F	C1_18_13_14_AGL	50	F	19_14_AB	F
B1_19_12_14_AFL	35	M	C1_24_16_14_DPD	50	F	19_14_CP	F
B1_18_11_14_MRC	36	F	C1_19_13_14_IGT	51	F	20_14_EES	F
B1_18_12_14_MGM	36	F	C1_21_13_14_ARP	51	F	20_14_PM	M
B1_19_11_14_NLLB	36	F	C1_21_15_14_PGM	51	M	20_14_SM	F
B1_18_15_14_CAM	37	F	C1_21_8_14_LAR	52	M	20_14_TK	F
B1_18_10_14_JAVG	38	M	C1_18_12_14_LBT	52	F	21_14_AF	F
B1_20_11_14_FER	38	M	C1_19_13_14_MHM	54	F	21_14_GLN	F
B1_18_12_14_CRM	39	F	C1_19_9_14_VMFV	54	F	21_14_TS	M

English of various L1s, and of comparable L1 English data (Lozano et al., 2020).[10] The data selected consist of written and spoken texts by the same participant from the L1 Spanish-L2 English and the L1 English corpus components. The same variables (task and procedures) were used in the collection of all the data (both native and non-native, written and spoken). The task was a story-retell based on a 4-min video-clip of the silent film *The Kid* by Charles Chaplin.[11] The participants were asked to watch the video-clip and then retell the story. They were allowed to watch the video-clip more than once and the researcher was never present during task completion. First, the written data along with the participant's demographical information and proficiency level were collected online. Then the spoken data were collected on site, at least 15 days after the written data collection, so as to avoid bias in the language produced.[12]

The L2 English data covers lower intermediate (B1) and lower advanced (C1) levels.[13] The reasons for this selection are twofold. First, the two selected levels provide the necessary distance for signs of language development to become evident, especially if the native speaker is taken as the final stage of development (L1 English component). Second, the beginner levels (A1/A2) were not included because beginners have difficulty producing data for the task in question.

There are 10 participants in each proficiency group (intermediate, advanced and native) (see Table 1). The L2 English participants were 14 female and 6 male university students in the English Studies or Translation and Interpreting undergraduate programmes at the University of Granada (Spain), between 18 and 24 years of age. The learner participants are cited in the table by partial reference to the file name of the corresponding texts stating, in this order, their proficiency level, their age, the number of years studying English (length of exposure), the task number and their first name and surname initials. Table 1 also shows the score obtained in the proficiency test[14] (Appendix 1 gives further information about the L2 participants). The L1 English participants were 8 female and 2 male American university students who were learning Spanish at the time of the data collection. The native participants are cited in the table by partial reference to the file name of their texts stating, in this order, their age, the task number and their initials. As shown, the native participants' ages range between 19 and 21.

Overall, 6 batches of data were analysed comprising the three proficiency levels (intermediate, advanced and native), and the two text types for each of the levels (written, WR, and spoken, SP). Since 10 participants (see above) represent each proficiency level, the dataset consists of 60 texts. A summary of the data is in Table 2.

Data annotation and statistical analysis was conducted using UAM CorpusTool (O'Donnell, 2009). Third person singular animate subjects in non-subordinate clauses, that is, in main clauses were analysed. The tagset used is in Fig. 1. The tagset categories (with examples) are explained below.

Table 2
Summary of the corpus data.

Sub-component	Min.-max. text words	Partial word count	Total word count
Intermediate WR	117–333	2090	6039
Intermediate SP	246–578	3949	
Advanced WR	135–288	1992	5834
Advanced SP	187–579	3842	
Native English WR	182–323	2208	6821
Native English SP	293–653	4613	
			18694

---

[10] The corpus and its documentation can be found at www.corefl.learnercorpora.com.

[11] The video-clip can be found at https://www.youtube.com/watch?v=4QkTNJFhu-g.

[12] A reviewer points out there may still be a task-order effect despite the 15-day gap between the two tasks. In response to this remark, and as suggested by another reviewer, no such effect has been noticed in this investigation, since the participants were actually more redundant in the spoken than in the written mode.

[13] The placement test used was the Oxford Quick Placement Test (2001).

[14] The proficiency test score ranges are: 30–39 for B1 and 48–54 for C1.
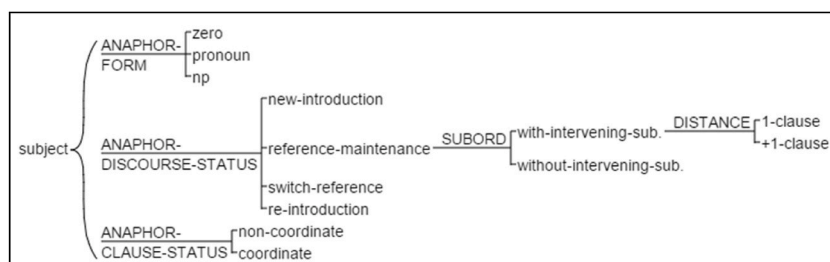
Fig. 1. Tagset: linguistic categories.

i) The anaphoric forms are analysed first into zero anaphors, anaphoric pronouns and NPs. No distinctions were made between stressed and unstressed anaphoric pronouns in the spoken data, so the results were comparable across the two modes of production.

ii) The discourse status of the subject is analysed into new introduction, when a participant is mentioned for the first time in the story (see "A woman with a baby carriage" in (7) below), reference maintenance, when reference to the same topic is maintained (see, for example, Ø in (7) and (8)), switch reference, when the topic is not maintained from the preceding subject (see "the man" in (7)), and re-introduction, when a referent is mentioned again later in the story (in our analysis, 4 clauses back).

iii) Distance is analysed next, in terms of whether or not there is intervening subordination between the subject under analysis and the antecedent co-referent in the parallel coordinate clause. The number of intervening clauses is also specified. For example, Ø in (7) is analysed as a non-distant RE (without intervening subordination) and Ø in (8) as a +1 clause distant RE (with intervening subordination).

iv) Finally, the grammatical subjects are analysed in terms of whether they stand in coordinate constructions or not. Please note that, for example, in (7) "A woman with a baby carriage" is analysed as non-coordinate and Ø and "the man" as coordinate. Only the subjects which stand in the 2nd and subsequent coordinate clauses are analysed as standing in coordinate clauses, because these are the positions where zero anaphors are licensed in English.

(7) ***A woman with a baby carriage**$_i$ comes by and **Ø**$_i$ appears to be looking for someone and **the man**$_j$ assumes the baby belongs to her.* (EN_WR_20_14_CP).

(8) ***Chaplin**$_k$ thought the baby he found could probably be hers and **Ø**$_k$ put him on that same buggy.* (ES_WR_C1_24_16_14_DPD).

The annotation of the 60 texts comprised the analysis of 1478 grammatical subjects. Partial counts are in Table 3.

The 60 texts were also tagged for their proficiency level (intermediate, advanced or native) and mode of production (written vs. spoken) into 6 different subsets, as shown below, for contrast analysis across the subsets.

Finally, after the grammatical subjects and the texts were tagged, the various texts subsets (see Fig. 2 above) were contrasted for the different linguistic categories and combination of categories. Descriptive and inferential statistics (χ^2) was performed using UAM corpus tool.[15]

The exploration of the tagged data and the presentation of the results in the next section take two different angles. First, intra-group written vs. spoken performance comparisons were made to investigate whether the written and spoken choices by a specific group differed or not. Second, for the written narratives, on the one hand, and for the spoken narratives, on the other, comparisons were made

Table 3
Grammatical subjects analysed.

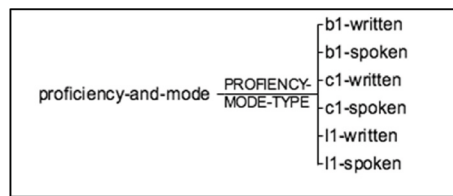| Sub-component	Grammatical subjects	TOTAL
Intermediate WR	209	472
Intermediate SP	263	
Advanced WR	171	423
Advanced SP	252	
Native English WR	193	583
Native English SP	390	
		1478

Fig. 2. Tagset: text subsets.

between the two L2 groups and between each L2 group and the native speakers. These comparisons were made so as to explore whether there was progress and eventually native-like behaviour in L2 English in each mode of production, and whether native-like behaviour revealed earlier in one of the two production modes.

6. Results

As to RQ1, which looked at a possible mode effect in the discourse-syntactic configuration of the narratives in L2 English, Fig. 3[16] (RQ1a) shows that the preferred discourse context in the analysed texts is reference maintenance for all language groups and for written and spoken discourse. Still, within group statistical analysis reveals significant differences between both modes of production in the advanced learners ($\chi^2 = 7.30$, p $= 0.0069$, h $= 0.267$ for reference maintenance; $\chi^2 = 5.43$, p $= 0.0198$, h $= 228$ for switch reference), using a higher proportion of reference maintenance contexts in their spoken narratives. No differences in the intermediate learners or in L1 English have been found across the two modes of production. Statistical analysis across proficiency groups reveals no significant differences in the written mode. Significant differences across proficiency groups are found only in the spoken mode between the intermediate learners and the native speakers ($\chi^2 = 3.99$, p $= 0.0459$, h $= 0.159$ for reference maintenance), the native speakers using a higher proportion of reference maintenance contexts in their spoken narratives.

As can be observed in Fig. 4 below (RQ1b), which shows the distribution of discourse contexts associated with coordinate clauses, coordination is typically used for reference maintenance in all the participants' narratives. Still, within group statistical analysis reveals significant differences in the advanced learners ($\chi^2 = 6.47$, p $= 0.0110$, h $= 0.400$ for reference maintenance), who markedly use coordinate constructions for reference maintenance in their spoken production. Across proficiency levels, statistical analysis reveals differences but only in the spoken narratives between the learner groups ($\chi^2 = 7.30$, p $= 0.0069$, h $= 0.395$ for reference maintenance; $\chi^2 = 7.68$, p $= 0.0056$. h $= 0.416$ for switch reference), where reference maintenance is higher in the advanced learners; and between the intermediate group and the native speakers ($\chi^2 = 6.95$, p $= 0.0084$, h $= 0.324$ for reference maintenance; $\chi^2 = 5.42$, p $= 0.0199$, h $= 0.285$ for switch reference), where reference maintenance is higher in the natives. No differences have been found across proficiency levels in the written narratives.

As to RQ2a, which looked at a possible mode effect on RE selection in reference maintenance coordination, Fig. 5 below shows that in L1 English zero is the most widely used anaphoric form in reference maintenance contexts in coordination, with a similar distribution both in their written and their spoken discourse. In contrast, this is not always the case in L2 English.

Within group statistical analysis reveals significant differences in the intermediate group ($\chi^2 = 8.15$, p $= 0.0043$, h $= 0.484$ for zero;
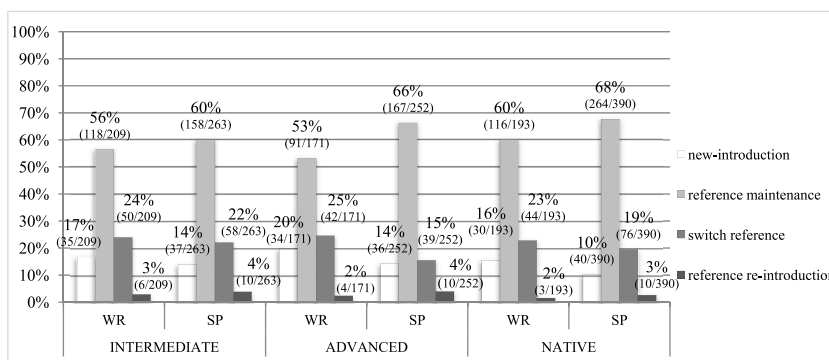

Fig. 3. Discurse configuration of narratives across language groups and text types.

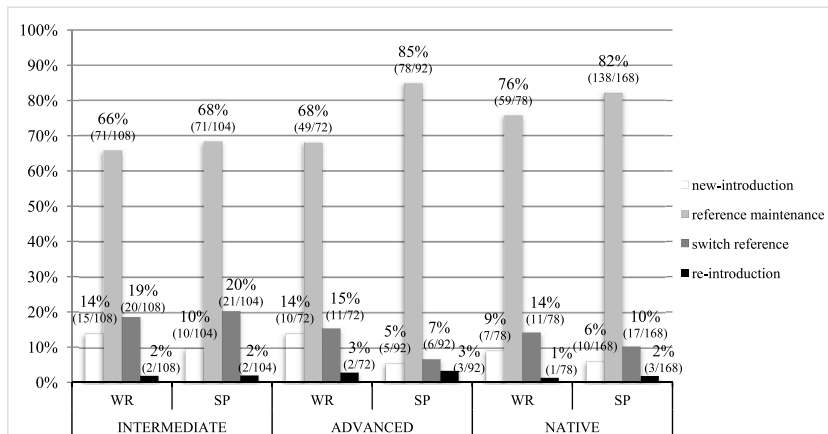[16] Percentages are shown without decimals. Appendix 2 shows percentages with decimals.

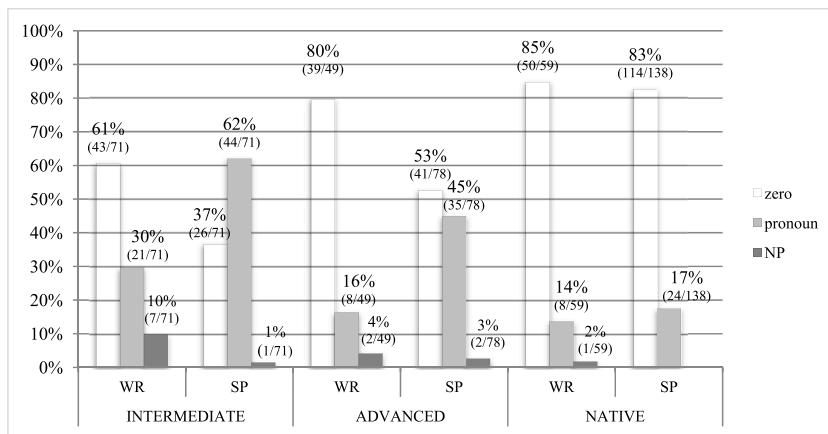**Fig. 4.** Discourse functions of coordination across language groups and text types.

Fig. 5. REs in reference maintenance coordination.

$\chi^2 = 15.01$, p = 0.0001, h = 0.663 for pronoun; $\chi^2 = 4.77$, p = 0.0290, h = 0.401 for NP), and in the advanced group ($\chi^2 = 9.43$, p = 0.0021, h = 0.582 for zero, h = 0.582; $\chi^2 = 10.95$, p = 0.0009, h = 0.636 for pronoun). In both cases, the proportion of fuller forms is significantly higher in their spoken narratives than in their written narratives. In contrast, there are no statistically significant differences between the written and spoken L1 English narratives. Across proficiency groups, in the written narratives there are significant differences between the two learner groups ($\chi^2 = 4.85$, p = 0.0276, h = 0.420 for zero), and between the intermediate learners and the native speakers ($\chi^2 = 9.25$, p = 0.0023, h = 0.555 for zero; $\chi^2 = 4.77$, p = 0.0290, h = 0.396 for pronoun). However, there are no significant differences between the advanced and the native choices in the written narratives. In spoken production, there are significant differences between the two learner groups ($\chi^2 = 4.36$, p = 0.0367, h = 0.345 for pronoun), and between both learner groups and the native speakers (vs. intermediates $\chi^2 = 44.83$, p = 0.0000 for zero, h = 0.981; $\chi^2 = 42.45$, p = 0.0000, h = 0.952 for pronoun; vs. advanced learners $\chi^2 = 22.20$, p = 0.0000 for zero, h = 0.659; $\chi^2 = 18.96$, p = 0.0000, h = 0.608 for pronoun). This means that, even if with proficiency the amount of zeros increase in both modes, the significant differences against native performance persist in advanced spoken English.

As to RQ2b, which looked at the effect of the production mode on RE selection in contexts of reference maintenance coordination of distant coreference, Fig. 6 shows the results. Statistical analysis yields no significant differences between written vs. spoken productions by any of the groups. Still, as can be observed from Fig. 6, the distribution of REs L2 performance is considerably different in their written and in their spoken texts.[17] Strikingly, in advanced performance, while zeros prevail in their written texts (53%), pronouns do in their spoken texts (65%). In L1 English, in contrast, the distribution of the percentages in their written and spoken productions seems similar.

---

[17] Marginal differences have been found in the intermediates ($\chi^2 = 3.49$, p = 0.0619, h = 0.652 for NP) and in the advanced learners ($\chi^2 = 3.11$, p = 0.0779, h = 0.551 for zero; $\chi^2 = 3.36$, p = 0.0669, h = 0.576 for pronoun). Frequencies are low at this point, which may have affected the statistical results. This is because, as will be shown later, intervening subordination between the subjects of coordinate clauses is scarce.
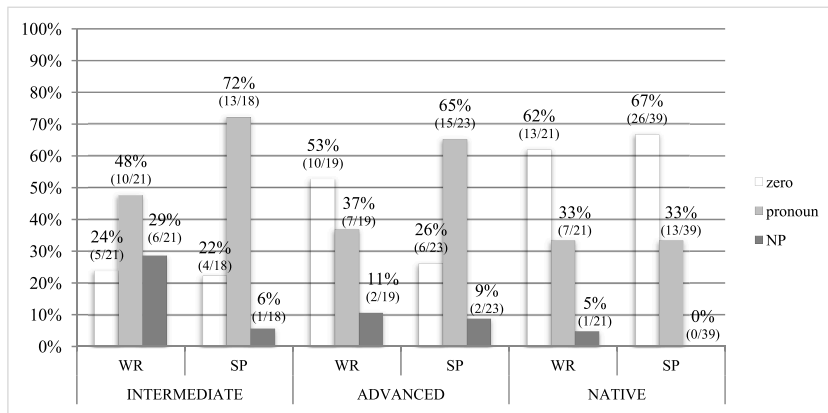
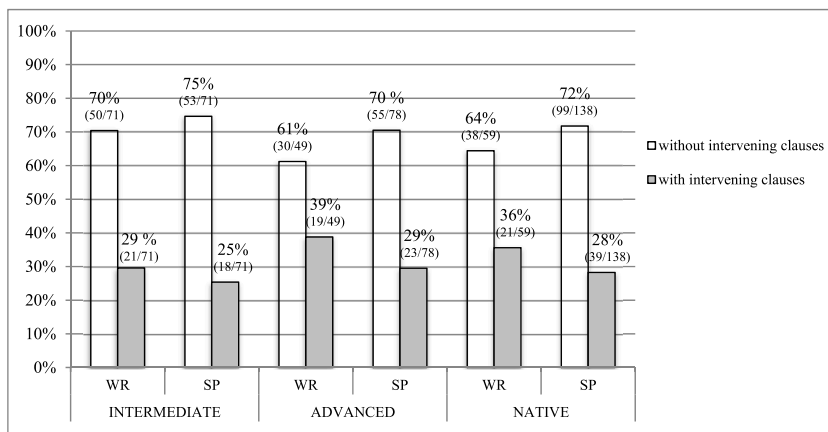**Fig. 6.** REs in reference maintenance coordination with intervening subordinate clauses.

Fig. 7. Intervening subordination in contexts of reference maintenance coordination.

Across proficiency groups, in the written narratives there are statistically significant differences between the intermediate learners and the native speakers ($\chi^2 = 6.22$, p = 0.0126, h = 0.792 for zero; $\chi^2 = 4.29$, p = 0.038, h = 0.688 for NP), the intermediate learners predominantly using pronouns and the native speakers predominantly using zeros. However, there are no significant differences between the advanced learners' and the native speakers' choices in their written productions. In the spoken narratives, there are statistically significant differences between the learners and the native speakers, the learners predominantly using pronouns and the native speakers zeros (vs. intermediates $\chi^2 = 9.76$, p = 0.0018, h = 0.929 for zero; $\chi^2 = 7.51$, p = 0.0061, h = 0.800 for pronoun; vs. advanced learners $\chi^2 = 9.54$, p = 0.0020, h = 0.839 for zero; $\chi^2 = 5.94$, p = 0.0148, h = 0.649 for pronoun).
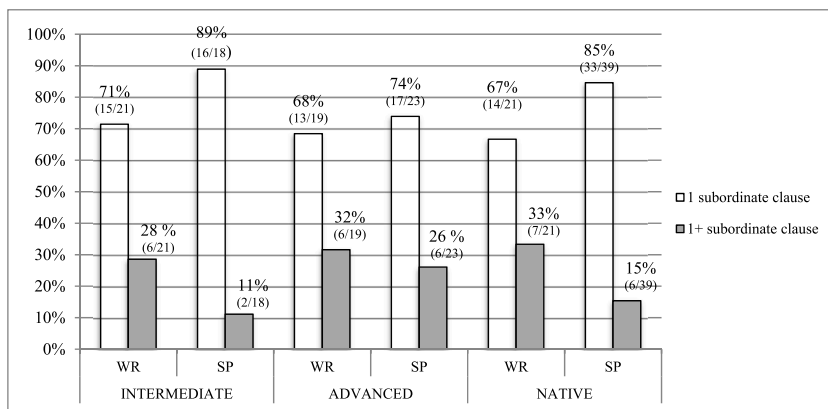

Fig. 8. Amount of intervening subordination in contexts of reference maintenance coordination.

As background information to RQ2b, it may be worth mentioning that in all the groups and in both modes of production absence of intervening subordination outweighs the presence of subordination between the anaphoric form and the co-referent antecedent (see Fig. 7). Also in contexts where there is intervening subordination, in all the groups and in the written and spoken narratives, there tends to be only one subordinate clause between the grammatical subjects under analysis (see Fig. 8).

7. Discussion

RQ1 looked at possible mode effects in the discourse-syntactic configuration of intermediate, advanced and L1 English written and spoken narratives. For RQ1a and RQ1b, as to within group mode effects, the results confirm no mode effects in L1 English. In L2 learners unexpected mode effects are found in advanced L2 English, where an even higher proportion of reference maintenance contexts in general and also in combination with coordination is found in their spoken narratives. No mode effects are found in intermediates' productions. However, in relation to this latter group, the comparisons across the various proficiency groups reveal attainment deficits in the intermediates' spoken productions in comparison with the native speaker's spoken productions: fewer reference maintenance contexts than in the native speakers' spoken productions are created overall in the intermediates' spoken productions, and also when reference maintenance is explored in combination with coordination. These results contrast with Ryan's (2015) results. Ryan (2015) found no differences in the distribution of the discourse contexts which favour light REs across the native and non-native oral narratives in his study. The slightly higher proficiency level of Ryan's (2015, p. 834) L2 English participants (of at least B2 level), in comparison to that of our intermediate participants (B1 level), may explain the diverging results for this proficiency group. In contrast, in our study native-like behaviour is found in the intermediates' written narratives as well as in the advanced written and spoken narratives. The native-like behaviour found in our advanced participants' spoken texts is according to Ryan's results.

Therefore, as predicted in RQ1, the results confirm no mode effects on reference maintenance coordination in L1 English. However, mode differences are found in the advanced group, and attainment deficits are revealed in the spoken productions by the intermediates. These two findings were unpredicted. The latter finding is the case, even if reference maintenance is supposedly the general tendency in human discourse (DuBois, 1987, p. 829; Givón, 1983, p. 8; Givón, 2001, p. 423), and coordination and zero subjects are also common cohesive devices in the learners' L1, i.e. Spanish. The attainment deficits revealed in the intermediates' spoken productions are resolved, however, by advanced level.

RQ2a looked at possible mode effects on RE selection in reference maintenance coordination across intermediate, advanced and L1 English written and spoken discourse. Within group results confirm the predictions for L1 English, for which no mode effects have been found on RE selection: zeros are preferred both in the native written and spoken narratives. Lack of mode effects is according to research findings for other native languages (Bel et al., 2010 for L1 Catalan acquisition; Perales & Portillo, 2007 for Spanish; Ngo et al., 2019 for Vietnamese; cf., however, Christensen, 2000 for Chinese). Evidence for L1 English did not seem to be available before this research, so the findings from this study should be taken as a relevant contribution. Overall, the L1 English preference for zero anaphors in reference maintenance coordination is consistent with the findings revealed by L1 English control groups used in L2 English research on either written or spoken performance (Quesada & Lozano, 2020, pp. 15–16; Leclercq & Lenart, 2013, p. 14, respectively). The findings also support L1 literature statements where zero is described as a device enhancing cohesion (Ariel, 1988; Givón, 1983, p. 18; Oh, 2006, p. 832).

As to L2 production, and unlike for L1 English, within group results confirm mode effects both in the intermediates and in the advanced learners, who select fuller forms in their spoken narratives than in their written narratives. In terms of language attainment, deficits are found in the intermediates' written and spoken productions, selecting a higher amount of fuller REs than the natives in both modes respectively. Statistically significant differences between both learner groups suggest that with proficiency, and in both modes, the learners seem to gradually become aware that zero is the target choice in this discourse-syntactic context. Still, native-like performance is only revealed in the advanced learners' written narratives. In spoken advanced production the amount of fuller REs is still significantly higher than in spoken L1 English. The developmental trend found in the written mode is as in Quesada and Lozano (2020) –a study of written narratives– but it differs from it in that in this other study the advanced learners select fewer zeros and do not show native-like behaviour. Quesada and Lozano (2020) also used written data from the COREFL corpus. However, first, the data in this other study comes from a picture-based task, while the present study data is from a video-based task, so the differences may be explained in terms of a task-effect. In relation to this, Kuiken and Vedder (2012, p. 365) explain that studies contrasting results from different tasks have revealed that picture-based narratives yield other results than video-based narratives. Second, Quesada and Lozano's (2020) analysis of topicality in main and subordinate clauses differs from our analysis. For them a different subject referent in intervening subordinate clauses breaks the referential continuity. In contrast, the present study focuses on main clauses and follows the suggestions by Gundel et al. (1993, p. 279) and Lozano (2016, pp. 258–259) and intervening subordinate clauses are analysed as not interfering with the topicality of parallel coordinate clauses. Therefore, a different approach in the analysis may also have yielded different results.

Therefore, as predicted in RQ2a the mode of production has no effects on L1 English RE selection in reference maintenance coordination but consistently reveals itself as a determining factor both in the intermediate and advanced levels. In L2 English the spoken

mode is associated with a higher amount of fuller forms compared both to their written choices, and also the choices by the native speakers in their spoken narratives.

RQ2b looked at mode effects on RE selection in reference maintenance coordination in contexts of distant coreference across intermediate, advanced and L1 English written and spoken performance. Interesting findings are revealed by some background results to this RQ: the occurrence of intervening subordination between the RE in question and the subject co-referent antecedent does not abound (it only occurs in around 30% of the cases for all the batches, see Fig. 7), and when there is, it is often limited to one intervening subordinate clause. This finding partially confirms that in contexts of reference maintenance coordination the subject co-referent antecedent of zeros is typically in the immediately preceding coordinate clause (see Williams, 1988, p. 356 for similar findings for L1 spoken English) – however, notice that zero is still the prevailing choice in L1 English in reference maintenance coordination contexts of distant reference (see Fig. 6). A methodological consequence of the low amount of intervening subordination is that frequencies are lower for this RQ.

In relation to the hypothesis for RQ2b, within group analyses yield no statistically significant differences (only marginal) across both mode of production neither for L1 English or for L2 English. The predictions are therefore not confirmed for the learners for which greater overexplicitness was expected in the spoken mode. Further investigation on a larger dataset is desirable, given that, as mentioned above, the frequencies are lower for this RQ. However, the results across proficiency groups confirm attainment differences across production modes in advanced L2 performance. No statistically significant differences were found in the written mode between advanced L2 English and L1 English: zeros are preferred by both groups, though they rank slightly lower in the advanced learners' written narratives. In contrast, statistically significant differences were found in the spoken mode between the advanced group and the native speakers: in spoken L2 English pronouns prevail over zeros, while in spoken L1 English zeros prevail, just as in spoken L1 English performance. The results for the intermediates show statistical differences in comparison with the L1 speakers' both in their written and their spoken performance respectively.

Therefore, as predicted in RQ2b the mode of production does not seem to have effects on L1 English RE selection in reference maintenance coordination in contexts of distant coreference. As to L2 English, no mode effects have been found for the intermediates in the intra-group analysis, and they show deficits in both modes of production in the comparisons against L1 English performance. The mode of production does not reveal itself as a factor in advanced L2 English in the intra-group statistical analysis either. However, higher amounts of fuller forms are used in the spoken productions by advanced L2 English learners in comparison to the choices by the native speakers in their spoken texts, and attainment deficits are statistically confirmed for the advanced learner group in the spoken mode and not in the written mode. Analysis of a larger dataset is desirable to confirm these findings.

The study then confirms an effect of the mode of production in L2 English (referential) cohesion, while no such an effect has been found in L1 English performance. In general, deficits in spoken production seem to persist while those in written production identified in intermediate performance seem to be resolved by advanced level. In particular, referential cohesion seems more challenging for L2 English learners than the discourse-syntactic configuration of their narrative texts. It may be argued that the hard-pressed conditions of speaking add a degree of complexity in L2 RE selection, resulting in the selection of a higher proportion of fuller REs in this mode. One important aspect in RE selection is memory: in order to select the appropriate anaphoric expression, a number of factors (distance, topicality, potential co-referents, character status, etc.) and of varied linguistic nature (syntactic, morphological, discursive, pragmatic) should be recalled and evaluated from the (preceding) context. In contrast, in the less taxing time conditions of the written mode, L2 learners can more carefully consider the linguistic aspects ruling the anaphor, and verify their selection during modulation. In relation to this, it has also been argued that the slower pace of the written mode offers L2 learners the opportunity to access their explicit knowledge and that the written mode promotes a focus on form (Williams, 2012, pp. 323, 328; see also Ellis, 2009, p. 12; Myles, 2015, p. 329 on explicit knowledge and written performance). Zero anaphors are syntactically (and discursively) granted and exceptional in English. Accordingly, it may be hypothesised that in their spoken performance learners readily apply the more general grammatical rule that grammatical subjects are generally overtly marked in English, as opposed to Spanish, where they are often unmarked (non-null vs. null subject language, respectively). This rule would possibly be part of their implicit knowledge and applied across the broad. In contrast, in written performance learners can more carefully retrieve the more specific rule for zero anaphors in English, which L2 English learners seem to be familiar with (zeros are selected in around 60% of the cases in the written narratives by the intermediates, Fig. 5), though they do not seem to master it until advanced stages of development. Further research is necessary, however, in order to confirm this hypothesis.

Importantly, the latter hypothesis coexists with L2 learners' general tendency to select fuller forms than those by the L1 users, as a pragmatic strategy to enhance clarity over redundancy (see Lozano, 2016; Ryan, 2015; Williams, 1988) rendering, as a result, less cohesive texts than those by the native users. In English reference maintenance coordination, where both zero anaphors and pronouns are licit, but pronouns favour discreteness and zeros enhance connectedness (Oh, 2006, pp. 831–832), L2 English participants have proved once again to preferably select the RE consisting of more coding material, the one which is more disruptive, but also the one that favours clarity. This paper has shown that this is specially the case in the L2 learners' spoken narratives, where cognitive demands are higher, as well as the risks of communication breakdowns.

8. Conclusion

The study has explored the effect of mode on (referential) text cohesion in intermediate (B1) and advanced (C1) L2 English and L1 English. It has focused on reference maintenance in contexts of syntactic coordination, and has looked at their distribution and at the REs used in these configurations, both in written and spoken narratives. The study confirms an effect of the mode of production in L2 English (referential) cohesion, while no such an effect has been found in L1 English performance. In general, deficits persist in spoken

production while those in written production seem to be resolved earlier. In relation to the distribution of reference maintenance coordination configurations, and while it seems that the maintenance of topical reference is typical of human discourse, intermediate L2 English learners show deficits in their spoken performance, creating an unexpectedly lower amount of reference maintenance contexts. Referential cohesion, on the other hand, is more challenging for L2 English learners than the discourse-syntactic configuration of their narratives. Deficits in RE selection, that is, higher amounts of fuller REs than in L1 English, are found in intermediate production regardless of the mode of production and also in advanced spoken production.

This paper confirms previous findings as to L2 RE selection in that L2 leaners' tend to be redundant in the REs they select and, importantly, throws light to a crucial aspect in language acquisition, that is, how L2 performance is modulated by the production mode. This paper supports claims that triangulation of written and spoken performance data can provide complementary insights as to L2 behaviour (Lozano, 2022). This methodological approach becomes even more promising when the data come from the same participant, task and corpus (Lozano, 2022, p. 967), which is the case of the present study. Additionally, written vs. spoken performance data have been associated with the manifestation of different types of knowledge (explicit vs. implicit knowledge, respectively) (Myles, 2015, p. 329). The findings in this paper provide further evidence for SLA theory to continue to investigate what type of knowledge (explicit or implicit, or a combination of both) is reflected in each production mode and, overall, whether or not one of the modes better reflects language acquisition.

Funding

CRediT author statement

Ana Díaz-Negrillo: Conceptualization, Methodology, Formal analysis, Data curation, Writing - Original Draft, Writing - Review & Editing, Supervision.

Mª Carmen Espínola Rosillo; Validation, Data curation, Writing - Original Draft, Writing - Review & Editing.

Appendix 1. L2 participants: further details

Intermediate participants (B1)							
File name	Test score (/60)*	Age	Age of onset	Length of exposure (years)	Stays in English-speaking countries	Length of stay (months)	Gender
B1_19_13_14_JMR	34	19	6	13			M
B1_22_14_14_GG	34	22	8	14	UK	10	F
B1_19_12_14_AFL	35	19	7	12	US	10	M
B1_18_11_14_MRC	36	18	7	11			F
B1_18_12_14_MGM	36	18	6	12			F
B1_19_11_14_NLLB	36	19	8	11			F
B1_18_15_14_CAM	37	18	3	15			F
B1_18_10_14_JAVG	38	18	8	10			M
B1_20_11_14_FER	38	20	9	11			M
B1_18_12_14_CRM	39	18	6	12			F

*B1: 30-39.

Advanced participants (C1)							
File name	Test score (/60)*	Age	Age of onset	Length of exposure (years)	Stays in English-speaking countries	Length of stay (months)	Gender
C1_18_13_14_RLR	50	18	5	13			F
C1_18_13_14_AGL	50	18	5	13			F
C1_24_16_14_DPD	50	24	8	16	UK	18	F
C1_19_13_14_IGT	51	19	6	13	Canada	10	F
C1_21_13_14_ARP	51	21	8	13			F
C1_21_15_14_PGM	51	21	6	15	Ireland	8	M
C1_21_8_14_LAR	52	21	13	8			M
C1_18_12_14_LBT	52	18	6	12			F
C1_19_13_14_MHM	54	19	6	13	UK	9	F
C1_19_9_14_VMF	54	19	10	9	UK	1	F

*C1: 48-54.

Appendix 2. Figures: further details

Fig. 3. Discursive configuration of narratives across language groups and text types.

	Intermediate written		Intermediate spoken		Advanced written		Advanced spoken		Native written		Native spoken	
	n	%	n	%	n	%	n	%	n	%	n	%
new-introduction	35	16.75	37	14.07	34	19.88	36	14.29	30	15.54	40	10.26
reference maintenance	118	56.46	158	60.08	91	53.22	167	66.27	116	60.10	264	67.69
switch reference	50	23.92	58	22.05	42	24.56	39	15.48	44	22.80	76	19.49
re-introduction	6	2.87	10	3.80	4	2.34	10	3.97	3	1.55	10	2.56

Fig. 4. Discourse functions of coordination across language groups and text types.

	Intermediate written		Intermediate spoken		Advanced written		Advanced spoken		Native written		Native spoken	
	n	%	n	%	n	%	n	%	n	%	n	%
new-introduction	15	13.89	10	9.62	10	13.89	5	5.43	7	8.97	10	5.95
reference maintenance	71	65.74	71	68.27	49	68.06	78	84.78	59	75.64	138	82.14
switch reference	20	18.52	21	20.19	11	15.28	6	6.52	11	14.10	17	10.12
re-introduction	2	1.85	2	1.92	2	2.78	3	3.26	1	1.28	3	1.79

Fig. 5. REs in reference maintenance coordination.

	Intermediate written		Intermediate spoken		Advanced written		Advanced spoken		Native written		Native spoken	
	n	%	n	%	n	%	n	%	n	%	n	%
zero	43	60.56	26	36.62	39	79.59	41	52.56	50	84.75	114	82.61
pronoun	21	29.58	44	61.97	8	16.33	35	44.87	8	13.56	24	17.39
NP	7	9.86	1	1.41	2	4.08	2	2.56	1	1.69	0	0.00

Fig. 6. REs in reference maintenance coordination with intervening subordinate clauses.

	Intermediate written		Intermediate spoken		Advanced written		Advanced spoken		Native written		Native spoken	
	n	%	n	%	n	%	n	%	n	%	n	%
zero	5	23.81	4	22.22	10	52.63	6	26.09	13	61.90	26	66.67
pronoun	10	47.62	13	72.22	7	36.84	15	65.22	7	33.33	13	33.33
NP	6	28.57	1	5.56	2	10.53	2	8.70	1	4.76	0	0.00

Fig. 7. Intervening subordination in contexts of reference maintenance in coordination.

	Intermediate written		Intermediate spoken		Advanced written		Advanced spoken		Native written		Native spoken	
	n	%	n	%	n	%	n	%	n	%	n	%
without intervening clause	50	70.42	53	74.65	30	61.22	55	70.51	38	64.41	99	71.74
with intervening clause	21	29.58	18	25.35	19	38.78	23	29.49	21	35.59	39	28.26

Fig. 8. Amount of intervening subordination in contexts of reference maintenance in coordination.

	Intermediate written		Intermediate spoken		Advanced written		Advanced spoken		Native written		Native spoken	
	n	%	n	%	n	%	n	%	n	%	n	%
1 subordinate clause	15	71.43	16	88.89	13	68.42	17	73.91	14	66.67	33	84.62
1+ subordinate clauses	6	28.57	2	11.11	6	31.58	6	26.09	7	33.33	6	15.38

References

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology Section A, 35*(3), 427–440.

Ariel, M. (1988). Reference and accessibility. *Journal of Linguistics, 24*(1), 65–87.

Ariel, M. (1990). *Accessing noun phrase antecedents*. Routledge.

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language, 56*(4), 521–536.

Ballester, E. (2013). Adult instructed SLA of English subject properties. *Canadian Journal of Linguistics, 58*(3), 465–486. https://doi.org/10.1017/S0008413100002668

Bel, A., Perera, J., & Salas, N. (2010). Anaphoric devices in written and spoken narrative discourse. Data from Catalan. *Written Language & Literacy, 13*(2), 236–259. https://doi.org/10.1075/wll.13.2.03bel

Carminati, M. N. (2002). *The processing of Italian subject pronouns*. Doctoral Dissertations. Available from Proquest.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (Ed.), *Subject and topic* (pp. 25–56). Academic press.

Chini, M. (2005). Reference to person in learner discourse. In H. Hendriks (Ed.), *The structure of learner varieties* (pp. 65–110). Mouton de Gruyter.

Christensen, M. (2000). Anaphoric reference in spoken and written Chinese narrative discourse. *Journal of Chinese Linguistics, 28*(2), 303–336.

Clark, H. H., & Sengul, C. J. (1979). In search of referents for noun and pronouns. *Memory & Cognition, 7*(1), 35–41.

Contemori, C., & Dussias, P. E. (2016). Referential choice in a second language: Evidence for a listener-oriented approach. *Language, Cognition and Neuroscience, 31* (10), 1257–1272. https://doi.org/10.1080/23273798.2016.1220604

Crosthwaite, P. (2011). The effect of collaboration on the cohesion and coherence of L2 narrative discourse between English NS and Korean L2 English users. *Asian EFL Journal, 13*. http://asian-efl-journal.com/264/quarterly-journal/2011/12/the-effect-of-collaboration-on-the-cohesion-and-coherence-of-l2-narrative-discourse-between-english-ns-and-korean-l2-english-users/.

Cunnings, I., Fotiadou, G., & Tsimpli, I. M. (2017). Anaphora resolution and reanalysis during L2 sentence processing: Evidence from the visual world paradigm. *Studies in Second Language Acquisition, 39*(4), 621–652. https://doi.org/10.1017/S0272263116000292

DuBois, J. W. (1987). The discourse basis of ergativity. *Language, 63*(4), 805–855.

Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, H. Reinders, R. Erlam, & J. Philp (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3–26). Multilingual Matters. https://doi.org/10.21832/9781847691767-003.

Fedorova, O. V. (2014). The role of potential referential conflict in the choice of a referring expression. *The Russian Journal of Cognitive Science, 1*(1–2), 6–21.

Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study* (pp. 1–41). John Benjamins.

Givón, T. (2001). *Syntax. An introduction* (vol. 1). John Benjamins.

Givón, T. (2017). *The story of zero*. John Benjamins.

Grabowski, J. (2007). The writing superiority effect in the verbal recall of knowledge: Sources and determinants. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and cognition: Research and applications* (pp. 165–179). Elsevier.

Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning, 56*(1), 155–196. https://doi:10.1111/j.0023-8333.2006.00344.x.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*, 274–307.

Hendriks, H. (2003). Using nouns for reference maintenance: A seeming contradiction in L2 discourse. In A. Giacalone (Ed.), *Typology and second language acquisition* (pp. 291–326). Mouton De Gruyter.

Kang, J. Y. (2004). Telling a coherent story in a foreign language: Analysis of Korean EFL learners' referential strategies in oral narrative discourse. *Journal of Pragmatics, 36*(11), 1975–1990. https://doi.org/10.1016/j.pragma.2004.03.007

Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental perspective. *Language & Cognitive Processes, 1*(1), 61–85.

Kibrik, A. A. (2011). *Reference in discourse*. Oxford University Press.

Kuiken, F., & Vedder, I. (2011). Task performance in L2 writing and speaking: The effect of mode. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 91–104). John Benjamins.

Kuiken, F., & Vedder, I. (2012). Speaking and writing tasks and their effects on second language performance. In S. M. Gass, & A. Mackey (Eds.), *The routledge handbook of second language acquisition* (pp. 364–377). Routledge.

Leclercq, P., & Lenart, E. (2013). Discourse cohesion and accessibility of referents in oral narratives: A comparison of L1 and L2 acquisition of French and English. *Discours. Revue de Linguistique, Psycholinguistique et Informatique, 12*. https://doi.org/10.4000/discours.8801

Lozano, C. (2009). Selective deficits at the syntax-discourse interface: Evidence from the CEDEL2 corpus. In N. Snape, Y. I. Leung, & M. Sharwood Smith (Eds.), *Language acquisition and language disorders* (vol. 47, pp. 127–166). John Benjamins. https://doi.org/10.1075/lald.47.09loz.

Lozano, C. (2016). Pragmatic principles in anaphora resolution at the syntax-discourse interface: Advanced English learners of Spanish in the CEDEL2 corpus. In M. Alonso Ramos (Ed.), *Spanish learner corpus research: State of the art and perspectives* (pp. 236–265). John Benjamins.

Lozano, C. (2018). The development of anaphora resolution at the syntax-discourse interface: Pronominal subjects in Greek learners of Spanish. *Journal of Psycholinguistic Research, 47*, 411–430. https://doi.org/10.1007/s10936-017-9541-8

Lozano, C. (2022). CEDEL2: Design, compilation and web interface of an online corpus for L2 Spanish acquisition research. *Second Language Research, 38*(4), 965–983. https://doi.org/10.1177/0267658321105052

Lozano, C., Díaz-Negrillo, A., & Callies, M. (2020). Designing and compiling a learner corpus of written and spoken narratives: COREFL. In C. Bongartz, & J. Torregrossa (Eds.), *What's in a Narrative? Variation in Story-Telling at the Interface between Language and Literacy* (pp. 21–46). Peter Lang. https://doi.org/10.3726/978-3-653-05182-7.

Martínez-Flor, A. (2006). Task-effects on EFL learners' productions of suggestions: A focus on elicited phone messages and emails. *Miscelanea: A Journal of English and American Studies, 33*, 47–64.

Mitkovska, L., & Bužarovska, E. (2018). Subject pronoun (non)realization in the English learner language of Macedonian speakers. *Second Language Research, 34*(4), 463–485. https://doi.org/10.1177/0267658317747925

Morrow, D. (1985). Prominent characters and events organize narrative understanding. *Journal of Memory and Language, 24*(3), 304–319.

Myles, F. (2015). Second language acquisition theory and learner corpus research. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 309–332). Cambridge University Press.

Ngo, B., Kaiser, E., & Simpson, A. (2019). Effects of grammatical roles and parallelism on referential form production in Vietnamese spoken and written narratives. In N. Duffield, T. Phan, & T. Trinh (Eds.), *Interdisciplinary perspectives on Vietnamese linguistics* (pp. 211–275). John Benjamins.

O'Donnell, M. (2009). The UAM Corpus Tool: Software for corpus annotation and exploration. In C. M. Bretones, et al. (Eds.), *Applied linguistics now: Understanding language and mind* (pp. 1433–1447). Universidad de Almería.

Oh, S.-Y. (2006). English zero anaphora as an interactional resource II. *Discourse Studies, 8*(6), 817–846.

Oxford Quick Placement Test (2001). Oxford University Press.

Perales, S., & Portillo, R. (2007). Sobre las propiedades referenciales de los sujetos nulos y pronominales del español oral y escrito. In E. Balmaseda Maestu (Ed.), *Las destrezas orales en la enseñanza del español L2-LE: XVII congreso internacional de la asociación del español como lengua extranjera (ASELE)* (pp. 889–900). Universidad de la Rioja.

Prentza, A. (2014). Pronominal subjects in English L2 acquisition and in L1 Greek: Issues of interpretation, use and L1 transfer. In L. Nikolaos, A. Thomaï, & A. Sougari (Eds.), *Major trends in theoretical and applied linguistics 2* (pp. 369–386). De Gruyter Open Poland.

Quesada, M. L. (2015). *The L2 acquisition of Spanish subjects: Multiple perspectives*. De Gruyter Mouton https://doi.org/10.1515/9781614514367.

Quesada, T., & Lozano, C. (2020). Which factors determine the choice of referential expressions in L2 English discourse? New evidence from the COREFL corpus. *Studies in Second Language Acquisition, 42*(5), 959–986.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A grammar of contemporary English*. Longman.

Ryan, J. (2015). Overexplicit referent tracking in L2 English: Strategy, avoidance, or myth? *Language Learning, 65*(4), 824–859. https://doi.org/10.1111/lang.12139

Santoro, M. (2020). The acquisition of English anaphoric expressions by adult Chinese speakers. *Journal of Psycholinguistic Research, 49*, 641–662. https://doi.org/10.1007/s10936-020-09717-4.

Schoonen, R., Snellings, P., Stevenson, M., & van Gelderen, A. (2009). Toward a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. In R. Manchón (Ed.), *Writing in foreign language contexts* (pp. 77–101). Multilingual Matters.

Van Dijk, T. A. (1981). Episodes as units of discourse analysis. In D. Tannen (Ed.), *Analyzing discourse: Text and talk* (pp. 177–195). Georgetown: Georgetown University Press.

Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning, 67*(2), 394–430. https://doi.org/10.1111/lang.12228

von Stutterheim, C. H., & Klein, W. (1989). Referential movement in descriptive and narrative discourse. In R. Dietrich, & C. Graumann (Eds.), *Language processing in social context* (pp. 39–76). North-Holland.

Vonk, W., Hustinx, L. G. M. M., & Simons, W. H. G. (1992). The use of referential expressions in structuring discourse. *Language & Cognitive Processes, 7*(3–4), 301–333.

Weissberg, B. (2000). Developmental relationships in the acquisition of English syntax: Writing vs. speech. *Learning and Instruction, 10*(1), 37–53. https://doi.org/10.1016/S0959-4752(99)00017-1

Williams, J. (1988). Zero anaphora in Second Language Acquisition. A comparison between three varieties of English. *Studies in Second Language Acquisition, 10,* 339–370.

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing, 21*(4), 321–331. https://doi.org/10.1016/j.jslw.2012.09.007