SSLA, **24**, 387–418. Printed in the United States of America.
DOI: 10.1017.S0272263102003029

TOPIC CONTINUITY IN L2 ENGLISH ARTICLE USE

Scott Jarvis
Ohio University

Much of the research on L2 article acquisition has investigated the effects of semantic, syntactic, and discourse universals on the systematicity and variability of learners' article use. The present paper looks at systematicity from the combined perspective of two putative discourse universals related to topic continuity (e.g., Givón, 1983) that have been addressed only separately in past studies of article acquisition: the tendency to mark the distinction between topics and comments (e.g., Huebner, 1983) and the tendency to mark the distinction between new, continuous, and reintroduced NP referents (e.g., Chaudron & Parker, 1990). The present study examines how well these discourse universals account for the patterns of article use and nonuse found in narratives written by 199 Finnish-speaking and 145 Swedish-speaking adolescent learners of English. The quantitative results of the study cast some doubt on learners' sensitivity to the topic-comment distinction and also suggest that learners' tendency to mark distinctions between new, continuous, and reintroduced NP referents is influenced by the prominence of such distinctions in the L1. The quantitative results are supported by a qualitative analysis of a subset of the data that suggests numerous other elements that are needed to characterize the systematicity of individual learners' interlanguage article systems.

Past research on L2 article acquisition, like so many other areas of L2 acquisition, has shown that learners seem to have an innate—and therefore universal—sensitivity to certain semantic, syntactic, and discourse distinctions and that they tend to mark such distinctions even from the early stages of language acquisition (see Young, 1995, p. 142). Semantic universals have received considerable attention (e.g., Huebner, 1983; Parrish, 1987; Thomas, 1989), but

Address correspondence to: Scott Jarvis, Dept. of Linguistics, Ohio University, Gordy Hall 363, Athens, OH 45701; e-mail: jarvis@ohio.edu.

© 2002 Cambridge University Press 0272-2631/02 $9.50

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

even more attention has been directed at discourse universals, particularly at the two that are the focus of the present investigation: (a) the tendency to mark the distinction between topic and comment NP referents (e.g., Huebner; Young); and (b) the tendency to mark the distinction between new, continuous, and reintroduced NP referents (e.g., Chaudron & Parker, 1990; Givón, 1984). These two putative discourse universals have been investigated only separately (i.e., in separate studies or as separate hypotheses) in past studies of L2 article acquisition, yet both are related directly to each other under the notion of topic continuity (e.g., Givón, 1983); together they form the following six-cell taxonomy (compare with the simpler taxonomy in Chaudron & Parker, p. 45):

1. New referent as (a constituent of the) topic (e.g., *A woman stole a loaf of bread.*);
2. New referent as (a constituent of the) comment (e.g., *A woman stole a loaf of bread.*);
3. Continuous referent as (a constituent of the) topic (e.g., *A woman stole a loaf of bread. Then she accidentally dropped it.*);
4. Continuous referent as (a constituent of the) comment (e.g., *A woman stole a loaf of bread. Then she accidentally dropped it.*);
5. Reintroduced referent as (a constituent of the) topic (e.g., . . . *A man picked up the bread and said that he took it, but an eyewitness came and explained what really happened. Eventually, the woman was arrested and taken to jail.*);
6. Reintroduced referent as (a constituent of the) comment (e.g., . . . *A man picked up the bread and said that he took it, but an eyewitness came and explained what really happened. Eventually, the woman was arrested and taken to jail. The man gave the bread back to the baker, but later he was arrested, too.*)

I will describe the differences between these six categories a little later on. For now, it is relevant to state that the purpose of the present study is to investigate whether learners differentiate between these six discourse conditions and, if so, how and to what degree articles are involved in the discourse distinctions learners make. The data examined in this paper, which relate to the examples given above, are narrative descriptions of a silent film that were written by Finnish-speaking and Swedish-speaking adolescent learners of English at three distinct levels of English instruction. Taking advantage of the empirical design of the study, I consider learners' sensitivity to topic continuity in conjunction with the potential moderating effects of the L1 and learners' general stage of acquisition—factors that were found in past studies to affect learners' patterns of article use (e.g., Huebner, 1985; Jourdenais, 1999; Master, 1987, 1997; Parrish, 1987; Young, 1995). The bulk of the empirical portion of this study is quantitative, but I also examine a subset of the data qualitatively to provide a fuller description of the nature of individual learners' interlanguage article systems.

DISCOURSE UNIVERSALS, TOPICALITY, AND TOPIC CONTINUITY

Bickerton (1981) appears to have been the first researcher to advocate a universalist position regarding the functions of the English articles (cf. Young,

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

1995, p. 138). His proposal includes two universals, which he called "primes," that are related to article use. The first, a semantic universal, is the distinction between specific and nonspecific reference, and several of the studies that have investigated whether learners structurally distinguish between specific and nonspecific NP referents in their use of articles have indeed found supporting evidence (e.g., Thomas, 1989; Young). In the present study, I focus almost exclusively on specific NP referents, given that this is the only type of indefinite (new or previously unmentioned) NP referent that can serve as topic (Reinhart, 1982, p. 12), so I am not concerned here with the effects of Bickerton's semantic universal. However, I am centrally concerned with his second universal, which involves discourse distinctions and is a precursor to the distinction between new, continuous, and reintroduced NP referents. Bickerton referred to his second universal as the distinction between presupposedness and nonpresupposedness (e.g., p. 251), and it has since been relabeled as [±Assumed Known to the Hearer], or simply [±HK] (e.g., Huebner, 1983; Master, 1987; Thomas). One could also refer to the distinction as new information versus old information, so it is not surprising that some researchers (e.g., Young, p. 172) have equated the [±HK] distinction with the topic-comment distinction. I will return to this point shortly.

While investigating the universal (or innate) nature of Bickerton's [±HK] distinction, Huebner (1983, 1985), Master (1987, 1997), Young (1995), and others have found that learners across proficiency levels do indeed show at least some sensitivity to this distinction, generally marking [+HK] NPs with the definite article and [–HK] NPs with Ø or the indefinite article. There are some caveats, however, such as: (a) the finding that, at the initial or slightly postinitial stage of article acquisition, learners tend to overgeneralize the use of the definite article to both [+HK] and [–HK] (e.g., Huebner, 1983; Master, 1987); and (b) what learners do during, before, and after that stage of overgeneralization also depends largely on whether their L1 has articles. For learners whose L1 lacks articles, Ø and the definite article often appear as competing options for the same NP contexts. This is where the distinction between topics and comments becomes relevant. In a VARBRUL analysis of article use by six speakers of Czech and Slovak (languages that lack articles) at progressive levels of English proficiency, Young found that the use of Ø was most strongly associated with comments (or rhemes) occurring in clause-final position. Young interpreted this result as suggesting that "since final rhemes are the preferred word order in Czech and Slovak, articles would be redundant in this position" (p. 168). Young suggested, in other words, that his learners were indeed sensitive to the distinction between topics and comments but that they tended to mark topic and comment NPs with overt articles (*the* or *a*) only when the status of the NP as topic or comment was not already indicated (as it would be in the L1) by its position within the clause. Later in the same study, however, he downplayed the effects of clause position by saying that topics (or themes) tended to be marked with *this* or *the* and comments tended to be marked with Ø, regardless of where they occurred in the clause. This is

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

critical because Young defined topics as [+HK] and comments as [–HK], and given that Young conflated the two types of distinctions, it is difficult to assess whether his learners really were sensitive to the topic-comment distinction, as opposed to simply the distinction between [+HK] and [–HK].

Huebner's (1983) longitudinal study of an adult Hmong speaker's acquisition of English provides somewhat more straightforward evidence of a learner's inherent sensitivity to the distinction between topic and comment NPs, but only in the earliest stage of acquisition. Unlike Young's (1995) study, Huebner's evidence does not confound the topic-comment and [±HK] dichotomies but instead applies only to [+HK] NP referents. Recognizing that [+HK] NPs can serve either topic or comment functions (see also Givón, 1983; Reinhart, 1982), Huebner found in his data recorded at time 1 that 74% of the specific [+HK] NPs that were marked with Ø were topics (or constituents of topics), whereas only 12% of the specific [+HK] NPs marked with *da* (his learner's interlanguage counterpart to *the*) were topics (pp. 140–142). These patterns strongly suggest that Huebner's learner was sensitive to the topic-comment distinction, but alternative explanations are also possible. For example, Parrish (1987) found similar patterns of competition between Ø and *the* in the early stages of English acquisition by an adult Japanese speaker, but Parrish interpreted her own and Huebner's results as reflecting learners' sensitivity to the distinction between subject and predicate positions (p. 368) instead of between topic and comment.

Settling this question once and for all would require an appeal to well-defined criteria for identifying topics and comments. Unfortunately, individual scholars differ in how they have defined the distinction. For example, Young (1995) maintained the traditional distinction between topics as old information ([+HK]) and comments as new information ([–HK]) (e.g., Chafe, 1976; Gundel, 1974), but Huebner (1983) assumed that the topic of a sentence is the [+HK] information that the sentence is about, whereas the comment is any information (either [+HK] or [–HK]) that is said about the topic (p. 58). Still other researchers have assumed that even the topic can be either [+HK] or [–HK] and that what is really essential to identifying the topic is understanding the "pragmatic aboutness" of the sentence (e.g., Reinhart, 1982). According to this view, the topic of a sentence can be new information ([–HK]) if that new information is what the sentence is saying something about (cf. the first example in the list presented earlier). This is the position adopted in the study by Chaudron and Parker (1990), who categorized topic NPs not just as new and known but as new, current, and known. Their three-way distinction was influenced by the work of Givón (1983), who argued that speakers inherently make certain assumptions "about topic-availability to the hearer" (p. 12) and recognize that a topic continuing from an immediately preceding clause will be more predictable and easier for the hearer to process. Topics that represent new information and topics that are reintroduced after other topics have entered the discourse are less predictable and more difficult to process. Because of these differences, Givón (1984) proposed that "more continuous,

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

predictable, non-disruptive topics will be marked by *less marking material*; while less continuous, unpredictable/surprising, or disruptive topics will be marked by *more marking material*" (p. 126, emphasis in original).

Chaudron and Parker (1990) tested this prediction in their investigation of 40 adult Japanese-speaking learners' use of English articles and other morphological and syntactic devices (e.g., pronouns, Ø anaphora, existential constructions, and left dislocation) as topic markers. Relabeling Givón's (1983) distinction between new, continuous, and reintroduced topics as the distinction between new, current, and known topics, Chaudron and Parker predicted that learners would structurally distinguish all three types of topics through the selection of different structural options in the different contexts. They interpreted the results of their oral picture-description task as generally supporting this prediction. The learners produced very few instances of Ø anaphora, existential constructions, and left dislocations, but their use of pronouns and English article options (*a*, *the*, and Ø) exhibited different overall patterns with new, current, and known topics. Although the patterns differed somewhat according to the learners' proficiency levels, the following general patterns could be seen: Topics involving new NP referents were marked mainly with *a* but also with Ø and *the*; current topics were expressed mainly through pronouns; and known topics were marked mainly with *the* but also largely with pronouns. Differences in the frequency of pronoun use accounted for the biggest differences across contexts, but new topics also stood out with respect to the use of *a*, and known topics stood out in terms of the use of *the*. These results generally support Givón's prediction (also known as the Quantity Universal) in the sense that they show that current topics are often marked with less marking material (e.g., pronouns) than are new and known topics (e.g., where overt article + NP is more frequent). However, a lot remains to be accounted for, such as the fact that the use of Ø was relatively high in all three contexts, especially with new topics, where more marking material would be predicted. Chaudron and Parker did not discuss this aspect of their results, presumably because they were less concerned with the details of Givón's Quantity Universal than with the more fundamental question of whether learners distinguish between the three contexts in the first place.

This more fundamental question is also the focus of the present investigation. However, I am concerned not only with learners' sensitivity to the distinction between new, current, and known topics but also with their sensitivity to the distinction between topic and comment. I have therefore expanded the three-way distinction made by Chaudron and Parker to a six-way distinction between topic and comment NPs in new, current, and known contexts. Examples of all six categories were given earlier, and I define them more explicitly in the following section. In the subsequent "Quantitative Results" section, prior to showing how these categories apply to learners' use of the English articles, I show the types of distinctions that are made by native speakers of English and native speakers of the learners' L1s. Against this backdrop I outline some predictions concerning Finnish-speaking and Swedish-

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

speaking learners' use of English articles and other structural devices in these six discourse contexts.

METHOD

Participants

The participants recruited for the study include 199 Finnish-speaking and 145 Swedish-speaking adolescent learners of English living in Finland who were at different levels of age and English instruction at the time of testing. Comparable L1-control groups comprising 108 Finnish speakers and 77 Swedish speakers, as well as a comparable native-control group consisting of 98 American adolescents, were also included. Table 1 shows the breakdown of participants according to L1, grade level, and years of English, Finnish, and Swedish instruction.

Note that all of the Finnish and Swedish groups had had at least 2 years of English instruction prior to the time of testing. Group F5 consists of 62 Finnish-speaking fifth graders who had had 2 years of English instruction and no Swedish instruction. Group F7 includes 46 Finnish-speaking seventh graders who had had 4 years of English instruction and were in their first year of Swedish instruction at the time of testing. The Finnish ninth graders, groups F9A and F9B, are interesting groups for comparison; group F9A had had 6 years of English instruction and 2 years of Swedish, whereas group F9B had had just the opposite: 2 years of English and 6 years of Swedish. The two Swedish-speaking groups, S7 and S9, had both had 2 years more of Finnish instruction than of English instruction, although their teachers commented that most of them displayed a better command of English than of Finnish, owing primarily to the language proximity between Swedish and English, on the one hand, and the language distance between Swedish and Finnish, on the other (see, e.g., Ringbom, 1987). Although both the Finnish-speaking and Swedish-speaking participants were natives of Finland, for purposes of convenience, I follow Ringbom in referring to them simply as Finns and Swedes.

Data Elicitation and Analysis

As part of a larger investigation (see Jarvis, 1998, 2000), the participants were shown the 8-minute "Alone and Hungry" segment of the silent Chaplin film *Modern Times*, in which a poor young woman attempts to steal a loaf of bread, is arrested, escapes, and is eventually befriended by Chaplin. The film is silent, but it includes nine intermittent titles (see the Appendix), whose relevance to the data will be mentioned toward the end of the "Qualitative Results" section. Because of time constraints, and in order not to give the participants too large a memory task, I divided the film segment into two parts. I showed the first 5 minutes of the film and then asked the participants to write a narrative description of it. Then, after a break lasting anywhere from 10 min-

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Table 1. Participant groups

	Groups													
	Experimental						Native			L1 control				
Variables	F5	F7	F9A	F9B	S7	S9	E5	E7	E9	F5c	F7c	F9c	S7c	S9c
n	62	46	46	45	68	77	24	29	45	29	38	41	40	37
Native language	Fin	Fin	Fin	Fin	Swe	Swe	Eng	Eng	Eng	Fin	Fin	Fin	Swe	Swe
Grade level	5	7	9	9	7	9	5	7	9	5	7	9	7	9
Years of English	2	4	6	2	2	4	—	—	—	2	4	6	2	4
Years of Finnish	—	—	—	—	4	6	—	—	—	—	—	—	4	6
Years of Swedish	—	—	2	6	—	—	—	—	—	—	—	2	—	—
Test language	Eng	Eng	Eng	Eng	Eng	Eng	Eng	Eng	Eng	Fin	Fin	Fin	Swe	Swe

Note. A dash represents inapplicable data.

Table 2. Taxonomy of NP referents categorized by continuity and topicality

	Continuity		
Topicality	New	Current	Known
Topic	NT	CT	KT
Comment	NC	CC	KC

utes to 2 days (depending on scheduling circumstances), I showed the participants the remaining 3 minutes of the film and asked them to write a narrative description of just the latter part, adding it to what they had already written about the first part of the film (i.e., the participants were normally given back the narrative that they had begun earlier). Later in this article I will show that the way the elicitation procedures were divided up had an effect on participants' use of articles, but only in one place in their narrative descriptions and only for some participants.

The written narratives constitute the data for the present investigation, and they were subjected to the following quantitative analysis. First, I identified each reference to the female protagonist (i.e., Paulette Goddard, the woman who stole the bread) in all of the participants' written narratives. I chose the female protagonist as the focus of analysis because she and Charlie Chaplin were the only referents in the film that were referred to multiple times throughout participants' narratives, and this condition was requisite for an examination of possible distinctions between new, current, and known NP referents. Chaplin was not chosen as the focus of analysis because most participants referred to him by name (instead of, e.g., as *a man*), and this would have limited my ability to investigate learners' knowledge and use of articles.

After identifying all references to the female protagonist in the data, I marked each instance according to the contextual category it represented: NT for new topic, NC for new comment, CT for current topic, CC for current comment, KT for known topic, and KC for known comment, as shown in Table 2. The relatively few instances that did not fit into any of these categories involved vocative, attributive, and first- and second-person pronoun references to the female protagonist; these were omitted from further analysis. With some modifications, I followed Reinhart's (1982) approach to distinguishing topic and comment NPs, and I followed Givón's (1983) approach to distinguishing between new, current, and known NP referents. Following Reinhart, I assumed that a sentence can have only one topic, which is not necessarily the grammatical subject. According to Reinhart, the subject position is the default topic position (p. 9), but a topic NP can occur in other syntactic positions as well, as long as it represents the most accessible information in the sentence and is really what the sentence is about. Both of the sentences in (1), for instance, can be said to be about the woman who stole the bread. In the first

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

sentence, the topic NP occupies the subject position, but in the second sentence it occupies the direct object position. Cases like this arise quite frequently when the subject is a new NP referent and the direct object is a current or known NP referent.

(1) *The woman stole a loaf of bread. Unfortunately, another woman saw her do it.*

One modification I made to Reinhart's (1982) description of topics is that, instead of assuming that the sentence or clause is the basic unit of discourse, I assumed that the T-unit (i.e., an independent clause with all of its accompanying dependent clauses; see, e.g., Hunt, 1965) is closer to a minimal discourse unit having its own topic. I further assumed that the topic of a T-unit is determined by the pragmatic aboutness of its independent clause and that a topic NP can occur in a dependent clause only when it is co-referential with the topic of the main clause (except in cleft constructions). The following definitions apply to the six categories of analysis in Table 2:

1. NT is a previously unmentioned NP referent that serves as the subject (whether in initial position or in an existential construction) of the main clause of a T-unit where no other NP referent has been mentioned previously;
2. NC is a previously unmentioned NP referent that does not meet the criteria for NT;
3. CT is an NP referent that was mentioned in the preceding T-unit or earlier in the current T-unit and is the subject of the main clause of the current T-unit or is the only NP referent in the main clause that was mentioned previously; an NP referent in a subordinate clause—though not in reported speech—is also defined as CT if it is coreferential with the topic (whether NT, CT, or KT) of the main clause;
4. CC is an NP referent that was mentioned in the preceding T-unit or earlier in the current T-unit but does not meet the criteria for CT;
5. KT is an NP referent that was mentioned earlier in the text, but not in the preceding T-unit or earlier in the current T-unit, and is the subject of the main clause or the only NP referent in the main clause mentioned previously;
6. KC is an NP referent that was mentioned earlier in the text, but not in the preceding T-unit or earlier in the current T-unit, and does not meet the criteria for KT.

After categorizing each reference to the female protagonist according to this scheme of discourse contexts, I tabulated the participants' use of the following morphological and lexical options in each context: articles (Ø, *a*(*n*), and *the*), pronouns, definite determiners (*this*, *that*, and possessive determiners), indefinite quantifiers (*one* and *some*), proper nouns, and marked syntactic structures (existential constructions, cleft constructions, left dislocation, right dislocation, fronting, inverted word order, postmodification, and overt markers of new information, e.g., *in the beginning*). The purpose of the quantitative analysis was to provide evidence necessary for determining whether the participants marked NP referents differentially across the six discourse contexts.

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

QUANTITATIVE RESULTS

L1 English

Results from the native English speakers' narratives are shown in Table 3. This table shows the distribution of articles and other types of morphological, lexical, and syntactic marking used in NPs denoting the female protagonist. The numbers represent percentages, with token counts given in parentheses. The morphological and lexical options listed to the left of the Total column are mutually exclusive in the present data, so the sum of the percentages across these columns is always 100 (with a few minor exceptions due to rounding of the percentages to the nearest integer). The right-most column, Syntax, indicates participants' use of marked syntactic structures in the six discourse contexts. The specific types of marked syntactic structures that occur in the data are as follows: Existential constructions (e.g., *There was a girl*) are involved in all of the marked syntax in NT contexts; the phrase *that one girl* accounts for one of the marked syntactic structures in KT contexts; and all other instances of marked syntax in the L1 English data involve relative-clause postmodifiers (e.g., *a girl who is alone and hungry*). Marked syntactic structures overlap with (i.e., occur only in addition to) the categories represented by the other columns, so the numbers in the Syntax column were not added to the totals, and likewise no percentages were calculated for the Syntax column. From the perspective of the discourse distinctions that they exhibit, the participants' use of marked syntactic structures turned out to be secondary to their use of articles and pronouns, so syntax will not be dealt with further in this section.

Table 3 shows essentially what one would expect native English speakers to do: In new contexts (NT and NC) they generally favor *a*(*n*), in current contexts (CT and CC) they favor pronouns, and in known contexts (KT and KC) they favor *the*. The native speakers in Chaudron and Parker's (1990) study exhibited similar patterns, and it was on the basis of such patterns that the researchers concluded that their learner and native-speaker participants maintained "a distinction between all three contexts (i.e., new, current, and known) in differing degrees of use of pronouns, bare nouns, definite articles, and indefinite articles" (p. 57). I suspect, however, that for native English speakers, these patterns arise not from a systematic three-way distinction but rather from a combination of the following discourse conventions: (a) an obligatory, almost categorical distinction between new versus not new (i.e., between indefinite and definite); and (b) a nonobligatory, noncategorical tendency for speakers to use pronouns in not-new contexts (e.g., to avoid repetition and to make communication more efficient) whenever the pronoun's referent is easily retrievable from the context, which is especially true of current contexts. This means that one should expect to find far more categorical distinctions in individuals' texts between new and not new than between current and known, and this is precisely what I found.

Over 40% of the participants in each of the L1 English groups (42% for E5c,

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Table 3. L1 English results

Group/ context		Ø		*a*(*n*)		*the*		Pronoun		Def. det.		Quant.		Prop. N		Total		Syn.
E5c	NT	—	—	41	(9)	41	(9)	9	(2)	9	(2)	—	—	—	—	100	(22)	(6)
	NC	—	—	100	(2)	—	—	—	—	—	—	—	—	—	—	100	(2)	(1)
	CT	—	—	1	(1)	28	(22)	69	(54)	1	(1)	—	—	—	—	100	(78)	—
	CC	—	—	—	—	25	(13)	74	(39)	2	(1)	—	—	—	—	100	(53)	—
	KT	—	—	—	—	75	(24)	16	(5)	9	(3)	—	—	—	—	100	(32)	(4)
	KC	—	—	—	—	79	(19)	17	(4)	4	(1)	—	—	—	—	100	(24)	(1)
E7c	NT	—	—	50	(14)	43	(12)	4	(1)	4	(1)	—	—	—	—	100	(28)	(3)
	NC	—	—	100	(1)	—	—	—	—	—	—	—	—	—	—	100	(1)	(1)
	CT	—	—	0	(1)	22	(44)	76	(155)	1	(3)	—	—	—	—	100	(203)	(2)
	CC	—	—	—	—	31	(45)	67	(96)	1	(2)	—	—	—	—	100	(143)	—
	KT	—	—	—	—	75	(43)	23	(13)	2	(1)	—	—	—	—	100	(57)	(6)
	KC	—	—	—	—	69	(45)	26	(17)	5	(3)	—	—	—	—	100	(65)	(1)
E9c	NT	—	—	71	(29)	12	(5)	2	(1)	15	(6)	—	—	—	—	100	(41)	(7)
	NC	—	—	100	(3)	—	—	—	—	—	—	—	—	—	—	100	(3)	(2)
	CT	3	(6)	—	—	23	(48)	75	(159)	—	—	—	—	—	—	100	(213)	(2)
	CC	4	(5)	1	(1)	35	(49)	59	(82)	2	(3)	—	—	—	—	100	(140)	—
	KT	5	(3)	—	—	64	(35)	25	(14)	5	(3)	—	—	—	—	100	(55)	(3)
	KC	3	(2)	—	—	69	(47)	25	(17)	3	(2)	—	—	—	—	100	(68)	(2)

48% for E7c, and 71% for E9c) exhibit a categorical distinction between new and not-new contexts in their use and nonuse of the indefinite article, whereas consistently fewer than 20% of the participants in the same groups (4% for E5c, 0% for E7c, and 16% for E9c) exhibit a categorical distinction between current and known contexts in their use of pronouns versus the definite article. This is hardly surprising given what is known about English grammar and discourse (see, e.g., Quirk & Greenbaum, 1973). If anything is surprising, it is probably the fact that there are not even more participants in groups E5c and E7c who make a categorical distinction between new and not-new contexts; the use of the definite article in NT contexts is high for both of these groups. I speculate that this is because they assumed that their intended audience (i.e., the researcher) had already seen the film and knew to whom they were referring. It is interesting, though, that the ninth graders did not make the same assumption or, alternatively, were more concerned with writing in conventional narrative format than with guessing whether their potential audience was likely to have seen the film. In any event, the high use of the definite article in NT contexts seems to be age related in the L1 English data, although in the next section I show that it may also be L1 related.

The English speakers' use of articles and pronouns across new, current, and known contexts seems relatively straightforward, but differences between topic and comment contexts are much less clear. On the one hand, the percentages in Table 3 show that participants exhibit the same overall preferences in NT and NC contexts, in CT and CC contexts, and in KT and KC contexts. On the other hand, inferential statistics suggest that there might be some important differences. A chi-square test could not be used to compare NT and NC, owing to the low token counts in NC contexts, but the same test did reveal a significant difference in the distribution of the definite article and pronouns across CT and CC contexts ($p < .05$ for group E7c; $p < .01$ for group E9c), though no significant differences between KT and KC contexts. The significant difference between CT and CC is somewhat difficult to explain in light of the fact that English is not a topic-prominent language, but it is important to recognize that the difference involves only a statistical trend, not a categorical distinction. Only one participant in group E5c and only two participants in group E9c exhibited a (probably coincidental) categorical distinction between CT and CC (i.e., by using only pronouns in one context and only the definite article in the other). The significantly stronger preference for pronouns in CT over CC contexts may result from the fact that topic referents tend to be more expected—that is, more retrievable from context—than comment referents, and this is one of the conditions that determine pronoun use.

One final point deserves attention with respect to the results in Table 3: Some participants in group E9c used Ø (or no article) in their NP reference to the female protagonist. This occurred in sentences such as *Policeman run after girl, grabs girl in the crowd and take her away*. (Chesterman, 1991, and Master, 1997, distinguish between the zero article [indefinite] and the null article [definite]. The example just given involves the null article, though this distinction

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

is beyond the scope of the present study.) All such instances were produced by just three individuals, but the frequency of Ø in their writing suggests that the instances are not merely mistakes. Even though telegraphic language of this kind is not usually characterized as nativelike, it is nevertheless true that nonpathological native speakers do produce simplified structures in certain contexts (e.g., Ferguson & DeBose, 1977), as seen in the present results. The same may also be true of L2 learners (e.g., Meisel, 1980, 1983), a point that I return to shortly.

L1 Swedish

Swedish is a northern Germanic language that is genetically and typologically quite closely related to English. For present purposes, it is relevant to mention that, like English, Swedish is a subject-prominent (not topic-prominent) language and that it has an article system whose functions and distribution in discourse largely overlap with those of English. The forms are somewhat different, though. The indefinite article in Swedish has two forms, *en* and *ett*, and the choice between these forms is determined by the grammatical gender of the noun (e.g., *en bil* "a car"; *ett hus* "a house"). The Swedish definite article is somewhat more complex, having both bound-morpheme and free-morpheme components. The bound morpheme is a suffix taking the form of *-en* or *-et*, depending on the grammatical gender of the noun (e.g., *bilen* "the car"; *huset* "the house"). The suffix is the default form of the definite article, but a prenominal definite determiner is additionally used when the noun is modified by an adjective (e.g., *den stora bilen* "the big car"; *det stora huset* "the big house") (see, e.g., Beite, Englund, Higelin, & Hildeman, 1963).

The patterns of discourse marking found in the L1 Swedish narratives are shown in Table 4, and many of these patterns are similar to those that were seen in the data from the native English speakers. For example, one sees that just a few of the Swedes use Ø in NPs referring to the female protagonist. This involves simplified language very similar to that just discussed relative to the American participants, and it again occurs in the narratives of only three of the L1 Swedish participants.

With respect to marked syntax, Table 4 shows that such patterns are found mainly in new contexts (NT and NC). Similar to the L1 English data, in the L1 Swedish data all of the marked syntactic structures in NT contexts involve existential constructions (e.g., *Det var en ung kvinna* "There was a young woman"), whereas all of the remaining instances of marked syntax involve the use of relative-clause postmodifiers (e.g., *Flickan som stal brödet* "The girl who stole the bread"). As before, syntax turns out to be only secondary to morphology in defining the discourse distinctions that the L1 Swedes exhibit. Of primary concern are the Swedes' use of pronouns and the Swedish indefinite and definite articles.

Like the Americans, the Swedes exhibit a preference for the indefinite article in new contexts, pronouns in current contexts, and the definite article in

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Table 4. L1 Swedish results

Group/ context		Ø		Ind. art.		Def. art.		Pronoun		Def. det.		Quant.		Prop. N		Total		Syn.
S7c	NT	—	—	95	(35)	5	(2)	—	—	—	—	—	—	—	—	100	(37)	(20)
	NC	—	—	100	(3)	—	—	—	—	—	—	—	—	—	—	100	(3)	(1)
	CT	—	—	0	(1)	34	(83)	66	(163)	—	—	—	—	—	—	100	(247)	(1)
	CC	—	—	1	(1)	46	(80)	53	(93)	—	—	—	—	—	—	100	(174)	(1)
	KT	1	(1)	—	—	83	(84)	16	(16)	—	—	—	—	—	—	100	(101)	(8)
	KC	—	—	2	(1)	89	(57)	6	(4)	3	(2)	—	—	—	—	100	(64)	(3)
S9c	NT	—	—	91	(30)	9	(3)	—	—	—	—	—	—	—	—	100	(33)	(6)
	NC	—	—	100	(5)	—	—	—	—	—	—	—	—	—	—	100	(5)	—
	CT	—	—	—	—	36	(82)	64	(144)	—	—	—	—	—	—	100	(226)	(1)
	CC	1	(2)	—	—	46	(89)	53	(102)	1	(1)	—	—	—	—	100	(194)	(2)
	KT	1	(1)	1	(1)	95	(69)	3	(2)	—	—	—	—	—	—	100	(73)	(8)
	KC	—	—	2	(1)	94	(62)	5	(3)	—	—	—	—	—	—	100	(66)	(1)

known contexts. The Swedes' preference for the indefinite over the definite article in new contexts is stronger than the corresponding preference by the Americans, and the difference between the Swedes and Americans is significant even when comparing only the seventh-grade groups ($\chi^2 = 14.67$, $p < .001$). This suggests that if the American seventh graders' high use of *the* in NT contexts is age related, it might simultaneously be L1 related. In any case, it is clear that the two Swedish groups make a clear distinction between new and not-new contexts, and in my examination of individual texts I found that over 85% of the individuals in both Swedish groups make categorical distinctions between new and not-new contexts in their use of the indefinite article. The story is quite different for the distinction between current and known referents, however: Fewer than 6% of the individuals in either group exhibit a categorical distinction between current and known contexts. Admittedly, at the group level, differences in Swedes' use of pronouns versus the definite article across current and known contexts are consistently significant at $p < .001$, but this appears to reflect only a statistical trend and not a categorical distinction.

Concerning whether the Swedes show a distinction between topic and comment referents in L1 Swedish, the present results suggest that they do not. A statistically significant difference in the distribution of pronouns and the definite article across topic and comment contexts was found only with respect to the KT versus KC distinction for group S7c. The difference is that pronouns are slightly (though significantly) more frequent with topics than with comments, which may, again, reflect the fact that topics tend to be more expected (and thus more retrievable from context) than comments are. Acknowledging this one significant difference, the overall results do not suggest that Swedes mark a distinction between topics and comments in their L1. Therefore, from the perspective of both L1 Swedish influence and L2 English input, one would predict that the topic-comment distinction will not show up in L2 English texts produced by Swedes. Concerning other discourse distinctions, however, one would predict that Swedish learners of English will exhibit categorical distinctions between new and not-new contexts in their use of the indefinite article and will also exhibit group-level statistical trends favoring pronouns in current contexts and the definite article in known contexts.

Swedes' L2 English

The Swedes' L2 English results, shown in Table 5, confirm these predictions. The percentages in Tables 4 and 5 are, in fact, so conspicuously similar that they leave little doubt concerning the (facilitating) influence of L1 Swedish in this area of interlanguage performance. The Swedish learners of English do indeed exhibit a clear distinction between new and not-new contexts, marking the former overwhelmingly with the indefinite article and marking the latter with either the definite article or a pronoun. From the perspective of categorical distinctions, I found that over 65% of the participants in each group (66%

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Table 5. Swedes' L2 English results

Group/ context		Ø		*a*(*n*)		*the*		Pronoun		Def. det.		Quant.		Prop. N		Total		Syn.
S7	NT	—	—	91	(61)	4	(3)	—	—	4	(3)	—	—	—	—	100	(67)	(22)
	NC	—	—	100	(1)	—	—	—	—	—	—	—	—	—	—	100	(1)	(1)
	CT	1	(3)	1	(5)	38	(207)	60	(326)	0	(1)	—	—	—	—	100	(542)	(2)
	CC	—	—	3	(8)	52	(156)	45	(134)	1	(2)	—	—	—	—	100	(300)	—
	KT	1	(2)	1	(2)	91	(128)	5	(7)	—	—	1	(1)	—	—	100	(140)	(7)
	KC	1	(1)	2	(3)	86	(120)	9	(12)	2	(3)	—	—	—	—	100	(139)	—
S9	NT	2	(1)	92	(61)	2	(1)	—	—	—	—	3	(2)	2	(1)	100	(66)	(22)
	NC	—	—	100	(11)	—	—	—	—	—	—	—	—	—	—	100	(11)	(4)
	CT	1	(5)	1	(6)	28	(145)	67	(348)	1	(4)	—	—	2	(8)	100	(516)	(3)
	CC	0	(1)	1	(3)	41	(133)	56	(182)	0	(1)	—	—	2	(7)	100	(327)	—
	KT	3	(4)	2	(2)	82	(108)	10	(13)	—	—	—	—	4	(5)	100	(132)	(18)
	KC	1	(1)	1	(1)	86	(125)	7	(10)	2	(3)	—	—	4	(6)	100	(146)	(5)

for group S7, 78% for group S9) use the indefinite article categorically only in new contexts, whereas fewer than 10% of the participants in these groups (1% for S7, 8% for S9) exhibit categorical distinctions in their use of pronouns in current contexts and the definite article in known contexts. At the same time, the group-level differences between their use of pronouns and *the* across current and known contexts are consistently significant at the level of $p < .001$, even though the use of *the* is unexpectedly high in CC contexts for group S7. In summary, the Swedish learners of English do seem to make a psychologically real distinction between new and not-new referents, and they additionally show a statistical preference for denoting current referents with pronouns and known referents with the definite article. Unlike the new versus not-new distinction, the current versus known distinction may not be psychologically real; it may simply be a condition-induced tendency, and this interpretation is supported by the fact that fewer than 10% of the Swedes were found to exhibit a categorical distinction between current and known contexts in their choices of linguistic options.

Regarding the distinction between topics and comments, significant differences (at $p < .001$) were found between CT and CC for both groups but were not found between KT and KC. The patterns seen in the CT and CC contexts continue to suggest that topic referents are more expected than comment referents are, which makes them more likely to be referred to with a pronoun. Nevertheless, the overall patterns in Table 5 (including all six discourse contexts) do not suggest that Swedish learners of English have an innate predisposition to mark the distinction between topics and comments (cf. Chaudron & Parker, 1990; Young, 1995).

Four other patterns in Table 5 deserve comment. The first is the learners' use of Ø with singular count nouns referring to the female protagonist. The use of Ø in such linguistic contexts would normally be regarded as erroneous, but the fact that the Swedes' use of Ø in the L2 is only slightly more frequent than it is in equivalent contexts in the L1 leads one to wonder whether all of the instances of Ø in Table 5 really do reflect errors (faulty competence) or even mistakes (slips in performance), or whether some of these instances may have been intentional uses of simplified (telegraphic) language, as was presumably the case with the corresponding patterns in the L1 English and L1 Swedish data (compare with Tables 3 and 4). Another pattern that seems to be an error but probably should not be regarded as such is the learners' occasional use of *a*(*n*) in current and known contexts. In most of these cases, the learners were reintroducing the female protagonist after a break in their writing. This is where the elicitation procedures had an effect on the data, as was alluded to earlier.

Another pattern in Table 5 that deserves comment is group S9's use of proper nouns (e.g., Ann, Cindy, Mariah) for referring to the female protagonist. All such instances were produced by just four participants, all of whom were students in the same English class, having the same teacher, so their use of proper nouns is probably an effect of instruction. The final pattern in Table

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

5 that I will address concerns the Syntax column. All but three of the instances of marked syntax in NT contexts involved existential constructions, and all but seven of the remaining instances of marked syntax in all other contexts involved relative-clause postmodifiers. The three exceptional instances in NT contexts involved two overt indicators of new information (e.g., *in the beginning*) and one instance of inverted word order (i.e., *In the street walk a girl*). The seven exceptional instances in other contexts involved two instances of dislocation in CT contexts (e.g., *This girl, she ran away*), one cleft construction in a CT (possibly CC) context (i.e., *Now was it she who had to help him*), and four pseudo existential constructions in a KT context (e.g., *There was also the girl*, meaning that the girl was also there).

L1 Finnish

Finnish is a Finno-Ugric language that is historically and typologically distinct from English. Finnish does not have an article system, although it does have optional means for indicating definiteness and indefiniteness. Finns normally do not mark nouns for either definiteness or indefiniteness (see Table 6), but when they do mark a specific, countable NP referent for indefiniteness (i.e., [–HK]), they often use the word for *one* (*yksi*) or an indefinite adjective or pronoun (e.g., *eräs*, *joku*, *semmonen*, roughly meaning "some") placed at the front of the NP (e.g., *eräs köyhä nainen* "a poor woman"). The use of such words is represented in Table 6 in the column labeled Quantifier. With respect to definiteness, Finns often use a third-person pronoun (*se* "it", *ne* "they") placed at the beginning of the NP (e.g., *se nainen* "the woman") (see, e.g., Chesterman, 1991; Laury, 1997). The use of this definite marker in the Finns' references to the female protagonist is represented in Table 6 in the column labeled *se*.

The most striking pattern in Table 6 is the high use of Ø across contexts. Markers of indefiniteness and definiteness do occur in the data but only infrequently. Indefinite quantifiers, as expected, occur mainly in new contexts (NT and NC), and the definite marker *se* occurs primarily in current and known contexts, as one would expect. Concerning the distinction between current and known, Table 6 indicates that pronouns are used primarily in current contexts, albeit infrequently in comparison with the use of Ø (i.e., bare NPs). All of these patterns together suggest that Finns do have a sense of the distinction between new and not new, and perhaps also between current and known, but their overwhelmingly high use of Ø across discourse contexts nevertheless suggests that they are not predisposed to marking such distinctions in their L1. This conclusion is supported by my examination of individual L1 Finnish texts, in which I found that fewer than 25% of the participants in each group (24% for F5c, 18% for F7c, 12% for F9c) exhibit a categorical distinction between new and not new, and fewer than 10% of the Finns in each group (0% for F5c, 8% for F7c, 0% for F9c) exhibit a categorical distinction between current and known.

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Table 6. L1 Finnish results

Group/ context		Ø		Ind. art.		Def. art.		Pronoun		Def. det.		Quant.		*se*		Prop. N		Total		Syn.
F5c	NT	79	(22)	—	—	—	—	—	—	—	—	21	(6)	—	—	—	—	100	(28)	(6)
	NC	—	—	—	—	—	—	—	—	—	—	100	(1)	—	—	—	—	100	(1)	—
	CT	64	(93)	—	—	—	—	30	(44)	—	—	—	—	6	(9)	—	—	100	(146)	(3)
	CC	75	(71)	—	—	—	—	17	(16)	5	(5)	—	—	3	(3)	—	—	100	(95)	(1)
	KT	91	(62)	—	—	—	—	1	(1)	1	(1)	—	—	6	(4)	—	—	100	(68)	(8)
	KC	86	(43)	—	—	—	—	—	—	8	(4)	—	—	6	(3)	—	—	100	(50)	(5)
F7c	NT	74	(25)	—	—	—	—	—	—	—	—	21	(7)	6	(2)	—	—	100	(34)	(6)
	NC	80	(4)	—	—	—	—	—	—	—	—	—	—	20	(1)	—	—	100	(5)	(1)
	CT	70	(113)	—	—	—	—	27	(43)	—	—	—	—	4	(6)	—	—	100	(162)	(1)
	CC	72	(96)	—	—	—	—	18	(24)	2	(3)	—	—	8	(10)	—	—	100	(133)	—
	KT	94	(58)	—	—	—	—	2	(1)	—	—	—	—	5	(3)	—	—	100	(62)	(7)
	KC	82	(61)	—	—	—	—	1	(1)	8	(6)	—	—	8	(6)	—	—	100	(74)	(4)
F9c	NT	83	(29)	—	—	—	—	—	—	—	—	17	(6)	—	—	—	—	100	(35)	(9)
	NC	83	(5)	—	—	—	—	—	—	—	—	17	(1)	—	—	—	—	100	(6)	(3)
	CT	66	(151)	—	—	—	—	33	(75)	—	—	0	(1)	—	—	0	(1)	100	(228)	(1)
	CC	80	(140)	—	—	—	—	17	(29)	1	(1)	1	(2)	2	(3)	—	—	100	(175)	—
	KT	89	(75)	—	—	—	—	1	(1)	—	—	—	—	2	(2)	7	(6)	100	(84)	(6)
	KC	97	(85)	—	—	—	—	—	—	—	—	—	—	3	(3)	—	—	100	(88)	(14)

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Concerning the distinction between topics and comments, it is relevant to point out that Finnish is a highly inflectional language in which subject NPs are marked overtly through the nominal case system. Based on this and other relevant criteria (see, e.g., Huebner, 1983, p. 69), Finnish is properly classified as a subject-prominent language. Even so, the flexible word order of Finnish does allow for topicalization through NP fronting (e.g., *Leivän varasti tyttö, eikä mies*, literally "Bread stole girl, not man", meaning "The bread was stolen by the girl, not by the man"). Despite the existence of this option in Finnish, however, only five instances of topicalization in the Finns' L1 reference to the female protagonist were found, all occurring in NT contexts. Most of the other marked syntactic structures occurring in all other discourse contexts involve relative-clause postmodifiers, which do not differentiate topics from comments. Again, the use of Ø dominates both topic and comment contexts in these results. The only noticeable difference between topic and comment contexts is that the Finns use pronouns somewhat more frequently in CT than in CC contexts.

From the perspective of the L1 Finnish data, one would predict that (a) Finnish learners of English will demonstrate frequent use of Ø in their reference to the female protagonist, (b) they will show some sensitivity to the distinction between new and not new and between current and known, but (c) few individual learners will demonstrate categorical distinctions between these discourse contexts. Based on the L1 Finnish data, one would also predict that Finnish learners of English will not overtly mark the distinction between topic and comment NP referents.

Finns' L2 English

The results in Table 7 provide qualified support for all but the last of these predictions. Even though the L1 and L2 results are not nearly as parallel for the Finns as they were for the Swedes, a comparison between the Finns' and Swedes' L2 results (i.e., Tables 5 and 7) makes it perfectly clear that L1 influence thrives in this area of L2 acquisition (see also, e.g., Master, 1987, 1997; Ringbom, 1987).

Most striking in Table 7 is the Finns' high use of Ø, which exceeds 15% of their morphological choices across groups and across contexts (except NC for groups F7 and F9B). Even though the use of Ø remains high even among the ninth graders, its use does diminish gradually by grade, though not by years of English instruction. Intriguingly, group F9B (2 years of English) uses Ø somewhat less frequently than does group F9A (6 years of English). The reason behind this is a bit elusive, but I suspect that it has to do with the fact that group F9B has studied Swedish for 6 years. The Swedish articles, which I described earlier, are phonologically and morphologically more salient than the English articles, exhibiting less phonological reduction and more formal redundancy through gender agreement and the existence of both free and bound forms. So, the learners in group F9B may be more aware of article func-

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Table 7. Finns' L2 English results

Group/ context		Ø		*a(n)*		*the*		Pronoun		Def. det.		Quant.		Prop. N		Total		Syn.
F5	NT	36	(21)	20	(12)	34	(20)	2	(1)	—	—	8	(5)	—	—	100	(59)	(7)
	NC	33	(1)	33	(1)	33	(1)	—	—	—	—	—	—	—	—	100	(3)	(1)
	CT	37	(114)	1	(3)	17	(52)	44	(134)	1	(2)	—	—	—	—	100	(305)	(4)
	CC	32	(35)	6	(7)	25	(27)	31	(34)	6	(7)	—	—	—	—	100	(110)	—
	KT	52	(48)	4	(4)	41	(38)	—	—	2	(2)	—	—	—	—	100	(92)	(4)
	KC	31	(23)	12	(9)	36	(27)	13	(10)	8	(6)	—	—	—	—	100	(75)	(1)
F7	NT	30	(14)	35	(16)	17	(8)	—	—	—	—	17	(8)	—	—	100	(46)	(11)
	NC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	CT	23	(47)	2	(4)	24	(48)	51	(103)	—	—	0	(1)	—	—	100	(203)	(2)
	CC	17	(16)	8	(7)	45	(41)	27	(25)	3	(3)	—	—	—	—	100	(92)	—
	KT	39	(27)	3	(2)	51	(36)	1	(1)	4	(3)	1	(1)	—	—	100	(70)	(5)
	KC	27	(22)	7	(6)	52	(42)	5	(4)	9	(7)	—	—	—	—	100	(81)	(2)
F9A	NT	24	(11)	53	(24)	9	(4)	—	—	9	(4)	4	(2)	—	—	100	(45)	(15)
	NC	100	(1)	—	—	—	—	—	—	—	—	—	—	—	—	100	(1)	—
	CT	19	(50)	—	—	26	(70)	54	(144)	1	(4)	—	—	—	—	100	(268)	—
	CC	20	(26)	5	(6)	40	(53)	29	(38)	6	(8)	—	—	—	—	100	(131)	—
	KT	38	(30)	3	(2)	58	(46)	1	(1)	—	—	—	—	—	—	100	(79)	(4)
	KC	31	(22)	4	(3)	56	(40)	4	(3)	6	(4)	—	—	—	—	100	(72)	—
F9B	NT	20	(9)	49	(22)	24	(11)	—	—	—	—	7	(3)	—	—	100	(45)	(9)
	NC	—	—	100	(1)	—	—	—	—	—	—	—	—	—	—	100	(1)	—
	CT	17	(55)	4	(13)	27	(88)	51	(165)	0	(1)	—	—	—	—	100	(322)	—
	CC	19	(38)	3	(6)	46	(91)	30	(59)	1	(2)	—	—	—	—	100	(196)	—
	KT	31	(26)	6	(5)	61	(52)	1	(1)	1	(1)	—	—	—	—	100	(85)	(1)
	KC	16	(14)	9	(8)	67	(57)	1	(1)	6	(5)	—	—	—	—	100	(85)	(1)

tions and of the fact that they are obligatory in certain contexts in Swedish, and they might transfer this awareness into their use of English (see Ringbom). This interpretation is admittedly speculative, but it is the only explanation I know of that accounts for group F9B's apparent advantage over F9A with respect to article omissions.

Regarding potential discourse distinctions that the Finns make in their use of English, Table 7 shows that their use of marked syntax is consistently higher in new than in not-new contexts, nearly always involving existential constructions in NT contexts and relative-clause postmodifiers elsewhere. The table also shows that the learners' use of *a*(*n*) and indefinite quantifiers (*one* or *some*) in new contexts is substantially higher than it is in not-new contexts (significant at $p < .001$ for all groups, using chi-square). However, while examining individual texts, I found that only 11% of the participants in group F5 categorically distinguish between new and not-new contexts in their use of *a*(*n*), although the percentage increases to 24% for group F7 and to 46% and 40% for groups F9A and F9B, respectively. Concerning the distinction between current and known, Table 7 shows that the Finns' use of pronouns in current contexts is substantially higher than in known contexts (significant at $p < .001$ for all groups, using chi-square), and over 20% of the individuals in each group (21% for F5, 43% for F7, 37% for F9A, and 44% for F9B) exhibit categorical distinctions in their use of pronouns in current contexts versus Ø and *the* in known contexts. These are the highest percentages of categorical distinctions between current and known found anywhere in the data and are far higher than predicted on the basis of the L1 Finnish data. Whether the distinction between current and known is psychologically real for the learners who produced this pattern is difficult to say. If it is psychologically real (e.g., in the form of a rule), the fact that the learners at the lowest stages of acquisition (group F5) show the weakest tendency to make this distinction casts doubt on the idea that the distinction is part of the innate knowledge that learners bring with them to the task of acquiring the target language.

With respect to the distinction between topic and comment NP referents, Ø, *the*, and pronouns are distributed significantly differently ($p < .001$, using chi-square) between CT and CC contexts for all groups except F5. Differences between KT and KC contexts, on the other hand, are significant only for group F5 (at $p < .001$). Table 7 shows that, in current contexts, pronouns are consistently more frequent for topic referents than for comment referents, and in known contexts the use of Ø is consistently more frequent with topics than with comments. Whether some learners use Ø as a topic marker is an intriguing question that harks back to Huebner's (1983) claims concerning the earliest speech production of his adult Hmong-speaking learner of English. Recall that Huebner's learner overwhelmingly preferred Ø with topic NP referents and *da* (the interlanguage equivalent of *the*) with comment NP referents. To see whether something similar may be occurring in the Finnish speakers' data, I examined individual texts for the categorical use of Ø in only topic (actually only CT and KT) contexts and indeed found this to be the case in the narra-

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Table 8. Numbers of texts where Ø and *the* occur in complementary distribution in competing contexts

	Groups			
Contexts	F5	F7	F9A	F9B
Topic versus comment (CT vs. CC and KT vs. KC)	8	4	2	1
New versus current (NT vs. CT)	9	3	2	1
New versus known (NT vs. KT)	6	6	2	3
Current versus known (CT vs. KT and CC vs. KC)	5	5	5	2

tives of 29% of the learners in group F5, 22% of the learners in both F7 and F9A, and 13% of the learners in F9B. I also examined whether some learners may use Ø categorically only in comment (CC and KC) contexts (cf. Young, 1995) and found this to be true in an additional 10% of the texts in F5, 4% of the texts in F7, 9% of the texts in F9A, and 7% of the texts in F9B. These results—and the fact that the categorical distinctions are highest among the lowest level group (F5)—suggest that Finnish-speaking learners of English may have an innate propensity to mark the distinction between topics and comments, a distinction that is not prominent in either their L1 or the target language. At the same time, one could anticipate potentially several possible factors that could affect the use of Ø in competing contexts, and I have consequently chosen this issue as the focus of my qualitative analysis in the following section.

QUALITATIVE EXAMINATION OF FINNS' USE OF Ø VERSUS *THE*

This section focuses on 40 narratives—all written by Finnish-speaking learners of English—in which a categorical complementary distribution was found in the use of Ø versus *the* between competing contexts involving learners' NP references to the female protagonist. The numbers of narratives per group in which such distinctions were found are shown in Table 8. Some texts fall into more than one category, so the grand total across rows and columns is 64, even though the actual number of texts represented is only 40: 15 texts from group F5, 13 from F7, 8 from F9A, and 4 from F9B. In two texts, categorical distinctions between topic and comment involve CT versus CC; in all other cases, the distinction between topic and comment involves KT versus KC. Concerning the distinction between current and known, 9 of the 17 cases involve CT versus KT, and the remaining 8 cases involve CC versus KC. Roughly half (i.e., 8 of 15) of the texts displaying a distinction between topic and comment show Ø in topic position and *the* in comment position, whereas the other half exhibit the opposite pattern. The same is true of the distinction between new and known, new contexts being marked with Ø in 9 of the cases, and *the* in the remaining 8 cases. With respect to the second and fourth rows,

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

there is a substantial tendency for current NP referents to be marked with Ø instead of with *the*. This is true in 12 of the 15 texts represented in the "New versus current" row and in 15 of the 17 texts represented in the "Current versus known" row.

The purpose of the present qualitative analysis is to examine more closely the 40 texts represented in Table 8 to verify whether the learners who wrote them really do use Ø and *the* to maintain a distinction between the discourse contexts in question. Because this requires taking the bigger picture into account, I will now expand my analysis of these 40 texts to include the use of Ø and *the* in all NPs, no longer just NPs referring to the female protagonist.

Topic versus Comment

As mentioned previously, Huebner (1983) and Young (1995) found that learners of vastly different L1 backgrounds seem to maintain a distinction between topics and comments through their use of Ø and *the*. Table 8 suggests that the same thing may be going on in several of the narratives in the present study as well. However, looking qualitatively at the use of Ø and *the* in all NP referents in these narratives, I found that only two narratives provide compelling evidence of the distinction between topic and comment. Both were written by participants in group F5, and both involve the frequent use of *the* in comment (NC, CC, and KC) positions and the infrequent use of *the* in topic positions. The narrative in (2), written by participant F5–57, is the clearest example. For convenience, I have listed separate sentences on separate lines, highlighted the use of *the*, and inserted the Ø symbol before all full NPs where no article occurs in the data. Original spellings have been preserved.

(2) *Ø Girl walking hungry and alone.*
Suddenly she sees a food.
And she stold that.
She hit Ø C.C.
Ø Policeman coming and he want restid the girl.
But Ø C.C. said: "I stould the bread."
Then Ø policeman restid Ø C.C.
Some woman coming and said: "It isn't the man—it was the girl that man laing."
Then Ø C.C. walking the restaurant.
He take very much food.
Then he doesnt bay the bild.
Ø Policeman restid Ø C.C.
He have the bee go the policecar.
Then the same girl coming a policecar. [Note: "the same girl" in this line is KT]
They jumping the car out.
And they walking happy a way.
Then they sitting on Ø gras.
And they dream a bout they dreamhouse.

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Then Ø C.C. said: "I'll do it."
But Ø policeman stand behind they.

Although the use of Ø in topic position and *the* in comment position is not completely categorical, the observed tendencies are clearly in these directions, which suggests that at least in this and the similar narrative mentioned earlier the learner may be attempting to maintain a distinction between topic and comment. Of course, alternative explanations are also possible. One possible explanation is that it is really the distinction between subject and nonsubject that matters here (see Parrish, 1987), and yet another possible explanation is that the crucial distinction for these learners is really the distinction between sentence-initial and noninitial NPs (see Young, 1995). In fact, in addition to these two narratives, 14 others (written mainly by participants in groups F5 and F7) suggest a distinction between initial and noninitial NPs. In one of these narratives, *the* is used in sentence-initial position and Ø is used in noninitial positions, and in the remaining 13 narratives, Ø occurs primarily sentence initially and *the* occurs only in noninitial positions. The narrative in (3), written by participant F7–07, is a good example of this pattern.

(3) *One woman was hungry and alone.*
Then she saw a bread car and she stoled some bread.
Ø Woman did saw it and she sayed it to the policeman.
Ø Police cacth the girl and arrested him.
Ø Charlie came by and he sayed that he did it not the girl.
Ø Police take him to a police car.
Ø Woman say that Ø Charlie didnt do it the woman did it and they took the girl in to a police car.
Ø Woman tryed to escape.
Ø Car fell of and the Charlie and the woman fly in to a cement and then they escaped. [Note: "the Charlie and the woman" is both subject and topic, but noninitial]
Ø Charlie and Ø woman sit on the crass.
Then they started to make Ø questions.
Ø Charlie asked to the woman that coud she imagine they in a house.
Then they boht imagine Ø Charlie and a woman in a house.
After that Ø Charlie sayed that he have to do it.

The occurs in this narrative in new, current, and known contexts, as well as in topic and comment positions. It also occurs in both subject and nonsubject positions. The only categorical pattern that *the* seems to exhibit in this and 12 other narratives is its categorical absence in sentence-initial positions. Given that the psychological reality of the distinction between initial and noninitial positions is widely accepted in phonology (see, e.g., Celce-Murcia, Brinton, & Goodwin, 1996) and is also considered relevant in discourse (e.g., Rutherford, 1987; Young, 1995), this explanation may indeed be valid and certainly seems to reflect the patterns in the present data. It may also account for the distribution of Ø and *the* in Huebner's (1983) study, which he interpreted as reflecting the topic-comment distinction.

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

New versus Current versus Known

In 10 of the 40 narratives, new NP referents are marked differently from current and known NP referents: In seven cases, new is marked with *the*, and current and known are marked with Ø; in one case, the opposite pattern holds; in the final two cases, new is marked with *a* and current and known are marked with either Ø or *the*. In 7 of the 40 narratives, the distinction is between current, on the one hand, and new and known, on the other. In all seven cases, current is marked with Ø, and new and known are marked with *the*. This is demonstrated in the narrative in (4), written by participant F5–68, in which the only NP referents marked with an overt article (usually *the*) are new and known referents, although, crucially, not all new and known referents are marked with an overt article.

(4) *There was the girl whose stole a bread.*
She run, and hit to Ø Chaplin.
Then Ø policeman chatch her.
But Ø Chaplin was realy Ø centleman, and he told, that he was stole the bread.
Ø Policeman belive him, and take Ø Chaplin.
But one people sed to Ø policeman: It wasen't Ø man, it was the girl!
Then Ø policeman go cool to Ø police station whit Ø Chaplin.
Ø Chaplin tryed smoking.
Then Ø policeman take Ø Chaplin in the police car.
Then Ø girl come in.
Ø Centleman Chaplin give his place to Ø girl.
Then Ø police car hit to another car.
Ø Girl and Chaplin run a way together, and they was wery happy.
Then they sat under the big tree.
Near the little house.
Then Ø Chaplin asked to Ø girl: "Where you live?"
Ø Girl answer: "No place, anywhere."
Ø Chaplin sed: "Do you want to live Ø little house?"
Then they dreaming somthing like this:
They are living in Ø little house . . .
But then Ø policeman sed them under Ø little tree.
Then they walkin away gen.
Finish.

Other Patterns

In several narratives, the Finns' use of Ø and *the* do not seem to reflect any of the discourse distinctions under investigation, and even in the narratives where discourse distinctions do seem to affect learners' use of Ø and *the*, additional factors also appear to be at play. In 10 narratives, for example, learners seem to associate one of the article options (usually Ø) with a prepositional function—or with a function that would be represented in Finnish with a locative case suffix. This can be seen in the narrative in (5), written by

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

participant F9B-14, where Ø occurs mainly after prepositions or in places where a locative suffix would be used in Finnish (and also before proper, plural, and noncount nouns).

(5) *A woman walked in the street and she was hungry.*
She noticed a bread car, and there was noone to watch it.
She stoled a bread and run away.
A man walked in the street and the woman run towards him.
A fine lady saw the hole thing and she told it to Ø bread cars owner.
They called Ø police. [Note: direct translation from Finnish "called to police"]
A man said he stoled the bread and then the police man took him.
But the fine lady said it was the woman who took the bread.
So the police took both of them.
the man and the woman were in a police car going to Ø jail.
the woman wanted away so she jumped out of the car.
the man jumped too.
the police man was in the street and the man knocked him with a stick.
Then a man and a woman escaped.
Ø Charlie and a girl walked the streets.
They sat on Ø crash [grass] and talked something.
the girl had no place to live.
They saw a couple in a houses yard and they seemed happy.
Ø Charlie started to imagine him and Ø girl in a house like that.
There were Ø apples and Ø cows and they ate Ø dinner.
Then he stopped dreaming.
She asked the girl if she'd like to live with him.
The girl said yes.
The police man told them to get out of the crash [grass], so they left.

Besides prepositional functions, nine narratives suggest word-specific effects on article use. In three of these cases, *the* occurs largely in phrases that seem to have been learned as prefabricated chunks (e.g., *the street, the grass*). In the six other narratives exhibiting word-specific effects, *the* is used frequently except with certain nouns such as *police*, *policeman*, *baker*, and *bread*, all of which seem to have highly institutionalized meanings—that is, even though they denote specific referents in the learners' narratives, their interpretation requires the reader to rely on world knowledge beyond what the current discourse context provides.

Three final factors deserve some mention. First, heavy NPs (i.e., nouns modified by adjectives or relative clauses) seem to be treated differently than unmodified nouns in some of the narratives (see Young, 1995, pp. 162–163). In three cases, heavy NPs are marked with Ø and light NPs tend to be marked with *the*, and in one additional case the opposite pattern holds. Second, the titles that occurred in the film (see the Appendix) also seem to have had an effect on some of the learners' use of articles. In at least two of the 40 narratives, the use of *the* is limited to almost only those phrases that occurred in the titles (e.g., *It was the girl—not the man*, *Remember me—and the bread?*). Finally, monitoring also seems to have had an effect on article use in at least

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

four of the narratives. This is illustrated in the narrative in (6), written by participant F9A-23, where the use of *the* is limited almost exclusively to three adjacent sentences, as if the learner were applying conscious rules concerning article use while constructing these particular sentences and applying (unconscious) internalized rules elsewhere.

(6) *Ø Girl is alone and hungry when she look Ø bread shops window so she stole Ø bread from Ø car which has just arriwed, but Ø woman saw what she did and tell it to Ø shop owner.*
Ø Owner run after Ø girl, but when he get Ø girl Ø Chaplin say to Ø police that he did it.
Ø Police take Ø Chaplin, but Ø women and Ø shop owner tell Ø police that Ø girl stoled the bread and Ø police let Ø Chaplin go.
Ø Chaplin go to Ø restaurant and take Ø lot of food from Ø table.
When Ø Chaplin has eaten, he tell Ø police that he hasn't paid.
Ø Police put him to Ø police car wich is almost full of Ø criminals.
Ø Few minutes later the girl who stoled the bread come inside the car.
Ø Chaplin stand up and let the girl sit on his site.
When Ø car almost hit another car, Ø Chaplin and the girl jump out of the car.
Ø Chaplin tell Ø girl to escape, but Ø girl want Ø Chaplin with her so theyr escape together.
Ø Chaplin and Ø girl sit down on Ø grass and they see when Ø man go to Ø work anh his wife is also out and they have Ø house.
Ø Chaplin ask Ø girl could she image that they have Ø own house.
They dream Ø little while about Ø own house.
Ø Chaplin promise that he will get for them altought he must go to Ø work.
They stand up when Chaplin see a policeman and they walk away.

DISCUSSION AND CONCLUSION

The primary issue addressed in this study is whether learners overtly mark distinctions between new, current, and known NP referents, on the one hand, and between topic and comment NP referents, on the other. The quantitative analysis focused primarily on distinctions involving new versus not new, current versus known, and topic versus comment. With respect to the first of these, I found that both Finns and Swedes are sensitive to the distinction between new and not-new NP referents, but the degree to which they mark this distinction appears to depend heavily on their L1 background. The distinction between new and not new is part of the obligatory grammar system of Swedish, and Swedish learners of English at different levels of proficiency show an overwhelming tendency to mark this distinction in English. The same distinction in Finnish is not obligatory, and Finnish learners of English correspondingly show a lower propensity for marking it, even though this tendency does increase in accordance with L2 proficiency.

The distinction between current and known NP referents is somewhat less straightforward. On the one hand, the quantitative results for all groups of Finns and Swedes do indicate significantly different preferences for marking

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

current versus known, with pronouns generally preferred in current contexts and *the* in known contexts. On the other hand, this distinction is rarely categorical in the Swedes' L2 narratives, and it is somewhat confounded by the widespread competition between Ø and *the* in the Finns' L2 narratives. Although the evidence in the present study is sufficient to confirm that the distinction between current and known is important at the level of language use, I would question whether the psychological reality of this distinction is comparable to that of the distinction between new and not new. Ultimately, it is a question of whether the new-current-known distinction (Chaudron & Parker, 1990; Givón, 1983) really does reflect a trichotomous system in learners' mental discourse grammars. I suspect that it does not, given that the new versus not-new distinction is reflected more in the choice between competing article options, whereas the current versus known distinction is reflected more in the choice between pronouns and full NPs.

The distinction between topic and comment is perhaps the most problematic, not least because of the widely differing views on how to define this distinction (see, e.g., Reinhart, 1982, for a discussion of this problem). In the present study I have assumed that a topic referent can be new, current, or known provided that it is the referent that a T-unit most centrally conveys information about. This is, to my knowledge, the only general view of the topic-comment distinction that is compatible with existing assumptions concerning topic continuity and article use (e.g., Chaudron & Parker, 1990; Givón, 1983). With the distinction defined in this way, the present results show very little consistent or compelling evidence of either learners or native speakers differentiating overtly between topic and comment NP referents. The most compelling evidence involves the relatively high percentages of Finns who use Ø only with topic or only with comment NPs denoting the female protagonist. However, when I analyzed the data qualitatively and examined learners' use of articles with all NP referents, I found only two narratives in which the distinction between topic and comment could not easily be discounted.

The present study therefore casts some doubt on the universality of the distinction between topic and comment, but I do acknowledge the limited generalizability of the present results, which should not be interpreted as refuting the findings of several past studies that have documented an early topic-comment stage in L2 acquisition (see Chaudron & Parker, 1990, p. 50, for a list of such studies). The learners in the present study had all had at least 2 years of prior English instruction at the time of testing, and it is possible that they had already passed through an earlier topic-comment stage. If they had not, then it is possible that this is because their L1s are subject prominent, unlike the L1s represented in most of the studies that have found a topic-comment distinction (e.g., Chaudron & Parker, 1990; Huebner, 1983). A final possibility is that the results of the present study may have shown a greater effect for the topic-comment distinction if the distinction had been defined differently. If I had followed Young (1995) in defining topics as old information and comments as new information, for example, then greater effects undoubtedly

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

would have been found. However, these effects would have overlapped completely with my findings concerning the distinction between new and not new, and thus the notion of topic would have been completely redundant. Future work is certainly needed to sort out exactly what a topic is and how useful this notion is (see, e.g., Hajicova & Kruijff-Korbayova, 1999, and Shi, 2000, for recent treatments of the topic-comment distinction).

Learners' use of Ø, which indicated a possible topic-comment distinction in the Finns' quantitative results, provides some of the most intriguing patterns in the present study. The frequency and distribution of Ø across discourse contexts in both the Finns' and Swedes' L2 data resemble the L1 patterns and therefore suggest that learners' use of Ø should not be interpreted simply as omission resulting from either carelessness or ignorance. Instead, the results of the present study suggest that learners' use of Ø is often—perhaps usually—quite intentional or at least rule-governed (cf. Master, 1987, 1997). In the Swedes' data, Ø is nearly as infrequent in the L2 English narratives as it is in the L1 Swedish narratives, and the few instances that do occur in both the L1 and L2 appear to result from an intentional use of a simplified register (cf. Ellis, 1982; Meisel, 1983). In the Finns' data, on the other hand, the use of Ø probably does not represent a simplified register as much as it represents the L1 Finnish convention of avoiding (what Finns perceive to be) redundant markers of definiteness and indefiniteness when these properties of an NP are already salient in a given discourse context (Ringbom, 1987; Sajavaara, 1983). The qualitative analysis in the present study indicated several factors that seem to affect Finnish-speaking learners' use of Ø versus *the*, and these factors may correspondingly reflect the conditions that determine whether Finns perceive *the* to be redundant. These factors include not only discourse distinctions, but also linguistic context (i.e., specific words, such as *police*, and specific grammatical constructions, such as pre- and postmodification and prepositional phrases), task-internal input (i.e., phrases from the titles in the film), and psycholinguistic context (e.g., the degree to which learners consciously plan and monitor their use of articles) (see Ellis, 1994, pp. 119–158, for a discussion of similar factors affecting variability).

The qualitative results therefore suggest that L2 article use is determined by several factors in addition to the discourse distinctions that have been the focus of the present investigation, and they also suggest that individual learners can simultaneously entertain multiple hypotheses about article use. I hope that future research will be able to catalog the full range of hypotheses that learners make about the functions of the English articles and will clarify the conditions that motivate such hypotheses. Future research in this area would also benefit from further clarification of the nature of topics and topic continuity, whether these discourse phenomena really are part of the human language endowment, and whether they truly represent universals of L2 acquisition. All of these suggestions are in line with the field's overarching goal of explaining the nature of interlanguage rules and determining their origins.

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

(*Received 10 July 2001*)

References

Beite, A.-M., Englund, G., Higelin, S., & Hildeman, N.-G. (1963). *Swedish grammar*. Stockholm: Almqvist & Wiksell.

Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. New York: Cambridge University Press.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–55). San Diego, CA: Academic Press.

Chaudron, C., & Parker, K. (1990). Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition*, *12*, 43–64.

Chesterman, A. (1991). *On definiteness: A study with special reference to English and Finnish*. New York: Cambridge University Press.

Ellis, R. (1982). The origins of interlanguage. *Applied Linguistics*, *3*, 207–223.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Ferguson, C. A., & DeBose, C. E. (1977). Simplified registers, broken languages, and pidginization. In A. Valdman (Ed.), *Pidgin and creole linguistics* (pp. 99–125). Bloomington: Indiana University Press.

Givón, T. (Ed.). (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: Benjamins.

Givón, T. (1984). Universals of discourse structure and second language acquisition. In W. E. Rutherford (Ed.), *Language universals and second language acquisition* (pp. 109–136). Amsterdam: Benjamins.

Gundel, J. (1974). *The role of topic and comment in linguistic theory*. Bloomington: Indiana University Linguistics Club Publications.

Hajicova, E., & Kruijff-Korbayova, I. (1999). On the notion of topic. In E. Hajicova, T. Hoskovec, O. Leska, P. Sgall, & Z. Skoumalova (Eds.), *Travaux du Cercle Linguistique de Prague [Prague Linguistic Circle papers]* (pp. 225–236). Amsterdam: Benjamins.

Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma.

Huebner, T. (1985). System and variability in interlanguage syntax. *Language Learning*, *35*, 141–163.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. Champaign, IL: National Council of Teachers of English.

Jarvis, S. (1998). *Conceptual transfer in the interlingual lexicon*. Bloomington: Indiana University Linguistics Club Publications.

Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, *50*, 245–309.

Jourdenais, R. (1999). *A pragmatic look at non-native article use*. Unpublished manuscript, Monterey Institute for International Studies, Monterey, CA.

Laury, R. (1997). *Demonstratives in action: The emergence of a definite article in Finnish*. Amsterdam: Benjamins.

Master, P. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system*. Unpublished doctoral dissertation, University of California, Los Angeles.

Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System*, *25*, 215–232.

Meisel, J. (1980). Linguistic simplification. In S. Felix (Ed.), *Second language development: Trends and issues* (pp. 13–40). Tübingen, Germany: Gunter Narr.

Meisel, J. (1983). Strategies of second language acquisition: More than one kind of simplification. In R. W. Andersen (Ed.), *Pidginization and creolization as language acquisition* (pp. 120–157). Rowley, MA: Newbury House.

Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, *37*, 361–383.

Quirk, R., & Greenbaum, S. (1973). *A concise grammar of contemporary English*. New York: Harcourt Brace Jovanovich.

Reinhart, T. (1982). *Pragmatics and linguistics: An analysis of sentence topics*. Bloomington: Indiana University Linguistics Club Publications.

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press

Ringbom, H. (1987). *The role of the first language in foreign language learning*. Clevedon, UK: Multilingual Matters.
Rutherford, W. E. (1987). *Second language grammar: Learning and teaching*. London: Longman.
Sajavaara, K. (1983). The article errors of Finnish learners of English. *Umeå Papers in English*, *4*, 72–87.
Shi, D. (2000). Topic and topic-comment constructions in Mandarin Chinese. *Language*, *76*, 383–408.
Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, *10*, 335–355.
Young, R. (1995). Form-function relations in articles in English interlanguage. In R. Bayley & D. R. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 135–175). Amsterdam: Benjamins.

APPENDIX

TITLES OCCURRING IN THE FILM SEGMENT

1. "She stole a loaf of bread." (said by the baker)
2. "No, she didn't—I did." (said by Chaplin)
3. "It was the girl—not the man." (said by a woman who witnessed the theft)
4. "Remember me—and the bread?" (said by Chaplin in the paddy wagon)
5. "Now is your chance to escape!" (said by Chaplin after the traffic accident)
6. "Where do you live?" (said by Chaplin after he and the woman had escaped)
7. "No place—anywhere." (said by the woman in answer to Chaplin's question)
8. "Can you imagine us in a little home like that?" (said by Chaplin)
9. "I'll do it! We'll get a home even if I have to work for it." (said by Chaplin)

https://doi.org/10.1017/S0272263102003029 Published online by Cambridge University Press