

## Semantic conditioning of syntactic rules: evidentiality and auxiliation in English and Dutch<sup>1</sup>

Pieter A. M. Seuren<sup>2</sup> & Camiel Hamans<sup>3</sup>

Ever since the category of evidentiality has been identified in the verbal grammar of certain languages, it has been assumed that evidentiality plays no role in the grammars of those languages that have not incorporated it into their verb morphology or at least their verb clusters. The present paper attempts to show that even if evidentiality is not visible in the verbal grammar of English and Dutch, it appears to be a motivating factor, both historically and synchronically, in the process whereby evidential predicates are made to play a subordinate syntactic role with regard to their embedded subject clause. This process, known as AUXILIATION (Kuteva 2001), appears to manifest itself in a variety of, often successive, grammatical processes or rules, such as Subject-to-Subject Raising (the subject of the embedded clause becomes the subject of the main verb, as in *John is likely to be late*), V-ING (as in *The man stopped breathing*), Incorporation-by-Lowering (the evidential main verb is lowered on to the V-constituent of the embedded subject clause, as in *John may have left*), or Incorporation-by-Raising (also known as Predicate Raising),

---

<sup>1</sup> We are grateful to those who attended our presentation of this paper at the 40th Annual Meeting of the Societas Linguistica Europaea held at the University of Joensuu (Finland) from August 29th till September 1st, 2007, and whose critical observations and suggestions have helped us to improve its quality. We must equally acknowledge our debt to two referees, who pointed out mistakes and weaknesses in the original paper, which we have done our best to remedy. All remaining errors, shortcomings and inadequacies are, of course, our own responsibility.

<sup>2</sup> Max Planck Institute for Psycholinguistics.

<sup>3</sup> Dutch Delegation PES-group.

not or hardly attested in English but dominant in Dutch. A list is provided of those English (and Dutch) predicates that induce one of the above-mentioned auxiliations and it is checked how many of those have an evidential meaning. This is set off against evidential predicates that do not induce an auxiliation rule. It results that, for English and Dutch, lexical evidentiality is a powerful determinant for the induction of syntactic auxiliations.

**Keywords:** auxiliation, Subject-to-Subject Raising, V-ING, Incorporation-by-Lowering, Incorporation-by-Raising

## 1. Introduction

Relatively little attention has been paid in the linguistic literature of the past few decades to the question of the semantic conditioning of syntactic rules. A basketful of syntactic phenomena have come to be more or less standardly accepted as being accounted for by “rules”, under different names and in a variety of theoretical contexts. One of these standardly accepted rules is what we call here SUBJECT-TO-SUBJECT RAISING (SSR), discussed in Section 2. Whilst these rules have been motivated partly on syntactic and partly on semantic grounds, their existence has simply been accepted as a fact of life. No rationale has been considered other than that they happen to be induced or “triggered” by specific predicates, which can be listed. Thus, the fact that SSR is allowed to apply in a sentence like *He is likely to win*, but not in *\*He is probable to win*, is attributed to what is seen as the more or less fortuitous lexical peculiarity that the predicate *likely*, but not the predicate *probable*, is “marked” for SSR.<sup>4</sup> A variant of SSR in English is the so-called V-ING construction, as in *John stopped going to work on his bike* (not to be confused with *John stopped to go to work on his bike*, which has nothing to do with auxiliations).

Another such rule schema is what we call here INCORPORATION, which manifests itself in two forms, either Incorporation-by-Lowering (LINC) or

---

<sup>4</sup> In certain dialects of American English *probable* does induce SSR, so that one finds sentences like *He is probable to be guilty*.

Incorporation-by-Raising (RINC) (also known as Predicate Raising or PR). Subject-to-Subject Raising, V-ING, Incorporation-by-Lowering and Incorporation-by-Raising are similar in that all four result in the subject of the embedded clause becoming the subject of the main verb, whose original S-clausal term (i.e. an argument term consisting of an embedded S), now deprived of its proper subject, is subordinated to the main verb. In Subject-to-Subject Raising, the subjectless embedded clause becomes a Verb Phrase (VP); in Incorporation-by-Lowering and Incorporation-by-Raising it is amalgamated with the main verb into a V-cluster. More details are given in Section 2.

Subject-to-Subject Raising, V-ING, Incorporation-by-Lowering and Incorporation-by-Raising fit into the historical process of AUXILIATION (Kuteva 2001): the well-known universal tendency in the languages of the world to re-analyse subject-clause embedding whereby full predicates become auxiliary verbs or preverbal particles, and in some cases even suffixes.<sup>5</sup> When a lexical predicate takes an embedded S-clause as subject term and is semantically “weak” in the sense that it modifies the meaning of the embedded S-clause, one often sees a tendency, in the grammars of the languages of the world, to tear this subject clause apart, so that either the lower subject is raised to the status of main subject, as with Subject-to-Subject Raising and V-ING, or the underlying “weak” main predicate is lowered into the subject-clause and amalgamated with its main predicate, as with Incorporation-by-Lowering, or the “strong” lower predicate is raised to the higher “weak” predicate and amalgamated with it to form a strong enough V-cluster, as with Incorporation-by-Raising.

---

<sup>5</sup> Needless to say, the rules and rule schemata are not presented as themselves representing historical developments. On the contrary, they are synchronic in the strict Saussurean sense, as opposed to diachronic. The “underlying structures” posited are hypothetical reconstructions of mental cognitive structures underlying the corresponding surface sentences and connected to them by means of transformational mapping rules. There may, in some cases, be a historical parallel in that underlying structures may also have been historical predecessors, but no such link is postulated *per se*. When we say that the rules and rule schemata in question fit into the historical process of auxiliation, what we mean is that they typically occur at certain stages of the historical (or prehistorical) auxiliation process.

In English, Subject-to-Subject Raising is the preferred strategy for a development towards auxiliation. Incorporation-by-Lowering, as induced by the modals with their well-known defective paradigm, appears to represent a later stage. In Dutch, Incorporation-by-Raising is the preferred starting point, probably followed by Incorporation-by-Lowering (see Section 6). Since, as is shown in Section 2, incorporation, whether of the lowering or of the raising type, makes for a tighter bond of the higher weak predicate with the lower, embedded predicate, there seems to be a tendency for auxiliation to proceed from Subject-to-Subject Raising to incorporation, as appears to have happened with the English modals known for their incomplete paradigm. A further development may consist in reducing the underlying weak predicates, once they have become auxiliaries, into particles or affixes attached to what has now become the main predicate of the superordinate sentence but was originally the main predicate in the embedded clause.

When we speak of underlying weak main predicates, we have in mind predicates whose semantic contribution to the sentence consists mainly in modifying, one way or another, the “strong” lexical predicate of the embedded S-clause. Tense predicates (as in Malay and Bahasa Indonesia), aspectual predicates (*begin, stop, continue, result, used to, do/happen repeatedly, point-event aspect*), epistemic modals (*may, must*) or causative predicates, are cases in point, as are predicates of achievement (*manage*), but also predicates like English *appear, seem, be reputed*, or Dutch *schijnen* ‘seem’, *blijken* ‘turn out’, *geacht worden* ‘be reputed’.

The general question raised in the present paper is whether there is a correlation that is striking enough for one to suspect that it is of a causal nature between, on the one hand, those predicates that are marked for a given rule and, on the other, some recognisable and coherent semantic category. The more specific question is whether there is such a correlation between syntactic rules involved in auxiliation processes on the one hand and the semantic category of evidentiality on the other. If there is reason to suspect that such a correlation exists, this should be sufficient grounds for a larger scale statistical breakdown of the lexicon as a whole. We look at this question specifically from the point of view of English and Dutch. Neither the general nor the specific question has so

far received much attention in the literature on the theory of grammar (though Noël 2001 made a hesitant start).

When one looks at possible correlations, one hopes, of course, that any such correlation turns out to be one of EQUIVALENCE, in the sense that the semantic category and the class of predicates inducing the rule in question coincide:

- A predicate P induces the rule R if and only if P belongs to the semantic category C.

Here, “predicate P belongs to the semantic category C” and “P induces the rule R” are both necessary and sufficient conditions for each other, in the standard sense of these terms.

This, however, may be asking too much. A lesser demand would be for there to be a ONE-WAY IMPLICATION, as in:

- If a predicate P belongs to the semantic category C, P induces the rule R.

Where “predicate P belongs to the semantic category C” is a SUFFICIENT CONDITION for the rule R to be induced by P, and where “P induces the rule R” is standardly called a NECESSARY CONDITION for P to belong to the category C. Or vice versa, with inversion of the sufficient and the necessary conditions:

- If a predicate P induces the rule R, P belongs to the semantic category C.

But even a one-way implication may be too strong a criterion. It may be the case that membership of a given semantic category is neither a necessary nor a sufficient condition for the applicability of a given syntactic rule, or vice versa, but that there still is a greater-than-chance likelihood, given all relevant factors involved, that in the grammar of a language L a predicate P belonging to the semantic category C induces the rule R, or that a predicate P inducing the rule R belongs to the semantic category C.

It is important to note that this is not about token frequencies in given text corpora, but about the chance of a lexical predicate of a given semantic category inducing a particular rule of grammar at the type-level at which a language is grammatically defined. Any reader who might reject the notion of a type-level grammatical definition of a language in terms of rule systems (and it seems that some members of the profession hold such a radical-pragmatics view nowadays) will, besides having to cope with numerous other incongruities, find the present article incomprehensible or pointless.

## 2. The syntactic rules of SSR and LINC

Subject-to-Subject Raising is a rule whereby the nominal subject term of an embedded subject clause is lifted, or raised, to the position of grammatical subject of the higher embedding verb, whereby the embedded S-clause is demoted to the status of Verb Phrase. Examples of cases in English where SSR is taken to have applied are given in (1):

- (1) a. She is likely to arrive at nine.
- b. She appeared to be nervous.
- c. She tends to be nervous.
- d. The boy was considered to have failed the test.
- e. The man was rumoured to have killed his wife.
- f. The ship is due to depart tomorrow.
- g. SSR is believed to have applied in this very sentence.

The reasons why SSR is taken to have applied in these cases are partly of a semantic nature. There is a strong intuition that the apparent or surface subject terms of the sentences in question are not the “real” or semantic subject terms. The real subject terms are felt to be clauses rather than the noun phrases parading as subject terms in (1a–g). Thus, (1a–g) are understood as surface forms for underlying structures corresponding to (2a–g), respectively, even if some of these underlying structures may be awkward or even ungrammatical when used as surface forms:

- (2) a. that she will arrive at nine<sub>Pred</sub>[is likely]  
 b. that she was nervous<sub>Pred</sub>[appeared]  
 c. that she is nervous<sub>Pred</sub>[tends]  
 d. that the boy (had) failed his test<sub>Pred</sub>[was considered]<sup>6</sup>  
 e. that the man (had) killed his wife<sub>Pred</sub>[was rumoured]  
 f. that the ship departs<sub>VP</sub>[<sub>Pred</sub>is due]<sub>Adv</sub>[tomorrow]]  
 g. that SSR has applied in this very sentence<sub>Pred</sub>[is believed]<sup>7</sup>

A further semantic consideration has been the fact that selection restrictions ignore the predicates that induce SSR and cross over to the verb phrase under the particle *to*. This fact stands out with particular clarity in the case of idiomatic expressions, as is illustrated in the following examples:

- (3) a. The penny appears to have dropped.  
 b. That chapter is now taken to be closed.  
 c. The chips are understood to be down.  
 d. Aspersions were rumoured to have been cast.  
 e. All hell used to break loose.

The assumption of semantically underlying forms of the kind illustrated in (2) would immediately account for facts of this nature.

Such semantic considerations, however, are insufficient to explain all the syntactic phenomena at issue, which are quite complex. One notes, for example, that, in many cases, a surface-form alternative for SSR is afforded by extraposition, as exemplified in (4a–g). But this is not always possible. The sentences (4c,d,f), for example, do not allow for extraposition, though they do allow for SSR:

<sup>6</sup> As a rule, perfective infinitives ('to have X-ed') contain an ambiguity, in that in any paraphrase where the infinitive is replaced with a finite verb form, they may correspond either to a simple past tense or to a (plu)perfect (see McCawley 1971). It is for this reason that the form *had* has been added between brackets in both (2d) and (2e).

<sup>7</sup> One notes, incidentally, that the belief expressed in (1g) is correct, while that expressed in (2g) and (4g) is not. The reason is, of course, that the reference object of the phrase *this very sentence* is different in the former from the latter two cases.

- (4) a. It is likely that she will arrive at nine.  
b. It appeared that she was nervous.  
c. \*It tends that she is nervous.  
d. \*It was considered that the boy (had) failed the test.  
e. It was rumoured that the man (had) killed his wife.  
f. \*It is due that the ship departs tomorrow.  
g. It is believed that SSR has applied in this very sentence.

In the theory of SEMANTIC SYNTAX (Seuren 1996: 144–149) these and related phenomena are accounted for, in the context of a general theory of the Auxiliary System in verbal complexes, by means of specific assignments of rule features and argument structures to lexical predicates, given well-defined structural patterns into which the predicates in questions are given a place. This system also accounts for the fact that predicates like *follow* or *make sense* do not take SSR but do allow for extraposition, as is illustrated in the following examples:

- (5) a. It follows that Tom is ill.  
b. \*Tom follows to be ill.  
(6) a. It makes sense that Tom is ill.  
b. \*Tom makes sense to be ill.

And it accounts for the fact that *that*-clauses in initial position are sometimes allowed and sometimes disallowed (they are allowed when dominated by an NP-node):

- (7) a. That Tom is ill follows.  
b. That Tom is ill makes sense.  
c. That she will arrive at nine is likely.  
d. \*That she was nervous appeared.  
e. \*That she is nervous tends.  
f. \*That the boy (had) failed his test was considered.  
g. \*That the man (had) killed his wife was rumoured.  
h. \*That the ship departs is due tomorrow.  
i. \*That SSR has applied in this very sentence is taken.



A further curious problem arises from the fact that a number of English transitive verbs allow for subject raising both in object and in subject position (in passive clauses). Thus, English has not only (8a) but also (8b), and analogously for, *inter alia*, the verbs *assume, consider, estimate, expect, feel, figure out, find (out), know, prove, reckon, show, suppose, take, understand*, all of which allow for both SSR in passive clauses and Subject-to-Object Raising (SOR) in active clauses:

- (8) a. The director believes me to be honest.  
 b. I am believed to be honest by the director.

This fact would seem to support the existence of a post-raising passivisation rule, whereby the grammatical object of the active clause is turned into the subject of the passive clause, while the verb is passivised. This assumption, however, leaves unexplained why, at least in current standard English, some verbs, which allow for Subject-to-Object Raising, do not allow for SSR, though they do allow for passives when no object clause is involved, as in the following c-sentences:

- (9) a. The director would like you to wait a little.  
 b. \*You would be liked to wait a little by the director.  
 c. √You are liked by the director.  
 (10) a. The director prefers you to wait a little.  
 b. \*You are preferred to wait a little by the director.  
 c. √You are preferred by the director.  
 (11) a. The director wants you to wait a little.  
 b. \*You are wanted to wait a little by the director.  
 c. √You are wanted by the director.  
 (12) a. The director wishes (for) these changes to be made.  
 b. \*These changes are wished (for) to be made by the director.  
 c. √These changes are wished (for) by the director.

The verbs of this category are typically verbs denoting a desire on the part of the person denoted by the subject term. This would suggest that any systematic passivisation process precedes subject raising and that subsequent SSR is

constrained by semantic factors, so that, for example, verbs of wanting or wishing cannot induce SSR, though they may induce Subject-to-Object Raising.

This fits in with the fact that, conversely, there are verbs which allow for SSR but not for Subject-to-Object Raising. Thus, (13a) is ungrammatical (for most speakers) but (13b) is grammatical:

- (13) a. \*The professor claims this to be true.  
b. This is claimed to be true by the professor.

The same holds, at least in standard modern English, for *deem*, *hear*, *hold*, *report*, *rumour*, *repute*, *say*, *see*, *state* and to some extent also for *think* (whereby one notes that *be rumoured* and *be reputed* only occur as passives: there are no active verbs \**rumour* or \**repute*). Interestingly, the passive forms of these verbs are typical evidential ones.

It should be noted that clearly nonevidential passive predicates like *be required*, *be requested*, *be told*, *be persuaded*, *be asked*, *be invited*, and a few more, are not cases of SSR but of SUBJECT DELETION (SD). A sentence like *John was told to leave* is the passive of underlying ' $\emptyset$  told John<sub>x</sub> S[x leave]', where the oblique argument term *John<sub>x</sub>* becomes the subject of the passive form and *x* in S[x leave] is deleted by SD, so that S[x leave] is demoted to VP[leave]. When the idiom test of (3) above is applied to such sentences, the idiom disappears and something like a literal interpretation forces itself upon the listener: *The penny was told to drop*.

Facts such as these fit well into a theory where passivisation is a very deep process, possibly even at the propositional level, preceding the system of syntactic transformations. Questions such as these, however, together with the details of the syntactic system, are of less concern in the present context (for a full discussion see Seuren 1996). All that needs to be stressed here is the fact that the syntactic system accounting for these and related phenomena is based on the assumption of given structural patterns in underlying or "deep" structures and the judicious assignment of rule features to predicates. The system as such is considered to be insensitive to semantic features of any kind. What we want to investigate here is whether some semantic unity can be detected in those

predicates in the English lexicon that optionally or obligatorily induce SSR, and if so, whether evidentiality might be a key semantic notion.

A word is in order, at this point, about the distinction between a RAISING ANALYSIS and an INCORPORATION ANALYSIS. Since Ross (1969) and McCawley (1971), it has been generally assumed that the well-known modal auxiliaries *may*, *must*, *can*, *will*, *need*, *dare* and *ought to* likewise induce SSR. The argument is that a sentence like (14a) may be considered to correspond to an underlying (semantic) (14b):

- (14) a. His position may have become untenable.  
 b. that his position has become untenable<sub>Pred</sub> [may]

However, the assumption that (14b) underlies (14a), and analogously for all other cases involving the auxiliaries in question, is not sufficient to conclude that SSR has applied, because SSR is not the only way to reduce (14b) to (14a). One may also consider a process whereby the higher predicate *may* is incorporated-by-lowering into the verbal complex of the embedded subject clause in a specific well-defined way.

In fact, a LINC treatment for the English modals is preferable to an SSR treatment, in that the LINC treatment explains their well-known defective paradigm. In the LINC treatment, the modals are so-called MIDDLE AUXILIARIES, which means that they are base-generated between the two tenses ±PRESENT and ±PERFECTIVE, with the result that they may be followed by a perfective infinitive but can themselves only occur as present or simple past tense finite verb forms, lacking infinitivals and participials of their own.<sup>8</sup>

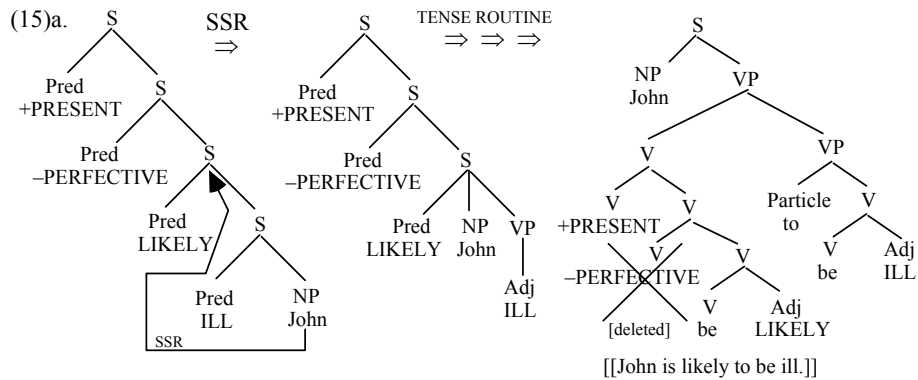
A third way of getting the lower subject into the position of higher subject consists in Incorporation-by-Raising, which would take the lower verbal cluster of (14b), *have become untenable*, and incorporate it into the higher predicate *may*, thus making the verbal cluster even more complex. This procedure is known as Predicate Raising, a rule widely attested in the complementation systems of natural languages but not for English. For a variety of reasons, however, Predicate Raising does not work for the English modals, mainly

---

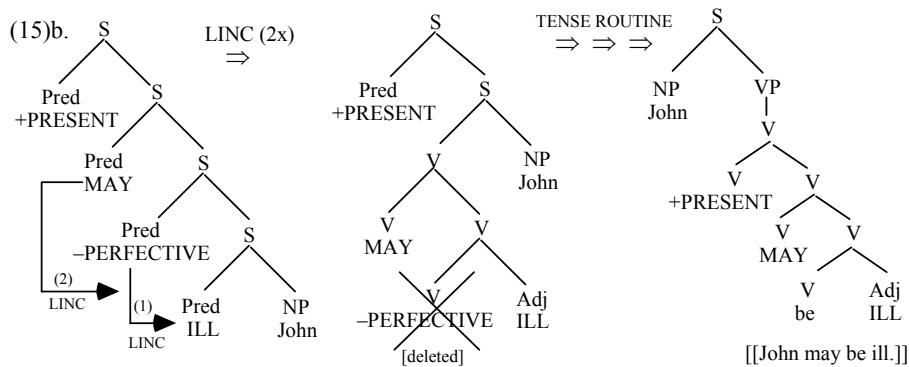
<sup>8</sup> For ample argumentation and a precise description, see Seuren (1996: 79–84, 91–156).

because it fails to explain their defective paradigm. Predicate Raising is discussed in greater detail in Section 6.

The two procedures, SSR and LINC, differ mainly in that SSR leads to a surface structure where the SSR-inducing predicate still functions as a predicate in its own right, followed by a verb phrase normally prefaced by the particle *to*, whereas LINC leads to one single clause without VP-embedding but with a composite verbal cluster in which the modal finds itself between the two tenses. The latter procedure thus results in a more tightly organised structure, a so-called V-island, which is closer to the morphology in that morphological processes prefer single categorial islands within which they can operate. The difference is visually shown in (15a,b).<sup>9</sup>



<sup>9</sup> How the surface NP-VP structure comes about, causing the subject-NP *John* to move from a bottom-right to a top-left position, as in (15a,b), is illustrated in (25a-f). Those who are interested in a more detailed description are referred to Seuren (1996), where they will find that what is called the TENSE ROUTINE is held responsible for this phenomenon.



It is argued in Seuren (1996: 111–116) that, for the English modals, a LINC-analysis is to be preferred over an SSR-analysis, mainly because a LINC-analysis, combined with a theory of Tense Incorporation, automatically accounts for the incomplete paradigms of the modals in question, which, as one knows, lack infinitives, participles and a third-person singular affix, and are followed by a so-called “bare” infinitive, that is, a, possibly perfective, infinitive without the particle *to* (except the idiosyncratic modal *ought to*).

Since the process whereby the paradigms of the modal auxiliaries in question became defective the way they are nowadays was not completed until around 1600 (Van Kemenade 1993, Barbiers 2002), one may surmise (though this hypothesis has never been investigated in any detail) that in older times these modals induced SSR and did not yet have the status of middle auxiliaries. They lost a part of their full morphological paradigms as a result of being regrammaticalised as middle auxiliaries inducing LINC.

A parallel development is found in the tense-modality-aspect systems of Creole languages, which are all of recent to very recent origin and, owing to the historical circumstances in which they came about, are subject to a constraint of semantic transparency (Seuren & Wekker 1985). Creole languages are characterised by the so-called TMA-system: they typically express tense, modality and aspect, in that order, through preverbal “particles” or markers which, in most cases, easily betray their lexical-verbal origin, much like English *be going to*, as in *The house is going to be sold* (Seuren 1981). Historically, these were then

united with the main (matrix) verb of the subject clause, probably by LINC, while, at the same time, they became obligatory elements of any finite sentence structure. These preverbal markers of tense, modality and aspect appear to occupy an intermediate position between, on the one hand, full words (“free forms” in the Bloomfieldian sense) and bound morphemes, in that they no longer function as independent words but have not reached the status of bound morphemes either, for which reason we call them particles or markers.<sup>10</sup>

An example is sentence (16), taken from Sranan, the English-based Creole of Surinam, and derived from an underlying (17a) via double cyclic LINC of the higher predicate into the predicate of the lower clause, as summarily shown in (17) (for technical details, see Seuren 1981):

- (16) Mi    ben    sa    gwe<sup>11</sup>  
      I     PAST  FUT    go away  
      ‘I would leave’

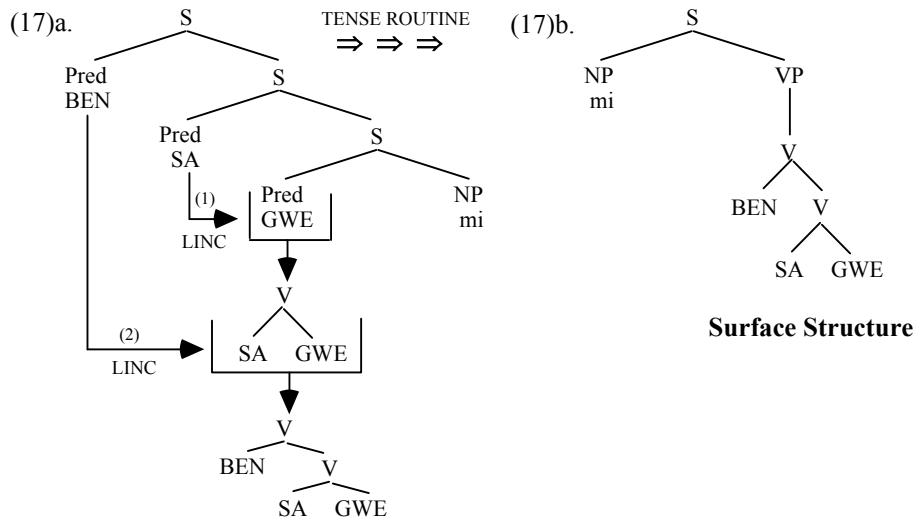
Similarly, one sees that the modality of futurity is expressed in English through the preverbal modal middle auxiliaries *shall* and *will*, whereas in many other languages, including the Romance languages, it has been incorporated into the affixal morphology of the verbal form. As regards the Romance languages, it is known that the affixes of futurity derive directly from the present and past tense paradigm of the Latin verb *habere* ‘have’, which, in one of its semantic specialisations, became the normal expression for futurity and was eventually regrammaticalised in the form of a set of affixes to be attached to the highest verb form in the auxiliary system of the sentence, contributing to the formation of the finite verb.<sup>12</sup>

---

<sup>10</sup> The switch from particle words to suffixes regularly involves postposition. Consider, for example, the postposed article morphemes in Danish and Rumanian, Latin *cum me* → *mecum*, English *towards home* besides *homeward*, Latin *habeo cantare* → *cantare habeo* → French *chanterai*, etc. etc. For more discussion, see Seuren (2009).

<sup>11</sup> *Ben* derives from English *been*, *sa* from *shall*, and *gwe* from *go away*.

<sup>12</sup> In another of its uses, Latin *habere* specialised for the aspectual tense of perfectivity, with the main verb as a past participle, in which function it has so far not developed beyond the stage of auxiliary verb, though it may well do so when one or more of the



Sometimes, predicates which induce Incorporation-by-Lowering or Incorporation-by-Raising and are thus united with the lower V-cluster see their original lexical content bleached in the sense that they, much in the way of satellites, become subservient to a lexically more central predicate to which they contribute aspectual, temporal or modal modifications. Well-known examples are the Dutch verbs *zitten* 'sit', *liggen* 'lie', *lopen* 'walk', *staan* 'stand', which are typically used as aspectuals but with meanings that are bleached to such an extent that their conditions of proper use have become hard to define.

Whether and, if so, to what extent auxiliatation — that is, the incorporation of semantically weak predicates into the main verbal cluster, as with Incorporation-by-Lowering or Incorporation-by-Raising, or their syntactic reinforcement by assigning them the subject term of the “strong” main predicate, as with Subject-to-Subject Raising or V-ING — is in any way functional is a question that is hard to answer. It would be gratifying, of course, to be able to maintain that such

---

languages in question lose their strict cultural and political regime and go through a lengthy period of overall instability. One knows that during such periods languages tend to be restructured, sometimes drastically.

processes spring from functional needs regarding the acquisition and storing of the language system or the production or interpretation of utterances in a social context. Yet any such claim is doomed to remain speculative as long as so little is known about what functionality actually implies and reliable experiments have not been carried out. What is functional for the processing involved in the production of an utterance is not necessarily also functional for the processing involved in comprehension or in language learning. For speakers, specific rules for the encoding of semantic content seem to be functional, whereas listeners and language learners are helped by semantic transparency — that is, by the avoidance of encoding rules. And other considerations may also count. But given our state of almost total ignorance regarding the ways utterances are produced or interpreted and regarding any possible influence of the social context, claims about functionality will, for the most part, have to remain either trivial or unsupported, despite the obvious validity of the functionality question and its importance for a true understanding of language and its use.

### 3. The notion of evidentiality

We now turn to the other half of the equation, the semantic category of evidentiality. Evidentiality is directly connected with speech acts of assertion. When uttered seriously, an assertion *A* creates a socially, and sometimes even legally, binding commitment on the part of the speaker, vis-à-vis those for whom *A* is meant, regarding the truth of the proposition expressed. Every new assertion establishes a new truth commitment and thus creates a new bit of social reality. Since, obviously, such truth commitments may have important consequences, it is in the nature of things that speakers should have means at their disposal to weaken or otherwise qualify the commitments entered into.<sup>13</sup> Evidentiality (French: *médiatisation*) is a semantic category whose expressions serve, in principle, to qualify the commitment assumed in virtue of the utterance of an assertion.

---

<sup>13</sup> Joseph aptly observes (2003: 324) that such commitment qualifications constitute a “cover one’s rear” strategy.



Every language has lexical predicates that specify, to a greater or lesser degree of precision, the limitation or degree of strength of the commitment at hand. In English, predicates of that kind occur in expressions like *probably*, *possibly*, *apparently*, *as far as I know*, *it is widely believed that*, *it is said that*, *it looks as if*, *one is inclined to believe that*, *it appears that*, *it is alleged that*, *it is rumoured that*, etc. Grammatically speaking, such commitment qualifications are not part of the speech-act component of the sentence underlying the assertive utterance but belong to the proposition with regard to which the truth-commitment is entered:

[<sub>speech act operator</sub> [I hereby assert that] <sub>proposition</sub> [it is rumoured that ...]]

Yet the semantic effect of such predicates amounts to a qualification of the commitment brought about by the speech act of assertion.

Although the question has no direct bearing on the analysis carried out in the present study, we limit ourselves to the expression of evidentiality of any kind in assertions, leaving other types of speech act out of account. It appears that in some languages evidentiality markers also occur in questions and even in commands (Aikhenvald 2003: 17). Yet such uses seem to be of a derivative nature, making for politeness or other forms of reservation in the sentences involved. Questions, commands and further speech-act categories have their own specific means of qualifying their socially binding speech-act force.

Ever since the grammars of non-European languages began to be studied, it has been known that in a great many languages of the world truth-commitment qualifications of a certain kind have become part of the grammatical machinery, in the sense that assertive sentences contain an obligatory slot specifying, in terms that are channelled into general semantic categories, the nature of the evidence on which the truth commitment is based. These slots tend to be incorporated into the verb morphology of the languages in question and to be manifest as bound morphemes. The morphological elements that qualify truth commitments as regards the nature of the evidence are called EVIDENTIALS or EVIDENTIAL MARKERS. The corresponding semantic category is generally called EVIDENTIALITY (Aikhenvald 2003: 1). Aikhenvald's (2003) typological survey shows that in languages which have them, evidentials tend to choose their

semantic load from the subcategories of  $\pm$ EYEWITNESS,  $\pm$ REPORTED,  $\pm$ INFERRED, in various combinations and sometimes with further subdistinctions.

The fact that evidentiality exists as a grammatical category has made us wonder if one may also distinguish a semantic category of evidentiality in the lexicon, with possible systematic consequences for the grammars of languages, not just in the morphology but also in the syntax. It would thus make sense to see if a class of EVIDENTIAL PREDICATES can be distinguished, which naturally lend themselves to some form of grammatical channelling and hence to grammaticalisation, whereby one obviously thinks of Subject-to-Subject Raising and incorporation (either by lowering or by raising).

For this to make sense, however, it is important that we be as precise as possible regarding the question of what we mean by evidentiality as a semantic category. We have seen that, as a morphological category, evidentiality is defined in terms of the criterial subcategories of  $\pm$ EYEWITNESS,  $\pm$ REPORTED,  $\pm$ INFERRED, sometimes with further subdivisions. In trying to identify evidentiality in full lexical predicates rather than verbal morphemes, we should stick to the same criteria, or else the enterprise loses its point.

So let us have a somewhat closer look at the semantics of evidentiality as a grammatical category. In doing so we do not want to be too restrictive, because then the analysis might lose relevance in the context of general theory. Nor do we want to be too lax in our criteria, because that might reduce the argument to vacuity.

According to Aikhenvald (2003: 11), “the semantic ‘core’ of evidentiality is source of information”. However, this “core” meaning is often extended to the epistemic modalities, in particular possibility, necessity and probability, as is evident from the data on verb morphology provided in the various contributions to Aikhenvald & Dixon (2003) and is fully acknowledged by Aikhenvald herself. So perhaps it would seem more appropriate to regard the NATURE OF THE EVIDENCE PROVIDED as the semantic core of evidentiality. This core then splits up into two subcategories, (a) a specification of the SOURCE of the information conveyed, which encompasses the criteria  $\pm$ EYEWITNESS and  $\pm$ REPORTED, possibly further subcategorised as hearsay, personal observation, experience, and other possible sources, and (b) the RATIONAL EVALUATION of the degree of

certainty supporting the truth commitment, corresponding to the feature  $\pm$ INFERRED. The value  $-$ INFERRED constitutes a plain, unqualified truth commitment; the value  $+$ INFERRED requires a further specification depending on whether the inference is one of necessity, possibility, or probability.<sup>14</sup>

#### 4. Epistemic modalities and other possible varieties of evidentiality

It is useful, at this point, to expatiate a little on the nature of epistemic modalities or inferences. Consider first the epistemic modality of possibility. Suppose a man called Smith utters (18), which is a clear case of epistemic possibility:

- (18) The fire may have been caused by witchcraft.

This implies two things. First, Smith asserts that, as far as his knowledge goes, it cannot be excluded that the fire was caused by witchcraft. But this is not enough, because (18) may well be compatible with Smith's knowledge at the moment of speaking, yet be false, in the ordinary sense in which we call assertions false if they are not true, on account of Smith's knowledge state being factually wrong. If, for example, Smith incorrectly thinks that witchcraft may cause what it says it does, then it is compatible with his imperfect knowledge state that the fire in question was caused by witchcraft. Yet (18) is still false, and the reason for its falsity seems to be the fact that the supporting knowledge state is factually incorrect. In general terms, POSSIBLE (Clause), in the epistemic sense of POSSIBLE, implies (a) that Clause is compatible with what the speaker knows

---

<sup>14</sup> Noël (2001) makes a distinction between SPECIFIC and NONSPECIFIC source of information and he seems to imply that only one of them makes for evidentiality, though his text is curiously evasive in this respect. But no matter whether Noël does or does not restrict the notion of evidentiality either way, such a restriction should be rejected. It would mean that *John is thought to be dishonest by his own wife* is, but *John is thought to be dishonest* is not a case of evidentiality (or the other way round). Such a view seems to us to be basically misguided, even though it is not to be excluded that some languages make a morphological distinction between specific and nonspecific source of information. But in such languages, both categories will count as evidential.

at the moment of speaking and (b) that the speaker's knowledge state is factually correct. Failing either or both of these two conditions, POSSIBLE (Clause) must be deemed false.

Analogously for epistemic necessity. When Smith utters (19), he implies (a) that, as far as his knowledge goes, it follows by logical, inductive or causal necessity that the fire was caused by witchcraft, and (b) that his knowledge state is factually correct. Failing either or both of these conditions, (19) is false:

(19) The fire must have been caused by witchcraft.

In general terms, NECESSARY (Clause), in the epistemic sense of NECESSARY, implies (a) that Clause logically, inductively or causally follows from what the speaker knows at the moment of speaking and (b) that the speaker's knowledge state is factually correct.

And similarly again for epistemic probability. (20) implies (a) that, as far as the speaker knows, the chance that the fire was caused by witchcraft is greater than the chance that it was not, and (b) that the speaker's knowledge state is factually correct.

(20) The fire was probably caused by witchcraft.

Here again, PROBABLE (Clause) implies in general (a) that it follows from the speaker's knowledge state that there is a considerably greater chance of Clause being true than of Clause not being true and (b) that the speaker's knowledge state is factually correct.

Prediction is a further epistemic category. In many languages, including English, prediction is regularly expressed through the modal auxiliary of futurity, as in:

(21) The fire will have been caused by witchcraft.

(21) is not a statement about the future but about the past, and it is used to express the speaker's prediction, based on available knowledge, that the fire was

caused by witchcraft. In this sense, *will* is clearly an evidential. (One notes that *shall* cannot be used in this sense.)

As has been said, we consider the epistemic modalities of possibility, necessity, probability and prediction, and their negations, to be instances of evidentiality, falling under the subcategory of rational evaluation of the degree of certainty supporting the truth commitment. This subcategory has further subdivisions, whereby inferences of necessity, possibility, probability and prediction form the ultimate subcategories.

One might wonder if “weak” or “ancillary” predicates like *be due to*, *tend to*, *happen to* or *used to* should also be reckoned among the evidentials. The predicate *be due to* might be taken to involve a degree of reportedness, and *tend to*, *happen to* and *used to* may be thought of as resulting from an inference to do with regularity of frequency. If these and related predicates are indeed taken to be of an evidential nature, the picture becomes even more pronounced, as one can see from Table 1 below, which contains a well-nigh complete list of all predicates in English that take either Subject-to-Subject Raising or some other auxiliatioin rule. Yet we have preferred to follow the, methodologically speaking, conservative policy of restricting the assignment of evidentiality to clear cases, thus making the argument more convincing.

## 5. The correlation between auxiliatioin (SSR/LINC) and evidentiality in English

We now come to the first half of our main question: is there, in English, a striking enough relation between lexical evidentiality and auxiliatioin – that is, Subject-to-Subject Raising, with Incorporation-by-Lowering and V-ING as its little brothers – for us to suspect that relation to be of a causal nature? (Remember that Incorporation-by-Raising does not occur in English subject clauses.) To answer this question, we consider predicates which satisfy both of the following conditions:

(a) They are lexically specified as taking embedded subject clauses in their underlying argument structure, which is thus of the form “P(Clause)”, where P is the predicate and Clause is the embedded subject clause.

(b) They induce a syntactic auxiliation rule of the kind specified above (SSR, LINC, or V-ING).

We check whether they are semantically evidential in that they express a degree of certainty regarding the correctness of either the source of information or the mental reasoning leading to the embedded subject clause. Then we try to determine how many evidential P(Clause) predicates do not induce a syntactic auxiliation rule (SSR, LINC, or V-ING).

Table 1 lists all English predicates with the underlying argument frame P(Clause) which induce an auxiliation rule and can thus be marked as +Aux. The list is well-nigh complete and has been compiled on the basis of many years of experience with English grammar and of the available literature on the subject. No list is given of semantically evidential predicates that do not induce an auxiliation rule, for the simple reason that no such list is available and it would transgress the boundaries of the present work to compile one. But we can follow the established practice in linguistics of going by example and, above all, by counterexample.

Table 1 is subdivided into three categories, adjectives, modals/aspectuals and passivised predicates, all of the underlying form P(Clause) and all inducing an auxiliation rule (they are +Aux). The adjectival predicates are given first, followed by the modal and other non-passive aspectual predicates, which are followed by the passive predicates (with the passive auxiliary *be*). It is indicated for each predicate which auxiliation rule is induced: SSR, LINC, or V-ING.

The number of P(Clause) adjectives in English which induce an auxiliation rule (always SSR) is extremely small: just four, or five if one includes American English *probable* (see note 4). Table 1 lists them all. It also lists all P(Clause) modals/aspectuals,<sup>15</sup> as well as all P(Clause) passivised predicates, give or take a few doubtful or dialectally varying cases, that have been found to induce an auxiliation rule.

One may hesitate in the case of the predicates *be hypothesised* and *be guessed*, which only doubtfully induce SSR in the sense that sentences like (\*?)*Harry was*

---

<sup>15</sup> The agentive modals *can*, *must*, *shall*, *will*, *ought to* have not been included, because it is not evident that they take a P(Clause) underlying argument structure. Nothing changes, however, in our overall conclusion if they are included.

*guessed/hypothesised to live in London* are only doubtfully grammatical. They are also doubtfully evidential. What appears to us to be their ambiguous status is probably due to the fact that evidentiality is not one monolithic notion but rather one that allows for degrees in that one assumes there to be core and peripheral evidentials. Speakers, or groups of speakers, may then differ according to whether they do or do not extend the assignment of a rule that is typically associated with core evidential predicates to predicates that are more peripherally evidential. In the light of our preference for a restrictive policy, we have excluded these two predicates from Table 1. The predicates that induce Subject-to-Object Raising besides SSR are marked as such.

As regards the adjectives, we see that there is, with one exception, an implicational relation from +Aux (+SSR) to +Evid, the one exception being the adjective *due*. This means that the small class of four (five) +Aux adjectives found in English is at least heavily coloured by evidentiality, since all but one are +Evid. But this one exception shows that, strictly speaking, there is no implicational relation of the kind described. The other way round there is nothing remotely like an implicational relation, given that there are quite a few +Evid adjectives with the required P(Clause) argument structure that do not induce any auxiliation rule (though many occur as high sentence adverbials), such as *apparent, clear, predictable, obvious, evident, necessary, possible* or (British) *probable*. All one can say, therefore, is that when an English adjective gets involved in auxiliation there is a fair chance that it will be an evidential.

certain	+SSR	+Evid	be declared	+SSR/SOR	+Evid
due	+SSR	-Evid	be deemed	+SSR	+Evid
likely	+SSR	+Evid	be demonstrated	+SSR	+Evid
probable (Amer.)	+SSR	+Evid	be disclosed	+SSR	+Evid
sure	+SSR	+Evid	be estimated	+SSR/SOR	+Evid
may	+LINC	+Evid	be expected	+SSR/SOR	+Evid
cannot	+LINC	+Evid	be felt	+SSR/SOR	+Evid
must	+LINC	+Evid	be figured out	+SSR/SOR	+Evid
will (prediction)	+LINC	+Evid	be found (out)	+SSR/SOR	+Evid
prove	+SSR	+Evid	be heard	+SSR	+Evid
seem	+SSR	+Evid	be held	+SSR	+Evid
turn out	+SSR	+Evid	be imagined	+SSR	+Evid
appear	+SSR	+Evid	be intended	+SSR/SOR	+Evid
promise	+SSR	+Evid	be judged	+SSR/SOR	+Evid
threaten	+SSR	+Evid	be known	+SSR/SOR	+Evid
ought	+LINC	-Evid	be meant	+SSR/SOR	+Evid
will/shall (fut.)	+LINC	-Evid	be presumed	+SSR	+Evid
happen	+SSR	-Evid	be proved(en)	+SSR/SOR	+Evid
tend	+SSR	-Evid	be reckoned	+SSR/SOR	+Evid
stop	+V-ING	-Evid	be reported	+SSR	+Evid
continue/go on	+V-ING/+SSR	-Evid	be reputed	+SSR	+Evid
start/begin	+V-ING/+SSR	-Evid	be revealed	+SSR	+Evid
be going to	+SSR	-Evid	be rumoured	+SSR	+Evid
used to	+SSR	-Evid	be said	+SSR	+Evid
be acknowledged	+SSR	+Evid	be seen	+SSR	+Evid
be alleged	+SSR	+Evid	be shown	+SSR/SOR	+Evid
be argued	+SSR	+Evid	be stated	+SSR	+Evid
be ascertained	+SSR	+Evid	be stipulated	+SSR	+Evid
be assumed	+SSR/SOR	+Evid	be supposed	+SSR/SOR	+Evid
be believed	+SSR/SOR	+Evid	be surmised	+SSR	+Evid
be claimed	+SSR	+Evid	be taken	+SSR/SOR	+Evid
be concluded	+SSR	+Evid	be thought	+SSR	+Evid
be considered	+SSR/SOR	+Evid	be understood	+SSR/SOR	+Evid

Table 1. Survey of correlation in English between auxiliation and evidentiality

It is perhaps relevant to note, in this connection, that English is the only European language that allows adjectives to induce an auxiliation rule (always SSR) and that in the rest of the world this phenomenon seems to be extremely rare. If this is correct, it suggests the possibility of an implicational scale for



languages in general.<sup>16</sup> As it is, it is unknown to what extent, or why, adjectives may be supposed to resist SSR in the languages of the world. If we look round among the European languages, we find that the phenomenon of adjectives inducing SSR or any other auxiliation rule is extremely rare and perhaps restricted to just English, as opposed to adjectives inducing OBJECT-TO-SUBJECT RAISING (OSR), as in sentences like (22 a), taken to be derived from an underlying (22 b):

- (22) a. This problem is hard to solve.  
 b. <sub>s</sub>[∅ solve this problem] <sub>pred</sub>[be hard]

In such sentences, the embedded subject clause is incomplete in that it lacks a subject term and is thus perforce reduced to Verb-Phrase status.

The picture is somewhat different for the twenty-two +Aux modal/aspectual predicates found in English. If the predicates *ought*, futurity *will/shall*, *happen*, *tend*, *used*, *stop*, *continue*, *go on*, *start*, *begin* and *be going to*, in their “impersonal” senses in which they take a subject S-clause as argument, are denied the status of evidentials, as they should be, one must conclude that lexical evidentiality is not a necessary condition for the induction of auxiliation (or that auxiliation-induction does not imply lexical evidentiality). This is clear from examples like those in (23), whose non-evidential predicates all induce SSR or LINC:

- (23) a. Dorothy will arrive at six.  
 b. Dorothy happened to be there at the right time.  
 c. Dorothy tends to forget birthdays.

---

<sup>16</sup> Unfortunately, linguistic typology has so far concentrated exclusively on mere surface phenomena, leaving out of consideration any possible theoretical machinery that might account for them. We believe this to be shortsighted, since whatever linguistic universals there are must primarily be properties of linguistic SYSTEMS, not of their PRODUCTS. The restriction to surface phenomena is also counterproductive, as is shown, for example, by the fact that typologists have so far been unable to present a workable definition of serial verbs, which are clearly definable in terms of rule systems, not in terms of surface structure alone (see Seuren 1990, 1991).

- d. Dorothy used to bake her own bread.
- e. Dorothy began being/to be in danger when she contradicted her boss.

But the “impersonal” SSR-inducing predicates *threaten* or *promise*, as in *Dorothy threatened to be killed by the avalanche* or *Dorothy promised to recover after the operation*, are good candidates for the status of evidential predicate. Their equivalents are found in many languages to be usable as impersonal intransitives and then to become subject to auxiliation in one form or another (see Cornillie: 2004, 2005; Heine & Miyashita 2008).

Given this, we see that roughly half the +Aux modals/aspectuals are +Evid. But are there other P(Clause) +Evid modals/aspectuals not inducing an auxiliation rule? The answer depends on whether one should consider predicates like *follow*, *emerge*, *result* or *make sense*, which can only occur with fully fledged *that*-clauses, to be evidentials.<sup>17</sup> If one does, which is reasonable, no implicational relation can be posited from +Evid to +Aux or vice versa. But it does look as if the number of P(Clause) +Evid verbs that do not induce an auxiliation rule is extremely small: the verbs just mentioned, *follow*, *emerge*, *result* and *make sense*, are the only ones we have been able to trace and they do not seem to be “core” evidentials the way predicates like *seem*, *appear* or the epistemic modals are. Therefore, unless we have been looking with our eyes closed, we may conclude that for any P(Clause) +Evid modal/aspectual verb, chances are that it will be a +Aux verb. Evidentiality is thus seen to be one of several motivating factors for auxiliation.

As regards the English passivised P(Clause) predicates the result is more pronounced. Here we see that all such predicates that induce an auxiliation rule (always SSR) are +Evid. We also find, however, that there are quite a few +Evid predicates of that category which do not induce any auxiliation rule. Examples are *be affirmed*, *be stated*, *be proclaimed*, *be observed*, *be asserted*, *be attested*, *be predicted*, *be guessed*, *be conjectured*, *be written*, and many more. Since no total count has been made of English +Evid but -Aux P(Clause) passivised predicates and since it is hard to make a reliable estimate of how many there are, we cannot

---

<sup>17</sup> *Make sense* also occurs with a subjectless subject clause, as in *It makes sense to inspect the attic*. This construction does not seem to be an instance of auxiliation but of the syntax of subjectless clauses (see also the example (22)).

now say precisely how strong the connection is between auxiliation and evidentiality for this class of predicates. But we can reliably say that if an English P(Clause) passivised predicate induces an auxiliation rule (SSR), it is an evidential.

The overall conclusion is that evidentiality is a strong motivating factor for the syntactic rules that are involved in the process of auxiliation. This conclusion is based on the fact that, according to Table 1, in just over 85 percent of the cases where a predicate induces an auxiliation rule that predicate is also an evidential.

## 6. The correlation between Predicate Raising and evidentiality in Dutch

Dutch, though similar to English in many ways, differs from English in that Subject-to-Subject Raising is not a productive rule in its syntax,<sup>18</sup> while Incorporation-by-Lowering, as far as it leads to a verbal cluster, seems to be restricted to the tenses (treated as predicates in this theory) and the passive auxiliary *worden* (see Seuren 1996: 219–269). In Dutch, the overwhelmingly dominant rule for the creation of verbal clusters, and hence also for auxiliation, is the rule of Predicate Raising (PR; see Seuren 1996, 2003). This rule is diagnosed (Evers 1975; Seuren 1972) on grounds of, sometimes lengthy, verbal clusters, all arranged towards the end of the clause, while the pertinent nominal arguments are arranged to the left, in serial order. Moreover, such verbal strings do not contain past participles, which are replaced by infinitives (so-called

---

<sup>18</sup> The rule of SSR is taken to occur, in the syntax of Dutch, in the tense routine, as explained above. Apart from that, SSR only occurs with the evidential passives *geacht worden* ‘be thought to’, *verondersteld worden* ‘be supposed to’ and *gezegd worden* ‘be said to’, as in (i)–(iii):

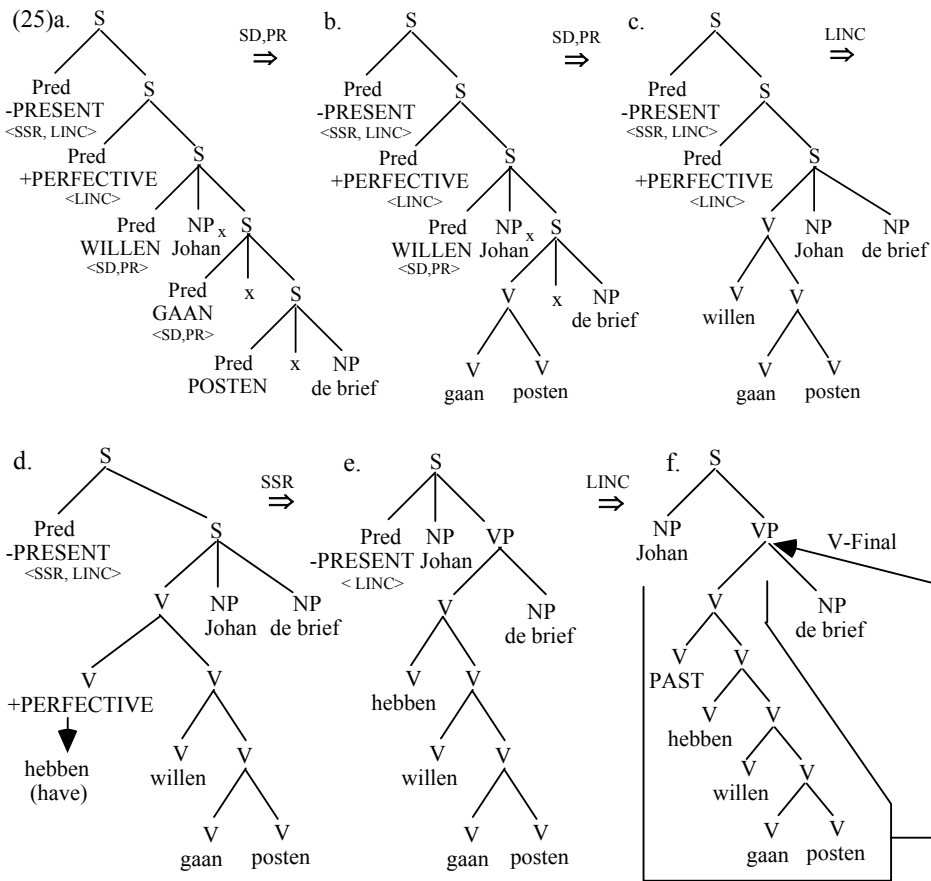
- (i) ... *dat Jan wordt geacht ziek te zijn* ‘... that Jan is thought to be ill’
- (ii) ... *dat Jan wordt verondersteld ziek te zijn* ‘... that Jan is supposed to be ill’
- (iii) ... *dat Jan wordt gezegd ziek te zijn* ‘... that Jan is said to be ill’

There are no signs that the rule is spreading to other predicates. On the contrary, *gezegd worden* is on the way out and *geacht worden* as well as *verondersteld worden* are restricted to the higher social registers.

infinitives *pro participio*, or, in German terminology, *Ersatzinfinitive*). Examples are:

- (24) a. ... *dat Johan zijn vriend de hond de krant* <sub>v</sub>[*heeft laten leren halen*]  
lit. '... that Johan his friend the dog the newspaper has let teach fetch'  
'... that Johan has let his friend teach the dog to fetch the newspaper'  
b. ... *dat Johan zijn dochter voortdurend* <sub>v</sub>[*heeft staan afluisteren*]  
lit. '... that Johan his daughter all-the-time has stand eavesdrop'  
'... that Johan has stood eavesdropping on his daughter all the time'

In terms of the theory of Semantic Syntax (Seuren 1996), the process giving rise to these verbal clusters is fairly simple. In this theory, the grammar consists in a machinery that transforms any given well-formed semantic input tree, or the SEMANTIC ANALYSIS (SA) into an appropriate surface structure. Given a well-formed SA, there is a cyclic process, starting from the most deeply embedded predicate, whereby the lowest Pred-constituent is right-adopted by the higher predicate, recursively, till all PR-inducing predicates have been processed, as shown in (25 a–f) for the clause ... *dat Johan de brief had willen gaan posten* '... that Johan had wanted to go post the letter'. All remaining material at any level is shifted upwards in the order given. Postcyclic V-Final for subordinate clauses, shifting the entire V-cluster to the far right, finishes the job. (For Dutch and German, subordinate clauses are standardly used for demonstration, because in main clauses only the nonfinite part of the V-cluster is moved to the right, which tends to mask the effects of PR.)



The cyclic rule of Predicate Raising applies, obligatorily or optionally, to a great many non-evidential transitive and intransitive lexical verbs taking an S-complement (see Seuren 1985: 184 for a complete list), such as *helpen* 'help', *leren* 'learn, teach', *zien* 'see', *horen* 'hear', *voelen* 'feel', *kijken* 'look', *laten* 'let, allow', *weten* 'manage', *vermogen* 'be able', *durven* 'dare', *wagen* 'dare', *verlangen* 'long for', *willen* 'want', *weigeren* 'refuse', *proberen* 'try', *trachten* 'try', *wezen* 'have gone to', *staan/zitten/liggen/lopen* 'stand/sit/lie/walk while ...', *komen*

'come to', *denken* 'think', *menen* 'believe', *doen* 'cause, make', *vergeten* 'forget', *hopen* 'hope', *(be)horen* 'should, ought to', *bedoelen* 'mean to', the agentive modals *moeten* 'must', *kunnen* 'can', *hoeven* 'need' and *mogen* 'be allowed', the aspectual verbs *beginnen* 'begin', *blijven* 'continue', *gaan* 'be going to', *plegen* 'do habitually', the futurity verb *zullen* 'will'.

The rule of Predicate Raising also applies to a number of core evidentials, such as *lijken* 'appear', *schijnen* 'seem', *blijken* 'turn out', *heten* 'be rumoured', predictive *zullen* 'will', and epistemic *moeten* 'must', *kunnen* 'may' and *hoeven* 'need', which all induce obligatory PR. The pair of evidentials *dreigen* 'threaten' and *beloven* 'promise, look like, be destined to' likewise belong to the class of P(Clause) predicates that induce PR.

As in English, there is a, probably very small, class of +Evid P(Clause) predicates that do not induce any auxiliation rule and thus require the equivalent of a full *that*-clause. Examples are *volgen* 'follow', *kloppen* 'be correct', *uitkomen* 'be disclosed'. Exactly how many more there are is not known, but the expectation is that their number will turn out to be extremely limited.

Remarkably, the core evidentials *lijken* 'appear', *schijnen* 'seem', *blijken* 'turn out' and *heten* 'be rumoured' are beginning to resist occurring in a perfective tense. A sentence like (26) is distinctly awkward if not already ungrammatical:

- (26) \*?... *dat Johan heeft schijnen te willen ontsnappen*  
lit. '... that Johan has seem to want escape'  
'... that Johan has seemed to want to escape'

This fact suggests that these verbs are going the way of the English modals, with their defective paradigm, and are beginning to turn into middle auxiliaries. This would place them in a higher position on the auxiliation cline than the others, which also cause verb clustering to take place, but from an ordinary subject-S-complementation position, which leaves their morphological paradigm unimpaired.

A further remarkable fact, unexplained in all theories of grammar except Semantic Syntax, is the undoubted grammaticality of the Dutch sentences (27a) and (28a), as opposed to the clear ungrammaticality of their literal German translations (27b) and (28b), respectively. In both cases, Dutch allows for an embedded infinitival *zullen* 'will', whereas German does not allow for an

embedded infinitival *werden*, even though *werden* is the literal German translation equivalent of Dutch *zullen*. This is remarkable because, as far as PR is concerned, German and Dutch run largely parallel (except that German has predominantly left-branching V-clusters, whereas the Dutch V-clusters are, in principle, right-branching).

- (27) a. ... *dat Johan het had zullen doen*  
 lit. '... that Johan it had will do'  
 '... that Johan would have done it'
- b. \*... *daß Johann es hätte tun werden*  
 lit. '... that Johan it had do will'  
 '... that Johan would have done it'
- (28) a. *Johan had beloofd het te zullen doen.*  
 lit. 'Johan had promised it to will do'  
 'Johan had promised to do it.'
- b. \**Johann hatte versprochen, es tun zu werden.*  
 lit. 'Johan had promised it do to will'  
 'Johan had promised to do it.'

This clearly shows that German *werden* has now fully developed into a middle auxiliary, to be lowered by Incorporation-by-Lowering (LINC) into a position between the two tenses, just like the English modals, whereas Dutch *zullen* has not (yet) been through this process and thus still takes Predicate Raising.

Dutch lacks the large class of passive evidentials inducing an auxiliiation rule that one finds in English. The only cases are the high social register passives *geacht worden* 'be thought' and *verondersteld worden* 'be supposed' and the now partly antiquated *gezegd worden* 'be said', which induce (optional) Subject-to-Subject Raising. The other passive evidentials do not induce any auxiliiation rule and thus require the equivalent of a full *that*-clause.

The overall picture for Dutch is thus fairly straightforward: most non-adjectival, non-passive single-word P(Clause) evidential predicates are subject to auxiliiation (Predicate Raising), while no adjectival evidential predicate is and only three passive evidentials are (involving SSR). Those non-adjectival, non-passive P(Clause) evidentials that are subject to auxiliiation all induce Predicate Raising, while of the passive evidentials only *geacht worden* 'be

thought', *verondersteld worden* 'be supposed' and *gezegd worden* 'be said' induce Subject-to-Subject Raising. Some evidentials, in particular *lijken* 'appear', *schijnen* 'seem', *blijken* 'turn out' and *heten* 'be rumoured', seem to be on their way towards becoming middle auxiliaries, which would imply that they will be LINC-inducers before long. Therefore, Dutch, like English, appears to display a strong bond between evidentiality and auxiliation, given the syntactic constraints that appear to hold for this language.

## 7. Discussion

It has been shown that, at least in English and in Dutch, there is a nontrivial relation between, on the one hand, those P(Clause) predicates that induce some auxiliation rule and evidentiality as a lexical semantic category. No precise statistical correlation has been computed, as this would require a breakdown of the entire lexicons of the languages concerned – an exercise the authors are happy to leave to others.

Whether this conclusion is in any way noteworthy depends on what one wants a grammatical theory to achieve. From a preliminary, global and intuitive point of view one would no doubt expect that those predicates that do not carry the main propositional information but place that information in a certain epistemic, temporal or aspectual light will be robbed of their high hierarchical status as first predicate in the underlying semantico-syntactic structure and be somehow placed in a syntactically ancillary position in the surface structure expressing the corresponding meanings. But such an *a priori* expectation must then be shown to be fulfilled in terms of a precise, formal description of the syntax of the languages involved, which, as one will realise, is far from a trivial undertaking.

Formal theories of syntax are a relatively recent phenomenon, originating from American structuralism in grammar as initiated by scholars like Edward Sapir and above all Leonard Bloomfield, and later turned into generative transformational grammar. In this tradition, the semantic aspect of language has, unfortunately, steadfastly been underemphasised, with the exception of the Generative Semantics movement, whose heyday, in the late 1960s and the early



1970s, was intense but short, swamped as it was by the successive versions of Chomskyan autonomous-syntax, which never had much of an eye for any form of semantic conditioning of syntactic rules and was thus not the proper vehicle for a formal or technical confirmation of preliminary, global intuitions regarding such conditioning. Generative Semantics, which has survived in the theory of Semantic Syntax, is a much more likely candidate for such an enterprise.

In any case, the formal confirmation of preliminary intuitions is noteworthy because it is highly relevant. It can only be taken to be irrelevant and not worthy of note if one rejects the entire notion of a formal theory of grammar or syntax, as, apparently, some linguists have recently taken to doing. These linguists may adopt the attitude that was current a century ago and be content with a linguistic description in intuitive, largely semantic, terms, but they will be unable to gain a deeper insight into the causal factors behind such systematic observations as are known from the formal syntactic literature and are also reported in the present article.

## References

- Aikhenvald, Alexandra. 2003. Evidentiality in typological perspective. In Aikhenvald & Dixon, eds. 1–31.
- Aikhenvald, Alexandra & Robert M. W. Dixon, eds. 2003. *Studies in evidentiality*. Amsterdam: Benjamins.
- Barbiers, Sjef. 2002. Current issues in modality. In Sjef Barbiers, Frits Beukema & Wim van der Wurff, eds. *Modality and its interaction with the verbal system*. Amsterdam: Benjamins, 1–18.
- Cornillie, Bert. 2004. *Evidentiality and epistemic modality in Spanish (semi-)auxiliaries. A functional-pragmatic and cognitive-linguistic account*. Leuven: University of Leuven PhD thesis.
- Cornillie, Bert. 2005. Subjectivity in *esperar*-based constructions. In Sophia Marmaridou, Kiki Nikiforidou & Elena Antonopoulou, eds. *Reviewing linguistic thought. Perspectives into the 21st century*. Berlin: Mouton de Gruyter, 77–96.
- Evers, Arnold. 1975. *The transformational cycle in Dutch and German*. Utrecht: University of Utrecht PhD thesis.
- Heine, Bernd & Hiroyuki Miyashita. 2008. Accounting for a functional category: German *drohen* 'to threaten'. *Language Sciences* 30: 53–101.

- Joseph, Brian D. 2003. Evidentials. Summation, questions, prospects. In Aikhenvald & Dixon, eds. 307–327.
- Kuteva, Tania. 2001. *Auxiliation: an enquiry into the nature of grammaticalization*. Oxford: Oxford University Press.
- McCawley, James D. 1971. Tense and time reference in English. In Charles J. Fillmore & Terence D. Langendoen, eds. *Studies in linguistic semantics*. New York: Holt, Rinehart & Winston, 96–113.
- Noël, Dirk. 2001. The passive matrices of English infinitival complement clauses: evidentials on the road to auxiliarihood? *Studies in Language* 25: 255–296.
- Ross, John R. 1969. Auxiliaries as main verbs. *Journal of Philosophical Linguistics* 1: 77–102.
- Seuren, Pieter A. M. 1972. Predicate Raising and dative in French and sundry languages. Oxford: Magdalen College (unpubl. ms; published in Seuren 2001: 139–184).
- Seuren, Pieter A. M. 1981. Tense and aspect in Sranan. *Linguistics* 19.11/12: 1043–1076.
- Seuren, Pieter A. M. 1990. Serial verb constructions. In Brian D. Joseph & Arnold M. Zwicky, eds. *When verbs collide: papers from the 1990 Ohio State Mini-Conference on Serial Verbs*. Working Papers in Linguistics No. 39. The Ohio State University Department of Linguistics, Dec. 1990: 14–33.
- Seuren, Pieter A. M. 1991. The definition of serial verbs. In Francis Byrne & Thom Huebner, eds. *Development and structures of creole languages. Essays in honor of Derek Bickerton*. Amsterdam: Benjamins, 193–205.
- Seuren, Pieter A. M. 1996. *Semantic syntax*. Oxford: Blackwell.
- Seuren, Pieter A. M. 2001. *A view of language*. Oxford: Oxford University Press.
- Seuren, Pieter A. M. 2003. Verb clusters and branching directionality in German and Dutch. In Pieter A. M. Seuren & Gerard Kempen, eds. *Verb constructions in German and Dutch*. Amsterdam: Benjamins, 247–296.
- Seuren, Pieter A. M. 2009. The clitics mechanism in French and Italian. *Probus* 21: 83–142.
- Seuren, Pieter A. M. & Herman Chr. Wekker. 1985. Semantic transparency as a factor in Creole genesis. In Pieter Muysken & Norval Smith, eds. *Substrata versus universals in creole genesis*. Amsterdam: Benjamins, 57–70.
- Van Kemenade, Ans. 1993. The history of English modals: a reanalysis. *Folia Linguistica Historica* 13: 143–166.

*Authors' addresses*

Pieter A. M. Seuren  
Max Planck Institute for Psycholinguistics  
PO Box 310  
6500 AH Nijmegen, The Netherlands  
e-mail: pieter.seuren@mpi.nl

Camiel Hamans  
Dutch Delegation PES-group  
European Parliament  
Brussels/Strasbourg  
Belgium/France  
e-mail: hamans@telfort.nl

received: 24 September 2007  
revision invited: 4 December 2007  
revision received: 5 May 2008  
accepted: 28 May 2008