Topics and centers
A comparison of the salience-based approach and the Centering theory[1,2]

Ivana Kruijff-Korbayová and Eva Hajičová

Abstract

We survey and compare the centering theory and the salience-based approach as two approaches that attempt to account for the relations between the changes of attention and the choices of referring expressions in discourse. We discuss the similarities and differences between the notions of backward- and forward-looking centers and the transitions from one backward-looking center to another in centering theory, and the notions of reference to an entity in the topic or in the focus and the changes in degrees of activation of the entities referred to in the salience-based approach.

1. Introduction

The utterances that constitute a coherent discourse are interrelated. One of the ways in which a discourse can hold together is that expressions contained in different utterances refer to the same or related entities. Usually, attention is concentrated on some entities and relations among them, and then moves on and on to others, sometimes returning back again, as the discourse unfolds. It is a task of discourse processing to determine what entities are being referred to at any given point, i.e., to track the changes of attention. In discourse production, one needs to plan what and when to refer to. In discourse understanding, one needs to resolve references.

The changes of attention throughout a discourse do not occur at random; rather, they comply to some general principles of communicative behaviour which are probably independent of a particular language, and they also follow the regulations of the particular language system in which the discourse is carried out, which determine functions that particular kinds of expressions fulfil. There may be similarities across languages as far as the functions and therefore the uses of certain kinds of expressions are concerned.

In the present paper, we survey and compare two approaches that attempt to account for the relations between the changes of attention and the choices of referring expressions in discourse. The two approaches considered are the centering theory and the salience-based approach.

In centering theory an utterance has its backward-looking center (of attention) which links it to the preceding utterance, and a partially ordered list of forward-looking centers. The ranking of the forward-looking centers affects the choice of the backward-looking center of the subsequent utterance. Centering theory expresses constraints on the realization of centers by linguistic expressions and on the transitions

[1]This paper is a substantially enlarged and modified version of the contribution presented at the International Conference at the Occasion of the 100th Anniversary of the Birth of Roman Jakobson and the 70th Anniversary of Prague Linguistic Circle, held in Prague, March 1996.

[2]The work was partially supported by Grant GAUK161/95 and on work sponsored by the U.S.-Czechoslovak Science Technology Joint Fund in cooperation with the U.S. National Foundation and the Czech Ministery of Education under Project Number 920-58. The research work of Ivana Kruijff-Korbayová has been subsidized within the Research Support Scheme of the Open Society Institute, grant No. 66/94.

from one backward-looking center to another across utterances within a discourse segment. The extent to which a discourse complies to the centering constraints is used as a measure of its coherence.

In the salience-based approach an utterance is seen as articulated into a topic part and a focus part. The principles based on the changes of salience in discourse relate reference to entities in either topic or focus to the degrees of activation (salience) of these entities in the stock of shared knowledge, i.e., a domain assumed by the speaker to be shared by the hearer. The degree of activation of an entity referred to in a discourse changes as the discourse unfolds, depending on whether the entity is referred to subsequently, and if so, whether it is referred to in the topic or in the focus. The salience-based approach expresses constraints on forms of referring expressions depending on the degrees of activation of their referents. Judgements about coherence of a discourse take into account whether or not the constraints based on salience are obeyed.

Considering, on the one hand, the notions of backward- and forward-looking centers of utterances and the transitions from one backward-looking center to another across utterances, and, on the other hand, the notions of reference to an entity in the topic or in the focus of an utterance and the changes in degrees of activation of the entities referred to utterance after utterance, one may start asking whether the centering theory and the salience-based approach have something in common. And if so, whether they are more or less terminological variants, or whether there are differences the analysis of which could yield fruitful extensions to either approach. The present paper is aimed to shed some light on this issue. We attempt to “align” the two approaches and to compare their assumptions, methods, claims and judgements.

Many intriguing questions arise from looking at the centering theory and the salience-based approach in the context of each other. As we cannot deal with them all, let alone provide a final answer to all of them, we hope at least to provide an interesting probe.

The rest of this paper is organized as follows. Section 2 provides an overview of the centering theory. Most of the text defining centering theory in Sections 2.1, 2.2, 2.3 is adopted from [Grosz et al. 1995] in order to avoid misformulations. Section 2.4 comprises our remarks on the centering theory rules and claims. Section 3 provides an overview of the salience-based approach based mainly on [Hajičová 1993]. Section 4 contains the core of the comparison of the two approaches. We discuss the two frameworks in detail, pointing out the similarities and differences between them. Section 5 rounds off the discussion with some more general remarks and a few suggestions for further research.

2. Centering Theory

2.1. Goals and method

Centering theory [Grosz et al. 1995][3] relates focus of attention, choice of referring expressions and the coherence of utterances within a discourse segment. It fits within a larger effort to provide an overall theory of discourse structure and meaning, as developed in [Grosz & Sidner 1986]. According to this theory, discourse structure has three components: linguistic structure, intentional structure and attentional state. At the level of linguistic structure, discourses consist of discourse segments, among which an embedding relation may hold. The intentional structure comprises intentions providing the basic rationale for the discourse, and relations among them. Attentional state models the discourse

[3]The name of the theory has originated in the works of Joshi & Kuhn [1979] and Joshi & Weinstein [1981] who introduced the centering terminology "to avoid confusion with previous uses of the term 'focus' in linguistics" [Grosz et al. 1995, p. 205], replacing thus Grosz' [1977] "focusing" terminology. For a discussion on the differences between the use of the term 'focus' in linguistics and in the domain of artificial intelligence, see Hajičová [1987].

participants' focus of attention at any given point in the discourse. Changes in attentional state depend on the intentional structure and on properties of the utterances in the linguistic structure.

Centering theory understands discourses as more than mere sequences of utterances: they must exhibit coherence. Moreover, some discourses appear more coherent than others.

According to centering theory, each discourse segment exhibits both local coherence (among the utterances in that segment) and global coherence (coherence with other segments in the discourse). Corresponding to these two levels of coherence are two components of the attentional state; the local level models changes in attentional state within a discourse segment, and the global level models attentional state properties at the inter-segmental level.

Centering is proposed as a model of the local-level component of attentional state[4]. Differences in coherence are argued to correspond in part to the different demands for inference made by different types of referring expressions, given a particular attentional state. Centering theory attempts to account for these differences by modelling the properties of the attentional state.

Centers are specified as those entities that link a given utterance to other utterances in the discourse segment that contains it. Centers are semantic objects rather than words, phrases or syntactic forms. They are related to linguistic expressions by the realize (directly realize) relation between an utterance of some phrase and a center (directly) realized in it. Centers are discourse constructs: it is an utterance and not a sentence in isolation that has centers; the same sentence when uttered in different discourse situations may have different centers.

The most fundamental claim of centering theory is that to the extent a discourse adheres to the centering constraints, its coherence will increase and the inference load placed upon the hearer will decrease.

Centering theory does not explicitly presuppose any particular theory of syntax. As long as expressions to which centers can be pertinent as the corresponding semantic objects identified by reference, can be discerned, centering theory can be applied. For a semantic theory to support centering it must provide an adequate basis for computing the relation of realization, which relates centers to linguistic expressions.

Centering theory is computational. The centering constraints are defined so that for the basic issues, the formulation of an algorithm that evaluates them is straightforward.

2.2. Rules and claims

The core of centering theory may be briefly reproduced as follows: each utterance U_n ($n \geq 1$) in a discourse segment is assigned a set of forward-looking centers $C_f(U_n)$; each utterance other than the segment initial utterance is assigned a single backward-looking center $C_b(U_n)$. The backward-looking center of U_{n+1}, $C_b(U_{n+1})$ connects with one of the forward-looking centers of U_n, $C_f(U_n)$, and it itself belongs to the set of forward-looking centers of U_{n+1}, $C_f(U_{n+1})$. The set of forward looking centers of an utterance contains the entities realized by the expressions that constitute that utterance. They are partially ordered to reflect the relative prominence of the elements in that utterance. Their ranking depends on grammatical roles, with SUBJECT>OBJECT>OTHER (having in mind, however, that other factors are also important, such as word order, clausal subordination, and lexical semantics). Lower-ranked elements of the C_f cannot be pronominalized unless higher-ranked ones are.

The basic constraint on center realization is given by Rule 1 of centering theory:

[4]See [Groz 1977] for the difference between the two levels of "focusing" in discourse, the global and the immediate one, and [Sidner 1979] for a detailed analysis of immediate focusing, with algorithms for tracking immediate focus.

Centering theory: Rule 1

If any element of $C_f(U_n)$ is realized by a pronoun in U_{n+1}, then the $C_b(U_{n+1})$ must be realized by a pronoun also.[5]

This constraint stipulates that no element in an utterance can be realized as a pronoun unless the backward-looking center is, too. It does not preclude using pronouns for other entities so long as the C_b is realized with a pronoun. Centering theory refers to psychological and cross-linguistic research supporting the position that the C_b is preferentially realized by a pronoun in English and by equivalent forms (e.g., zero pronouns) in other languages, that there is a single C_b in an utterance and that grammatical role plays a determining role in identifying the C_b.

The connection between the backward-looking center of utterance U_{n+1} and the forward-looking centers of utterance U_n may be of several types:

1. CENTER CONTINUATION: $C_b(U_{n+1}) = C_b(U_n)$, and this entity is the most highly ranked element of $C_f(U_{n+1})$. In this case, $C_b(U_{n+1})$ is the most likely candidate for $C_b(U_{n+2})$; it continues to be C_b in U_{n+1}, and continues to be likely to be C_b in U_{n+2}.
2. CENTER RETAINING: $C_b(U_{n+1}) = C_b(U_n)$, but this entity is not the most highly ranked element of $C_f(U_{n+1})$. In this case, $C_b(U_{n+1})$ is not the most likely candidate for $C_b(U_{n+2})$; although it is retained as C_b in U_{n+1}, it is not likely to be C_b in U_{n+2}.
3. CENTER SHIFTING: $C_b(U_{n+1}) \neq C_b(U_n)$.

The coherence of a discourse segment is affected by the kinds of centering transitions engendered by a speaker's choices of linguistic realizations in the utterances constituting that segment.

The basic constraint on center movement is given by Rule 2 of centering theory:

Centering theory: Rule 2

Sequences of continuation are preferred over sequences of retaining; and sequences of retaining are preferred over sequences of shifting.

In particular, a pair of continuations, represented as $Cont(U_n,U_{n+1})$ and $Cont(U_{n+1},U_{n+2})$, is preferred over a pair of retentions, $Ret(U_n,U_{n+1})$ and $Ret(U_n+1,U_{n+2})$; and a pair of retentions, $Ret(U_n,U_{n+1})$ and $Ret(U_{n+1},U_{n+2})$, is preferred over a pair of shifting, $Shift(U_n,U_{n+1})$ and $Shift(U_{n+1},U_{n+2})$.

This rule reflects the intuition that continuation of the center and the use of retentions when possible to produce smooth transitions to a new center provides basis for local coherence. In a locally coherent discourse segment, shifts are followed by a sequence of continuations.

Rule 2 provides a constraint on speakers, and on discourse generation systems. It does not have the same direct implementation for interpretation systems: rather, it predicts that certain sequences produce a higher inference load than others. However, there are as yet no validating psycho-linguistic results referred to by centering theory.

Centering theory makes the following claims:

(i) **A unique** C_b: Each utterance (other than the segment initial one) has exactly one backward-looking center.

(ii) **Ranking of** C_f: The C_f elements are partially ordered according to a number of factors. Ranking of elements in $C_f(U_n)$ guides determination of $C_b(U_{n+1})$; because $C_f(U_n)$ is only partially ordered, additional factors may constrain the choice.

---

[5]Rule 1 ignores the cases when a forward-looking center is realized by a deictic pronoun.

(iii) **Centering constrains realization possibilities**: Rule 1 stipulates one constraint centering imposes on realization. Other such constraints are expected to exist.

(iv) **Preferences among sequences of center transitions**: Rule 2 hypothesizes a preference among types of transitions.

(v) **Primacy of partial information**: The information needed to compute a complete unique interpretation for an utterance may not be available until subsequent utterances are produced. Thus, a semantic theory must support the construction of partial interpretations, in particular for elements of C_f.

(vi) **Locality of** $C_b(U_n)$: The choice of a backward looking center for an utterance U_n is from the set of forward-looking centers of the previous utterance U_{n-1}. In this sense the C_b is strictly local. $C_b(U_n)$ cannot be from $C_f(U_{n-2})$ or prior sets of forward-looking centers.[6]

(vii) **Centering is controlled by a combination of discourse factors**: Center determination is not solely a syntactic, semantic, or pragmatic process.

It is important to notice that Rule 1 constrains the realization of the most highly ranked element of the $C_f(U_n)$ that is realized in U_{n+1} given that pronominalization is used. Entities realized in U_n and not realized in U_{n+1} do not affect the applicability of Rule 1. Likewise, if no pronouns are used, Rule 1 is not applicable. In particular, a non-pronominal expression can be used for the $C_b(U_n)$ as long as there are no other [weak, KH] pronouns in U_n. In such cases, centering theory asserts that the choice of $C_b(U_{n+1})$ should reflect the ranking of elements in $C_f(U_n)$. The higher the ranking of the intended $C_b(U_{n+1})$ in $C_f(U_n)$, the lower inference load should be required to interpret U_{n+1}.

2.3. Examples

In this section, we illustrate the application of the centering rules to sample discourses adapted from [Grosz et al. 1995]. We briefly reproduce the authors' comments.

Example 1[7]

(a) *John has been having a lot of trouble arranging his$_{John}$ vacation.*
C_f={John>trouble>vacation}

(b) *He$_{John}$ cannot find anyone to take over his$_{John}$ responsibilities.*
C_b=John, C_f={John>responsibilities}, Cont(U_a,U_b)

(c) *He$_{John}$ called up Mike yesterday to work out a plan.*
C_b=John, C_f={John>Mike>yesterday,plan}, Cont(U_b,U_c)

(d) *Mike has been acting quite odd recently.*
C_b=Mike, C_f={Mike}, Shift(U_c,U_d)

(e) *He called John at 5 AM on Friday last week.*
C_b=Mike, C_f={Mike>John>5 AM,Friday,last week}, Cont(U_d,U_e)

(f) *He$_{Mike}$ wanted to meet him$_{John}$ urgently.*
C_b=Mike, C_f={Mike>John}, Cont(U_e,U_f)

[6] It may appear that $C_b(U_n)$ is from $C_f(U_{n-2})$ or prior sets of forward-looking centers, but then it is only because it is in $C_f(U_{n-1})$ also.

[7] We employ the additional notation of identifying the intended referent of a pronoun by a subscript. We also depict the assumed ordering of the elements of the C_f's by ">" for strict order and "," for partial order.

(g) *John was annoyed by his/Mike's call.*
C_b=Mike, C_f={John>Mike>call}, Ret(U_f,U_g)

This discourse obeys the centering rules and is therefore considered coherent. Slight modifications of the utterances lead to violations of the centering rules:

(d') *He$_{Mike}$ has been acting quite odd recently.*
C_b=John, C_f={John}, Cont(U_c,$U_{d'}$)

According to the authors, if (1d) is replaced by (d'), where the intended referent of 'he' is Mike, we get an incoherent discourse, because the pronoun will be straightforwardly interpreted as referring to John, which is misleading. Centering theory predicts this correctly; John is the most highly ranked element in $C_f(U_c)$, and therefore it will be the preferred choice for $C_b(U_{d'})$. The incoherence is thus explained by violation of Rule 1 according to which lower-ranked elements of $C_f(U_n)$ cannot be pronominalized in U_{n+1} unless higher-ranked ones are.

Another violation of Rule 1 is obtained if (1f) is replaced by (f'):

(f') *Mike wanted to meet him$_{John}$ urgently.*
C_b=Mike, C_f={Mike>John}, Cont(U_e,U_f)

According to Rule 1 it has to be the pronoun that realizes the C_b; in this case, 'him' should refer to 'Mike', but then there would be a clash because the subject and the object would be the same, or the person called 'Mike' in (1f') would have to be someone else than the person called 'Mike' in (1c-e), which would be quite odd. So, centering theory considers this example as incoherent, and explains the incoherence by violation of Rule 1.

A violation of Rule 2 can be illustrated on the following example:

Example 2

(a) *John went to his favourite music store to buy a piano.*
C_f={John>music store>piano}

(b) *It was a store John had frequented for many years.*
C_b=store, C_f={store>John>many years}, Shift(U_a,U_b)

(c) *He was excited that he could finally buy a piano.*
C_b=John, C_f={John>piano}, Shift(U_b,U_c)

(d) *It$_{store}$ was closing just as John arrived.*
according to centering: C_b=piano, C_f={piano>John}, Shift(U_c,U_d)

The problem with this discourse according to centering theory is that it flips the center of attention back and forth. In terms of centering theory every pair of utterances constitutes a center shift. Furthermore, (2d) violates the locality constraint on the choice of C_b, because the store is not realized in (2c) , which means that it is not present in C_f, and therefore it cannot be the C_b of U_d. Centering theory would predict piano as the referent of 'it' in U_d, because it is an element of $C_f(U_c)$, but piano is not a suitable referent for 'it' given the propositional content of U_d. So, the whole discourse is judged incoherent.

On the other hand, the following re-formulation of "the same" discourse would be coherent according to centering theory, as it comprises only center continuations:

Example 3

(a) *John went to his favourite music store to buy a piano*
C_f={John>music store>piano}

(b) *He$_{John}$ had frequented the store for many years.*
C_b=John, C_f={John>store>many years}, Cont(U_a,U_b)

(c) *He$_{John}$ was excited that he$_{John}$ could finally buy a piano.*
C_b=John, C_f={John>piano}, Cont(U_b,U_c)

(d) *He$_{John}$ arrived just as the store was closing for the day.*
C_b=John, C_f={John>store}, Cont(U_c,U_d)

The affect of the violation of the centering rules and of the ranking can be illustrated Example 4:

Example 4

(a) *Susan gave Betsy a pet hamster.*
C_f={Susan>Betsy>hamster}

(b) *She$_{Susan}$ reminded her$_{Betsy}$ that such hamsters were quite shy.*
C_b=Susan, C_f={Susan>Betsy>hamsters}

(c_1) *She$_{Susan}$ asked Betsy whether she$_{Betsy}$ liked the gift.*
C_b=Susan, C_f={Susan>Betsy>gift}, Cont($U_{b,c1}$)

(c_2) *Betsy told her$_{Susan}$ that she really liked the gift.*
C_b=Susan, C_f={Betsy>Susan>gift}, Ret($U_{b,c2}$)

(c_3) *Susan asked her$_{Betsy}$ whether she$_{Betsy}$ liked the gift.*
C_b=Betsy, C_f={Susan>Betsy>gift}, Shift($U_{b,c3}$)

(c_4) *She$_{Betsy}$ told Susan that she$_{Betsy}$ really liked the gift.*
C_b=Betsy, C_f={Betsy> Susan>gift}, Shift($U_{b,c4}$)

The sentences ($4c_1$) through ($4c_4$) are each examined as a continuation after (4a,b). The authors claim that there is a decrease in acceptability from (4a,b,c_1) to (4a,b,c_4) (for some native speakers (4a,b,c_4) is even an unacceptable discourse), which can be explained by the preferences stated in Rule 2. (4a,b,c_4) is even worse then (4a,b,c_4) because in addition to the center shift, there is no continuity of grammatical subject and object.

Example 5[8]

(a) *Have you seen the new toys the kids got this weekend?*
C_f={toys,kids}

(b) *Stuffed animals must really be out of fashion.*
C_b=stuffed animals, C_f={animals>fashion}

(c) *Susie prefers the green plastic tugboat to the teddy bear.*
C_b=bear, C_f={Susie>tugboat>bear}, Ret(U_b,U_c)

(d) *Tommy likes it$_{tugboat}$ better than the bear too,*
C_b=tugboat, C_f={Tommy>tugboat>bear}, Shift(U_c,U_d)

(e) *but only because the silly thing$_{tugboat}$ is bigger.*
C_b=tugboat, C_f={tugboat}, Cont(U_d,U_e)

The authors argue that the most likely referent of 'the silly thing' in (5d) is the boat, since the boat is the most likely C_b (due to the assumed (syntactic) ranking of object before other relations as preferred

---

[8] The determination of C_b and C_f in Example 5 is ours. Grosz et al. [1995] do not provide such an explicit analysis of this example.

candidates for C_b). It follows that according to centering theory the sequence (5a-e) is a more coherent discourse than if (e) were replaced by (e'):

(e') *although the silly thing$_{bear}$ is bigger.*
C_b=bear, C_f={bear}, Shift(U_d,$U_{e'}$)

Since the authors assume that a child (or at least Tommy) likes bigger things more than smaller things, pragmatic considerations in (d') lead to the preference of 'bear' rather than 'the boat' as the referent of 'the silly thing'. But this is not what would be expected according to the ranking of $C_f(U_e)$. This conflict explains the lower degree of coherence.

The C_b can also be realized by a non-pronominal expression, provided there is no other pronoun in the utterance. The authors discuss the following example:

Example 6

(a) *The house appeared to have been burgled.*
C_f={house}

(b) *The door was ajar.*
C_b={house}, C_f={door>house}

(c) *The furniture was in disarray.*
C_b={house}, C_f={door>house}, Cont(U_b,U_c)

According to the authors, the house is realized but not directly realized in (6b) as well as in (6c), and it ranks above the directly realized centers, the door and the furniture, respectively.

2.4. Our remarks

Let us first remark on Example 4 and Example 5 from the preceding section. It should be noted that 'she' in (4b) could also be interpreted as referring to Betsy:

(b') *She$_{Betsy}$ reminded her$_{Susan}$ that such hamsters were quite shy.*
C_b=Susan, C_f={Susan>Betsy>hamsters}

This is a case not examined by the authors. Perhaps (4'b) below yields more naturally the alternative interpretation of the anaphoric pronouns:

Example 4'

(a) *Susan gave Betsy a pet hamster.*
C_f={Susan>Betsy>hamster}

(b) *She$_{Betsy}$ thanked her$_{Susan}$.*
C_b=Betsy, C_f={Betsy> Susan}

(c_1) *She$_{Susan}$ asked Betsy whether she$_{Betsy}$ liked the gift.*
C_b=Susan, C_f={Susan>Betsy>gift}, Shift($U_{b,c1}$)

(c_2) *Betsy told her$_{Susan}$ that she$_{Betsy}$ really liked the gift.*
C_b=Susan, C_f={Betsy>Susan>gift}, Shift($U_{b,c2}$)

(c_3) *Susan asked her$_{Betsy}$ whether she$_{Betsy}$ liked the gift.*
C_b=Betsy, C_f={Susan>Betsy>gift}, Ret($U_{b,c3}$)

(c_4) *She$_{Betsy}$ told Susan that she$_{Betsy}$ really liked the gift.*
C_b=Betsy, C_f={Betsy> Susan>gift}, Cont($U_{b,c4}$)

Applying the centering rules and assumptions, (4'c_4) would be the most acceptable, and acceptability would be expected to decrease from (4'c_4) to (4'c_1), because (4'c_4) exhibits both center continuation and subject continuation, (4'c_3) exhibits center retention with a change in subject, (4'c_2) exhibits center shift with subject continuation, and (4'c_1) exhibits center shift with a change in subject. However, in our opinion, the sequence (4'a,b,c_3) is equally acceptable as that of (4'a,b,c_4). It seems then that other factors are at play besides the preference of 'center continuation' over 'center retention' (and – since also (c_1) or (c_2) can follow (a) and (b) - also of these two over 'center shifting').

In Czech, as a pro-drop language, there would be no pronominal subject in the equivalent of (4b):

[$\varnothing_{Susan/Betsy}$] Připomněla jí$_{Betsy/Susan}$, že taková morčata jsou velmi plachá.

The zero subject would not be obligatorily interpreted as anaphoric to the subject of the preceding utterance. The lexical setting would play a decisive role in the interpretation.[9] The cases where the lexical setting is really neutral are very rare. We have as yet not seen a convincing psycholinguistic study to the effect that continuation of the grammatical subject is preferred in lexically neutral contexts.

Our conjecture is that the preference for seeing (4b) as a subject continuation of (4a) is due to the fact that the context prompts for their parallelism, especially by the similar Agent-Patient structure of the two utterances together with the use of pronouns to refer to both of them, and in the absence of information to the contrary originating, e.g., in the lexical setting of (4b). In (4b'), on the other hand, the parallel interpretation is contradicted by the content[10]. Centering theory does not block such interpretation, but assigns a lower priority to it. We are not sure whether centering constraints should take such strong precedence over constraints based on something like semantic compatibility in general.

It should be noted in this respect that two strategies similar to the distinction between "center retention/continuation" and "center shift" have been already noted by Weil [1844], who speaks about *marche des idées*, distinguishing a *parallel pattern* (similar to center retention/continuation) and a *progressive pattern* (similar to center shifting). This distinction is also made in the writings of Daneš on the so-called *thematic progressions* [Daneš 1970]. Neither Weil nor Daneš, however, makes any claims about preferences between the two strategies.

Let us now turn to some of the essential claims of centering theory:

A unique C_b: The authors themselves point out (p. 215) that this claim has a particular significance in generation of coherent discourses, but that with respect to discourse understanding, centering theory does not offer general criteria to discern the C_b of an utterance: If there is a single weak pronoun in the utterance U_{n+1}, then in accordance to Rule 1 it is the $C_b(U_{n+1})$.

The question what is this pronoun's referent remains. Should it be predicted that the pronoun, i.e. $C_b(U_{n+1})$, realizes the highest ranked element of $C_f(U_n)$ realized in U_{n+1}, unless the contents of U_{n+1} lead to some other referent? This would amount to a preference for center continuation over retaining over shift, which is probably not correct. Moreover, one cannot say which elements of $C_f(U_n)$ are realized in U_{n+1} and which ones are not prior to reference resolution. Once reference has been resolved, it is questionable, whether one still needs to determine the $C_b(U_{n+1})$.

[9] However, if the speaker wants to make the anaphoric reference to an item other than the grammatical subject clear, she would use a demonstrative pronoun 'ta' (for feminine gender, 'ten' for masculine, and 'to' for neuter), meaning "the latter", in the subject position:

Ta$_{Betsy}$ jí připomněla, že taková morčata jsou velmi plachá.

The use of 'ten' in such a situation is not related to the syntactic function of the anaphor, but it usually refers back to an item identified by an expression with a higher degree of communicative dynamism than another possible "competitor" for reference.

[10] Usually, the one who received a present thanks the one who gave it, not the other way round.

If there is more than one pronoun in U_{n+1}, the determination of the $C_b(U_{n+1})$ is beyond the constraints of centering theory. Since there is always only one backward looking center in U_{n+1}, what about if two elements of $C_f(U_n)$ are realized in U_{n+1} by a pronoun as in the following piece of discourse:

Example 7[11]

(a) *John called Mary a REPUBLICAN.*
C_f={John>Mary}

(b) *Then she$_{mary}$ INSULTED him$_{john}$.*
C_b=John, C_f={Mary>John} or C_b=Mary, C_f={Mary>John} ?

(c_1) *He$_{john}$ asked Mary why she$_{mary}$ feels OFFENDED.*
C_b=John, C_f={John>Mary}, Cont(U_b,U_{c1}) or Shift(U_b,U_{c1}) ?

(c_2) *John$_{john}$ asked her$_{mary}$ why she$_{mary}$ feels OFFENDED.*
C_b=Mary, C_f={John>Mary}, Shift(U_b,U_{c2}) or Ret(U_b,U_{c2}) ?

(c_3) *He$_{john}$ asked her$_{mary}$ why she$_{mary}$ feels OFFENDED.*
Alternative one: C_b=John, C_f={John>Mary}, Cont(U_b,U_{c3}) or Shift(U_b,U_{c3}) ?
Alternative two: C_b=Mary, C_f={ John>Mary}, Shift(U_b,U_{c3}) or Ret(U_b,U_{c3}) ?

(d) *He$_{john}$ felt very sorry.*
C_b=John, C_f={John}, Cont(U_{c1},U_d), Shift(U_{c2},U_d)

(7a,b) can be continued, at least in our opinion, both by (c_1) and (c_2), either of which can then be continued by (d). Since centering theory does not determine which of two pronouns is the $C_b(U_{n+1})$, we have two possibilities in (7b): the C_b can be either John or Mary. But depending on the choice we make, the prediction should be that (7a,b,c_1,d) is either a pair of continuations or a shift followed by a continuation. The discourse (7a,b,c_2,d), on the other hand, would be either a sequence of shifts, or a retention followed by a shift. So, according to centering theory, if we take the choice $C_b(U_b)$=John in (7b), we get a clear preference for (7a,b,c_1,d) over (7a,b,c_2,d), because a sequence of continuations is preferred to a sequence of shifts. If we take C_b=Mary, though, we get no such clear prediction. Should the possibility of judgement about coherence of a discourse depend on the choice of C_b? In our opinion, similar considerations hold for the sequence (7a,b, c_3).

Moreover, we wonder whether under $C_b(U_b)$=John, the sequence (7a,b,c_1,d) is really to be judged as more coherent then (7a,b,c_2,d). At least, we, as non-native speakers of English, do not perceive any difference between them.

Maybe the following rule can be added: take the highest ranked element of $C_f(U_n)$ realized by a pronoun in U_{n+1} to be the C_b. With such an approach the C_b would indeed depend on the preceding context. But it would not help us to resolve pronouns, because in order to decide which one realizes the C_b, we would already need to know what it co-refers with. But one would have to look at what effect this additional rule would have on the coherence judgements. In any case, such preference for the C_b would amount to a preference for center continuation over retention over shift, which is probably not correct. Certainly not in a general case, i.e., in utterances with just one pronoun or with no pronoun at all.

Likewise, if there is no pronoun in U_{n+1}, centering theory does not help to determine which expression realizes the $C_b(U_{n+1})$.

---

[11] We mark the intonation center by CAPITALS.

Example 8

(a) *The store in our street is very small but popular.*

(b) *Refilling the shelves keeps the staff busy all the time.*

Again, it appears that it is first necessary to know what the expressions in U_{n+1} refer to and then it can be said which one realizes the $C_b(U_{n+1})$, for instance on the basis of the $C_f(U_n)$ ranking. But then, the centering constraints could not be used any more to help reference resolution, in particular to help to decrease the inference load needed in this process. If centering were to be used to guide the choice of "candidates for antecedents", in the sense of preferences as hinted at above, it is questionable whether this would make the process of reference resolution more efficient at all, given that the preferences could be overridden any time on the basis of contents, and that evaluation of the preferences would take some processing as well.

Ranking of C_f:

Given that the ranking of elements of $C_f(U_n)$ determines their likelihood to be the $C_b(U_{n+1})$, the $C_b(U_{n+1})$ will be as first choice taken to correspond to the highest ranked element of $C_f(U_n)$ realized in $U_{n+1,}$ (additional factors may constrain the choice, but centering theory does not discuss them). Due to the partial ordering of the forward-looking centers, there may be more than one element in C_f(Un) with the same ranking; since there is always only one backward looking center in U_{n+1}, centering will not give a clear prediction of the $C_b(U_n)$ in cases where at least two of the elements with the same ranking are realized in U_{n+1} and no higher ranked elements are realized.

We agree with centering theory in this respect. In accordance with what we have said earlier, we do not think that it is possible to predict which elements will be subsequently referred to in a discourse.

There is, however, one objection to the claims of centering with respect to the role of the $C_f(U_n)$ ranking and its effect on the choice of the $C_b(U_{n+1})$: centering theory when applied to discourse understanding will systematically make wrong predictions for utterances with a single pronoun (in any position, subject or not) referring anaphorically to anything but the highest ranked element of $C_f(U_n)$ realized in $U_{n+1,}$ because, if we understand well, the first-choice candidate for the antecedent of a pronoun will be the highest ranked element of $C_f(U_n)$. If that one turns out not suitable, the lower ranked one will be tried, as long as there is one. If none is suitable, the utterance U_{n+1} will be judged incoherent (observing the locality principle). In such cases, the inference load needed to resolve pronominal reference will increase, because the wrong choice(s) will have to be rejected by reasoning about the content, like in Example 9 and Example 10 below:

Example 9

(a) *Jim received some new piece of information from that remote country.*
Cf={Jim>piece_of_information>country}

(b) *It$_{country}$ was reported to be a republic.*

If we let the centering constraints guide the reference resolution of 'it' in (9b), the first candidate for antecedent would be 'Jim' as the highest ranked element in $C_f(U_a)$. But it would be rejected as unsuitable on the basis of gender disagreement between 'it' and 'Jim'. The next candidate would be 'piece of information'. But it should be rejected under normal circumstances due to semantic incompatibility of 'piece of information' and 'republic'. The third candidate could be 'country' and it would be acceptable. The modifier 'information' could also provide a candidate, probably with the same ranking as 'country'. It is not clear which one of the two candidates would/should be tried first, probably the choice would be random as far as centering is concerned.

Another example of the same issue is the following:

Example 10

(a) *Susan spends hours every day talking to her baby-sister.*
C_f={Susan>hours>sister}

(b) *She$_{sister}$ does not understand yet, of course.*
first choice: C_b=she=Susan, but pragmatically, this does not make sense, so
second choice: C_b=she=sister, C_f={sister>word}

Another issue could be raised with respect to which elements are to be included in the $C_f(U_n)$ and their ranking: Is the prominence in U_n related only to the immediate co-text, and in no way to the (broader) context? There are examples in the original article of cases where elements mentioned in U_n and not mentioned in U_{n+1}, are kept in the $C_f(U_{n+1})$, if they are associated with elements mentioned in U_{n+1}. What are the constraints on including not mentioned elements? This issue seems to be related also to the issue of locality of C_b (see below).

Preferences among sequences of center transitions:

Rule 2 hypothesizes a preference among types of transitions. We would like to remark that a discourse generated strictly in accordance to Rule 2 would be monotonous and boring. It is in fact difficult to find stretches of discourse with sequences of continuations[12] in real texts (at least in the Czech, mostly technical, texts that we analyzed with centering in mind).

Locality of $C_b(U_n)$:

The backward looking center for an utterance U_n is chosen from the set of forward-looking centers of the previous utterance U_{n-1} which are claimed not to be constrained by the features of any previous utterance in that segment. If $C_b(U_n)$ comes from some utterance U_{n-k}, where k >1, then this is possible only in case that this item is also included in $C_f(U_{n-1})$ (see our comment above concerning ranking and elements kept implicitly in C_f). The strict locality principle indicates that entities from U_{n-1} not realized in U_n get 'forgotten', at least for centering; they remain in the focus space [Grosz & Sidner 1986], but centering says nothing about the way in which their "recovery" at a discourse segment pop is done.

Let us consider the following modification of Example 2:

Example 2'

(a) *John went to his favourite music store to buy a piano.*
C_f={John>music store>piano}

(b) *The store had a choice of valuable antique instruments.*
C_b=store, C_f={store>instruments}, Shift(U_a,U_b)

(c) *He was excited that he$_{john}$ could finally buy the piano of his$_{john}$ dream.*
intuitively: C_b=John, C_f={John>piano>dream}, Shift(U_b,U_c)

We consider this discourse to be perfectly coherent, the antecedent of the pronoun 'he' – which we presume would be the $C_b(U_c)$ – is not an element of the $C_f(U_b)$ and we do not think it can be assumed to be among them "implicitly".

What should be said in such cases? Does this signal a segment boundary? Seen from the perspective of discourse relations (e.g. [Mann & Thompson 1987], [Asher 1993]), the utterance U_b appears to provide a background to U_a, which is an asymmetric relation where U_a is a nucleus and U_b a satellite, and thus U_b can be considered embedded with respect to U_a. U_c does not seem to relate (at

[12] A sequence of continuations according to centering theory requires at least three utterances with the same center one after another.

least in a significant way, if we ignore the relation between *valuable* and *of his dream*) to U_b, but it can be seen as an elaboration with respect to U_a. Perhaps also at the intentional level [Grosz & Sidner 1986] one would say that U_a starts a discourse segment DS_1, U_b constitutes an embedded discourse segment DS_2, and U_c, after a discourse segment pop belongs into DS_1 again. Given these considerations of discourse structure, it would not be surprising that U_c seems to disobey the locality constraint. It is a question whether the locality constraint applies in this case, and how. Are embedded discourse segments skipped over as far as the determination of the C_b is concerned?

Last but not least, a remark on the notion of a center: According to centering theory, centers are semantic objects and discourse constructs. The authors do not compare their notion of center to the notions of discourse markers, discourse entities or discourse referents, used by others (e.g. [Heim 1983], [Webber 1983], [Kamp & Reyle 1993]). As we understand their explanations, the notion of center is at least similar to these notions, especially because centers are not words, etc. It is then obvious that it is an utterance and not a sentence in isolation that has centers and that the same sentence when uttered in different discourse situations may have different centers, because determination of centers is related to reference resolution. One can ask, however, whether some expressions in a sentence bear some features reflecting or indicating they realize (possible) centers. For instance, if an expression is the only weak pronoun in a sentence, it seems that centering theory will always predict it to be the expression which realizes the backward looking center. The center itself, of course, will be a different semantic object in different discourse situations (for more about this, see Section 4.2).

3. Salience-based approach

3.1. Goals and method

The salience-based approach (as presented in [Hajičová 1993]) represents an attempt to account for the relations between the changes of salience and the choices of referring expressions in discourse developed on the basis of the Praguian framework of Functional Generative Description of language (FGD) (cf. [Sgall 1967], [Sgall et al. 1986]).

FGD works with several levels of representation of a sentence, ranging from the phonemic (graphemic) level which represents the external shape of a sentence, to the tectogrammatical (underlying) level which represents its linguistic meaning. The representation of a sentence at the tectogrammatical level (TR) consists of nodes, corresponding to lexical items, carrying complex labels encoding various kinds of information about the lexical items. The TR has three ingredients: the structural relations between semantemes (reflecting their dependency and coordination relationships), the scale of communicative dynamism (reflecting "deep word order") and the distinction between contextual boundness and non-boundness for each semanteme (reflecting its position in the linguistic patterning of the opposition between "given" and "new"[13]). A suitable operational test for this distinction can be based on question-answer pairs or on negative and other continuations.

Based on the primitive notion of contextual boundness (CB), the TR of a sentence is divided into a topic part and a focus part. This topic-focus articulation of a sentence is defined as follows (see Sgall et al. [1986]):

The focus of a sentence contains all contextually non-bound nodes depending directly on the verb together with all nodes subordinate to them (i.e., depending on them directly or indirectly), be they

[13]A cognitively "given" entity can be referred to by a contextually non-bound item if it is being chosen from a set of alternatives. On the other hand, a contextually bound item need not always be coreferential with a co-textual antecedent; it may refer to an entity made salient by the situation of the discourse.

contextually bound or non-bound; the verb itself is contained in the focus if it is non-bound. Whenever the verb as well as all nodes directly dependent on it are contextually bound, the focus is defined likewise with respect to a deeper embedded node. *The topic of a sentence* consists of those nodes that are not contained in the focus.

TFA is not only a matter of utterances but also of sentences: two utterances differing only in the assignment of TFA are occurrences of two different sentences with different meanings.

It is one of the fundamental aspects of the process of communication that one of the participants, the speaker, attempts to cause (one or more of) the other participants, the hearer(s), on the basis of the rules of the given language, to modify some entities in his (their) respective cognitive state(s).[14] For this purpose, the speaker produces utterances containing expressions that refer to the entities that should be modified, and expressions that refer to what the modifications are to be. The hearer, on the other hand, interprets the speaker's utterances as containing expressions of these two types. The entities that should be modified are referred to in the topic, whereas the modifications are referred to in the focus.

We assume that before the initiation of communication, the memory of each communication participant is in a certain state, which is not empty. The memory states of the individual participants usually differ during real-life communication[15]; for the sake of any communication to be possible, the participants assume that their memory states have something in common at every point during their communication, notably also at its beginning.

Within the FGD framework, the term *stock of shared knowledge* (SSK) has been introduced to refer to the "common ground" that the speaker assumes to be in the hearer's mind. The stock of knowledge an individual communication participant assumes to share with the other(s) at the initial point of a discourse is based on his/her individual memory state, including, e.g., the recollections from their previous encounters and the reflections of their current situational context(s). As their discourse progresses, the stock of shared knowledge pertaining to a participant A with respect to a participant B changes according to what A considers to be in the center of attention at the given point, i.e., what A assumes to be most activated (salient, foregrounded) in B's memory state (cf. [Hajičová et al. to appear]).

The SSK has a dynamic character. As a discourse progresses and attention moves from some entities to others, the set of entities and relations among them included in the SSK changes. The entities in the SSK are assigned different *degrees of activation*, i.e., some of them are easier to access for the interlocutors and thus more readily available for reference by linguistic means. The changes in the activation degrees of entities in the SSK thus reflect the changes in attention as a discourse progresses. The changes in the degree of activation depend on whether the entity is being referred to subsequently, and if so, by what form of expression, and whether it is rendered as contextually bound or non-bound, in the topic or in the focus. Associative links through which an entity can be accessed indirectly (implicitly) are also taken into account.

According to the salience-based approach, one of the main characteristics of a smooth discourse is that the speaker chooses only those items to be rendered by contextually bound expressions which she assumes to have a high degree of activation in the SSK, because then the hearer should be able to identify the referents easily. Entities with a low degree of activation can only be referred to by contextually non-bound expressions. Concentrating on the verb and the lexical items directly dependent

---

[14]For the sake of clarity, the following convention of pronominal reference to the communication participants is henceforth employed: the forms 'she', 'her' are used for the speaker, whereas the forms 'he', 'his' are used for the hearer.

[15]An important portion of communication seems to be devoted to the participants' efforts to find out where they diverge and their attempts to change one another's memory state, so that they (become to) converge on the matter of the communication.

on it in the TR of a sentence (leaving thus aside embedded lexical items), one can say that non-activated entities can be referred to only in the focus part of the sentence, while activated entities can be referred to in the topic- as well as in the focus part of the sentence.

The salience-based approach to modelling attention changes in discourse specifies assignments of activation degrees to entities in the SSK and states heuristics which relate the degree of activation of an entity and the acceptable forms of a linguistic expressions referring to it, in particular in terms of being rendered as contextually bound or non-bound.

This approach has been formulated on the basis of the FGD framework and it relies on its notions. It can be adapted for another theory of grammar as long as counterparts of such notions as contextual boundness and TFA can be defined in that theory. The salience-based approach does not make any explicit claims on a suitable semantic theory. It seems that any theory could be employed as long as it supports the relation between a linguistic expression and a (representation of an) entity in the (model of) the SSK.

3.2. Rules

Up to now (i.e. before getting an insight into the possibilities of how to handle the impact of the situation and of the lexical and clausal factors within the sentence), the essential idea underlying the salience-based approach can be understood as follows:

- to define assignments of activation depending on whether an entity is referred to in the topic or in the focus of an utterance in a discourse and to keep track of activation changes after every sentence as a discourse progresses
- to formulate principles of reference relating the activation degree of an entity in the SSK and the form of the expression referring to it and to check whether linguistic expressions comply to the principles

The activation hierarchy is used as one of the clues for resolution of anaphoric reference, or in the case of production, for selection of appropriate forms of referring expressions.

It has been proposed to represent the degrees of activation of entities in the SSK by natural numbers assigned to these entities in such a way that the most activated entities obtain the smallest number and the lower the activation of an entity the greater is the number assigned to it (cf. [Hajičová & Vrbová 1981; 1982], [Hajičová 1987; 1993], [Hajičová et al. 1995]). Denoting by a^n an entity a referred to in a sentence with degree of activation n, the following rules for transitions (denoted by $a^n \rightarrow a^m$) in degree of activation triggered by sentence S_i can be formulated (similarly to [Hajičová 1993][16]):

1. if a^n is referred to within the focus part of S_i, then $a^n \rightarrow a^0$
2. if a^n is referred to by a weak (unstressed) anaphoric pronoun or a zero-pronoun within the topic part of S_i, then $a^n \rightarrow a^n$
3. if a^n is referred to by a definite noun phrase within the topic part of S_i, then $a^n \rightarrow a^1$
4. if $a^n \rightarrow a^m$ then $b^r \rightarrow b^{m+2}$ for all entities b associated with a and not referred to in S_i
5. if an entity a^n is neither referred to nor included among the entities associated with some entity that is referred to in S_i, then $a^n \rightarrow a^{n+2}$ for a^n that has only been referred to in the focus of the sentence(s)

[16]Some amendments to the activation assignment rules dealing with complex sentences have been proposed in [Hajičová et al. 1995]. We have not taken them into account for the sake of our current comparison, however. The main reason is that centering theory does not include a refined treatment of complex grammatical structures either, and thus there would not be a ground for a proper comparison.

preceding S_i, and $a^n \rightarrow a^{n+1}$ for an entity that has been referred to in the topic of one of the preceding sentences

We are aware that also the subject position contributes in and of itself to the likelihood an entity will be referred to in the topic proper of the next utterance and of the preference of parallel syntactic structures, but we do not suppose these factors to be among the prominent ones.

The degrees of activation impose a partial ordering on entities in the stock of shared knowledge: there can be entities having the same degree of activation, which is represented by assigning the same number to them.

It is assumed that in a co-operative communication setting, the participants are aware of certain principles of "good behaviour" and adhere to them[17]. The speaker obeys these principles in order for her utterances to be felicitous, that is, not deceiving the hearer and reasonably easy for him to interpret. The hearer, on the other hand, presumes the speaker wants to be understood and therefore presumes her to undertake this very effort. This enables him to make use of the same or corresponding principles in order to understand the speaker's utterances.

The principles of co-operative communication in the salience-based approach relating the degree of activation of an entity and the form of linguistic expression referring to it, are stated as follows:

Salience-based approach: Principle 1

(a) The speaker constructs her utterances in such a way that:

- contextually bound expressions refer to entities activated above a certain threshold in the SSK
- contextually non-bound expressions refer to activated as well as non-activated entities in the SSK

(b) The hearer interprets the speaker's utterance in such a way that:

- the entities referred to by contextually bound expressions are to be identified among the entities activated above a certain threshold in the SSK

Principle 1 relates activation of entities in the SSK and the rendering of linguistic expressions referring to them as contextually bound or non-bound. Principle 2 below relates TFA and the structuring of information in discourse:

Salience-based approach: Principle 2

The speaker constructs and the hearer interprets the utterances in a discourse in such a way that:

- the topic part refers to the entities in the SSK that are to be modified or introduced into new relations with some other entities
- the focus part refers to the modifications of entities in the SSK or new relations to be achieved

These principles correspond to the assertions concerning the relation between activation, TFA and contextual boundness made in earlier works within the FGD framework.

In regard of the use of pronouns in discourse and the determination of their referents, Hajičová [1993] states that when looking for an antecedent of a pronoun one should take into account the following aspects:

(a) relative activation of entities in the SSK:

[17] Such principles are known especially from Grice [1975], who calls them the maxims of communicative behaviour. Our principles are more specific, but cover a narrower range of aspects.

- ♦ the most activated entities have precedence over the less activated ones (esp. for pronouns like 'this')
- ♦ the entity referred to by a noun phrase in the subject position in the preceding utterance has precedence over the entities referred to by noun phrases in other syntactic positions (esp. for weak pronouns like 'he')

(b) factual knowledge, which excludes interpretations that are counter to the hearer's knowledge and intuitions about the world

(c) grammatical structure: especially the tendency towards a parallel patterning of subsequent utterances, embedded entities being less prominent, or the subject being retained

3.3. Examples

In this section we first reproduce one example from [Hajičová 1993], and then we analyze the examples from Section 2.3 according to the salience-based approach. We denote the entities referred to by subscripts inserted into the input sentences. The state of the SSK after processing an utterance is depicted as an unordered list whose members are the entities referred to with their degrees of activation indicated by subscripts. The underlining marks items included in the focus. Intonation centre is marked by CAPITALS. For the sake of the current discussion we do not take into account associative links among entities in the SSK when determining the activation degrees (i.e., activation assignment rule 4 is being omitted). We show a schematic picture of the activation changes through this text in the Appendix, Figure 1.

Example 11

(a) *The school garden*$_g$ *was full of CHILDREN*$_c$.
$\{g^1, c^0\}$

(b) *they*$_c$ *talked NOISILY,*
$\{c^0, g^2\}$

(c) *but the teachers*$_t$ *did not REPROVE them*$_c$*,*
$\{t^1, c^0, g^3\}$

(d) *because they*$_c$ *were so EXCITED*.
$\{c^0, t^2, g^4\}$

(e) *Outside PARENTS*$_p$ *were waiting*.
$\{p^0, c^1, t^3, g^5\}$

(f) *One of them*$_p$*, a father*$_f$*, stood in front of a MICROPHONE*$_m$
$\{p^0, f^1, m^0, c^2, t^4, g^6\}$

(g) *as if he*$_f$ *were prepared to TALK*.
$\{f^1, p^1, m^2, c^3, t^5, g^7\}$

(h) *The pupils*$_c$ *got CALM*
$\{c^1, f^2, p^2, m^4, t^6, g^8\}$

(i) *and their*$_c$ *teachers*$_t$ *lined them UP*.
$\{c^1, t^1, f^3, p^3, m^6, g^9\}$

(j) *Both pupils and teachers were in a festive MOOD*$_d$.
$\{c^1, t^1, d^0, f^4, p^4, m^7, g^{10}\}$

This discourse is quite smooth, providing that reference to school in (11a) activates via associative link the entity referred to by the CB expression 'the teachers' in the topic of (11b). The expression 'the pupils' should be interpreted as co-referential with 'children' on the basis of factual knowledge and the fact that the entity referred to by 'children' in (11a) and co-referred by weak pronouns in (11b), (11c) and presumably also (11d) is highly activated (degree 3).

However, the utterance (11d), in particular the resolution of the pronoun 'they' in its topic, poses a problem: the entity c^0 is the most activated one, but the entity t^1 was referred to in the subject of the immediately preceding utterance, (11c), and should therefore be also preferred. Moreover, the difference in the activation degree between c^0 and t^1 is very small and therefore indecisive.

In our opinion it is correct that our approach does not provide a strong preference for either *c* or *t* as the entities referred to by 'they' in (11d). The pronoun seems to be used improperly in that utterance and even a human interpreter might find its reference ambiguous.

We now provide an analysis of the examples presented in Section 2.3 according to the salience-based approach.

Example 1 revisited

(a) *John*$_j$ *has been having a lot of trouble*$_t$ *arranging his*$_j$ *VACATION*$_v$.
$\{j^1,t^0,v^0\}$

(b) *He*$_j$ *cannot find anyone to take over his*$_j$ *RESPONSIBILITIES*$_r$.
$\{j^1,r^0,t^2,v^2\}$

(c) *He*$_j$ *called up Mike*$_m$ *yesterday to work out a PLAN*$_p$.
$\{j^1,m^0,p^0,r^2,t^4,v^4\}$

(d) *Mike*$_m$ *has been acting quite ODD recently.*
$\{j^2,m^0,p^2,r^4,t^6,v^6\}$

(e) *He*$_m$ *called John*$_j$ *at 5*$_5$ *AM on Friday*$_f$ *last WEEK*$_k$.
$\{j^0,m^0,5^0,f^0,k^0,p^4,r^6,t^8,v^8\}$

(f) *He*$_m$ *wanted to meet him*$_j$ *URGENTLY.*
$\{j^0,m^0,5^2,f^2,k^2,p^6,r^8,t^{10},v^{10}\}$

(g) *John*$_j$ *was annoyed by his*$_m$*/Mike*$_m$*'s CALL*$_c$.
$\{j^0,m^0,c^0,5^4,f^4,k^4,p^8,r^{10},t^{12},v^{12}\}$

Note that the entity *j* is introduced with the one-but-highest degree of activation in utterance (1a) through reference in the topic, and preserves it there after (1b,c) thanks to the pronominal co-reference in the topics. The entity *m* is introduced with the highest activation degree in (1c) through reference in the focus. In (1d), entity *m* preserves its activation degree thanks to the co-reference in the topic, while *j* is not referred to and therefore fades away by one degree (because it has been referred to in the topics of the preceding utterances). A difference of two activation degrees is thus created between *m* and *j* before the processing of (1e) which enables to resolve the pronoun 'he' as co-referring to the entity *m*.

Note also that after processing utterance (1e), the entities *j* and *m* are both activated to the highest degree. The only possibility to resolve the pronouns in (1f) is then the application of a preference for parallelism between the two utterances.

Example 2 revisited

(a) *John*$_j$ *went to his favourite music store*$_s$ *to buy a PIANO*$_p$.
$\{j^1,s^0,p^0\}$

(b) *It*$_s$ *was a store*$_s$ *John*$_j$ *had frequented for many YEARS*$_y$.
$\{j^1,s^1,p^2,y^0\}$

(c) *He*$_j$ *was excited that he could finally buy a PIANO*$_p$.
$\{j^1,s^2,p^0,y^2\}$

(d) *It*$_{?p}$ *was closing just as John*$_j$ *ARRIVED*.
according to activation degrees only: $\{j^1,s^3,p^1,y^4\}$, but really: $\{j^1,s^1,p^2,y^4\}$

Note that whereas the activation degree of the entity *j* remains 1 through utterances (2a,b,c), the activation degree of the entity *p* oscillates between 0 and 2, and that of the entity *s* is 0 after (2a) and then decreases to 1 after (2b,c). The oscillation of the activation degree of *p* indicates that this discourse is not really smooth. Even worse so, the resolution of the pronoun 'it' in (2d) is problematic. According to the preferences of the salience-based approach, it should refer to the entity *p* which has the highest degree of activation. However, factual knowledge contradicts this choice, because among the entities that are activated at that point it is more likely that the store is closing rather than the piano. This discourse is thus judged bad also under the TFA approach.

Example 3 revisited

(a) *John*$_j$ *went to his favourite music store*$_s$ *to buy a piano*$_p$
$\{j^1,s^0,p^0\}$

(b) *He*$_j$ *had frequented the store*$_s$ *for many years*$_y$.
$\{j^1,s^1,y^0,p^2\}$

(c) *He*$_j$ *was excited that he*$_j$ *could finally buy a piano*$_p$.
$\{j^1,s^2,p^0,y^2\}$

(d) *He*$_j$ *arrived just as the store*$_s$ *was closing for the day*$_d$.
$\{j^1,s^1,d^0,y^4,p^2\}$

Note that also in this variant of the discourse the activation degree of the entity *j* remains 1 through utterances (3a,b,c), the activation degree of the entity *p* oscillates between 0 and 2, and that of the entity *s* is 0 in (3a) and then decreases to 1 after (3b,c). But unlike in (2d), the entity *s* is not referred to by a pronoun but by a definite noun phrase in (3d), which makes it much easier to interpret. Clearly, the discourse in Example 3 exhibits also more parallelism than the variant in Example 2, from which one may want to conclude that it is smoother.

4. A confrontation

4.1. Goals, method and models

The main goal of centering theory is to model the local component of the attentional state within a discourse segment, with the aim of reducing inferences needed in order to determine the referents of linguistic expressions, in particular pronouns. After processing a non-initial utterance, the model consists of a partially ordered list of forward-looking centers and a backward-looking center. The ordering of the forward-looking centers is essential for determination of the most likely candidate(s) for pronominal reference in the next utterance. This ordering is determined within the utterance, and depends on factors like grammatical roles, word order, fronting, etc. Centering theory imposes constraints on the realization of the backward-looking center of an utterance with respect to the list of forward-looking centers of the immediately preceding utterance, and on the transitions from one backward-looking center to another across utterances within the same discourse segment. According to centering theory, compliance to these constraints serves as a measure of discourse coherence.

The salience-based approach attempts to model the concentration of attention throughout a discourse, as one aspect determining the reference of nominal expressions, besides factual knowledge and relations between sentences, e.g. syntactic parallelism. Attention is modelled by activation of entities in the stock of knowledge shared between a speaker and a hearer. The entities in the stock of shared knowledge are partially ordered reflecting their relative degrees of activation. The relative degrees of activation contribute to the determination of candidate(s) for reference by nominal expressions in subsequent utterance(s). An activation degree is assigned to an entity referred to in a sentence on the basis of its previous activation and depending on where and how it is referred to, i.e. whether it is referred to in the topic or in the focus, and whether it is referred to by a weak pronoun or by some other means. The salience-based approach formulates principles relating the form of referring expressions and their occurrence in the topic or in the focus in relation to the degree of activation of the entities they refer to. Compliance to these constraints determines whether a given discourse is judged to be smooth or not.

4.2. Terminology

In order to be able to compare the centering theory and the salience-based approach, we propose the following rough "alignment" of the notions they employ, without claiming that there is an equivalence, or that there is a straightforward correspondence at all (see remarks below):

	Centering theory	*Salience-based approach*
1.	centers of an utterance i.e. semantic objects, discourse constructs	entities in the SSK referred to in an utterance i.e. the mental objects or similar entities topic and focus are parts of the linguistic meaning of a sentence (tectogrammatical, underlying structure), contextual boundness (non-boundness) is ascribed to the tectogrammatical structures of the expressions (linguistic items)
2.	backward looking center	the entity referred to by the topic proper
3.	determination of the expression realizing the $C_b(U_n)$	determination of the TFA of U_n (communicative dynamism, contextual boundness)
4.	determination of the elements in $C_f(U_n)$ and their order	assigning activation degrees to entities in SSK yielded by U_n
5.	ranking of the elements in $C_f(U_n)$ determined primarily by syntactic roles	hierarchy of activation of entities in SSK after utterance U_n determined on the bases of TFA and by what means they are referred to, if at all in U_n
6.	centering is a discourse matter, centers appertain to an utterance in a discourse	assigning activation degrees is a discourse matter, TFA appertains to a sentence (and is „inherited“ by an utterance of the sentence)
7.	centering constraints applied across an adjacent pair of utterances within a discourse segment	salience-based principles applied to the state(s) of the SSK

8.	compliance to centering constraints is a measure of local coherence of a discourse segment	compliance to principles is a measure of discourse smoothness (and helps to resolve reference assignments)
9.	inferencing is to verify the reference resolution hypotheses made according to centering rules	inferencing and activation assignments are two (complementary) aspects (in reference resolution)
10.	discourse segment boundaries determined at another, i.e. the intentional, level of discourse structure	possible locations of discourse segment boundaries correlated with points of changes in constellations at the highest degrees of activation in SSK (locations of segment boundaries "arise" from activation changes)

Our comments:

One important difference of the two approaches examined in this paper concerns the status of the centers: Centers cannot be compared to the contextually bound or non-bound items, or to the topic or focus of a sentence, simply because centers, unlike the other items, are semantic objects. We think they can be compared to the entities represented in the SSK. In the current version of the salience-based approach, such entities are created for noun phrases, but it is assumed that verbs and other items also refer to entities in the SSK.

Centering theory uses the relation of realize and directly realize to capture the link between an expression occurring in an utterance and the semantic object of the center corresponding to it. As far as we understand it, the realization relation is capable of reflecting associations between semantic objects, e.g. the associative link between a house and a door. A similar idea is present in the salience-based approach where direct reference to an entity can activate an associated entity.

According to the FGD framework, as we said above, a proper uttering of a sentence in discourse is characterized by its topic being linked to the context. So, we can investigate whether there are any correlations between the centers of utterances and the entities referred to in the topic.

As far as the examples we analyzed up to now are concerned, we conclude that the backward-looking center seems to correspond to the entity referred to by the topic proper (i.e., linguistically anchored as the item carrying the lowest degree of communicative dynamism)[18]. Our conclusion can be also formulated as follows: the expression realizing the backward looking center seems to be identifiable with the topic proper, with the following provisions:

(i) centering theory applies the locality restriction on C_b, but the topic proper can link to any adequately activated entity in the SSK

(ii) in centering theory the first utterance of a discourse segment does not have a C_b, whereas even the first utterance of a discourse (segment) has a topic proper which refers to some entity which the speaker presumes to be sufficiently activated in the SSK (unless the first utterance is an occurrence of a topicless, "out of the blue", sentence)

Although the backward looking center itself is a discourse construct, there can be some indication as for which expression realizes it in the linguistic meaning of the sentence. For instance, in the following sentence, which occurred as an utterance (b) in Example 3, the pronoun 'he' is unambiguously the topic proper:

[18] See also the observation made by Hoffman [1996, p. 557].

(b) *He had frequented the store for many years.*

On the other hand, in a modification of this sentence with two pronouns, either of them could be the topic proper, which would correspond to two different orderings of communicative dynamism (deep word order [Sgall et al. 1986]):

(b') *He had frequented it for many years.* CD(he)<CD(it)

(b") *He had frequented it for many years.* CD(it)<CD(he)

Thus two utterances differing only in that their C_b's would be realized by different expressions could be understood as two different sentences, i.e. two sentences with different linguistic meanings.

Similarly to the backward looking center, the forward looking centers of an utterance can be compared to certain entities referred to in an utterance of a sentence. There is an important difference between the two approaches concerning what elements are considered: In centering theory, on the one hand, the list of forward looking centers contains only the entities realized and directly realized in the given utterance, U_{n-1}. Due to the requirement of locality on the $C_b(U_n)$, the $C_b(U_n)$ can (coherently) only link to an element of $C_f(U_{n-1})$, that is, to an entity realized or directly realized in U_{n-1}. On the other hand, in the salience based approach, the topic proper (corresponding to the $C_b(U_n)$ as we suggested above), links to an entity with some degree of activation in the SSK which has resulted from processing utterances up to U_{n-1}. The salience-based approach therefore does not have a locality restriction on the entity referred to by the topic proper (or anywhere else in U_n). It would be desirable then to combine this approach with further considerations of the relation between attention changes in discourse and the choice of referring expressions. However, the centering approach as it is too strict.

Another important difference between the centering theory and the salience-based approach concerns what entities are included in the $C_f(U_n)$ and in the SSK, and therefore which entities are "eligible" for reference in U_{n+1}. As we said above, the $C_f(U_n)$ ranking only applies to objects (directly) realized in U_n, whereas the activation hierarchy of entities in SSK contains all entities referred to (directly or indirectly, by association) in the preceding discourse, unless they have faded away (below a certain "oblivion" threshold).

Both the $C_f(U_n)$ ranking and the activation hierarchy in SSK are partial orders. The difference is in the prerequisites of the ranking of forward-looking centers in centering theory and the hierarchy of activation degrees in the salience based approach. In the centering approach, it is essentially based on grammatical roles (SUBJECT > OBJECT > OTHERS)[19], whereas in the salience-based approach it is in principle the TFA (being referred to by a pronominal or non-pronominal expression in the topic or in the focus)[20]. The main difference is that whereas centering theory still has to define the impact of phenomena like word order, intonation patterns or "focussing" syntactic constructions, the TFA already reflects these phenomena. TFA allows for a relatively uniform description of the ordering that should result from an utterance with a particular TFA.

We suggest that the ranking of $C_f(U_n)$ elements based on grammatical roles (with further elaboration) in centering theory could be replaced by a partial order of $C_f(U_n)$ based on TFA.

It is tempting to examine how the $C_f(U_n)$ ranking compares to the hierarchy of activation degrees in the SSK. Let us illustrate this point on Example 2 and Example 3.

The centering theory explains the fact that (3) is more coherent than (2) by the fact that all the utterances in this segment have the same C_b. In our approach, the difference in coherence is ascribed to the difference in the degrees of activation: the reference by a noun ('the store') in (2d) is preferred since

---

[19] But compare Hoffman's [1996] postulation of ranking in terms of theta roles.

[20] Ultimately, the activation degree assignments might be based on contextually bound and non-bound items occurring as more or less deeply embedded and on degrees of communicative dynamism rather then on topic-focus dichotomy only.

'the store' has a lower activation than 'the piano' in this segment after the utterance of (c) (in other words, in the hierarchy of activation 'a piano' would be more 'at hand' for the reference by a pronoun than 'the store'). The difference in the degrees of activation also may explain why 'a piano' is referred to by a pronoun in (c) neither in (2) nor in (3).

Similar considerations hold when examining utterances in (12) (quoted here from [Grosz et al. 1995], their ex. (15), which is assumed to be a part of a longer segment currently centered on John); 'he' in (a) and (b) refers to John):

Example 12

(a) *He$_{john}$ has been acting quite odd.*
C_b=John, C_f={John}
{$John^1$}

(b) *He$_{john}$ called up Mike yesterday.*
C_b=John, C_f={John>Mike}, Cont(U_a,U_b)
{$John^1$,$Mike^0$}

(c) *John wanted to meet him$_{mike}$ urgently.*
C_b=Mike, C_f={John>Mike}, Shift(U_b,U_c)
{$John^1$,$Mike^0$}

The backward-looking center of (a) is the referent of 'he' (=John); due to the Rule 2, the C_b of (b) is also 'he' referring to John; due to the assumption that if there is a single pronoun in the utterance this pronoun is the C_b of the given utterance, the C_b of (c) should be 'him'; this leads to an incoherence because 'him' does not refer to John but to Mike (a center shifting). The authors claim that the only possible interpretation of 'John' in (c) would be that this is another John. They admit, however, that such a claim may be too strong and we agree with them. We would ascribe the oddity of (c) in the context of (12) to the fact that 'John' preserves a high activation for a longer stretch of discourse (being referred to by a pronoun in (a), John either was introduced in the discourse before (a) was uttered, or has been activated by the situation of the discourse; in any case, 'John' is present 'on the scene' for the rest of the segment), but our framework would not exclude the interpretation of 'John' in (c) as being coreferential with 'he' in (a) and (b).

To summarize our comparison of the backward- and forward-looking centers and the TFA, we can say that two layers of description have to be distinguished: that of language structure (in which the backward looking center can be understood as the topic proper, i.e. as linguistically anchored) and that of a cognitively based hierarchy (the scale of forward looking centers seems to come closer to degrees of activation in the SSK than to a hierarchy of syntactic relations).

Centering theory relies on the theory of discourse structure proposed by Grosz & Sidner [1986] as for the definition of segment boundaries. Within this theory, the segments are defined on the basis of intentions, and segment boundaries are boundaries between parts of discourse answering different purposes. The centering rules only operate within a discourse segment. As far as we understand the authors, centering theory is not meant to contribute to the determination of segment boundaries.

Whereas centering constraints concern the relation between an expression in U_{n+1} and the $C_f(U_n)$, or they concern sequences of center transitions, the salience-based principles concern the relation between items in the underlying structure of U_{n+1} and the state of the SSK.

In the salience-based approach, on the other hand, a discourse segment is understood as a portion of a discourse which is "about the same or related entity(ies)". Discourse segmentation reflects the changes in SSK: roughly, as long as the same entities remain at top of the SSK hierarchy, the corresponding portion of the discourse may constitute a segment.

5. Concluding remarks

We have contrasted the two frameworks for discourse modelling, each stemming from a different theoretical background, but converging in their search for principles guiding discourse coherence. The centering theory certainly is one of the most penetrating and at the same time one of the most influential approaches to discourse structure. Our analysis has pointed to one possible issue of further investigation: a deeper collaborative study systematically analyzing the ranking of forward-looking centers with the hierarchy of the degrees of activation in the stock of knowledge assumed by the speaker to be present at the hearer's memory and to be shared by both the interlocutors.

References

Asher Nicholas [1993]. Referring to abstract objects in discourse. Kluwer Academic Publishers.

Cole P. and Morgan J. (eds.) [1975] Syntax and semnatics 3: Pragmatics. New York: Academic Press.

Daneš František [1970]. Zur linguistischen Analyse der Textstruktur. In: Folia linguistica, No. 4, pp. 72-78.

Daneš František [1985]. Věta a text [A sentence and a text]. Studie a práce lingvistické, Vol. 21.Prague.

Grice H.P. [1975]. Logic and conversation. In: Cole & Morgan [1975].

Grosz Barbara J. [1977]. The representation and use of focus in dialogue understanding. Technical note 151, SRI Project 5844, Stanford.

Grosz Barbara, and Sidner Candy L. [1986]. Attention, intentions and the structure of discourse. In: Computational linguistics, Vol. 12: pp. 175-204.

Grosz Barbara, Joshi Aravind and Weinstein Scott [1995]. Centering: a framework for modeling the local coherence of discourse. In: Computational linguistics, Vol. 21, No. 2, pp. 203-226.

Hajičová Eva [1987]. Focusing - A meeting point of linguistics and artificial intelligence. In Artificial Intelligence: Methodology, Systems, Applications, Jorrand, Ph., Sgurev, V. (eds.), Elsevier Science Publishers B.V., North-Holland.

Hajičová Eva [1993]. Issues of sentence structure and discourse patterns. Theoretical and computational linguistics, Vol. 2. Prague: Charles University.

Hajičová Eva and Vrbová Jana [1981]. On the salience of the elements of the stock of shared knowledge. Folia linguistica 15: 291-303.

Hajičová Eva and Vrbová Jana [1981]. On the role of the hierarchy of activation in the process of natural language understanding. In: Proceedings of Coling 1982, Horecký Jan (ed.), Prague: Academia, pp. 107-113.

Hajičová Eva, Hoskovec Tomáš and Sgall Petr [1995]. Discourse modelling based on hierarchy of salience. In: Prague Bulletin of Mathematical Linguistics: 64, pp. 5-24. .

Hajičová Eva, Korbayová Ivana and Kruijff Geert-Jan [to appear]. Salience in dialogues. In: Proceedings of the Dialogue in the heart of Europe conference, April, Prague.

Heim Irene [1983]. File change semantics and the familiarity theory of definiteness. In Meaning, use and interpretation of language, Bauerle, R., Schwarze, C., von Stechow, A. (eds.), Walter de Gruyter.

Hoffman Beryl [1996]

Joshi Aravind K. and Kuhn Steve [1979]. Centered logic: The role of entity centered sentence representation in natural language inferencing. In: Proceedings of the International Joint Conference on Artificial Intelligence, pp. 435-439.

Joshi Aravind K. and Weinstein Scott [1981]. Control of inference: Role of some aspects of discourse structure - centering. In: Proceedings of the International Joint Conference on Artificial Intelligence, pp. 385-387.

Kamp Hans and Reyle Uwe [1993]. From discourse to logic. Kluwer Academic Publishers.

Mann William C. and Thompson Sandra A. [1987]. Rhetorical Structure Theory: A theory of text organization. ISI Reprint Series, ISI/RS-87-190, USC Information Sciences Institute, University of Southern Carolina, Carolina, USA .

Sgall Petr [1967]. Functional sentence perspective in a generative description. Prague studies in mathematical linguistics, Vol. 2: pp. 203-225. Prague.

Sgall Petr, Hajičová Eva and Panevová Jarmila [1986]. The meaning of the sentence in its pragmatic aspects. Mey, J.L. (ed.), Reidel.

Sidner Candace L. [1979]. Towards a computational theory of definite anaphora comprehension in English discourse. Ph.D. Thesis, Artificial IntelligenceLaboratory, Massachusetts Institute of Technology, June. Technical Report 537.

Webber Bonnie L. [1983]. So what can we talk about now?. in Computational Models of Discourse, Brady, M., Berwick, R.C. (eds.), MIT Press, Cambridge, USA.

Weil Henri [1844]. translation in Weil [1887].

Weil Henri [1887]. The Order of Words in the Ancient Languages compared with that of the Modern Languages. Translated by Super, Charles W. Boston: Ginn & Company, Publishers. New edition with an introduction by Aldo Scaglione. John Benjamins b.v. (1978). Amsterdam, The Netherlands.

Appendix: Activation schemata

In the Figures 1 through 4 below, we provide schematic illustrations of the activation degree changes, according to the salience-based approach as discussed earlier, for the entities referred to in the texts in Examples 11, 2, 3 and 12, respectively.

In each picture, the markers of the entitities are the same as in the respective text. The highest degree of activation (0) is depicted at the left hand side, the lower degrees (1 to 8) successively to the right of it. Degress lower than 8 are not included in the picture. Each line corresponds to the state of activation in the SSK after processing of the respective utterance, i.e., the n-th line represents the state after the utterance of the n-th sentence in the text.

Figure 1
Activation schema for the discourse in Example 11.

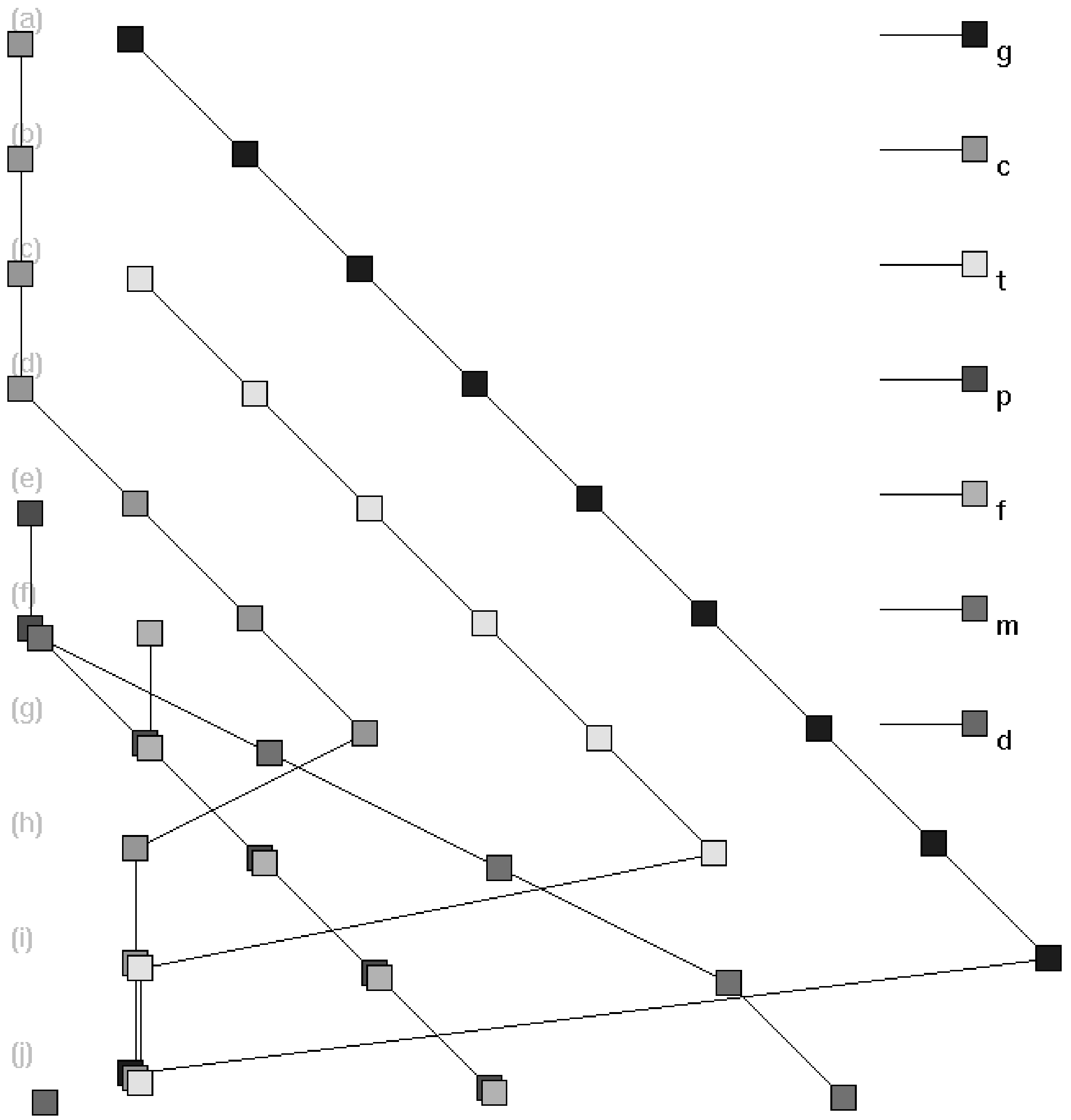

Figure 2
Activation schema for the discourse in Example 2.

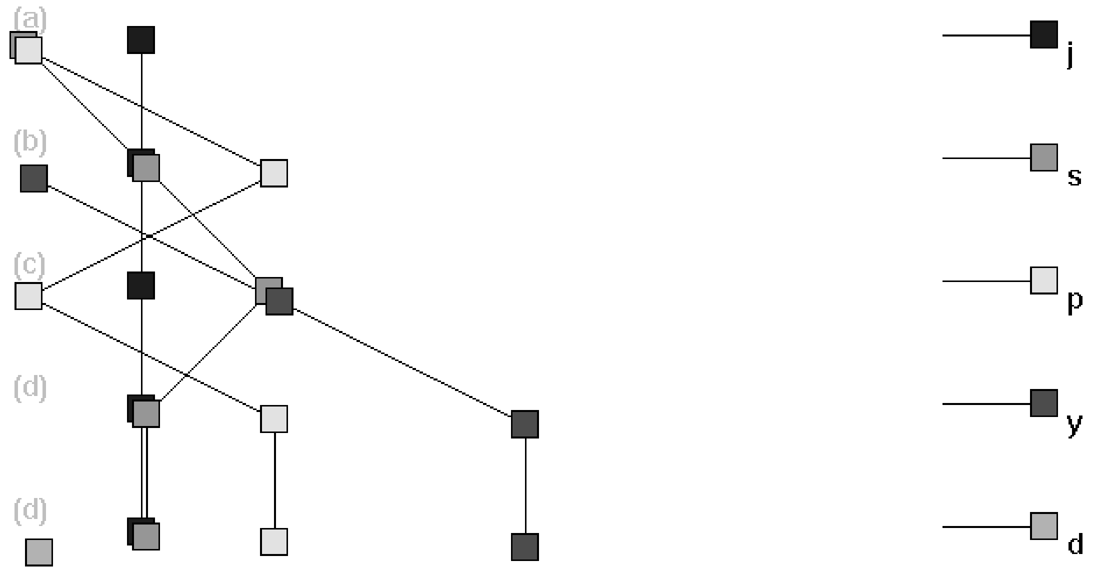

Figure 3
Activation schema for the discourse in Example 3.

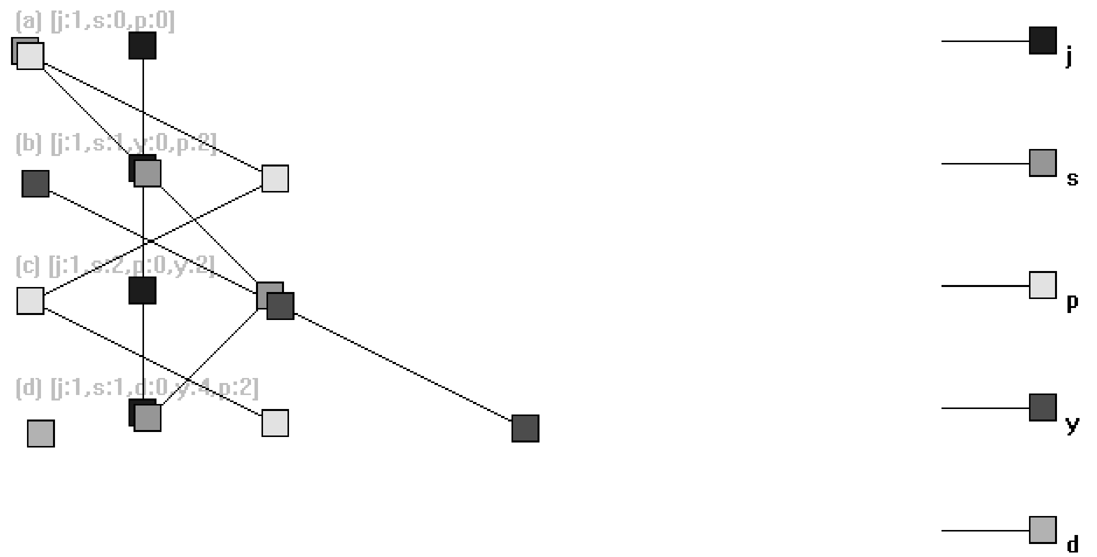

Figure 4
Activation schema for the discourse in Example 12.