On Lexical Cohesive Behavior of Heads of Definite Descriptions: A Case Study

Beata Beigman Klebanov and Eli Shamir

School of Computer Science and Engineering,
The Hebrew University of Jerusalem, 91904, Israel,

Abstract. This paper uses materials from annotation studies of lexical cohesion (Beigman Klebanov and Shamir, 2005) and of definite reference (Poesio and Vieira, 1998; Vieira, 1998) to discuss the complementary nature of the two processes. Juxtaposing the two kinds of annotation provides a unique perspective for observing the workings of the reader's common-sense knowledge at two levels of text organization: in patterns of lexis and in realization of discourse entities.

1 Introduction

Introducing the notion of **cohesion** of a text, Halliday and Hasan [6] detail numerous ways in which textual elements connect with each other, providing the perception of the unity of the text. Among other kinds of texturizing devices, they cite **lexical cohesion** – use of words with related meaning, and **referential cohesion** – the tendency of a text to refer repeatedly to the same set of entities.

Subsequent research discovered that whereas the cohesive devices are often identifiable on the basis of their form, the patterns they form with each other – what connects to what – are by no means easy to establish by some systematic, algorithmic 'resolution' procedure [4, 10, 12]. Readers were thought to do it effortlessly; however, reader-based studies show cases of disagreement and difficulty [9].

In this paper, we will attempt to analyze the relationship between lexical and referential cohesion: whether and when they reinforce each other, and what can be learned from their divergence. Human-annotated data is used to provide information about each of the phenomena: The Wall Street Journal 1989 article 'Computers Start to Get Personal' (shown as appendix A) was used both in Poesio and Vieira's seminal annotation study regarding referential behavior of definite noun phrases [9, 11] and in Beigman Klebanov and Shamir [1–3] lexical cohesion annotation experiment.[1]

In what follows, we introduce the two annotation schemes, and sketch out our predictions of the interaction between the two types of cohesion (sections 2, 3). We then proceed to testing the prediction on the basis of the annotated data.

2 Lexical Cohesion through Anchoring

In Beigman Klebanov and Shamir's rendering, lexical cohesion between items in a text arises on the basis of stereotypical, common-knowledge-based connections between the

---

[1] We will use another annotated text which will be introduced in due course.

relevant concepts [1–3]. Sensitive to text dynamics, their question to the subjects was: "For every concept first mentioned in the text, which previously mentioned concepts help the easy accommodation of the current concept into the evolving story, if indeed it is easily accommodated, based on the common knowledge as perceived by the subject" [3]. The preceding helper concept is called an **anchor**, and the relation is marked *anchored→anchor*.

Only first mentions of every lexical item in the text were subjected to anchoring annotation. There were no limitations on the number of anchors per item or on the textual distance between the anchored item and its anchor. Subjects were made aware of the notion of referential cohesion through examples, and were asked not to mark connections solely on the basis of identity of reference in the given text.[2]

Beigman Klebanov and Shamir's experiment resulted in 10 texts annotated for anchoring relations by 22 readers. Detailed statistical analysis of the data identified 2 outliers, and set annotator agreement thresholds that ensure high degrees of reliability [1]. In particular, an item is **reliably anchored** if it was given some anchor by at least 13 out of 20 non-outliers[3]; an anchoring pair $a→b$ is **core**, or **strong**, if the anchored item a is reliably anchored, and the specific anchor b features in at least 6-7 annotations.[4] Examples will be given during case-study presentation; also see [2].

3   Definite Reference

The phenomenon of definite reference concerns both the text and the reader's general knowledge; it is thus particularly interesting to investigate its relationship with anchoring. We will concentrate on one type of definite references - noun phrases quantified by 'the'. Other definite references include demonstratives and possessives; however, Poesio and Vieira's study [9] only considered 'the'-definites.

According to the discourse analysis literature, 'the'-definites are used to mention entities that should be **uniquely identifiable** by the hearer/reader on the basis of the nominal alone in her mental representation of the current discourse augmented with her general knowledge, as construed by the speaker/writer [5, 8].

Unique identifiability is a cognitive status commanding an intermediate degree of **givenness** in the following hierarchy [5]: In Focus > Activated > Familiar > Uniquely Identifiable > Referential > Type Identifiable.

The sources for this degree of givenness are variable, and include linguistic context, situational context, and general knowledge:

Linguistic: My sister has two children, a boy and a girl. *The boy* is an excellent student.
Situational: The boy sitting in the front row is misbehaving.
General Knowledge: When John came home, he saw that *the door* was already open.

In Poesio and Vieira's study [9], definite descriptions were classified according to the following taxonomy:

---

[2] The gist of the example was that if in a given story a child's father went out to the sea because he is a sailor, then *sailor→sea*, but not *sailor→father*.

[3] This corresponds to 99% reliability threshold.

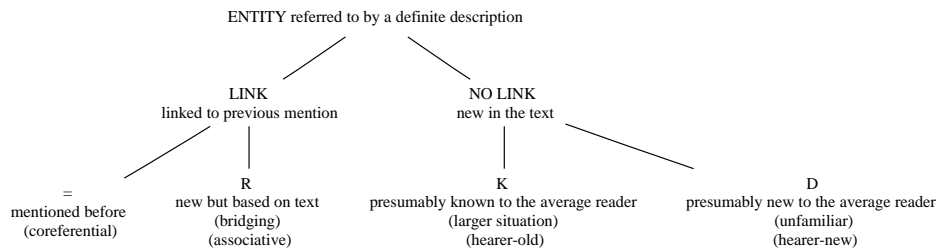[4] The exact threshold depended on the text.

Fig. 1. Poesio and Vieira's classification of Definite Descriptions. Category descriptions are shown according to Vieira's [11] summary of instruction given to the annotators. Additional category titles shown in parenthesis are used by Poesio and Vieira in their article, highlighting the correspondence with other concepts commonly employed in studies of the phenomenon.

3.1 'D'-type definites

Given the position of definites in the givenness hierarchy, the D type (no previous knowledge of the entity) is not expected to be abundant. However, about 20% of the definite descriptions annotated in Poesio and Vieira's study belong to this type [11]. A cursory inspection of cases unanimously classified as D yields numerous examples of long definites with lexically rich post-modifiers of the head:

– the **class** of asbestos including crocidolite
– the **unit** of New York-based Loews Corp that makes Kent cigarettes
– the **fact** that New England proposed lower rate increases

In these cases, the reader probably does not know about the aforementioned classes, units, and facts, but the rich description given on this first mention is sufficient for securing the unique identifiability. Thus, the identifiability is not based on either knowledge or *prevoius* linguistic information, but on linguistic information within the NP itself.

In terms of lexical anchoring, we expect heads of D definites to be left un-anchored, as they are judged to be new and not based on previous text. This is even more so for the 'heavy-tail' definites, where identifying information follows the head.

3.2 '=' type Definites

Definites that repeat an already mentioned entity constitute less than half of all 'the'-definites in Poesio and Vieira's study. A repeated mention of an entity can be done with the same lexical head (as in . . . *a boy* . . . *The boy* . . .), or with a different lexical head (as in . . . *a boy* . . . *the kid* . . .). In the first case, anchoring information will not be available, as the relevant item would constitute a lexical repetition.

In the different-head case where the current head is first mentioned within the given definite, we ask whether this head is anchored at all, and if yes, whether the anchor coincides with the previous mention of the same entity (this would be the case if *kid→boy* in the last example). In such cases, lexical and referential cohesion would go hand-in-hand,

intensifying the connectedness between the two items. Cases of divergent anchoring and reference annotations would constitute examples of the item's different positioning at the two planes of textual structure.

3.3 'R' type Definites

These are cases where the entity is new but based on text, or, as Poesio and Vieira's guidelines [9] elaborate, "based on, dependent on, related to some other idea or thing in the text", exemplifying by "The Parks wanted to buy *an apartment* but *the price* was very high", where *the price* should be classified as R with *an apartment* as its trigger/antecedent. This class is the smallest in their data, only 6-11% of all definites.

This class – often termed **bridging references**, or indirect anaphora – bears the most similarity to anchoring. We point out that in reference annotation, the trigger/antecedent itself is taken to be a noun phrase,[5] whereas anchoring is at the word, rather than phrase, level, and the part-of-speech of the anchor is not constrained in any way.

Furthermore, anchoring guidelines allowed multiple anchors, whereas reference guidelines ask for "*the* previous related mention", i.e. a single antecedent. Indeed, some cases of annotator disagreement were due to different choices of antecedent [9].

Poesio and Vieira note that this was by far the most difficult category to distinguish from the rest. A point of interest is whether the difficulty was in the initial judgement of relatedness to something, or in pinpointing a previous noun phrase that functions as the relevant previous item. In the first case, anchoring annotators might have a similar difficulty; in the second, we expect to find robust anchoring decisions often going to something other than the head of the triggering/antecedent noun phrase.

3.4 'K' type Definites

K definites refer to an entity that "was not mentioned in the text, not related to something in the text, but it refers to something which is part of the common knowledge of the writer and readers in general" [9]. This class covers 20-25% of definites.

Similarly to R type, non-relatedness is likely to be understood as non-relatedness to some other *entity* mentioned in the text. Hence, K class is not necessarily exclusive of anchoring, as long as the anchor is not a head of a noun phrase.

Alternatively, K definites could be references to entities that come-out-of-the-blue as far as the text is concerned, but have a natural reference point in the reader herself, as in Poesio and Vieira's example: "During *the last 15 years* housing prices increased nearly fivefold." In such cases, the text would not provide any material for anchoring.

Table 1 summarizes prediction of anchoring according to reference classification.

4 Reference vs. Anchoring - A Case Study

In the article "Computers Start to Get Personal" – henceforth, **PC-text**, there are 6 examples of LINK-ENTITY (=/R) and 10 examples of NO-LINK (K/D) definites.

---

[5] Poesio and Vieira speak about "... the head noun of the definite description and *the head noun* of its antecedent", p. 200, our italics.

Table 1. Predicted relationship between referential and anchoring behavior.

Anchoring / Reference	Anchored	Anchor Identity
= (Coref) Same Head	—	
Other Head	Yes	Possibly same as coreferent
R (Bridging)	Yes	Same or other than the trigger/antecedent
K (Knowledge)	Yes	Not a head of an NP
	No	
D (Unfamiliar)	No	

In our analysis of LINK-ENTITY definites, we expand the data with coreferential definites from a NYT article "Inspectors: TWA Explosion Did't Originate Near Cockpit" (appendix B) – henceforth, **TWA-text**. This text was annotated for coreference chains within MUC initiative [7], and used for anchoring annotation as well.

4.1 LINK-ENTITY Definites

Out of the 6 LINK-ENTITY definites in PC-text, 4 are unanimously classified as repeated reference. However, we do not get a chance to see anchoring at work, since in none of the 4 cases is the head a newly introduced lexical item.

Table 2. LINK-ENTITY definites with lexically new heads. The heads are boldfaced. Antecedents are in square brackets. We show cases with majority =/R classification. The last column shows the head's anchors produced by the parenthesized number of people (out of 20).

Definite NP	P&V class	Anchors
the Homebrew	D	hobbyists (6) homebrew (1) garage (1)
Computer **Club**	R [hobbyists]	
	= [hobbyists]	
the world **leader** in	= [ibm]	pioneer (4) led (3) world (3) new (2) ibm (2)
computers	= [ibm]	chairman (1) team (1) owners (1) after (1)
	D	example (1) triggered (1) market (1)

The two cases with judgement variability are shown in table 2. The definites are embedded in a larger NP with which they co-refer: an appositive (*IBM, the world leader in computers*) and a list construction with a single member (*for hobbyists such as the Homebrew Computer Club*). There is an uncertainty in reference annotations - some annotators marked embedded NP as coreferential with the whole of the appositive/list construction, whereas others apparently decided that the embedded NP does not constitute a second mention, but rather is part of the ongoing first mention of the entity (D); the single R annotation seems to hit the middle ground between the two solutions.

There is no reliable anchoring of the head of the definite, contrary to our prediction for =/R types. In one case there is a tendency to anchor the head in its antecedent's head

(*club→hobbyists* by 6 people), which is virtually nil in the other case (*leader→ibm* by 2 people). In the latter case, there are better anchors which are not heads of noun phrases: *pioneer* is a modifier inside the NP *many pioneer PC contributors*, and *led* is a verb.

For additional information on coreference and anchoring behavior of definites, we turn to the TWA-text, and seek coreference chains that include coreferential definites with first-mention heads. There are four such chains; the relevant heads are boldfaced:

Crash-chain: the crash – the **accident** – the crash – the plane crash
Plane-chain: Trans World Airlines Flight 800 – the **Boeing 747** – the **jet** – the **plane** – the plane – plane – Flight 800 – plane – the **airplane**
Day-chain: Tuesday–the 20th **day** since Flight 800 exploded in midair off Long Island
Man-chain: James K. Kallstrom–the assistant **director** of the FBI's New York office

Table 3 presents the anchoring data. Let us consider the last two chains first. For *day* and *director*, the situation is similar to *leader* in PC-text: the items are embedded in a coreferential appositive construction. *Day* shows negligible anchoring; *director* has a stronger anchoring pattern, though it misses the core-reliability threshold.

Table 3. Anchoring behavior of coreferential definites, TWA-text.

Same-chain lexical heads	Anchoring Patterns	
	Within the Chain	Outside the Chain
crash – – accident	accident→crash (17)	accident→ explosion (11) catastrophic (9) wreckage (7) survived (4) destroyed (2)
flight 800 – – Boeing 747– – jet – – plane – – airplane	boeing → flight (15) jet → flight (17) boeing/747 (16) plane → jet (17) flight (15) boeing/747 (16) airplane → plane (19) jet (18) boeing/747 (11) flight (10)	boeing → airlines (14) twa (11) cockpit (11) jet → airlines (16) cockpit (12) twa (9) plane → airlines (12) cockpit (12) twa (10) airplane → airlines (10) cockpit(9) twa(7) transportation (3) hangar (3) altitude (2)
Tuesday–day	day → tuesday (2)	
Kallstrom – – director		director → chairman (7) assistant (4) board (3) vice_chairman (2) senior (1)

While the agreement on the best anchor for *director* is not high, *chairman* is the preferred choice. This case shows clearly the distinction between referential and lexical structures: the NP headed by *chairman* is not related referentially to the one headed by *director*, as they refer to different people in different organizations. However, the director is the second senior official named and quoted in this short article, so the items form a pattern, for some readers. On the other hand, *director* is not cohesive anchoring-wise with *Kallstrom*, as, presumably, this name does not feature in the readers' knowledge.

We now turn to the crash- and plane-chains in table 3. None of the 7 definites in question is embedded in its coreferent NP. Their behavior is radically different: all items are anchored, with numerous strong anchors. Thus, *accident* in strongly anchored in its coreferent's head *crash*, but also in other items.

The pattern in the second chain is interesting: each member of the chain is strongly anchored in each of the preceding members; all members of the chain[6] are strongly anchored in the same additional items: *airlines, cockpit, twa*. We note that within-chain anchors tend to be somewhat stronger than out-of-chain ones; this might mean that coreference intensifies the perceived lexical connectedness. The anchoring support given to the referential structure makes it different from *Kallstrom–director* chain, in that the latter is an accidental referential connection, whereas the items in plane-chain have a lot of associative commonality. Anchoring also shows additional 'networking' of the chain in the given text, connecting it to other aviation-related things.

To summarize the analysis of LINK-TYPE entities with respect to our predictions: heads of coreferential definites that are not embedded within their coreferent NP confirm our expectation – they are anchored, and there is substantial anchoring texture accompanying the coreference links. Anchoring provides additional information, showing common-knowledge-based connection with items outside the coreference structure.

For heads of definites embedded in their coreferent NP, the anchoring support tends to be weak or lacking. These are cases with an *explicit syntactic construal* of co-referentiality, through an appositive. In the Gricean framework, this means that the reader is not likely to be able to work out the connection without the explicit help; *Kallstrom-director* is a good example. Hence, in terms of the involvement of the reader's knowledge, these cases are akin to D-definites – self-identifying lexically elaborate introductions of unfamiliar entities. In fact, in both such cases in PC-text, there was a minority D annotation. The anchoring behavior of these definites matches what we would expect of D-types, and is quite unlike that of other coreferential definites.

4.2 NO-LINK Definites

Table 4 shows the K/D definites from PC-text. Let us consider the 5 cases of majority K annotations. Apart from *Journal*, all other items are reliably anchored, and have at least one very strong anchor: *century→ centennial*, *office → business*, *drives→ computers*, *drives→ disk*, *telephone → modems*.

It is predicted that anchors should not be nominal heads in the text, so they are not perceived as discourse entities. *Centennial* is an adjective; *business* and *disk* are non-head noun modifiers. However, *computers* and *modems* are NP heads. It seems that reference and anchoring are at odds regarding the relatedness of the entity to the text.

The definite *the telephone* is embedded in 'the internal modems that allow PCs to share data via the telephone', headed by the anchor *modems*. This is reminiscent of the problem with embedded appositives – perhaps the annotators thought that an entity which is still being introduced could not be used as a basis for an R-type connection.

The annotators knew about *the disk drives for PCs*, but took them to be unrelated to previously discussed computers. Indeed, the three computers launched in 1977 did not have disk drives, as the text implies ('could store about two pages of text in their memories' carries a scalar implicature that they could not store more). From the modern perspective, these are rather non-typical computers. Perhaps, the reference annotators assumed the 1977 perspective from which disk drives were a new development, whereas

---

[6] including *flight*, which is not listed as it was not introduced inside a coreferential definite NP

Table 4. NO-LINK definites. The heads are boldfaced. For minority =/R annotations, the antecedent is in square brackets. Dots . . . indicate ≥3 additional anchors, each marked by 1 person.

	Definite NP	P&V class			Anchors
1	The Wall Street **Journal**	K	K	K	wall_street (5)
2	the past **century**	K	K	K	centennial (19) year (7) past (2)
3	the **face** of personal computing	K	D	D	personal (1) changed (1)
4	the **Apple** II	R [three computers]			computers (17) . . .
		K	D		
5	the home and **office**	K	K	K	business (11) computers (5) home (4) pcs (2) desktop (4) owners (2) . . .
6	the **Altair**	= [b.-f.-kit types]			pcs (1) types (1) apple_ii (1)
		K	D		commodore (1) tandy (1)
7	the **team** that developed the disk drives for PCs	D	D	D	led (2) chairman (1)
8	the disk **drives** for PCs	K	K	K	computers (10) disk (9) keyboards (2) pcs (2) data (2) technology (2) . . .
9	the internal **modems** that allow PCs to share data via the telephone	K/D	D	D	computers (19) pcs (10) keyboards (6) disk (4) technology (4) screens (4) drives (3) . . .
10	the **telephone**	K	K	K	modems (14) television (4) technology (3) computers (2) . . .

anchoring annotators took their own current perspective, judging disk drives *generally* related to computers. One difficulty for this explanation is the perspective of the text: it mentions *current*, i.e. 1989, computers, which, presumably, did have disk drives.

Turning to the 3 cases where D type predominates, we note that two of them – items 3 and 7 – bear out our prediction of lack of anchoring. Additionally, case 6, coming from 'earlier built-from-kit types such as the Altair, Sol and IMSAI' is a by now familiar case of a definite embedded inside a NP with overlapping reference; the reference annotation shows confusion, but the lack of anchoring places this case firmly in 'D'-type company.

Case 4 also shows uncertainty regarding reference. Apple II is one of the *three computers* mentioned in the preceding sentence. The anchoring pattern is extremely strong, suggesting that (some) computers are the entity to which Apple II is related. A minority vote indeed opts for R; the K/D decisions are puzzling.

However, perhaps the most surprising case is that of *modems* (case 9): A robust D type reference-wise, with an overwhelming anchoring connection to *computers*. The discussion of perspective taken by annotators is possibly relevant here: The three computer brands repeatedly mentioned in the text did not have modems. Still, surely the connection between modems and computers is of readers' knoweldge!

We suggest that the solution in the wording of instructions to reference annotators: ". . . the D[efinite]D[escription] is self-explanatory or it is given together with its own identification. In these cases it becomes clear to the general reader what is being talked about even without previous mention in the text or without previous common knowledge of it" [9]. Thus, D-types are not necessarily unfamiliar to the reader; rather, the reader does not have to use the familiarity in order to achieve unique identification. In

this case, the 12-word definite headed by *modems* is squarely within this category, even though the common knowledge could be in place, too. The mere existence of cases like this is surprising Grice-wise, as they seem to provide superfluous information that the reader could have recovered on the basis of her knowledge. Modems are introduced as an important invention of the past, as are disk drives; however, there is no elaboration about what disk drives are. Possibly, the 'modems' case is slightly over-indulging for an up-to-date 21st-century reader; perhaps modems were still a rarity in 1989.

We thus see that D-class is not homogenous with respect to the role of knowledge: it contains elements that are unfamiliar, so the reader has to use the material inside the definite to reach the unique identifiability, and elements that are in fact quite familiar in the reader's common knowledge, but their presentation in the text is such that the reader does not have to use her knowledge to interpret them.

In contrast, anchoring sides completely with the reader's knowledge: a putative connection is either supported by it or not, irrespective of what the text says about each entity. This is because anchoring asks for intuitive stereotypical judgment, for the shallow but robust load brought into the text by the mere use of a word, like *modems*. Anchoring is meant to uncover how such 'loads' organize into structures in the text, below the level of the discourse-entity-based who-did-what-to-whom stories where reference operates.

5 Conclusion

This paper reported a case study of the relationship between referential behavior of definite NPs and lexical anchoring of their heads, on the basis of juxtaposed relevant annotations of two texts.

We observed a tendency for definites whose referent repeats or relates to some previous textual entity (=/R) to be anchored in their antecedents, as well as in other things, providing additional text-based connections (plane vs. airlines, cockpit).

Even when an entity is judged to be referentially unrelated to the text, but of reader's knowledge (K-type), the anchoring pattern shows what could have triggered the relevant knowledge earlier in the text (*disk→computers*, *century→centennial*). Often, the anchor is not a nominal head, although we saw cases of heads as well.

When the entity is judged new in the text (D-type), we discerned two sub-types. In case the entity is genuinely unfamiliar, the lexical anchoring texture for the head is indeed meager. In case the entity is familiar, but in the current case could as well be uniquely identified by the current description, the anchoring pattern reveals the familiarity. Such discrepancies could be interesting from a historical perspective, as, assuming Gricean cooperative framework, they detect potential knowledge mismatches between text-creation-time and current audience.

In the other direction, lack of anchoring tends to corresponds to D-class, but also to cases where the definite is embedded, through an appositive or a list structure, inside an NP with overlapping reference. These cases often show disagreement in reference type classification, possibly reflecting confusion as to the availability of a referent-in-the-making as an antecedent. There is usually a minority D annotation in these cases.

Clearly, additional parallel annotation is needed to check out these trends. Such work is promising in exposing the intricate, multi-level workings of the reader's knowl-

edge upon the text: not only does it help to consolidate the plot of the story by tracing repeated and related referents, but also to prepare subtle, associative ground for introduction of new things and ideas, and to strengthen the perceived unity of the text by enriching the network of connections between its elements.

Acknowledgments

References

1. Beigman Klebanov, Beata: Using Readers to Identify Lexical Cohesive Structures in Texts. In Proceedings of ACL Student Session (2005) 55-60
2. Beigman Klebanov, Beata, and Shamir, E: Lexical Cohesion: Some Implications of an Empirical Study. In Proceedings of 2nd International Workshop on Natural Language Understanding and Cognitive Science, Miami, USA (2005) 13-21
3. Beigman Klebanov, Beata, and Shamir, E: Guidelines for Annotation of Concept Mention Patterns. Technical Report 2005-8, Leibniz Center for Research in Computer Science, The Hebrew University of Jerusalem, Israel (2005)
4. Ge, Niyu, Hale, John, and Charniak, Eugene: A Statistical Approach to Anaphora Resolution In Proceedings of the 6th Workshop on Very Large Corpora (1998) 161-170
5. Gundel, J.K., Hedberg, N., and Zacharski, R: Cognitive Status and The Form of Referring Expressions in Discourse. Language **69**(2) (1993) 274-307
6. Halliday, M.A.K., and Hasan, R.: Cohesion in English. Longman Group Ltd. (1976)
7. Hirschman, Lynette: MUC-7 Coreference Task Definition, version 3. In Proceedings of the 7th Message Understanding Conference (1997)
8. Lyons, Christopher: Definiteness. Cambridge, UK: Cambridge University Press (1999)
9. Poesio, Massimo, and Vieira, Renata: A Corpus-based Investigation of Definite Description Use. Computational Linguistics **24**(2) (1998) 183-216
10. Tetreault, Joel: Analysis of syntax-based pronoun resolution methods. ACL (1999) 602-605
11. Vieira, Renata: Definite Description Processing in Unrestricted Text. PhD Thesis, University of Edinburgh (1998)
12. Vieira, Renata, and Poesio, Massimo: An Empirically-based System for Processing Definite Descriptions. Computational Linguistics **26**(4) (2000) 539-593

Appendix A: PC text

Computers Start to Get Personal, 1977

1989

Wall Street Journal

(During its centennial year, The Wall Street Journal will report events of the past century that stand as milestones of American business history.)

Three computers that changed the face of personal computing were launched in 1977.

That year the Apple II, Commodore Pet and Tandy TRS-80 came to market. The computers were crude by today's standards. Apple II owners, for example, had to use their television sets as screens and stored data on audiocassettes. But Apple II was a major advance from Apple I, which was built in a garage by Stephen Wozniak and Steven Jobs for hobbyists such as the Homebrew Computer Club. In addition, the Apple II was an affordable $1,298.

Crude as they were, these early PCs triggered explosive product development in desktop models for the home and office.

Big mainframe computers for business had been around for years. But the new 1977 PCs – unlike earlier built-from-kit types such as the Altair, Sol and IMSAI – had keyboards and could store about two pages of data in their memories. Current PCs are more than 50 times faster and have memory capacity 500 times greater than their 1977 counterparts.

There were many pioneer PC contributors. William Gates and Paul Allen in 1975 developed an early language-housekeeper system for PCs, and Gates became an industry billionaire six years after IBM adapted one of these versions in 1981. Alan F. Shugart, currently chairman of Seagate Technology, led the team that developed the disk drives for PCs. Dennis Hayes and Dale Heatherington, two Atlanta engineers, were co-developers of the internal modems that allow PCs to share data via the telephone.

IBM, the world leader in computers, didn't offer its first PC until August 1981 as many other companies entered the market. Today, PC shipments annually total some $38.3 billion world-wide.

Appendix B: TWA text

Investigators: TWA explosion didn't originate near cockpit
1996
New York Times News Service

After picking apart some of the wadded remains of the cockpit of Trans World Airlines Flight 800, investigators concluded Tuesday that the catastrophic explosion that destroyed the Boeing 747 most likely did not originate inside the cockpit or in the electronics bay beneath it.

They were partly persuaded by a surprising discovery found in the ton of wreckage that had been the jet's cockpit: The circles of glass that cover many of the cockpit dials, and even a light bulb above a staircase that led to the plane's upper deck, had somehow survived the crash intact.

"You have this mass of wreckage and yet things from that area are relatively the way they were before the accident," said Robert Francis, vice chairman of the National Transportation Safety Board. "There is no indication at this point of anything in that area that would give cause for concern in terms of something having initiated there."

A senior investigator who looked at the cockpit wreckage Tuesday said that one of the plane's altimeters — instruments that show the plane's altitude — was frozen with a reading of 13,100 feet. Altimeters are mechanically driven instruments that do not depend on electricity to work, so the finding suggests that the mechanics continued working for several seconds after the initial explosion, at about 13,700 feet.

Federal investigators continued their search for the cause of the crash Tuesday, the 20th day since Flight 800 exploded in midair off Long Island and plunged into the Atlantic Ocean, killing all 230 people on board. On the seas and on the shore, investigators said they made a modest amount of progress, though they still have not determined if the plane crash was caused by a bomb, a missile attack or a mechanical malfunction.

At the former Grumman hangar in Calverton, investigators on Tuesday began piecing together the fractured parts of the airplane. They also pulled about one-third of the cockpit wreckage off the one-ton ball of metal, essentially unwrapping it.

James Kallstrom, the assistant director of the FBI's New York office, said he had sent many agents who had been working in Suffolk County back to their home offices, mostly in New York City. Criminal investigators are anxious for the cause to be determined, he said, adding: "We are in a bit of a waiting pattern."