

Discourse coherence and pronoun resolution

Florian Wolf and Edward Gibson

Massachusetts Institute of Technology, Cambridge, MA, USA

Timothy Desmet

Ghent University, Belgium

This paper used self-paced reading to test processing preferences in pronoun interpretation in English two clause sentences. The results demonstrate that people's preferences can be reversed by changing the coherence relation between the clauses. The results are not compatible with the existence of a single all-purpose strategy in pronoun resolution. Rather, the results support Kehler's (2002) hypothesis that the processing patterns observed in pronoun processing are a byproduct of more general cognitive inference processes underlying the establishment of coherence, such that discourse coherence guides pronoun reference, and pronoun reference guides discourse coherence.

INTRODUCTION

An important component of language comprehension in most natural language contexts involves connecting clauses and phrases together in order to establish a coherent discourse. One critical way in which coherence can be established between clauses is by the use of referring expressions, such as pronouns (Garnham, 2001; Halliday & Hassan, 1976; Johnson-Laird, 1983; Kintsch & Van Dyke, 1978; Sanford & Garrod, 1989). Thus an important part of discourse comprehension involves discovering how antecedents for pronouns are resolved. One well-known account of discourse processing with implications on pronominal resolution is Centering Theory, which predicts that pronouns prefer to have

Correspondence should be addressed to Florian Wolf, MIT NE20-448, 3 Cambridge Center, Cambridge, MA 02139, USA. Email: fwolf@mit.edu

Timothy Desmet is a Research Assistant of the Fund for Scientific Research – Flanders (Belgium) (F.W.O. Vlaanderen).

antecedents in subject position (Brennan, Friedman, & Pollard, 1987; Grosz, Joshi, & Weinstein, 1995; cf. also Wundt, 1911). In support of Centering Theory, Gordon, Grosz and Gilliom (1993) found that there is a preference to use pronouns to refer to entities in subject position, but not for entities in object position. Consider the sentences in (1):¹

- (1) a. Fiona complimented Craig, and she congratulated James.
 b. Fiona complimented Craig, and he congratulated James.

Intuitively, (1a) is easier to process than (1b). Centering explains this preference because the pronoun 'she' in (1a) refers back to the subject of the preceding clause, whereas the pronoun 'he' in (1b) refers back to the object of the preceding clause. A problem for Centering Theory is provided by the contrast in (2):

- (2) a. Fiona complimented Craig, and James congratulated her.
 b. Fiona complimented Craig, and James congratulated him.

Contrary to Centering Theory's subject preference prediction, Chambers and Smyth (1998) found in a self-paced reading experiment that sentences like (2b) were read faster than sentences like (2a). This pattern of results motivates the Parallel Preference account (Chambers & Smyth, 1998; see Lappin & Leass, 1994, for a combination of Centering Theory and Parallel Preference). Under the parallel preference account (Smyth, 1994), pronouns are argued to prefer antecedents in a parallel position when the pronoun- and the antecedent-containing sentence have the following properties: (a) both sentences have the same global constituent structure, (b) the thematic roles of the verbs in both sentences concur. When these conditions are met, subject pronouns should prefer subject antecedents, and object pronouns should prefer object antecedents. This is the case in (1) and (2) above. In (1), people prefer the preceding clause's subject as the referent for the subject pronoun, whereas in (2) people prefer the preceding clause's object as the referent for the object pronoun.

Although a parallel preference account can explain the preferences in (1) and (2), it does not explain the preferences in (3), from Winograd (1972):

- (3) a. The city council denied the demonstrators the permit because they advocated violence.
 b. The city council denied the demonstrators the permit because they feared violence.

¹ This paper follows the processing literature in focusing on the interpretation of unstressed pronouns. See Akmajian and Jackendoff (1970) among others for a discussion of the interpretation of stressed pronouns.

In sentence (3a) the pronoun ‘they’ refers to ‘the demonstrators’, whereas in sentence (3b) it refers to ‘the city council’. Neither sentence seems particularly difficult to process. Notice, however, that both Centering Theory and Parallel Preference predict a preference for ‘they’ to refer to the subject, ‘the city council’ – Centering Theory because it predicts a preference for subject antecedents, and Parallel Preference because it predicts a preference for an antecedent in a parallel position. Examples like (3) motivate causal-inference-based accounts of pronoun processing (Hobbs, 1979; Hobbs, Stickel, Appelt & Martin, 1993; Kehler, 2002). According to such accounts, ‘they’ refers to ‘the demonstrators’ in sentence (3a) because advocating violence is assumed to be a good reason for being denied a permit. In sentence (3b) ‘they’ refers to ‘the city council’ because fearing violence by demonstrators is a good reason for denying a permit to these demonstrators.

Experimental evidence relevant to causal-inference-based accounts of pronominal resolution is provided by Ehrlich (1980), who used an off-line questionnaire to investigate people’s preferred pronoun resolution. Ehrlich found that pronoun resolution is only driven by causal inferences (cf. Caramazza, Grober, Garvey, & Yates, 1977; Stewart, Pickering & Sanford, 2000) when the clauses containing pronoun and antecedent respectively are in a causal relation. When there is no such causal relation, Ehrlich found that people prefer antecedents in topic or subject position (cf. Centering Theory, Grosz et al., 1995).

Although causal-based strategies can explain the effects in (3), they do not explain the patterns in (1) and (2), because there is no causal connection between the two clauses in each of these sentences. Furthermore, resorting to a topic-based strategy like Centering Theory as suggested by Ehrlich makes the right prediction for (1), but not for (2), where the pronoun with an object antecedent is easier to process.

Based on Hobbs (1979), Kehler (2002) provides a hypothesis that aims to explain all of these patterns of pronoun resolution. Instead of arguing for pronoun-specific processing mechanisms, Kehler (2002), like Hobbs (1979), argues that pronoun resolution is a byproduct of establishing coherence. Kehler (2002) extends Hobbs’ (1979) key insight that the establishment of coherence guides pronoun resolution and vice versa, noting that discourse coherence and pronoun resolution mutually constrain each other: Pronoun resolution guides coherence, but coherence also guides pronoun resolution. Thus he hypothesises that how a pronoun is resolved may depend on the coherence relation between the clauses.

Two classes of coherence relations that are particularly relevant to the examples that have been discussed in the pronoun resolution literature are *cause-effect* and *resemblance*. A cause-effect relation holds between two clauses if a plausible causal relation can be inferred to hold between the

events described by the two clauses. The event described by one clause is the cause for the event described by the other clause, as in (3a). Because the demonstrators advocated violence, the city council denied them a permit to demonstrate. Kehler (2002) argues that the pronoun is interpreted such that a plausible cause-effect relation between the two clauses can be established. Pairing ‘they’ with ‘the demonstrators’ provides a more plausible interpretation for (3a) than pairing ‘they’ with ‘the city council’. A similar analysis applies to the pronoun resolution of ‘they’ in (3b).

The resemblance discourse relation is relevant to explaining the pattern of preferences in (1) and (2). A resemblance relation holds between two clauses if the events described by the two clauses are in a similarity or in a contrast relation, as in the following examples from Kehler (2002):

- (4) a. Resemblance, Similarity
 Gephardt organized rallies for Gore, and Daschle distributed pamphlets for him.
- b. Resemblance, Contrast
 Gephardt supported Gore, but Armev opposed him.

Kehler (2002) hypothesizes that the first step in establishing a resemblance relation between clauses is to find parallel corresponding entities and events. Then these entities and events are put into similarity or contrast relations. For example, in sentence (4a), ‘organised rallies’ is parallel and similar to ‘distributed pamphlets’ (both predicates describe actions of supporting a political candidate), and ‘Dick Gephardt’ is parallel and similar to ‘Tom Daschle’ (both are American politicians that are similar in that they support Al Gore). Then, Kehler (2002) argues, the pronoun ‘him’ is paired with its parallel preceding element, ‘Gore’. In sentence (4b), ‘supported’ is parallel and in contrast to ‘opposed’. ‘Gephardt’ is parallel and in contrast to ‘Armev’ (both are politicians that are in contrast in that one of them supports Gore and the other one opposes him). Then, as in sentence (4a), the pronoun ‘him’ is paired with its parallel preceding element, ‘Gore’. Thus, in both sentences (4a) and (4b), the pronoun is bound to its antecedent during the establishment of a resemblance coherence relation, when parallel entities are matched.

The resemblance relation is the most plausible coherence relation between each of the clauses in the sentences in (1) and (2). In particular, the use of the similar verbs ‘complimented’ and ‘congratulated’ in the absence of any other cues induces a resemblance-similarity relation between each pair of clauses. Kehler’s theory then predicts that a parallel preference strategy would be in effect under the resemblance relation, which has been observed experimentally in such sentences (Chambers & Smyth, 1998).

A strong prediction of Kehler's theory is that pronoun resolution preferences can be altered depending on the coherence relation between clauses. The experiment presented here tests this prediction directly.

EXPERIMENT

Method

Participants. Forty participants from the MIT community were paid for their participation. All were native speakers of English and were naive as to the purpose of the study.

Materials. Twenty sets of sentences were constructed, each with four conditions in a 2×2 design: coherence relation (resemblance, cause-effect) \times parallel reference (parallel, nonparallel). An example item is presented in (5):

- (5) a. Resemblance, Parallel Reference
Fiona complimented Craig and similarly James congratulated him after the match but nobody took any notice.
- b. Resemblance, Nonparallel Reference
Fiona complimented Craig and similarly James congratulated her after the match but nobody took any notice.
- c. Cause-Effect, Parallel Reference
Fiona defeated Craig and so James congratulated him after the match but nobody took any notice.
- d. Cause-Effect, Nonparallel Reference
Fiona defeated Craig and so James congratulated her after the match but nobody took any notice.

Each sentence consisted of three clauses. The second clause was the target clause which consisted of the same words across the coherence manipulation. We manipulated the coherence relation between resemblance and cause-effect by making two changes to the items: (1) by using different connectives between the clauses ('and similarly' vs. 'and so'), (2) by using a different verb in the first clause. For resemblance, the verbs in the two clauses were semantically similar according to the WordNet lexical database (Fellbaum, 2001), e.g., 'compliment' and 'congratulate' in (5). For the cause-effect conditions, the verb of the first clause in the cause-effect condition was chosen so that there was a plausible causal relation between the two clauses such that the object pronoun referred to the subject of the first clause, e.g., 'defeat' and 'congratulate' in (5). The first clause verb in the cause-effect conditions always differed from the first clause verb in the resemblance conditions. The remainder of the sentences consisted of a prepositional phrase and a third clause. This portion of the

items was the same across the four conditions. Overall, the only differences between the resemblance and cause-effect conditions were the verbs of the first clause and the connectives relating the two clauses.

Notice that this experiment did not explore the relative contribution of different coherence cues to changing pronoun interpretation preferences. This does not diminish the point of the current design, which is simply to show that changing the coherence relation by using one or more cues may alter pronoun interpretation preferences.

The target sentences were combined with 76 fillers of various types in four lists balancing all factors in a Latin Squares design. Appendix A provides a complete list of the stimuli. The stimuli were pseudo-randomised separately for each participant, so that at least one filler item intervened between two targets.

Procedure. The task was self-paced word-by-word reading with a moving window display (Just, Carpenter, & Woolley, 1982) using Linux computers running software developed in our lab. Each trial began with a series of dashes marking the length and position of the words in the sentences, printed approximately a third of the way down the screen. Participants pressed the spacebar to reveal each word of the sentence. As each new word appeared, the preceding word disappeared. The amount of time the participant spent reading each word was recorded as the time between key-presses. After the final word of each item, a question appeared which asked about information contained in the sentence (e.g. 'Did James congratulate Fiona?'). Participants pressed one of two keys to respond 'yes' or 'no'. After an incorrect answer, the word 'INCORRECT' flashed briefly on the screen. No feedback was given for correct responses. Participants were asked to read sentences at a natural rate and to be sure that they understood what they read. They were told to answer the questions as quickly and accurately as they could and to take wrong answers as an indication to read more carefully.

Before the main experiment, a short list of practice items and questions was presented in order to familiarise the participant with the task. A session averaged 25 minutes.

PREDICTIONS

The predictions are made in terms of reading times on the pronoun plus the next word, because in self-paced reading, effects often spill over to the next word (Sanford & Garrod, 1989). Faster reading times are assumed to reflect easier processing of the pronoun.

Centering Theory predicts that pronouns referring to antecedents in subject position should always be read faster. Thus, the pronouns in

sentences (5b) and (5d) should be read faster than those in sentences (5a) and (5c).

Parallel Preference makes the opposite prediction for sentences (5a) and (5b). Because the pronouns in the experimental items are in object position, Parallel Preference predicts that pronouns referring to antecedents in (parallel) object position should be read faster. Thus, the pronouns in sentence (5a) should be read faster than those in sentence (5b). Parallel Preference does not apply to sentences (5c) and (5d), because these sentences do not meet Smyth's (1994) criteria for parallelism.

Causal-inference-based accounts do not apply to sentences (5a) and (5b), because the events described by the clauses are not causally related. Causal-inference accounts predict that the pronoun in sentence (5d) should be read faster than in (5c), because a causal inference to resolve the pronoun is much easier to establish in (5d); in (5c), it is hard to see why James should congratulate Craig, because he lost the match. Ehrlich's (1980) proposal that a topic-based strategy applies when there is no causal relation predicts that the pronoun in (5b) should be read faster than the pronoun in (5a).

Kehler's (2002) coherence-based theory predicts that the cues in sentences (5a) and (5b) will indicate a resemblance relation between the clauses, so that a parallel preference strategy will be in effect. Thus the pronoun in sentence (5a) should be read faster than the one in sentence (5b). Kehler's account furthermore predicts that the cues in sentences (5c) and (5d) will indicate a causal relation between the clauses, with the consequence that the pronoun in sentence (5d) should be read faster than the pronoun in sentence (5c) because of the more plausible causal inference for sentence (5d). Thus Kehler's account predicts an interaction between the coherence relations and the pronominal reference.

RESULTS

Table 1 shows the question answering performance for the experiment. A 2×2 ANOVA, coherence relation (resemblance, cause-effect) by reference (parallel, nonparallel), revealed an interaction by subject, $F_1(1, 39) = 8.150$, $MSe = 1210$, $p < .01$; $F_2(1, 19) = 3.385$, $MSe = 605$, $p = .08$. Pairwise comparisons by subject showed that under resemblance, question answering performance was better under parallel than under nonparallel reference, $F_1(1, 39) = 5.354$, $MSe = 845$, $p < .05$. There was no significant difference under Cause-Effect, $F_1(1, 39) = 2.395$, $MSe = 405$, $p = .13$.

Only reading times for items for which the comprehension question was answered correctly were analysed. Reading times beyond 3 SD from the mean for a given condition and position were excluded from the analysis.

TABLE 1
Question answering performance in per cent correct

Pronoun reference	Coherence relation	
	Resemblance	Cause-effect
Parallel	86	80.5
Nonparallel	79.5	85

This affected 2.79% of the data. Mean word-by-word reading times by subject are shown in Figure 1.

A 2×2 ANOVA, coherence relation by reference, was computed for the region including the pronoun and the following word. It showed a significant interaction of coherence relation and reference, $F_1(1, 39) = 14.669$, $MSe = 103997$, $p < .001$; $F_2(1, 19) = 13.398$, $MSe = 67545$, $p < .005$. There was also a main effect of coherence-relation cause-effect items were read faster than resemblance items, $F_1(1, 39) = 4.431$, $MSe = 40563$, $p < .05$; $F_2(1, 19) = 3.898$, $MSe = 22222$, $p = .06$. For the region containing the pronoun and the region before that region, there was a significant three-way-interaction of coherence relation, reference, and region, $F_1(1, 39) = 12.111$, $MSe = 64630$, $p < .005$; $F_2(1, 19) = 20.126$, $MSe = 44344$, $p < .0005$. There were no other significant effects.

Pairwise comparisons showed that under resemblance, parallel was read faster than nonparallel, $F_1(1, 39) = 7.849$, $MSe = 60866$, $p < .01$; $F_2(1, 19) =$

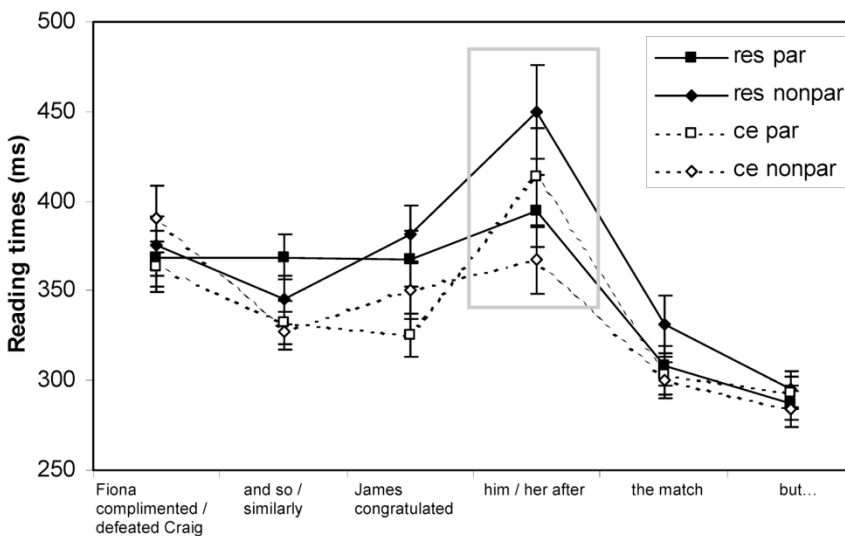


Figure 1. Mean word-by-word reading times by subject.

5.785, $MSe = 40196$, $p < .05$. Under cause-effect, nonparallel was read faster than parallel, $F_1(1, 39) = 4.822$, $MSe = 43829$, $p < .05$; $F_2(1, 19) = 5.785$, $MSe = 27907$, $p < .05$.

DISCUSSION

The results of the experiment showed that under a resemblance discourse relation, pronouns with an antecedent in parallel object position were read faster than pronouns with an antecedent in subject position. This is predicted by the Parallel Preference account as well as by Kehler's (2002) account. By contrast, Centering Theory and Ehrlich's (1980) account would have predicted a subject antecedent preference. Causal-inference-based accounts make no prediction for pronoun preferences in the absence of causal relations between the clauses containing pronoun and antecedent.

Under the cause-effect discourse relations in our items, pronouns referring to a subject antecedent were read faster. This is predicted by causal-inference-based accounts as well as Kehler's (2002) account, but not predicted by the Parallel Preference account. Centering Theory does predict this preference, but not as a part of a causal inference process.

To summarise, the only account that makes the correct predictions for all conditions is Kehler's (2002). It predicts different preferences in pronoun resolution, depending on the coherence relation between the clauses containing the pronoun and the antecedent.

CONCLUSIONS

The results from the experiment reported here support the idea that the preferences observed in pronoun processing depend on the coherence relation between the clause containing the pronoun and the clause containing the antecedent (Kehler, 2002). However, this is not to say that other factors such as focusing attention on specific discourse elements (cf. Grosz et al., 1995; Wundt, 1911) play no role in pronoun processing. In fact, Kehler (2002) points out that in narratives, shifting attention to different discourse entities is an important factor in pronoun processing preferences. He argues that under such circumstances the observed preferences may be more like predicted by accounts such as Centering Theory. Notice, however, that of the accounts considered here, Kehler's (2002) is the only one that predicts all observed preferences not as a result of the operations of pronoun-specific mechanisms, but as a byproduct of more general cognitive mechanisms and their interaction-establishing coherence and focusing attention.

Manuscript received October 2003

Revised manuscript received February 2004

REFERENCES

- Akmajian, A., & Jackendoff, R. (1970). Coreferentiality and stress. *Linguistic Inquiry*, 1, 124–126.
- Brennan, S. E., Friedman, M. W., & Pollard, C. J. (1987). A Centering approach to pronouns. In *Proceedings of the 25th Meeting of the Association for Computational Linguistics*, pp. 155–162. Stanford, CA.
- Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, 16, 601–609.
- Chambers, G. C., & Smyth, R. (1998). Structural parallelism and discourse coherence: A test of Centering Theory. *Journal of Memory and Language*, 39, 593–608.
- Ehrlich, K. (1980). Comprehension of pronouns. *Quarterly Journal of Experimental Psychology*, 32, 247–255.
- Fellbaum, C. (Ed.). (2001). *WordNet – An electronic lexical database*. Cambridge, MA: MIT Press.
- Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Hove, UK: Psychology Press.
- Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17, 311–347.
- Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2), 203–225.
- Halliday, M. A. K., & Hassan, R. (1976). *Cohesion in English*. London: Longman.
- Hobbs, J. R. (1979). Coherence and coreference. *Cognitive Science*, 3, 67–90.
- Hobbs, J. R., Stickel, M. E., Appelt, D. E., & Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence*, 63, 69–142.
- Johnson-Laird, P. (1983). *Mental models*. Cambridge: Cambridge University Press.
- Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2), 228–238.
- Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
- Kintsch, W., & van Dyke, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363–394.
- Lappin, S., & Leass, H. (1994). An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20(4), 535–561.
- Sanford, A. J., & Garrod, S. C. (1989). What, when, and how? Questions of immediacy in anaphoric reference resolution. *Language and Cognitive Processes*, 4(3/4), 235–262.
- Smyth, R. H. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, 23, 197–229.
- Stewart, A. J., Pickering, M. J., & Sanford, A. J. (2000). The role of implicit causality in language comprehension: Focus versus integration accounts. *Journal of Memory and Language*, 42, 423–443.
- Winograd, T. (1972). *Understanding natural language*. New York: Academic Press.
- Wundt, W. (1911). *Völkerpsychologie*. Leipzig: Engelmann Verlag.

APPENDIX A: ITEMS FOR THE EXPERIMENT

1. Charles commended/saved Harriet and similarly/so Richard praised her/him in the newspaper but everything was just a big show.
2. David reprimanded/betrayed Sarah and similarly/so Helen chastised her/him after the holidays but all the criticism showed very little effect.
3. Michael disciplined/attacked Shirley and similarly/so Leonard punished her/him two days ago but in the end they reached an agreement.
4. Peter questioned/assaulted Julie and similarly/so Carol interrogated her/him for an hour but a few moments later the police arrived at the scene.
5. Stuart honored/liberated Martha and similarly/so Joseph admired her/him a great deal but unfortunately the feeling was not mutual.
6. Nathan disliked/abandoned Allyssa and similarly/so Nicole hated her/him for a while and in the end they all avoided each other.
7. Ryan safeguarded/feared Emma and similarly/so Adam protected her/him in the evening but all their caution would probably not have been necessary.
8. Kevin rebuked/kicked Claire and similarly/so Grace scolded her/him in the house but nobody else cared about all these quarrels.
9. Erik embraced/rescued Lisa and similarly/so Liam hugged her/him with great enthusiasm and everybody was a little bit relieved.
10. Brian scolded/harassed Cathy and similarly/so Scott lectured her/him after the meeting and everybody ended up hating each other.
11. Fiona complimented/defeated Craig and similarly/so James congratulated him/her after the match but nobody took any notice.
12. Christina lectured/pestered Christopher and similarly/so Stephanie reprimanded him/her for one hour although nobody thought it would have any effect.
13. Jonathan despised/denounced Madeline and similarly/so Patricia scorned her using harsh language but after a while everybody was reconciled again.
14. Rebecca interrogated/punched Anthony and similarly/so Suzanne cross-examined him/her for a while but nothing interesting was said.
15. Melissa suspected/deceived William and similarly/so Natalie distrusted him/her in the end and the whole working atmosphere was spoiled.
16. Tina thanked/supported Robert and similarly/so Fred acknowledged him/her at the conference but nobody seemed to be sincere.
17. Sophia admired/outdid Joshua and similarly/so Gloria respected him/her in the beginning but very soon things changed.
18. Melanie hired/impressed Bradley and similarly/so Malcolm recruited him/her after the interview but not all of the co-workers were satisfied with the situation.
19. Heather hit/insulted Aaron and similarly/so Caitlin punched him/her in the nose and the result was a big fight.
20. Hannah appointed/outperformed Michael and similarly/so George nominated him/her for the job although some people were not happy with the decision.