A corpus-based analysis of evidentiality in popularised scientific discourse

Master paper submitted in partial fulfilment of the requirements for the degree of “Master in de Taal- en Letterkunde: Engels – Spaans”

Annelien Siau

Supervisor: Dr. Julie Van Bogaert
A corpus-based analysis of evidentiality in popularised scientific discourse

Master paper submitted in partial fulfilment of the requirements for the degree of “Master in de Taal- en Letterkunde: Engels – Spaans”

Annelien Siau

Supervisor: Dr. Julie Van Bogaert
Acknowledgements

I would like to thank the following people for their support during the composition of this dissertation.

First and foremost, I want to thank my supervisor Dr. Julie Van Bogaert. She was always there to answer my numerous questions, provided helpful remarks and made time to correct my writing.

I would also like to thank Mien Stoffels, a great friend who supported me from beginning to end. I am very grateful for her encouraging words and useful advice.

Finally, a great thank you goes to my family and friends, who always supported me and never lost their confidence in me.
# Table of contents

ACKNOWLEDGEMENTS ............................................................................................................. 4

LIST OF FIGURES .................................................................................................................. 7

LIST OF TABLES ..................................................................................................................... 8

INTRODUCTION ...................................................................................................................... 9

1 EVIDENTIALITY .................................................................................................................. 12

1.1 EVIDENTIALITY IN NON-EUROPEAN LANGUAGES ......................................................... 12

1.1.1 Defining the concept: evidentiality and modality .......................................................... 13

1.1.2 Types of evidence ........................................................................................................... 16

1.2 EVIDENTIALITY IN ENGLISH ........................................................................................ 18

1.2.1 Discussion ....................................................................................................................... 18

1.2.2 Types ............................................................................................................................... 20

1.3 EVIDENTIALITY IN DUTCH ............................................................................................ 23

1.4 IMPORTANCE OF EVIDENTIALITY ................................................................................. 24

1.4.1 Functional motivations for evidentiality ....................................................................... 25

1.4.2 Evidentiality in scientific discourse ............................................................................. 26

2 AUDIOVISUAL TRANSLATION ....................................................................................... 28

2.1 SUBTITLING ..................................................................................................................... 29

2.1.1 Definition ......................................................................................................................... 29

2.1.2 Strategies ......................................................................................................................... 30

3 METHODOLOGY ............................................................................................................. 34

3.1 DATA COLLECTION ....................................................................................................... 34

3.2 CRITERIA .......................................................................................................................... 35
List of figures

FIGURE 1. COMBINATION OF THE TYPES OF EVIDENCE IN NON-EUROPEAN LANGUAGES .................................. 18
FIGURE 2. LEXICAL EXPRESSIONS OF ENGLISH EVIDENTIALITY .................................................................. 22
FIGURE 3. RESPONSIBILITY FOR THE STATEMENT EXPRESSED BY THE SPEAKER ...................................... 56
FIGURE 4. FROM CONTRACTION TO EXPANSION .................................................................................. 56
List of tables

TABLE 1. INCIDENCE OF TYPES OF EVIDENCE................................................................. 39
TABLE 2. FREQUENCY OF THE STRUCTURES NP + VERB, PN + VERB AND PRON + VERB................................................................. 49
TABLE 3. CHANGE OF TYPES FROM ENGLISH TO DUTCH........................................ 60
TABLE 4. TIME ANALYSIS OF A 'DELETION' EXAMPLE............................................ 70
TABLE 5. TIME ANALYSIS OF A 'DELETION' EXAMPLE............................................ 71
In Introduction

Evidentiality is a linguistic phenomenon acknowledged in the 20th century. The term appears for the first time in Boas' posthumously published *Kwakiutl Grammar* (1947), in which he recognises "a small group of suffixes [which] expresses source and certainty of knowledge" (Boas et al. 1947 qtd. in Jacobsen 1986: 4). He uses the term 'evidential' to refer to one of the suffixes. Boas' students (Edward Sapir, Morris Swadesh, and Harry Hoijer) also recognise the concept of evidentiality in their works on Indian languages, but it was a friend of Boas, Roman Jakobson, who gave evidentiality its basic definition. Jakobson (et al. 1984: 46) defines 'evidential' in 1957 as "a tentative label for the verbal category which takes into account three events - a narrated event, a speech event and a narrated speech event (Ens), namely the alleged source of information about the narrated event." Furthermore, he suggests four possible sources of evidential information: someone else's report (quotative, i.e. hearsay evidence), a dream (relative evidence), a guess (presumptive evidence) or one's own previous experience (memory evidence) (Jakobson et al. 1984). This definition, which soon became widely spread, introduced the concept of evidentiality in the field of linguistics. Twenty-four years after Jakobson's publication, in 1981, several linguists gathered at a symposium in Berkeley with the intention to compare evidentiality in a variety of languages and to explore such general questions as the areas of epistemology for which different languages provide evidential markings, the nature of such markings, and the ways in which they arise and spread.

(Chafe & Nichols 1986: viii)

This conference resulted in an edited volume entitled *Evidentiality: the Linguistic Coding of Epistemology* (Chafe & Nichols 1986). This collection of articles firmly established the notion of evidentiality in linguistic circles, and soon put it at the centre of various discussions. Starting out as the topic of typological studies, evidentiality spread to grammaticalisation, cognitive linguistics, syntax and pragmatics (Dendale & Tasmowski 2001). In addition, with scholars increasingly trying to identify the evidential systems of various languages, it spread from American Indian to European languages as well.

Following the example of those scholars, this dissertation deals with evidentiality in English. This subject is of particular interest and complexity because of the disagreement on how to define the concept in languages – like English – lacking grammatical evidentiality
markers. In fact, some scholars deny that evidentiality exists at all in such languages. To refute this statement, this dissertation examines how evidentiality can be expressed in English. The delineation of a possible English evidentiality system is done in the literature review. Subsequently, I investigate the features of the English evidentiality used in a popularised form of science: the television documentary. The corpus consists of English evidentiality expressions and their concomitant Dutch subtitles, taken from the first season of the American science documentary How the Earth Was Made. This corpus-based approach is adopted to examine the effect of the specific genre on the evidential forms and types. Additionally, to allow for an adequate amount of results, the Dutch subtitles of the evidentiality expressions are studied as well. This comparative interlingual study serves to gain more insight into the multifunctional meaning of evidential expressions.

Since the dissertation tackles two notions, i.e. the English evidentiality expressions and their Dutch subtitles, the theoretical framework provided in the literature review equally consists of two parts. First, the sources about evidentiality are compared and discussed, with the aim to establish a workable definition to rely on when constituting the corpus and investigating the data. Subsequently, evidentiality in English and in Dutch are explored, followed by a section about the functional motivations for evidentiality. In particular, the section about English evidentiality is based on the evidential forms provided by Ifantidou (2001), while the sections about Dutch evidentiality and functional motivations are inspired by one of the few scholars who have studied evidentiality in Dutch, Ferdinand de Haan (2000). In the second part, the main subtitling techniques used in an interlingual medium are discussed, as it is important to be aware of how the multimodal genre may influence the translation of the evidentiality expressions.

The literature review is followed by the results of the corpus analysis, which again utilised the twofold structure of the theoretical introduction. The first study defines evidentiality for the particular English data. The first section of this study classifies the data into the four evidentiality types proposed in the literature review (i.e. the indication of direct evidence, reported evidence, inferential evidence and the expression of (un)certainty), and investigates their specific characteristics. The second section focuses on the formal encoding of the evidential data, while the third examines the functional motivations of the speaker.

Following the investigation of the English evidential data, the second study is concerned with the subtitles. These may provide additional information about the
meanings of English evidentials when used in scientific discourse. The interlingual comparative analysis of the subtitles puts emphasis on the differences between the source sentences and the subtitles.

Finally, I discuss how these differences come into existence, and whether or not they are the consequence of a different functional motivation. By providing an answer to these questions, this study doubles as a study on subtitling and the ways in which subtitles may affect the English evidentiality expressions.
I. LITERATURE REVIEW

1 Evidentiality

The current part tackles the notion of evidentiality in four sections. The primary concern of this part is the delineation of the domain of evidentiality for this particular dissertation. First of all, to get acquainted with the concept of evidentiality, section 1.1 presents evidentiality in non-European languages by outlining various definitions and the most generally acknowledged types. The findings of section 1.1 are used as a basis in order to create a suitable evidentiality system for English (section 1.2) and for Dutch (section 1.3). These systems are used as a starting point for the composition of the corpus and the analysis of the English evidentiality data. Finally, in the last part, by discussing the functional motivations for evidentiality, I emphasise the significance of evidentiality in general and in particular in scientific discourse.

1.1 Evidentiality in non-European languages

As mentioned in the introduction, the first acknowledgment of an evidential structure was done in a study about Kwakiutl. The fact that evidentiality was first brought to light in studies of American Indian languages is not surprising given that in those languages "the marking of evidentiality through verb suffixes is widespread" (Chafe & Nichols 1986: viii). This observation explains why the first studies on evidentiality deal with evidentiality as a grammatical phenomenon and focus on languages spoken in various parts of North and South America (cf. Jacobsen on Makah; Oswalt on Kashaya; Schlichter on Wintu; Whistler on Patwin; Gordon on Maricopa). Defining evidentiality for European languages, which generally lack evidential verb suffixes, is a difficult task. In order to arrive at an appropriate overview of English evidentiality markers and a functional definition of English evidentiality, I begin from the starting point: the grammatical evidentiality in non-European languages.
1.1.1 Defining the concept: evidentiality and modality

One of the pioneers of the investigation of evidentiality as a grammatical category is Alexandra Y. Aikhenvald. In her book *Evidentiality* (Aikhenvald 2004), she gives the following definition:

> In about a quarter of the world's languages, every statement must specify the type of source on which it is based - for example, whether the speaker saw it, or heard it, or inferred it from indirect evidence, or learnt it from someone else. This grammatical category, whose primary meaning is information source, is called 'evidentiality'.

(Aikhenvald 2004: 1)

This definition deals with the conceptual notion of evidentiality. Aikhenvald states that the main meaning of evidentials is the indication of the source of information on which a statement is based. This is possible via visual, auditive, inferred, or hearsay evidence. Similarly, in the oft-cited definition by Anderson (1986: 274) it is said that evidentials give "the kind of justification for a factual claim which is available to the person making that claim" and this indication of evidence has to be the primary meaning of the evidential structure. Nevertheless, not all linguists agree with this one and only meaning given to evidentiality.

On the one hand there exists a consensus that "[t]he basic characteristic of linguistic evidentiality is the explicit encoding of a source of information or knowledge (i.e. evidence) which the speaker claims to have made use of for producing the primary proposition of the utterance" (Diewald & Smirnova 2010: 1). According to Willett (1988), this view corresponds to evidentiality in the narrow sense, because an explicit relationship between evidentiality and modality is denied. On the other hand, there are various opinions regarding evidentiality in the broad sense. For instance, one can deduce from the earliest explanation of the concept of evidentiality by Boas, that evidentiality not only refers to source of knowledge, but also to certainty of knowledge. This expression of (un)certainty is generally denoted as epistemic modality. According to Dendale and Tasmowski (2001: 343), "the forms marking the source of information also mark the speaker's attitude", which explains the difficulty with the interface between evidentiality and modality when trying to define the concept of evidentiality.
As Dendale and Tasmowski (2001: 341) note, one of the main difficulties raised by researchers concerning the conceptual problem of the term evidentiality is "the question of the scope and definition of the terms evidentiality and evidential and their relation to the terms epistemic modality and epistemic modal marker." To answer this question they suggest three relations between evidentiality and modality, being disjunction (where they are conceptually distinguished from each other), inclusion (where one is regarded as falling within the semantic scope of the other), and overlap (where they partly intersect) (Dendale & Tasmowski 2001). Disjunction is considered as a characteristic of evidentiality in the narrow sense, while inclusion and overlap refer to a broad definition. As noted above, Aikhenvald (2004: 5) argues for a relation of disjunction, denying "any reference to validity or reliability of knowledge or information." In addition, according to a more recent study, "there has been growing acceptance of the assumption that evidentiality is a semantic-functional domain in its own right, and not a sub-division of epistemic modality" (Diewald & Smirnova 2010: 2). This narrow view is typical of linguists who look at languages which encode evidential information in their inflectional morphology (Ifantidou 2001). In other words, they investigate evidentiality as a grammatical phenomenon. By contrast, according to Willett (1988), the relation between evidentiality and modality is one of inclusion. He clarifies this by saying "that evidential distinctions are part of the marking of epistemic modality, even though evidentials as such are seldom explicitly mentioned in theoretical treatments of modality" (Willett 1988: 52).

Other broad views on the concept of evidentiality are proposed by Chafe and Mithun. Chafe (1986), as one can see in the title of the Berkeley edition Evidentiality: Linguistic Coding of Epistemology, considers evidentiality as an equivalent for epistemology. He includes everything that involves attitudes towards knowledge. The diagram below explains his approach.

<table>
<thead>
<tr>
<th>Source of knowledge</th>
<th>Mode of knowing</th>
<th>Knowledge matched against</th>
</tr>
</thead>
<tbody>
<tr>
<td>??? --</td>
<td>Belief</td>
<td>Reliable</td>
</tr>
<tr>
<td>Evidence --</td>
<td>Induction</td>
<td>Knowledge</td>
</tr>
<tr>
<td>Language --</td>
<td>Hearsay</td>
<td>Knowledge</td>
</tr>
<tr>
<td>Hypothesis --</td>
<td>Deduction</td>
<td>Knowledge</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Chafe 1986: 263)
In Chafe's analysis, evidentiality refers not only to the source of knowledge and the mode of knowing but also to the reliability of that knowledge and that knowledge matched against verbal resources and expectations. Thus, Chafe regards evidentiality as "an indication of the source and reliability of a speaker's knowledge" (Willett 1988: 55). Mithun (1986) does the same. She indicates three ways to qualify the reliability of an utterance: specify the degree of precision or truth, specify the probability of its truth, or specify expectations concerning their probability (Mithun 1986). These broad views typically include both grammatical and lexical encoding of evidentiality.

Finally, when modality and evidentiality partly intersect, one speaks of an overlapping relation. The interface is then expressed by evidential inference, "which [van der Auwera and Plungian] claim to be identical to the modal value of epistemic necessity" (Dendale & Tasmowksi 2001: 342).

This dissertation applies the concept of evidentiality in its broadest sense because of the following reasons. First of all, in order to create an appropriate evidentiality system for languages that lack grammatical evidentiality markers, it is easier to start from a broad point of view since in those languages "the reason for indicating a source of information is often to give the interlocutor an idea of the degree of certainty or reliability of a given piece of information (Frajzyngier 1985; Willett 1988; Dendale 1991; Bybee et al. 1994; Boye & Harder 2009)" (Van Bogaert & Dendale 2013: 4). When dealing with English, for instance, it is necessary to include non-grammatical forms of evidentiality due to the fact that it hardly has any grammaticalised evidential markers. Secondly, in the light of this specific study, it is recommended to include the epistemic attitude of the speaker. The necessity of this inclusion of (un)certainty is further explained in section 1.4, where the importance of evidentials in a scientific discourse is discussed. Connected to this is the opinion of Van Bogaert and Dendale, who argue for a relativistic approach to the definition of evidentiality:

For us, how one delineates evidentiality and whether one can assign evidential status to a given linguistic expression and designate it with the term evidential (instead of something like information source marker) is contingent on the aim of the study at hand and on the research paradigm in which it inscribes itself.

(Van Bogaert & Dendale 2013: 24)

---

¹ For a discussion on must see section 1.2.1.
One of the aims of this study is to investigate the occurrence of English evidentiality in a scientific genre, namely television documentary. Consequently, it is not useful to adopt the strict definition of Aikhenvald, who is interested in typological studies of grammatical evidentiality. As will be shown in sections 1.2, 1.3 and 1.4, this study is more in line with discourse analyses and defines evidentiality in a functional light. Therefore, "it seems perfectly reasonable to incorporate a large variety of expressions, whatever status they have, whether grammatical or lexical and to include pragmatically inferred meanings" (Van Bogaert & Dendale 2013: 24).

1.1.2 Types of evidence

The indication of the information source, being the basic semantic function of evidentials, can be expressed in various ways. Scholars discovered different categories while studying American Indian languages. A vital and frequently cited example of such a study is Barnes’s (1984) work on Tuyuca. Barnes distinguishes five evidentiality types based on how the information is received. Below, examples are given for each evidential category, viz. (1) visual evidence, (2) nonvisual evidence (to indicate any of the senses other than visual), (3) apparent evidence, (4) secondhand evidence and (5) assumed evidence.

(1) *diiga ape-wi*

'He played soccer'. (I saw him play.)

(2) *diiga ape-ti*

'He played soccer'. (I heard the game and him, but I didn't see it or him.)

(3) *diiga ape-yi*

'He played soccer'. (I have seen evidence that he played: his distinctive shoeprint on the playing field. But I did not see him play.)

(4) *diiga ape-yigi*

'He played soccer'. (I obtained the information from someone else.)

(5) *diiga ape-hiyi*

'He played soccer'. (It is reasonable to assume that he did.)

(Barnes 1984: 257)

Another well-known classification based on the study of several American Indian languages is Willett’s (1988). The main parameter of his categorisation is direct versus indirect evidence. Willett speaks of direct or attested evidence when the speaker was a direct
witness to the source of information. Attested evidence can be obtained by the visual sense, the auditory sense or one of the other three senses. Indirect evidence implies that the source of the speaker's information is of secondary nature and includes reported (evidence via verbal report) and inferring (based on inference) evidence. Reported evidence can be second-hand or third-hand. The former denotes hearsay information received from a direct witness, while the latter represents hearsay information passed on from one person to another and finally to the speaker. The third reported evidence option, namely folklore, refers to information that is part of the oral literature, like myths and legends. Finally, the inferring evidence may involve results, i.e. observable evidence, or a mental construct, named reasoning. (Willett 1988)

A more recent model, partly based on Willett, is the one proposed by Aikhenvald:

1. VISUAL (SENSORY): information acquired through seeing
2. (NON-VISUAL) SENSORY: information acquired through other forms of sensory perception (smell, taste, touch)
3. INFERENCE: conclusion based on visual or tangible evidence
4. ASSUMPTION: based on indications other than visible evidence (logic, supposition, general knowledge), with a strong reasoning component
5. HEARSAY: reported information without making reference to the person from whom the information was acquired
6. QUOTATIVE: reported information with explicit mention of the source

(Aikhenvald 2004: 63-64)

Categories 1 and 2 coincide with the direct evidence category of Willett, even though they lack the explicit mention of the auditory evidence. The others are part of indirect evidence. Aikhenvald's term 'assumption' is comparable to Willett's 'reasoning'. Among the differences are the exclusion of the first-hand versus second-hand hearsay distinction, the deletion of the notion folklore, and the addition of a second type of reported evidence, i.e. quotative. The term quotative is also mentioned by Anderson (1986). In the general category of reported evidence he distinguishes at least four subdivisions: "(a) hearsay, (b) general reputation, (c) myth and history (these three being evidentials), and (d) 'quotative' (marginally an evidential)" (Anderson 1986: 289). Even though he calls the quotative evidence marginally an evidential, he follows Aikhenvald by saying that it stands for "This is what X said" (Anderson 1986: 289). Other linguists use different terms. Chafe (1986: 269)
calls all reported evidence hearsay evidence and states that the construction used to cite a reference, i.e. quotative, is a "hearsay evidential expressed in [its] most precise and deliberate form." Regarding the reported evidence, I will use the more general terms defined by Aikhenvald: hearsay and quotative. The distinction is made between mentioning or not mentioning of the source, which means that these categories can also comprise the subdivisions made by Willett and Anderson. For instance, the sentence The Romans believed that Romulus and Remus founded Rome, refers to a myth and is analysed as indicating quotative evidence, while Romulus and Remus are believed to have founded Rome denotes the same myth but refers to hearsay.

As a summary, and to conclude this section of evidential subcategories based on American Indian languages, the three taxonomies are combined in Figure 1.

<table>
<thead>
<tr>
<th>Types of evidence</th>
<th>Direct</th>
<th>Indirect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Attested</td>
<td>Reported</td>
</tr>
<tr>
<td></td>
<td>VISUAL</td>
<td>HEARSAY</td>
</tr>
<tr>
<td></td>
<td>NON-VISUAL</td>
<td>QUOTATIVE</td>
</tr>
</tbody>
</table>

**FIGURE 1.** Combination of the types of evidence in non-European languages

### 1.2 Evidentiality in English

#### 1.2.1 Discussion

In the majority of the studies on evidentiality it has been assumed that European languages lack "grammatical markers and grammatical systems of evidentiality" (Diewald & Smirnova 2010: 2). Aikhenvald (2004) addresses this problem for English by saying that English does not have pure evidential markers and therefore compares evidentiality with other categories, like modality. She also draws a distinction between pure evidential markers, which are grammatical, and evidential strategies, which are lexical or pragmatic, and concludes that in most European languages the evidential structures are 'mere' evidential strategies. Even though Aikhenvald states that some languages lack a specific evidential grammatical category, she does admit that referring to a source of information is universal.
A lot of linguists agree with this statement and apply it to English, as can be seen in the following overview of citations:

Although English does not have a specific grammatical category of evidentials, a variety of optional, nonpropositional constructions can function as evidentials.
(Barton 1993: 746)
The definition of evidentials in English thus has to be a functional one.
(Barton 1993: 746)
English has a rich repertoire of evidential devices. It expresses evidentiality with modal auxiliaries, adverbs, and miscellaneous idiomatic phrases, although not, for example, with a coherent set of verb suffixes like those in some California Indian languages.
(Chafe 1986: 261)
One undercover evidential in English is the inferential value of polysemous must, distinct from its obligational one, as nicely delineated by Chafe (1970:179-84).
(Jacobsen 1986: 7)

There is general agreement as to the notion that English does not have a well-defined grammatical system to express evidentiality. Therefore it is advisable to define English evidentiality in a functional light, by using lexical means. However, saying that English does not at all have grammatical evidential markers needs some consideration since there is one borderline case; the modal must.

The criteria to distinguish grammatical evidentials from lexical evidentials are not clear-cut. According to de Haan (2000), a grammaticalised evidential morpheme is characterised by the following four criteria:

1. Evidentials are not themselves the main part of the clause.
2. Evidentials do not show agreement with the speaker.
3. The morphemes have the expression of evidentiality as their primary meaning.
4. Evidentials cannot be in the scope of a negative element.
(de Haan 2000: 75-76)
The first and second criteria are responsible for the exclusion of sentences like *it is evident that* and *I see that* (de Haan 2000). The fourth criterion is the reason why de Haan considers the Dutch verb *moeten* as an evidential (cf. section 1.3), whereas the third contributes to the elimination of *must* from the evidential system. Criterion 3 "is used to distinguish between true evidentials and those elements for which evidentiality is only inferentially present" (de Haan 2000: 75). To illustrate these inferential evidentials, Anderson (1986: 275), who also mentions the difference, gives four examples: (1) *the toast is burnt*, (2) *the toast burned*, (3) *the toast has burned*, (4) *the toast must have burned*. De Haan (2000) says that from the third statement *the toast has burned*, the second *the toast burned* may be inferred. In that case "the perfect denotes, secondarily, the evidential notion of 'evidence for an action'", but cannot be considered an evidential itself (de Haan 2000: 76). According to de Haan (2000), the modal *must* is a similar example. Anderson (1986: 275), by contrast, considers the fourth example as a true evidential because, "[w]hen the present state is used as circumstantial evidence for inferring an unwitnessed past event, English normally adds the epistemic *must* of logical inference." Consequently, *must* indicates that there is evidence for the action expressed, for instance the smell of the burnt toast. De Haan argues that this is not enough proof for the verb *must* to be treated as a grammaticalised evidential, since the expression of evidentiality is not its primary meaning. Following Anderson, this dissertation proposes to treat the epistemic modal *must* as a grammatical inferential evidential.

In addition, Ifantidou (2001) notices that sometimes a structure may function as an evidential without this information being linguistically encoded, which is called pragmatic inference. In this view, the sentence *John is feeling miserable today*, can be based on observation (the speaker has seen his miserable expression), hearsay (the speaker repeats what John told him), or inference (the speaker deduces the information from John's behaviour) (Ifantidou 2001).

1.2.2 Types

A good overview of the lexical structures that express English evidentiality is given by Ifantidou (2001). Considering the definition of an evidential as a functional one, Ifantidou organises the categories around the two main functions (in the broad sense). The first function, *indication of source of knowledge*, can be obtained by observation (sensory/perceptual evidence), by hearsay (from other people), by inference and by memory. The observational evidence is mainly expressed by perception verbs like *I see*, *I*
hear, I feel, it tastes, or by verbs which express less reliability like looks like, sounds like, feels like, smells like. Hearsay can be indicated by the expressions tells me, I hear, people say, he is said, he is reputed, allegedly, reportedly. Ifantidou also includes less direct hearsay devices which primarily perform other functions like it seems, it's supposed to, apparently. The structures presumably, seems to/must be, must have, I gather are typical of inferential evidence and even though not frequently treated as evidentials, Ifantidou adds so, I deduce, and consequently. Finally, the information source can also be one's own memory, expressed by I remember, I recall, as I recollect. (Ifantidou 2001: 5-7)

Secondly, Ifantidou's evidentiality system for English includes the speaker's degree of certainty, characterised by propositional attitude and parenthetical expressions (I think, I know, I suspect, I guess, I suppose), adverbials (probably, certainly, possibly, undoubtedly, surely, evidently, obviously) and epistemic modals (may, might, can, could, must, will, ought to/should). (Ifantidou 2001: 5-7)

The above-mentioned categorisation represents a broad view on evidentiality, based on the suggestion that "[i]n its broadest sense, an utterance has an evidential function if and only if it overtly communicates evidential information, whether this information is linguistically encoded or pragmatically inferred" (Ifantidou 2001: 161). This broad view is suggested by the inclusion of expressions which are not standardly treated as evidential. These are lexical expressions not included by the majority of the scholars, but for Ifantidou important enough to mention. Ifantidou adds and at the same time deletes some terms previously seen. On the one hand, she takes into account the definition by Jakobson, who includes memory as a possible evidential. She states that "[s]ince memory is variably reliable, such expressions have a claim to be considered as evidentials" (Ifantidou 2001: 7).

On the other hand, she leaves out Aikhenvald's 'quotative' and 'assumption' categories. As already explained, this study will adopt Aikhenvald's terms for reported evidence.

To conclude, I give a scheme of the evidentiality types seen in section 1.1.2 (cf. Figure 1) combined with the English lexical expressions proposed by Ifantidou. The overview represented in Figure 2 serves as a basis for the composition of the corpus and the English evidentiality discussed in the corpus analysis.
# A. INDICATION OF SOURCE OF KNOWLEDGE

## 1. Direct evidence

### a. Visual

| I see, looks like | I see him swimming. |

### b. Non-visual

| I hear, I feel, it tastes, sounds like, feels like, smells like. | I hear that the train is coming. |

### c. Memory

| I remember, I recall, as I recollect | I remember that he was an excellent professor. |

## 2. Indirect evidence

### a. Reported

| Hearsay | I hear, he is said, he is reputed, allegedly, reportedly, it seems, it's supposed to, apparently |
| Quotative | X tells me, people say, they suggest |

| He is said to have committed the crime. |

| Tom tells me John is the burglar. |

### b. Inferring

| Inference | Presumably, seems to, must be/must have, I gather, guess, suppose, so, I deduce, consequently |

| There is a car on the driveway. |

| I gather that Tom is in town |

| Assumption | Must, I assume |

| I assume that Tom is on holiday. |

## B. SPEAKER’S DEGREE OF CERTAINTY

### a. Propositional attitude and parenthetical expressions

| I think, I know, I suspect, I guess, I suppose | I think he is a very clever person. |

### b. Adverbials

| probably, certainly, possibly, undoubtedly, surely, evidently, obviously. |

| Obviously he did not kill the cat. |

### c. Epistemic modals

| may, might, can, could, must, will, ought to/should |

| It may be possible that he ran away. |

---

**FIGURE 2.** Lexical expressions of English evidentiality
1.3 Evidentiality in Dutch

In order to be able to analyse the Dutch subtitles of the English evidentiality expressions, a concise overview of evidentiality in Dutch is appropriate. Nevertheless, studies about a possible evidentiality system in Dutch are rather scarce. De Haan is one of the few linguists who examine the properties of evidentiality in Dutch and presents five forms:

a. Modal verb moeten 'must'
   De film moet uitstekend zijn
   'The film is said to be excellent'

b. Past tense modal verb zou(den) 'should'
   Bij de brand zouden alle bewoners zijn omgekomen
   'All inhabitants are said to have perished in the fire'

c. Raising verb schijnen 'seem'
   Jan schijnt ziek te zijn
   'John seems to be ill'

d. Complements of perception verbs
   Ik hoor, dat Jan ziek is
   'I hear that John is ill'

e. "Quotative"
   "Douane controleert El Al niet"
   'Customs [officers] do not check El Al [planes]'

(de Haan 2000: 74)

Dividing these Dutch expressions into the evidential categories seen above is not an easy task. First of all, the modal verbs moeten and zouden express either modality or evidentiality (see below). With their evidential meaning they can have both a quotative and an inferential interpretation. For instance, sentence a. can be interpreted as My friends say it is a good film or I like the director's work and therefore this film must be good as well (de Haan 2001: 214). The raising verb schijnen is also marked by ambiguity. Example c. Jan schijnt ziek te zijn can refer to a quotative, People say that John is ill, or an inferential, I infer from his absence that he is ill. The perception verbs in Dutch refer to indirect evidence or in the case of example d. to hearsay evidence (de Haan 2000). Finally, de Haan uses 'quotative' to denote the literal citation of someone's words indicated by means of quotation marks. In the corpus analysis I
will examine whether these expressions are (frequently) used in the translation of the English evidential constructions.

While de Haan (2000) excludes the English epistemic modal *must* from the grammaticalised evidential system, he opts to treat the Dutch modal *moeten* as a true evidential. He adopts this approach because of the agreement with criterion (4) *Evidentials cannot be in the scope of a negative element* (cf. section 1.2.1). When the verb *moeten* co-occurs with a negative element, *moeten* has scope over the negation (de Haan 2000). This is only possible when dealing with the evidential interpretation of *moeten*. If *moeten* is replaced by another modal verb (*hoeven, kunnen*) the evidential interpretation gets lost and the modal verb is in the scope of the negative element (de Haan 2000). Secondly, the other criteria also hold (cf. section 1.2.1). The verb *moeten* is not the main part of the clause nor does it show agreement with the speaker. In isolation the verb *moeten* accepts different interpretations. As illustration, sentence a. *De film moet uitstekend zijn* can have a deontic reading (*it is required to be a good film*), an epistemic interpretation (*it is probable that it is a good film*), and an evidential interpretation (*there is evidence that it is a good film*) (de Haan 2000). However, de Haan argues that, seen in the context, the main interpretation of the verb *moeten* is the evidential one. It has indication of evidence as its primary meaning, which coincides with criterion 3. (de Haan 2000)

The three-way ambiguity of *moeten* is caused by grammaticalisation. Diachronically, epistemic modality (Stage II), which expresses probability based on evidence, is developed from deontic modality (Stage I). Evidential *moeten* (Stage III), which expresses only evidence, arose from epistemic *moeten* by grammaticalisation (de Haan 2000). De Haan (2000) claims that the English epistemic modal *must* has not (yet) arrived at the third stage and therefore cannot be considered as a true evidential. As outlined in section 1.2.1, in this dissertation I will not agree with de Haan, because I consider *There is evidence that the toast has burned; the house is full of smoke* as to have an evidential value comparable to *The toast must have burned*.

### 1.4 Importance of evidentiality

The classification of the English evidentiality system (cf. Figure 2) is based on the main meanings of evidentiality: expressing the source of knowledge and the certainty of knowledge. Inside this division several functional motivations for evidentiality can be found, especially in scientific discourse.
1.4.1 Functional motivations for evidentiality

Considering direct evidence, the speaker is a witness of the actions described and therefore makes "him/herself into a reference point" (de Haan 2001: 217). According to de Haan (2001), receiving the information from a first-hand source implies high reliability and responsibility. Reportives, on the other hand, remove the responsibility from the speaker. Because of the introduction of another person, the speaker outsources the responsibility for the information expressed, which has two consequences. On the one hand, by acknowledging that the statement is not witnessed by the speaker, he/she disavows responsibility for a claim he/she does not agree with or is not certain of. In English this may be expressed by *it is said, people think*, etc. On the other hand, it is a useful way to assign authority to the statement. By adding a second convincing party, like experts in the field, the statement can be considered more valuable and truthful. (de Haan 2001)

Another way to reinforce this authority is to substantiate the argument with clear evidence. For instance in scientific or academic writing, it is appropriate to ask how one has obtained particular information or whether there exists proof for a certain statement. These questions can be answered by reportive markers, as explained above, or by inferential evidentials. De Haan (2001: 193) defines inference as "the grammaticalized way of showing that the speaker makes his/her statement based on a deduction from facts, and not on a direct observation of the action itself." Examples of these facts can be footprints in the snow indicating the presence of a human or animal, or bread crumbs as the evidence of a loaf of bread which once existed. The statement can be inferred from the available evidence (1a) or the evidence can be expressed overtly (1b):

(1a) bóahō-yu
rot- INFER: OTHER: PAST
'It rotted' (said of a plant after pulling it up to examine it)

(1b) Dáman dara-t vac dyi turasno.
shallow sit- PI INFER ART peach
Pui' cu-t va' ia'ray gu jivihl.
Thus so- PI then. fell ART wind
'These peach trees must have been planted shallowly. That's why the wind blew them over'.
(de Haan 2001: 193)
Thus, the significance of evidentials lays firstly, in the fact that they permit the speaker to acknowledge or deny responsibility for a statement, and secondly, in the possibility to assign authority to the statement in order to make it more trustworthy. The latter can be realised by adding a second persuasive information source, or evidence from which the information is inferred. Additionally, dealing with evidentiality in English, I have included the function of expressing (un)certainty towards an utterance via epistemic modals, adverbials or parenthetical expressions. In the following section I try to prove that these functional motivations are also highly valuable in scientific writing.

1.4.2 Evidentiality in scientific discourse

Ahmad (2012) examines the stylistic features of scientific English and distinguishes it from literary language because of the following features:

The scientific language is accurate, precise, and detached from individual impulse. It aims to inform about an important issue and what particular approach is taken up to investigate that issue. It is an objective interpretation of facts and findings. It contains such components and findings that need external and experimental evidence to consolidate their validity. (...) The pursuit of universal generalization in scientific texts enables the author to signal credibility, reliability, objectivity and ultimately authority to their readers and the research community.

(Ahmad 2012: 47, emphasis added by author)

On the one hand, the citation above is significant because it defines scientific English as detached from individual impulse and objective, which is a statement I discuss in the second part of this section. On the other hand, it emphasises the need to provide evidence, which can be linked to the main role of evidentiality (information source), the possibility to signal (un)certainty, which is the second function of English evidentiality (cf. Figure 2) and the assignability of authority, one of the motivations outlined in the previous section. In spite of these observations which stress the evaluative and informative nature of scientific writing, little is written about the application of evidential markers in a scientific discourse. Consequently, this part deals with a concept closely related to evidentiality, i.e. hedging. Hedging, "the expression of tentativeness and possibility" (Ahmad 2012: 52), is one of the features of scientific English delineated by Ahmad. Ahmad (2012: 52) incorporates in the
notion of hedging, linguistic devices such as epistemic main verbs (to indicate, to suggest, to propose, to tend, to seem, to appear), epistemic modal auxiliaries (may, might, can, could) and epistemic adverbials like hypothetically, probably, likely. Accordingly, from the overlap between the epistemic hedging expressions mentioned by Ahmad and the English evidential expressions categorised above (cf. Figure 2), one can deduce that evidentiality also is a vital feature of scientific writing.

In view of the close connection between hedging and evidentiality, I refer to another source in which Loredana Frățilă (2007) associates hedging with objectivity. As claimed by Ahmad, scientific discourse is assumed to be objective. Frățilă (2007) criticises this statement by signalling two different functions of hedging that challenge this objectivity. First, she says that "[the authors] use hedges to distance themselves from their statements and thus to reduce the risk of opposition and minimize the threat to face that may be part of any act of communication" (Frățilă 2007: 89). Frățilă (2007: 89) sees hedges as threat-to-face minimisers, which are useful techniques to express mitigation of responsibility and uncertainty towards a proposition "in order to place their users in a secure corner, in case other people have different opinions of the same issue and to somehow negotiate some degree of flexibility for their claims." The distancing from a statement is reminiscent of the negation of responsibility, one of the functional motivations for evidentiality described above. In addition to the use of hedges to forestall criticism, hedges are also treated as a politeness strategy. They help to build a polite relationship between the author and the reader since hedges soften strong statements and/or straightforward opinions and at the same time reduce possible disagreement. Moreover, since adopting a polite attitude towards fellow scientists is essential, hedges are useful to maintain a good relationship between scientists in the scientific community. This politeness can be expressed by suitable quotatives when referring to colleagues' opinions. Thus, according to Frățilă, the two functions of hedges express subjectivity, which undoes the myth of objectivity in a scientific discourse. Additionally, it is impossible to detach the author and his/her personal opinion from his/her writing. (Frățilă 2007: 89-91)

In sum, if one evaluates hedges as some kind of evidentials, one may assume that the functional motivations for evidentiality expressed in the previous section are also applicable to, and even more meaningful in scientific writing.
The aim of this dissertation is twofold. First of all I investigate the notion of evidentiality (as discussed in section 1) in a corpus consisting of English data from the documentary *How the Earth Was Made*. Secondly, the Dutch subtitles of those English evidential statements are studied. In order to come to the investigation of the evidential translation from English into Dutch, one has to overcome two difficulties associated with the chosen medium.

In contrast with written texts, the present corpus offers an extra dimension. There is no one-to-one relationship, like for instance with a written text which is translated into a target language and again published in a written form. A documentary starts with a script which is interpreted (mainly) by a narrator. This first transformation, from the script into the narrator's voice, is what the viewers hear during the documentary. Secondly, this aural information is then, after being translated, transferred into visual information via subtitles. Thus, with this multimodal medium a lot more aspects have to be taken into account, which will be outlined in section 2.1.1.

Because of this multimodality, the second part (section 2.1.2) of this main section is not a mere study of how statements can be translated, but rather of how this audio-visual transfer (into subtitles) influences the translation of the evidential expressions. Because of the value of evidentials in scientific texts, one would expect a faithful translation of these markers. Surprisingly, this is not always the case. We will see that it is important to differentiate an infelicitous translation due to the different expressions of evidentiality in the source and target languages from an infelicitous translation due to the techniques of subtitling. The former is illustrated in example (1), taken from Taylor's (2004) paper about subtitling. He uses the example to explain the term decimation and at the same time shows (unconsciously) the effect subtitling can have on evidential expressions. Decimation occurs when part of the semantic content of the original expression gets lost.

(1) You will have heard on the news that all the passengers were killed.  
*Lo sai che tutti passegeri sono morti.*  
'You know all the passengers are dead.'  
(Taylor 2004: 159)

The original English sentence which expresses a quotative evidential (*they said it on the news*) is subtitled into Italian. In Italian the quotative has disappeared and by using *Lo sai/you know*, the indication of how the information is received has become insignificant. The
example above shows how the translation strategies used in subtitles, in this example decimation, can have an influence on the expression of evidentiality and consequently on the investigation of the corpus. Section 2.1.2 elucidates more strategies of how subtitles can affect the (evidential) structure and/or message of the source language.

2.1 Subtitling

2.1.1 Definition

All studies (e.g. Díaz Cintas 2008; Gambier & Doorslaer 2010; Gottlieb 2004; Ivarsson & Carroll 1998; de Linde & Kay 1999; Orero 2004; Taylor 2004) agree that subtitling is a relatively new practice which suffered an increase in the late 20th century when audiovisual media boomed. It is defined by Gottlieb as

the rendering in a different language of verbal messages in filmic media, in the shape of one or more lines of written text, presented on the screen in synch with the original verbal message.

(Gottlieb 2004: 86)

This definition is only relevant for interlingual subtitling, i.e. subtitling between different languages, "as opposed to the intralingual type aimed at deaf and hard-of-hearing audiences" (Gottlieb 2004: 86). Accordingly, the intralingual type is not taken into account in this investigation.

In addition, there is an agreement that subtitling has to be treated differently from normal standard translation, due to the multimodal nature of subtitling. Multimodality is introduced by Jakobson as intersemiotic translation and is characterised by

the fact that the meaning created in one modality (e.g. the visual) may be translated in another modality (e.g. the written language) or even vice versa in this digitally manipulative age, but it also simply means that the source and target texts consist of a number of interacting semiotic modalities.

(Jakobson 1966 qtd. in Taylor 2004: 157)
One can see that the intersemiotic nature of subtitling, from speech to writing, contrasts with the isosemiotic nature of 'normal' translation, from writing/speech to writing/speech (Gottlieb 2004: 86).

Next to the intersemiotic difference, subtitling has another distinctive feature. With the translation of a polysemiotic text one has to be aware of the non-verbal channels: the visual and audio components like music and effects, picture and writing (displays and captions) (Gottlieb 2004; de Linde & Kay 1999; Taylor 2004). If these additional components carry (part of) the meaning, omissions in the target language are possible. Consequently, the semiotic modalities other than the verbal can facilitate the processes of deletion.

2.1.2 Strategies

It is now clear that subtitles can be treated as a particular form of translation. The special strategies used by subtitlers are discussed in this section.

According to Jan Ivarsson and Mary Carroll (1998: 85-91), a subtitler uses some specific techniques being (1) condensing the text, (2) omission or paraphrase, (3) muddled speech, (4) ellipsis, (5) merging short dialogues, (6) simplifying the syntax, (7) simple vocabulary, (8) subtitle breaks, (9) borrowing time. The explanation of the techniques starts with (3) muddled speech and (5) merging short dialogues. When dealing with dialogues, it is advisable to merge two or three phrases expressed by the same person into one single subtitle. Muddled speech occurs most of the times with interviews. It may happen that an interviewee speaks very unclear, with incoherent phrases and bad articulation. In such case "the best approach is to (...) translate relatively freely, giving the gist of what the person has said in a reasonably coherent form" (Ivarsson & Carroll 1998: 87).

Nevertheless, even though the speakers explain things as accurately as possible, this does not mean they use simple vocabulary. It is obvious that "it is easier for viewers to absorb and it takes them less time to read simple, familiar words than unusual ones" (Ivarsson & Carroll 1998: 89). That is why, generally, a subtitler opts for the more common word when dealing with a choice of two synonyms. But this is only done when it does not affect the style and spirit of the genre. In the corpus analysis it will become clear whether the informative and scientific style of the documentary influences the vocabulary of the subtitles. (Ivarsson & Carroll 1998)

The other techniques can be subsumed under deletion strategies or time strategies. To begin with, according to Ivarsson and Carroll (1998) condensing the text is the most
difficult task of a subtitler. When the characters in, for instance a movie, talk a lot more than there is space for to translate in the subtitles below the picture, the text has to be condensed, which means that the subtitler has to decide what to translate and what to delete. This decision should be based on what he thinks is the essential content for the viewer to understand the message. This might be the first reason why evidential markers disappear in subtitles, as illustrated in Taylor’s example in 2.1.1.

After one has decided which content should be transmitted, one has to choose which technique to adopt: (2) omission (cf. example 2) or paraphrase (cf. example 3).

(2) "Well, it's just that this morning Mr. Smith came into the office and told us that he has heard that we will all be fired some time during next week."
   *This morning Mr. Smith told us that we will all be fired next week.*
   (Ivarsson & Carroll 1998: 86)

(3) "You should have heard what Doris said. It appears that 'She goes out with American soldiers' said behind your back means you're well on the way to Hell and Damnation!"
   *'Going out with an American soldier', said Doris, 'is a shortcut to hell.'*
   (Ivarsson & Carroll 1998: 86)

According to Ivarsson and Carroll (1998), there is no strict rule to adopt the one or the other. On the one hand they prefer omitting part of the text, because they believe "[o]mission is less intrusive than paraphrase, a point worth bearing in mind, especially if the original is an artistic work or a statement by a person responsible for the opinions expressed" (Ivarsson & Carroll 1998: 86). Taylor (2004) follows the idea that deletion is the most commonly used technique by subtitlers, but he attributes this to the intersemiotic nature of subtitles. On the other hand, Ivarsson and Carroll (1998: 87) admit that "[m]ost subtitlers do not consciously choose between these approaches: omitting part of the dialogue usually makes it necessary to paraphrase the rest to a certain extent."

Another important factor, analogous with the previous one, is ellipsis. Ivarsson and Carroll literally say the following:

Words whose main purpose is to keep the conversation ticking over ("well", "you know"), tautologies and repetitions can safely be omitted. But this does not mean to say that subtitlers should ignore those little
words that often make all the difference or give the lie to a person's character.
(Ivarsson & Carroll 1998: 87)

They warn for the danger of ellipsis. Bearing in mind the functions of evidentials, one could almost interpret 'words that often make all the difference or give the lie to a person's character', as a feature of evidentials. Like in the constructed example of (4), the ellipsis of evidential markers has to be taken into account, because otherwise the subtitle is a lie uttered by the speaker.

(4) The ancient Greeks believed that the earth was flat.
De aarde was plat.
'The earth was flat.'

Apart from the techniques that deal with omission, one also has to consider the syntax and vocabulary. The sixth technique, simplifying the syntax, can also have an influence on whether or not one expresses evidentiality. It is clear that "[s]imple syntactic structures tend to be shorter than complex ones, and the difference in terms of meaning is sometimes negligible" (Ivarsson & Carroll 1998: 88). By re-evaluating example (4), it seems obvious that the sentence with an embedded clause is longer, due to the complex syntactic structure, than the simple sentence used in the subtitle. Even though the latter might fit more into the limited space of the subtitle area, it does not give an accurate representation of the connotation (it's their opinion) meant by the speaker.

The penultimate technique mentioned by Ivarsson and Carroll is subtitle breaks. This is an important notion for the subtitler him/herself. In short it means that the subtitler has to avoid making subtitle breaks between structures that belong together. The goal should be to "make the breaks coincide with the beginnings and ends of phrases, so that each subtitle forms a coherent unit"(Ivarsson & Carroll 1998: 90), as seen in example (5).

(5) Wrong Right
She had furnished the room well. She had furnished the room well
The interior was mainly in red
and green, as these were her
favourite colours. The interior was mainly in red
and green, her favourite colours.

(Ivarsson & Carroll 1998: 91)
In the explanation of the last technique, **borrowing time**, Ivarsson and Carroll admit again that evidential markers can get lost in subtitling. The instances they denote as small talk may express (un)certainty, which is central in English evidentiality:

Where absolutely necessary, subtitlers can "borrow time" from the next sentence (...). This means letting the subtitle remain on the screen during the first part of the next sentence. The dialogue that is masked in this way should of course be of minor importance. Fortunately, sentences often start with small talk such as "Well, I don't know really, I suppose you could say ...". Such phrases can be omitted or abridged.

(Ivarsson & Carroll 1998: 91)

Apart from all the previously mentioned techniques of subtitling, another notion to bear in mind is **punctuation**. Especially in the case of evidentiality, the right use of quotation marks is essential. Also in subtitles, they cannot be omitted, "[t]hey should be repeated in each of the subtitles, otherwise viewers tend to forget that they are still reading a quotation" (Ivarsson & Carroll 1998: 114). Quotation marks are necessary when quoting a particular source, being in most cases a person. However, in a documentary (or television in general) it is possible to view the person who is speaking. Consequently, literal evidentiality markers, like quotation marks, disappear because of the combination of the verbal with the visual medium, like I have already pointed out in the beginning. (Ivarsson & Carroll 1998)
II: CORPUS ANALYSIS

3 Methodology

3.1 Data collection

The corpus used for this investigation is made up of sentences found in the American television documentary *How the Earth was Made*. The data are taken from the first season, which comprises thirteen forty-five minute episodes. This popularised genre of science was chosen because of the following reasons. First, evidentiality is centered around the expression of the source of information. Acknowledging the source to explain certain findings is an important feature of science as well. Second, the television documentary offers an additional advantage, since the Dutch subtitles can help to define the English evidentiality. The corpus was created as follows. Each episode was watched attentively and every evidential-like expression was written down together with its concomitant Dutch subtitles. American English evidentiality thus served as starting point. Afterwards, the utterances were re-examined and only the true evidential statements were retained. As a result, the corpus consists of 512 tokens, 255 of which are English sentences containing an evidential marker, spoken by the narrator or persons in the documentary.

The episodes of the documentary follow a fixed format, which emphasises the scientific nature of this popularised form of science. According to Meyers (2003: 273), "[t]he dominant model of popularization assumes that the aim of the process is to convey scientific knowledge to a wider audience." The composition of each episode contributes to the achievement of this goal. For instance, the second part of the fourth DVD deals with the origin of Hawaii. As an introduction, the narrator gives a description and the geological facts of Hawaii, containing the biggest volcano on earth. Next, the first question, “Could this huge island have been built entirely by lava?”, is asked, followed by clues, research and opinions of scientists, which lead to an answer. When the mystery is solved, the question is answered and followed by the enumeration and appearance of the most important clues on the screen. Subsequently, the next question, “What force on earth could create the most active volcano in the world?”, is uttered and the process of investigation is recapitulated. At the end, the key information is repeated in order to offer a complete account of how the

---

1 The criteria used to compose the corpus are outlined in section 3.2.
Hawaiian chain was created. In this way, all the episodes are clues to answer the main question of how the earth was made. This approach of repetition, summaries, question-answer, and experts who search for and provide evidence, is reminiscent of scientific investigation and serves the aim of popularised science. Consequently, in this scientific context the indication of the source of information and the expression of (un)certainty plays a crucial role.

3.2 Criteria

The selection of the data was based on the functional definition of English evidentiality, outlined in the literature review. The point of departure for the recognition of the evidential forms were the formal devices expressed in Figure 2 (cf. section 1.2.2), which stand for the two meanings, i.e. indication of information source and expression of (un)certainty. While watching the documentary, the question was asked whether certain utterances indicated direct evidence, reported evidence or inferential evidence, or whether they expressed certainty or uncertainty.

The recognition of sentences that signal direct and reported evidence turned out to be uncomplicated, whereas the inferential evidentiality type caused some difficulties. For this particular study it is advisable to reconsider English inferential expressions. As explained previously, while watching the documentary it becomes clear that all the episodes are centred around evidence. Evidence is given in order to answer the main question asked at the beginning, which deals with the origin of a certain natural phenomenon. This evidence is presented by means of evidential markers, which are especially based on inference. In the literature review, de Haan (2001: 193) defines inference as "the grammaticalized way of showing that the speaker makes his/her statement based on a deduction from facts, and not on a direct observation of the action itself." Examples (1a) and (1b) are the sentences that accompanied this definition. In (1a), which is an instance of Tuyuca, the speaker inferred the statement from the available evidence whereas in (1b) the evidence on which the inference is based is expressed overtly. Example (1b) is an instance of Tepehuan, a Uto-Aztecan language which uses particles to express inference.

(1a) bóahō-yu
   rot-INFER:OTHER:PAST
   'It rotted' (said of a plant after pulling it up to examine it)
(1b) (de Haan 2001: 193)
These examples demonstrate two kinds of inference. In (1a) the evidence is unspecified, whereas in (1b) the speaker refers to a specific inferential source. In addition, de Haan focuses on inference as a grammatical phenomenon. Thus, in the American Indian sentences above the fact that the speaker has evidence is expressed either by a verbal suffix (1a) or by a particle (1b). In English, by contrast, it is necessary to express evidentiality lexically in order to obtain an accurate representation of the connotation meant in the grammaticalised evidentiality language. Therefore, it is impossible to recognise the kind of evidentiality expressed by example (1a) in English. The second kind of inference, by contrast, is used abundantly in the documentary. In example (1b) above the available evidence (the fact that the wind blew them over) is expressed. In other words, example (1b) can be interpreted as: The wind blew the peach trees over. This is evidence that the peach trees must have been planted shallowly. In this view, I would like to examine the following four tokens:

(2) Gray gneiss rocks at the top of the Matterhorn prove that Africa collided with Europe, creating the Alps.

(3) The crystals are clear evidence that under the springs is a volcano.

(4) Diatomite reveals the climate was once wetter.

(5) What's striking about these samples is that the minerals inside are elongated. It is a clue that these rocks must once have been crushed by massive forces.

In example (2) the statement is: Africa collided with Europe, creating the Alps. The speaker has inferred this statement from the available evidence and the evidence is expressed overtly, namely the gray gneiss rocks at the top of the Matterhorn. Therefore, Gray gneiss rocks at the top of the Matterhorn are evidence that Africa collided with Europe, creating the Alps is a fitting paraphrase. Similarly, following the interpretation of (1b), example (3) can be paraphrased as I see crystals under the springs. I have evidence that under the springs is a volcano. The same goes for example (4): The presence of diatomite is evidence that the climate was once wetter.
Finally, sentence (5) coincides almost completely with the translation of (1b) because of the use of the evidential must. Consequently, these sentences are included in the corpus as true evidential expressions, in particular as evidentials that denote inference.

3.3 Methods

The corpus analysis consists of two studies. The aim of the first analysis is twofold. The study intends to shed light on the features and the occurrence of evidentiality types in a specific genre, namely scientific narrative. In this way, this corpus-specific investigation offers a broader view on the concept of evidentiality in the English language on the one hand, and serves as a basis for the comparative study that follows, on the other hand.

The first study deals with the English evidential expressions in a qualitative and in a quantitative manner. First of all, the data were classified with regard to the type of evidence they express, i.e. direct evidence, reported evidence, inferential evidence and additionally, the expression of (un)certainty. Consequently, the categorisation was made with the following questions in mind: Did the speaker witness the event? Did the speaker hear the statement? Was the statement inferred from available facts? Does the speaker express (un)certainty concerning his/her statement? Section 4.1.1 presents the results of the classification by outlining the difficulties encountered when analysing the data and by discussing specific features of each type. Once the classification was done, the frequency of each type was calculated and presented in Table 1. Thirdly, as opposed to the semantic approach in the first classification, section 4.1.2 focuses on the evidential expressions from a formal point of view. In this section the evidentials are divided according to their internal structure, with an emphasis on the verb phrase (VP). Finally, to conclude the first analysis the question is raised what are the functional motivations for the evidentiality data (section 4.1.3). When analysing the genre-specific data, it was noticed that the functional motivations expressed in the literature review were unusable. Therefore, the broad discourse-oriented approach of Martin and White (2005) was adopted. Via their engagement system, the functional motivations, which are led by the communicative needs of the genre, are discovered and exemplified.

When the delineation of the English scientific evidentiality system and the examination of the peculiarities were done, the second analysis took place. The second study looks at the Dutch translations of the evidentiality expressions in order to gain more insight into the English concept of evidentiality. This goal is achieved by an interlingual
comparative analysis of the original spoken English sentences and their concomitant subtitles (section 4.2). To begin with, the subtitles which translate the English evidentiality expressions literally and the ones which are marked by a difference were counted. From the 137 literal translations, the typical Dutch evidentiality translations were deduced. The other 118 tokens were further analysed and two main classes were discovered. In the first place, it seemed that some Dutch subtitles express a change in type. Secondly, it was noted that the subtitles are characterised by specific formulations. Explanations for these different formulations were found and outlined in section 4.2.2 and 4.2.3. The former consists of a detailed subtitle analysis in order to examine the role of the subtitling techniques on the drastic differences in formulation. The latter investigates whether or not the changes are functionally motivated. These last sections show that the comparison between the Dutch subtitles and the English expressions does not only help to describe English evidentiality in scientific discourse, but also to give an insight into the work of the subtitler.
4 Results

4.1 Classification and investigation of the English data

The current section deals exclusively with the 255 English evidentiality expressions and contains three main subsections. The general numbers of the division into the different evidentiality types are followed by a detailed analysis of those types. This analysis uses the various kinds of verbs as a starting point. The pure evidentiality types, i.e. those which indicate direct evidence (section 4.1.1.1), reported evidence (section 4.1.1.2) and inferential evidence (section 4.1.1.3), are completed with the analysis of expressions that indicate (un)certainty (section 4.1.1.4). After the semantically-oriented part, the second part focuses on the formal encoding (section 4.1.2), whereas the third examines the functional motivations for evidentiality which are at work in the documentary (section 4.1.3).

4.1.1 The English evidentiality types

| TABLE 1. Incidence of types of evidence |

The 585 minutes of television resulted in a corpus of 255 American English evidential expressions and their subtitles, which gives an average of 19.62 examples of evidentiality in each episode. The current section presents the numbers of the different evidentiality types, which are visualised in Table 1. The diagram represents all data, including the exception
A tsunami isn't over in just a few seconds.

Men denkt dat een tsunami zo voorbij is.

Firstly, the indication of direct evidence, the distinction between visual and non-visual evidence not taken into account, represents 52 occurrences (20.39%) on a total of 255. The two main categories are the expressions that indicate reported and inferential evidence, which contain 78 (30.59%) and 81 (31.76%) examples respectively. Both types can be divided into two subcategories. Following Aikhenvald (2004), the reporting evidentials distinguish between quotative and hearsay evidence, while the inferentials indicate inference and assumption. Quotative evidentials refer overtly to the quoted source whereas hearsay evidentials do not. The reported evidence type is almost equally divided into quotative (37/78) and hearsay (41/78) evidentials. By contrast, within the inferential evidence type one finds four assumptions in comparison with 77 inference examples. Secondly, the expressions of (un)certainty count 38 tokens. Finally, five sentences represent more than one type.

4.1.1.1 Direct evidence

The first type to be described is the one that indicates direct evidence, which means the speaker was a direct witness of the event. This direct observation can be visual, non-visual or based on memory. The corpus comprises 51 tokens of direct evidence, all signalling visual evidence. Normally, the visual direct evidence is expressed by perception verbs, but in the corpus the most frequently used direct evidence verbs are to show and to reveal. The former is a special type of perception verb, namely a causative perception verb: 'To show something to someone’ means 'to make someone see something'. Similarly, the latter has a causative connotation when used with a clause, being 'cause or allow (something) to be seen'. Example (2) illustrates the visual meaning literally by using the subject footage. In the translations also, a link between to show/to reveal and to see is established by translating them as laten zien or onthullen, which makes the causative meaning explicit.

(2) This incredible time-lapse footage shows that because the top of the flow is still exposed to the air, it cools and crusts over.

---

3 Definition is quoted from the Oxford Dictionary, consulted online: http://oxforddictionaries.com/definition/english/reveal?q=to+reveal 25/04/2013
Only five expressions that indicate direct visual evidence are characterised by the verbs to see and to notice. In the corpus, to show is exclusively used with an inanimate subject, while to notice and to see require a human subject. This distinction is further developed in section 4.1.2.

4.1.1.2 Reported evidence

78 out of 255 evidentiality expressions mark reported evidence, including both quotative and hearsay evidence. The present section introduces the findings noticed when analysing this second largest group of data. Specifically, it discusses the difficulties observed when categorising the reporting evidentials, as well as the specific features of these evidentials in the genre of television documentaries.

First of all, while dividing the data into the four categories, a case of ambiguity occurred. Data (3), (4), (5) and (6) serve as illustration:

(3) **Samurai writings speak** of people living on the coast running for higher ground as soon as they felt an earthquake.
(4) **Native American myths explain** the earthquake as shocks from a battle between warrior spirits.
(5) **Research suggests** we might not be as safe as once thought.
(6) **This rock is telling** us that we had basically wet land.

The examples above challenge the notion of reported evidence on the one hand, and inference on the other hand. When evidentials signal reported evidence they usually indicate a human source, whereas inferential evidentials refer to a deduction from facts. As seen in the literature review (cf. Figure 2), the former is expressed in English by *people say, they suggest, he is reputed*, etc., in other words, by verbs of communication. When looking at sentences (3) to (6), the condition of ‘saying verbs’ is fulfilled, which, at first sight, justifies a classification as evidentials that indicate reported evidence. However, in none of the examples does a person occur who reported the statement; all the subjects designate inanimate things. Consequently, the question is raised whether these statements may be based on inference. If the answer to this question is positive, the *Samurai writings*, the *Native American myths*, the *research* and the *rock*, were investigated by the speaker and he/she inferred the statements from them. The two options, i.e. reported evidentials or inferential evidentials, are plausible but unsatisfactory, since a closer look reveals significant
distinctions between the subjects. From a poetic point of view, the lifeless *Samurai writings* that *speak* of something are an instance of personification. In this view, data (3) to (6) are all examples of personification. However, the attribution of human qualities to something nonhuman is more acceptable with examples (3) and (4). In sentence (3), for instance, the subject *Samurai writings* implies a human activity, i.e. someone writing something. Accordingly, the subject does express a person, albeit implicitly. As a result, this example of personification can be rephrased as: *The Samurai speak (in their writings) of people living on the coast running for higher ground as soon as they felt an earthquake*. The subtitler agrees with this interpretation and transfers the nonhuman subject into the *Samurai* and the saying verb into a verb of description: *De Samurai beschrijven dat kustbewoners naar hogere gebieden vluchtten, zodra ze een aarschok voelden* (*The Samurai describe that people living on the coast running for higher ground as soon as they felt an earthquake*). The Native American myths of illustration (4) require a similar interpretation because the source of myths is typically human, and because this subject is reminiscent of Anderson’s subdivisions (cf. section 1.1.2). He included myth and history into the type of reported evidence.

Because of the implication of a human (activity), this dissertation prefers to treat examples (3) and (4) as belonging to the type of reported evidence. Examples (5) and (6) need a different analysis. The former can be interpreted as partly reporting, partly inferential whereas the latter only accepts an inferential reading (*From the features of this rock we deduce we had basically wet land*) due to the totally nonhuman nature of the subject *this rock*. The following sentences illustrate the double interpretation of sentence (5):

(5) **Research suggests** we might not be as safe as once thought.

'Researchers suggest we might not be as safe as once thought.'

'From the facts represented in the research I deduce we might not be as safe as once thought.'

Both analyses are valid because a research implies the investigators who did the research, as can be seen literally in example (7), and at the same time the research represents facts deduced from a scientific investigation.

(7) **New research by geophysicist Ares Rosakis** suggests that the San Andreas may offer a new and even more deadlier threat.
Due to the scientific features of a research, the inferential interpretation is more sufficient than with the writing of myths. As a consequence, evidentials (5) and (6) are considered in this dissertation as inferentials. To sum up, similar instances of personification are evaluated according to the level of humanity or investigation implied in the source.

After having delineated what constitutes the reported evidence type, the remaining part of this section focuses on the subdivisions of the reporting evidential, i.e. hearsay and quotative evidence. Of the 78 data of reported evidence, 37 mark quotative evidence while 41 denote hearsay evidence. Given that the aim of popularised scientific discourse is to involve people from outside the scientific community, this almost equal division between referring to the specific source and not specifying the source, is not remarkable. The laypeople who are watching the documentary, are interested in the results of the research itself, not in the names of the scholars behind it of whom they have never heard. This explains the high number of hearsay uses. Another possible explanation for these almost equal numbers may be the frequent use of generic nouns. The narrator in the documentary gives a lot of information via the use of the plural, which refers to a whole group or class. Example (8) shows this generic approach and contrasts with example (9):

(8) Many geologists believe the exact same volcanic action accounts for the formation of the fifth and final lake.
(9) Geologist Mike Poland believes they're major clues in the mystery of Iceland's formation.

Since generic nouns refer to a whole class (of geologists), the subject of sentence (8) is considered vague enough to be analysed as hearsay evidence. In example (9), by contrast, the source cannot be described more specifically. This specific description is due to the fact that, in most of these quotative cases, the person mentioned by the narrator actually features in the documentary. In this view, example (9) serves as an introduction of the geologist Mike Poland.

Finally, seven of the 78 reporting evidentials are realised by alternative means, like for instance, quotation marks. With regard to quotation marks both quotative (cf. example 10) and hearsay evidence (cf. example 11) are represented.

(10) The pilot said: "It seems that we're in some sort of a chimney."
(11) The question then became: "What drives the hotspot? What causes the hotspot?"
The small number of quotation marks reveals a tendency towards indirect speech in science documentaries and is a consequence of the multimodal medium. Next to the introduction of an expert by a quotative sentence uttered by the narrator, the makers of the documentary also acknowledge the source by showing the name and profession of the speaker on the screen. Because the speaker is shown and the name is presented, quotation marks are not needed. These examples are not included in the corpus.

4.1.1.3 Inferential evidence

In view of the criteria proposed in section 3.2, most inferential evidentials4 are characterised by verbal expressions such as to prove, to reveal, to turn out, to indicate, to conclude and this is proof, this is evidence, these are clues, these are indications. Yet, none of these devices was mentioned by Ifantidou (cf. Figure 3), who indicates I gather, I guess, I suppose, I deduce as typical inferential verbs. The lack of these inferential verbs in the corpus is due to the distinct genre. As can be deduced from the use of the first person singular, Ifantidou's verbs were derived from spontaneous conversations. The verbs, which emphasise the speaker, show that the direct speaker-hearer interaction is central. In the corpus, by contrast, sentences like the constructed example (12b) are not found.

(12a) Those high pressures indicate that these rocks formerly were produced at depths of 20 to 25 miles.

(12b) 'I deduce from those high pressures that these rocks formerly were produced at depths of 20 to 25 miles.' (constructed example)

The verb to deduce, together with to guess, to gather and to suppose, only allows a person as subject of the sentence. But, in the corpus, human subjects followed by an inferential verb are non-existent. The subjects of this inference group refer to things, certain tests, experiments or investigations on which the statements are based. The preference for these inanimate things as subjects, which is the emphasised part of the sentence, has two explanations. First of all, the scientific nature of the expressions plays an important role. With a documentary which is centred around seeking and providing evidence, it is obvious that the evidence needs most attention, and this is achieved by putting it at the beginning of the sentence. Secondly, most of the inferential data are spoken by the narrator, who does not act in the 'story' he is telling. It is his job to stay in the background, so he only gives an

4 inferential here means the evidentials based on inference, not on assumption.
impersonal objective account of the facts. Thus, the conversational interaction which marks Ifantidou's verbs does not take place between viewer and documentary maker nor between viewer and narrator.

Nevertheless, the narrator does introduce names of scientists, but only as the authors of their research or findings. Examples are the following:

(13) Bauder's measurements reveal this glacier moved over ten feet every month.
(14) Morgan's research proves that geysers were exploding around the lake.
(15) Doctor Roberts' data reveal that this whole glacier is sliding forward at an astounding rate of two feet per day.

In sum, Ifantidou's verbs require a person who acts in a conversation, while the inferential data show a tendency towards impersonalisation due to the evidence-based and objective nature of the science documentary.

In contrast with the inference group, the second subdivision of the inferential type does accept human subjects. The information above discusses inferential evidentials which indicate inference, while the part that follows deals with assumption. Only four cases of assumption are comprised in the corpus:

(16) Scientists assumed the destruction must have been caused by an asteroid of at least 100 feet.
(17) Throughout most of recorded history man has just assumed that beyond certain level the sea was pretty flat, and fairly lifeless.
(18) Gilbert assumed that if the crater was caused by an asteroid, he should find a giant alien rock in the middle of it.
(19) The assumption was there had to be another volcano these days.

In examples (16), (17) and (18) above, the person who assumes is expressed. This is obvious given that the verb to assume follows the pattern 'someone assumes something'. Alternative patterns exist. To assume can be replaced by the assumption is which again expresses the typical scientific impersonalisation through nominalisation, as noted in example (19).

Both inference and assumption may be expressed by the modal must. The inferential meaning of must have/must be is illustrated clearly in the following data:

(20) The grooves in the rocks could mean just one thing. The glacier must have been thousands of feet thick.
(21) Yellowstone's hot water features all point to one conclusion. Yellowstone must be powered by the heat of a volcano.

These data are reminiscent of example (1b) of de Haan, mentioned in section 3.2. First the evidence is given and in the second sentence the findings are inferred and expressed by use of the modal must. This observation is in favour of the proposition made in the literature review: There is smoke in the kitchen. The toast must have burned. Consequently, there is no doubt that modal must may be treated as an evidential. In addition, all the examples expressed above have a high level of certainty because of the use of must. More about the second meaning of evidentials can be found in section 4.1.1.4

Although the verbs proposed by Ifantidou are not present in the corpus, the devices treated by Ifantidou as borderline cases, are. For instance, the conjunction so expresses an evidential meaning in examples (22) and (23). Its meaning is often reinforced by the evidential must.

(22) They are made of coral which can only grow under water. So some immense force, unleashed by Krakatoa, must have put them here.

(23) Today, you see here hills and trees. So, obviously, there's not a big hole left in the ground from the eruption of the Yellowstone volcano.

Finally, a small number of the inferential verbs is represented by the verbs to tell and to confirm. These verbs denote cases of personification with subjects such as these bubbles, ice cores from Antarctica, thick layers of debris, etc. As already discussed when dealing with reported evidence, they are analysed as inferential evidentials because of their nonhuman or scientific features.

4.1.1.4 Expression of (un)certainty

Next to verbs of seeing, communication or inference, the corpus consists of a vast amount of verbs that express (un)certainty such as to suspect, to know, to doubt, to think, to have a hunch. After having examined the basic evidentiality meaning, i.e. indicating information source, which is outlined above via the three types, this section turns to the second meaning, the expression of certainty or uncertainty. Particularly, these expressions represent 38 occurrences of the complete 255 tokens and are characterised by a 'verb of doubt', a 'verb of certainty', an epistemic modal, or adverbials like obviously.
As explained in section 3.1, the television documentary consists of a combination of scientists talking about their research (e.g. I believe), and a narrator who summarises their findings (e.g. scientists think). Usually the narrator, via a voice-over, starts with the doubts or beliefs of one or more scientists, which he expresses with sentences like (24). Afterwards, their ideas are substantiated with evidence and the uncertainty verbs change into verbs expressing certainty, often accompanied by the adverb of time now:

(24) Scientists had a hunch that this colossal ridge might help explain how the trench was born.

(25) We now know, by looking at meteorites like this one for example, that in fact the solar system was formed 4,567 billion years ago.

Most of the (un)certainty types establish a link with one of the information source types. By analysing example (26) below, one notices a combination of inference (so) and certainty (obviously) while sentence (25) is a clear example of how the certainty expression (we now know) is based on observation (by looking at meteorites). Examples (27) and (28) show how (un)certainty expressions may also indicate hearsay evidence, mostly expressed by a passive voice. What is more, as already mentioned, modal must is used both for the indication of certainty as well as for the indication of inference, as is illustrated by example (29).

(26) Today, you see here hills and trees. So, obviously, there's not a big hole left in the ground from the eruption of the Yellowstone volcano.

(27) Originally these were thought to be related to volcano activity.

(28) It is now known what a geological wonder the Mariana trench is.

(29) Based on the size of the boulders that we see here, we know that that float must have discharged down the order of 1.5 million cubic feet per second.

These (un)certainty types which also express the source of information reinforce the statement made in the literature review, which said that in non-grammatical evidentiality languages the expression of evidentiality frequently coincides with epistemic expressions. These findings justify the inclusion of the expression of (un)certainty as second meaning of evidentiality.

In addition, the adoption of the pragmatic inference technique is possible on most (un)certainty data. In sentence (30) the narrator says that Tankersley is convinced.
Tankersley is convinced catastrophe drove the mammoths to extinction.

If the pragmatic inference technique is adopted, one may deduce that the speaker knows Tankersley is convinced because Tankersley himself told it to him (reported evidence), the narrator inferred this conviction from research by Tankersley (inferential evidence) or the narrator deduced it from his convincing behaviour (direct evidence). Sentences (24) and (25) allow this pragmatic inference technique as well.

4.1.2 Formal encoding

As indicated before, English expresses evidentiality lexically, by using devices such as verbs (I hear, he says, I deduce, he knows), modal verbs (must, may), adverbs (allegedly, consequently, probably), adjectives (evident, possible), nouns (assumption, evidence, proof), conjunctions (so) and lexical phrases (according to). The speaker thus has a variety of choices, but this variety is not represented in the corpus. Full verbs account for the highest incidence (213/255), and are commonly accompanied by the following options:

1. Noun phrase (NP) + verb
The frequent earthquakes on Anak Krakatau indicate the volcano is still dangerous.

2. Proper noun (PN) + verb
Menzies believes the core sediments the river left behind, dramatically accelerated the ice sheet's flow.

3. Pronoun (Pron) + verb
They believe the shape of the descending tectonic plate may hold the answer.

The first category (NP + Verb) represents most occurrences (140/213). The possibility pronoun + verb (Pron + Verb) occurs in 43 cases, followed by the proper nouns + verb (PN + Verb) with 30 examples. All three combinations are used in the four evidentiality types, as visualised in Table 2 below. Note that the most frequent combination (NP + Verb) comprises both inanimate things (theory, studies) and persons (scientists, geologists). The former may express inference or direct evidence. As explained in section 4.1.1.3, almost all inference statements are based on evidence, facts, studies or research, and direct evidence is based on what is shown or revealed by a certain study. The nouns denoting humans of this specific majority combination may refer to reported evidence or the expression of (un)certainty.
The proper nouns, obviously, are almost always linked with quotative evidence, as a consequence of the definition outlined above. A smaller amount expresses (un)certainty, while direct or hearsay evidence is absent. Two exceptional cases of inference are found with the *Proper noun + verb* structure. The first one is one of the examples of assumption, repeated here as example (31). The second one, sentence (32), contains the proper name of a submarine, analysed according to its nonhuman features as an inferential example.

(31) Gilbert assumed that if the crater was caused by an asteroid, he should find a giant alien rock in the middle of it.

(32) The Trieste dived to the bode of the trench and confirmed that it is the deepest point on the planet.

Thirdly, the pronouns may refer to NPs as well as to proper nouns mentioned in the previous sentences, which explains the occurrence of this structure in all types.

**TABLE 2. Frequency of the structures NP + Verb, PN + Verb and Pron + verb**

<table>
<thead>
<tr>
<th></th>
<th>(Un)certainty</th>
<th>Direct</th>
<th>Reported</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Quotative</td>
<td>Hearsay</td>
<td></td>
</tr>
<tr>
<td>NP + Verb</td>
<td>16</td>
<td>40</td>
<td>10</td>
<td>29</td>
</tr>
<tr>
<td>PN + Verb</td>
<td>7</td>
<td>/</td>
<td>17</td>
<td>/</td>
</tr>
<tr>
<td>Pron + Verb</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>4</td>
</tr>
</tbody>
</table>

The three frequent structures as evidential markers are followed by the statement. This combination has as a consequence that the majority (83.53%) of the data are constituted by a subject, a main verb and a *that*‐clause as object. Even though the conjunction *that* is not always expressed, all the example presented above follow this pattern.

Among the less frequent formal structures, one finds nouns (15 occurrences), modal *must* (4 occurrences), conjunctions (3 occurrences), adjectives (4 occurrences), passive constructions (8 occurrences), and quotation marks (7 occurrences). Quotation marks are a non‐linguistic way to express evidentiality. Obviously, when dealing with spoken language, one cannot see the quotation marks. Nevertheless, the seven occurrences were detected because of the typical direct speech structure of the spoken sentence, the intonation of the speaker and the quotation marks used in the subtitles. The quotation marks in example (34)
indicate quotative reported evidence and are accompanied by the linguistic evidential markers NP + verb.

(33) The pilot said: "It seems that we're in some sort of a chimney."

4.1.3 Functional motivations for evidentiality

As mentioned in the literature review, the indication of the source of evidence serves the speaker’s communicative purposes. Below, the three proposed functional motivations from the perspective of the speaker are summed up, accompanied by prototypical data:

1. Acknowledge responsibility for a statement
   We see that the secret of the slipping San Andreas Fault is actually the rock itself.

2. Deny responsibility for a statement
   Experts warn if this warming continues, the Alps will by ice-free by the end of the century and they fear these great mountain peaks will tumble down even faster.

3. Assign authority to the statement
   Da Vinci proposed that these fossils were formed under the ocean. Marine fossils are evidence that these rocks were once covered by a tropical sea.

When applying these functional motivations, which were proposed in the literature review, to the data, a problem occurred. It was noticed that these functional motivations are not sufficient when dealing with the data of the specific scientific corpus. For instance, they do not take into account inferential sentences like example (34). Following the functional motivations expressed above, the speaker of the sentence would be classified as denying responsibility for the statement because he is no direct witness. At the same time, the sentence itself implies a high degree of certainty and conviction expressed by the speaker, which means he does act as responsible for the statement. Because of this flaw in the categorisation, it is recommended to start from a broader point of view, namely the engagement types proposed by Martin and White (2005).
The theory showed that the earth's crust consists of separate moving plates on which the oceans and continents sit.

The first part of the analysis (section 4.1.3.1) sheds light on the engagement categories which use an evidential marker. The second (section 4.1.3.2) tries to find the 'engagement motivations' of the documentary maker and/or the narrator.

4.1.3.1 Engagement positions

To fully understand the communicative purpose of the speaker and the documentary maker to use evidential expressions, a broad interpretation of functional motivations is needed. Martin and White (2005) deal with an engagement system which is dialogically orientated. They adopt a broad discourse-oriented approach and focus on how the speaker 'engages' with "prior utterances, alternative viewpoints and anticipated responses" (Martin and White 2005: 97). Their engagement system consists of several dialogistic positions. First a distinction is made between dialogically contractive and dialogically expansive locutions. The former contains the categories, 'disclaim' and 'proclaim', while the latter consists of 'entertain' and 'attribution'. The diagram below presents the categorisation and the main definitions.
CONTRACTION

Disclaim: The textual voice positions itself as at odds with, or rejecting some contrary position.

Proclaim: By representing the proposition as highly warrantable (...) the textual voice sets itself against, suppresses or rules out alternative positions.

Entertain: by explicitly presenting the proposition as grounded in its own contingent, individual subjectivity, the authorial voice represents the propositions as but one of a range of possible positions - it thereby entertains or invokes these dialogic alternatives.

Attribution: By representing proposition as grounded in the subjectivity of an external voice, the textual voice represents the proposition as but one of a range of possible positions - it thereby entertains or invokes these dialogic alternatives.

EXPANSION

Deny

Counter

Concur

Pronounce

Endorse

Acknowledge

Distance

(Martin and White 2005: 97-89)

The analysis starts with the explanation of the engagement resources, which can all be exemplified by the evidential data. Therefore, they can be seen as functional motivations to use the kind of evidentiality expressions described in the corpus. The data which illustrate the category 'entertain' and the subcategories of the categories 'disclaim', 'proclaim', and 'attribution' are presented below.
The first group of examples demonstrates the dialogically **contractive** wordings, which means "they close down the space for dialogic alternatives" (Martin & White 2005: 103).

(35) A tsunami isn't over in just a few seconds.

(36) This would suggest that volcanic eruptions still happen, **even** under the enormous weight of ice.

(37) Today, you see here hills and trees. So, **obviously**, there's not a big hole left in the ground from the eruption of the Yellowstone volcano.

(38) **To me** that's not something that can be explained by something that's just near the surface.

(39) Contemporary reports **show** that most of Krakatau victims were not killed by these deadly avalanches.

Example (35) is the only instance of **denial** found in the corpus. The denial is directed to the addressee and anticipates the "beliefs which they (sic) speaker/writer assumes that at least some members of his/her mass audience will be subject to" (Martin & White 2005: 119). The following extract from an advertisement placed in magazines by the British Heart Foundation is given as example:

> We all like something to grab hold of. But sometimes you can have too much of a good thing. And a man whose tablet diet consists of double cheeseburgers and chips can end up looking like a tub of lard. **There's nothing wrong with meat, bread and potatoes.** But how about some lean meat, wholemeal bread and jacket potatoes?
> (Martin & White 2005: 118)

The denial, **There's nothing wrong with meat, bread and potatoes**, includes opinions or claims by people who believe that 'there is something wrong with meat, bread and potatoes' (Martin & White 2005). Similarly, in example (35) the narrator responds to his audience who may believe 'a tsunami is over in just a few seconds'. This example is included in the corpus because its subtitle expresses evidentiality, a point which will be discussed in section 4.2.3 which deals with the Dutch subtitles. In addition, denials of this corrective type "present the addressee as having greater expertise in some area than the addressee and as, on that
basis, acting to correct some misunderstanding or misconception on the addressee's part" (Martin & White 2005: 120). This goal coincides with the specific aim of popularised science, which is educational.

In the same category of denial, the countering proposition is found. In example (36) the scientist/speaker counters the proposition that would have been expected. The proposition that eruptions still happen is in counterrelationship with the enormous weight of ice. Countering examples do not use evidential markers to express their purpose, and are therefore excluded from the analysis. Example (37), which belongs to the inference and certainty type, is an instance of concurrence, which "overtly announce[s] the addressee as agreeing with, or having the same knowledge as, some projected dialogic partner" (Martin & White 2005: 122). With the adverb of certainty obviously, the speaker makes clear that the statement that follows is the only correct interpretation of the findings.

A reporting example where the speaker is explicitly present denotes pronunciation. In example (38) the speaker gets involved and shares his opinion, which is clearly by to me. According to Martin and White (2005: 128) "such insistings or emphasising imply the presence of resistance, some contrary pressure of doubt or challenge against which the authorial voice asserts itself." By doing so, he reduces the possibility to express alternative opinions. The speaker has also a particular role with endorsement. Endorsements are "those formulations by which propositions sourced to external sources are construed by the authorial voice as correct, valid, undeniable" (Martin & White 2005: 126). Consequently, as illustrated by the inferential example (39), the speaker takes over responsibility for the proposition and expresses his confidence towards the cited source. Endorsement is typically represented by the verbs to show, to prove and to demonstrate. (Martin & White 2005)

The following group of examples belongs to the dialogically expansive ones. With these examples "the effect is to invite or at least entertain dialogic alternatives and thereby to lower the interpersonal cost for any who would advance such an alternative" (Martin & White 2005: 103). The inferential example (40) belongs to the category of entertain, which contains "those wordings by which the authorial voice indicates that its position is but one of a number of possible positions and therefore, to greater or lesser degrees, makes dialogic space for those possibilities" (Martin and White 2005: 104). This category is considered the default category for modality and evidentiality. By expressing uncertainty (I think) or surmise (apparently, to suggest, it seems) the speaker presents the propositions as one among
potential alternatives (Martin & White 2005). In example (40) the verb suggest indicates that the speaker is willing to hear alternative options.

(40) New research by geophysicist Ares Rosakis suggests that the San Andreas may offer a new and even more deadlier threat.

(41) People say that more than half a million people died from the drought in India.

(42) The politician claims that only 2,000 people died from the drought in India. (constructed example)

Note that I believe refers to 'entertain', while they believe or he believes is seen as an acknowledgement. In contrast with endorsement, the author of an acknowledgement is considered neutral. The speaker only reports, by means of reporting verbs such as to say, to believe, to announce, without adding a personal value to the proposition. As a consequence, the writer is presented as an "informational fair trader" who simply reports the opinions of others without taking a stand (Martin & White 2005: 113). Nevertheless, the reader is able to value the statement as reliable and credible or as unreliable and doubtful (Martin & White 2005). For example, the subject people in the reporting example (41) may be seen as less reliable than for instance the subject The United Nations.

Another subcategory of the 'attribution' category is distancing, which is the complete opposite of endorsement. The speaker 'distances' "him/herself from the responsibility for what is being reported" (Martin & White 2005: 113). Obviously, this maximises the space for dialogistic alternatives. Example (42) is a constructed example, since the corpus does not contain an example of distancing. An explanation for this will be given in the following section.

4.1.3.2 Summary and interpretation

The outsourcing or acknowledgement of responsibility on behalf of the speaker is thus more complicated than initially thought. The main engagement types to express responsibility are represented on the axis below, from low to high responsibility acknowledged by the speaker.
Because of the dialogistic approach, the types also take into account the response by a putative dialogic partner. The possibility of dialogic alternatives is represented by Figure 4, from contractive, i.e. excluding dialogic alternatives, to expansive, i.e. invoking dialogic alternatives.

**Figure 3.** Responsibility for the statement expressed by the speaker

![Figure 3](image_url)

**Figure 4.** From contraction to expansion

![Figure 4](image_url)

As seen in the previous section, (almost) all engagement subcategories make use of evidential expressions. Now it is important to discover which engagement options, which are seen as functional motivations, are at work in the popularised form of science, and are consequently responsible for the division of the different evidentiality types in the corpus. To find this, we need to begin at the starting point: the genre.

The documentary makers work with a narrator who talks most of the time. One may say that he is moved by two communicative purposes. On the one hand, it is his job as an objective narrator to stay in the background and to explain things or to report what other people say. Objectivity corresponds to neutrality. So, when looking at figure 3, this first goal is reminiscent of acknowledgement. When applying the category 'acknowledgement', the narrator acts indeed as an 'informational fair trader'.

On the other hand, the narrator has to bear in mind the aim of the documentary, which is explaining scientific facts to laypeople and convincing them of certain scientific findings. For this task he needs undeniable facts and a low input of contradictory opinions. The combination of Figure 3 and Figure 4 shows that endorsement corresponds to these needs. The propositions are presented as true and convincing, because the narrator himself
takes the responsibility for it, which means he trusts the source, and therefore not allows contractive alternatives.

The specific aim of the genre is also responsible for the small role of distancing. The narrator and the scientists try to convince the audience, not to distract them or make them doubt. In addition, because of the important role of the narrator, the categories concurrence and pronouncement are of a minor importance in the documentary. They imply the intervention of a person, and therefore they will only influence the scientists who feature in the documentary. This is also the reason why, next to the modality adverbials and the suggestions or beliefs expressed by scientists, the representation of entertain examples will be scarce in the documentary. Finally, denial and countering are of no importance because they do not invoke an evidentiality structure, unless in the subtitles.

In conclusion, when starting from the needs and aim of the genre, endorsement and acknowledgement are the main motivations to use evidentials in a documentary. This observation explains the high incidence of reported evidence (people say that) and inferential evidence (research shows) found in the corpus.
4.2 English speech versus Dutch subtitles

This comparative study focuses on the subtitles of the English evidentiality expressions. Via the comparison between the Dutch subtitles and the English original sentences, the present part aims to shed light on the multifunctionality of the English evidential markers. Furthermore, the investigation of the translations offers an insight into the evidentiality system of Dutch and the motives of the subtitler. First of all, section 4.2.1 provides the differences between original and translated text in terms of expressions of evidentiality. The section is divided into two parts, which outline the distinctions between the English and Dutch evidentiality types on the one hand, and the differences in formal encoding on the other hand. Subsequently, the reason for the drastic differences is searched in the subtitling restrictions (section 4.2.2.) and finally, it is examined whether the differences may be due to different functional motivations of the subtitler.

4.2.1 Differences

A slight majority (53.70%) of all the Dutch subtitles offers a specific and literal translation of the English evidentiality expressions, as can be seen in example (1):

(1) The repeated ultra precise measurements reveal that land here on the surface hardly moves at all.

De herhaalde superprecieze metingen onthullen dat de grond aan het oppervlak amper beweegt.

As a result, the following Dutch evidentiality verbs are derived from the translations. Direct evidence, which is indicated by English verbs such as to show, to notice and to reveal, is expressed in Dutch by the verbs (aan)tonen, laten zien, (op)merken and onthullen. Secondly, the verbs of saying which mark reported evidence are translated into Dutch as: vragen (to ask), begrijpen (to understand), bevestigen (to confirm), geloven (to believe), zeggen (to say), voorstellen (to propose), waarschuwen (to warn). The Dutch verbs that indicate inference are: bewijzen (to prove), bewijs leveren (reveal evidence), aannemen (to assume), moeten (must) and aangeven (to indicate). Lastly, the Dutch verbs weten (to know), denken (to think, to have a hunch) and vermoeden (to suspect) express certainty or uncertainty.

All the literal translations are characterised by an evidentiality verb. The verbs mentioned above are thus the recurring Dutch evidentiality verbs in the science documentary. In comparison with the evidentiality forms proposed by de Haan (modal verb
moeten, modal verb zouden, raising verb schijnen and the perception verbs), the current corpus provides a broader view on the Dutch evidentiality verbs.

Beside the literal translations, 118 cases of the 255 tokens (46.27%) show a difference between the original English evidentiality expression and its Dutch subtitle. These differences range from minute alterations, like a distinct verbal expression, to more significant ones, like the change of evidentiality type. Accordingly, the categorisation that follows encompasses two classes. The first one (section 4.2.1.1) discusses the distinctions with regard to the types of evidentiality, i.e. either the type is changed or not expressed in Dutch, while the second one (section 4.2.1.2) deals with the formal possibilities of evidentiality in English and in Dutch.

4.2.1.1 Types

This section starts with examples of the most drastic change found among the 118 subtitles which differentiate from the English evidentiality expressions. The subtitles of ten English sentences demonstrate a complete deletion of the evidential meaning by reducing the English evidentiality expression to a pure declarative statement. This omission occurred with direct evidence types, as well as with reported and inferential evidence types. Examples (2) to (4) illustrate the removal of the underlined evidentiality markers for the three types:

(2) We see that the secret of the slipping San Andreas fault is actually the rock itself. (Direct evidence)
Het schuiven van de San Andreasbreuk wordt veroorzaakt door de steen zelf.

(3) People agree that only a massive force could have created such a huge crater. (Reported evidence)
Zo'n krater kon alleen door een enorme kracht zijn veroorzaakt.

(4) It turns out there's a really strong relationship between the age of the seafloor and its depth in the water. (Inferential evidence)
Er is een sterk verband tussen bodemouderdom en waterdiepte.

Another example, in contrast with the data above, is the unique occurrence that is marked by the addition of an evidential meaning. The English source sentence does not express evidentiality, but the subtitler adds a hearsay evidence marker (men denkt) which signals uncertainty.
A tsunami isn’t over in just a few seconds.

Men denkt dat een tsunami zo voorbij is.

The reason for the addition or deletion of the evidential meaning in the subtitles will be discussed in section 4.2.2 and 4.2.3, where the subtitles are analysed more closely.

Next to the deletion or addition of an evidential meaning, 52 Dutch subtitles contrast clearly with the original expressions because they belong to a different evidentiality type. The numbers are represented in Table 3.

**TABLE 3.** Change of types from English to Dutch

<table>
<thead>
<tr>
<th>English/Dutch</th>
<th>Direct</th>
<th>Reported</th>
<th>Inferential</th>
<th>(Un)certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td></td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reported</td>
<td></td>
<td></td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Inferential</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>(Un)certainty</td>
<td>5</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

25 evidentiality expressions change from the English reported evidence type *(to believe)* to the Dutch expression of uncertainty *(denken)*. 19 examples indicate in English direct evidence *(to show, to reveal)* whereas their subtitles indicate inferential evidence *(bewijzen, blijken)*. The English expression of (un)certainty *(to know, to think)* is represented by six tokens, five of which transfer to reported evidence *(volgens)*, one to inference *(moeten)* and two go the other way around, i.e. from inference *(to suggest)* to uncertainty *(denken)*.

The largest group (25 occurrences) establishes a change from the reported evidence type *(R)* to the expression of (un)certainty *(U/C)*. 19 occurrences of this group are characterised by the verb *to believe*. Originally, *to believe* was considered a marker of reported evidence, whereas in Dutch it is 19 times translated as the expression of uncertainty with the verb *denken* *(‘to think’)*. What is more, in the complete corpus only six subtitles translate *to believe* as indicating reported evidence. This is obviously due to the double meaning of the verb *to believe*, which is first ‘accept that (something) is true, especially without proof’ and second ‘hold (something) as an opinion; think’. The comparison below illustrates the two meanings respectively:

5 Definitions are quoted from the Oxford dictionary, consulted online: http://oxforddictionaries.com/definition/english/believe?q=believe, 14/05/2013
Lyell believed the world wasn't shaped in a few days or even years. (R)
Lyell geloofde dat de wereld niet in enkele dagen geschapen was. (R)
Lyell believed this process had been going on for thousands of years. (R)
Lyell dacht dat dat proces al duizenden jaren gaande was. (U)

In a similar way, the second largest group (19 occurrences) deals with the verbs to show and to reveal. Even though I have treated to show as a direct evidence marker because of its causative perceptive character explained in section 4.1.1.1, some subtitles (9 occurrences) translate to show as indicating inference by using the verbs: uitwijzen, bewijzen, blijken. Example (8) illustrates the change from direct (D) to inferential evidence (I).

(8) Extinct volcanoes show that this process began on the coast. (D)
Dode vulkanen bewijzen dat dit proces aan de kust begon. (I)

In comparison, only ten tokens of the whole corpus contain subtitles which translate to show as a direct evidence marker with the verbs (aan)tonen and laten zien.

Similarly, the verb to reveal was also treated as indicating direct evidence, but is sometimes translated as indicating inference. Of the 24 direct evidence examples with to reveal, ten express inference in Dutch with verbs such as bewijzen ('to prove'), erop duiden ('to indicate') and blijken ('to become clear'). Example (9) shows a translation that indicates direct evidence, whereas in examples (10) and (11) the subtitles indicate inference. The translation with blijken will be discussed in the following section. When to reveal occurs in combination with the noun evidence, it is considered as indicating inference in English as well as in Dutch, as illustrated by example (12):

(9) Doctor Roberts' data reveal that this whole glacier is sliding forward at an astounding rate of two feet per day. (D)
De metingen van dokter Roberts laten zien dat deze gletsjer met een verbazingwekkende snelheid van 60 cm per dag verschuift. (D)

(10) The pumice from 1883 reveals that the Krakatau was a stratovulcano. (D)
Het puimsteen uit 1883 bewijst dat de Krakatau een stratovulkaan was. (I)

(11) Microscopic analysis reveals the grains have been cemented together by chemicals and pressure. (D)
Uit analyse blijkt dat de korrels samengeklonterd zijn door chemicaliën en druk. (I)
The nest reveals strong evidence that plants and mammals did exist here. (I)
Het nest bewijst dat hier vroeger planten en zoogdieren leefden. (I)

These observations, which are based on the subtitles, prove that, in contrast with the assumptions expressed in the first study, the three English verbs to believe, to show and to reveal are able to express more than one type of evidence.

After having mentioned the translation of English reported evidence into the Dutch expression of (un)certainty, this paragraph deals with the opposite. Examples include:

(13) Merguerian is convinced that the same thing happened in Central Park. (C)
Volgens Merguerian is dit ook in Central Park gebeurd. (R)

(14) Tankersley is convinced catastrophe drove the mammoths to extinction. (C)
Volgens Tankersley zijn de mammoeten door een ramp uitgestorven. (R)

(15) Geologist David Franzi knows that only a raging torrent could have shifted them. (C)
Volgens David Franzi kan alleen een woeste stroom ze hebben verplaatst. (R)

Example (13) is analogous with example (24) discussed in section 4.1.1.4 of the first study and here recapitulated as example (14). When discussing example (14) in the first analysis, the pragmatic inference technique was applied. It was said that, next to its expression of certainty, one may interpret the sentence as indicating direct evidence (the narrator deduced the statement from Tankersley's convincing behaviour), reported evidence (Tankersley himself told it to the speaker), as well as inferential evidence (the narrator inferred this conviction from research by Tankersley). As a result, instead of the literal translation Tankersley is ervan overtuigd dat, the subtitler decides to ignore the certainty expression and translates the sentence with the structure volgens X (‘according to X’), which denotes that the statement is derived from something X said (reported evidence). Similarly, in example (13) and (15), the subtitler preferred to ignore the certainty expression and indicates, in these cases, quotative reported evidence. This and other uses of the preposition volgens will be examined concisely in the section that follows.

Finally, in example (16) the subtitles add an evidential type to the original one:

(16) These bubbles suggest to geologists that the Krakatau was a stratovulcano. (I)
Hierdoor denken geologen dat de Krakatau een stratovulkaan was. (U/I)
At first sight, example (16) expresses a change from inference (these bubbles suggest) to uncertainty (geologen denken), but the inferential meaning still remains present through the word hierdoor ('because of this').

In sum, one notices a variety of differences in type between the English and the Dutch evidentiality expressions. Nevertheless, the use of these particular subtitles are justified by the multiple evidential meanings of certain English verbs, the application of the pragmatic inference technique or the addition of an epistemic marker beside the original evidentiality type.

4.2.1.2 Formal encoding in Dutch

After having treated the semantic differences, the current section zooms in on the formal differences between the English evidentiality expressions and their Dutch subtitles. The data which represent a change in type are re-evaluated in terms of their formal encoding. This analysis is followed by the subtitles which differ in formulation but indicate the same type of evidence.

The arguments of the 'difference in type' examples mentioned in the previous section are characterised by a change of function. For the formal analysis, additional examples, which represent the deletion of the evidential meaning, the addition, and the change of type respectively are given:

(17) He noticed below the falls was a great gorge which locals said was steadily increasing in length as the water wore away the edge of the falls.
    Onder de waterval lag een grote kloof die steeds langer werd naarmate de rand van de waterval wegsleet.

(18) A tsunami isn't over in just a few seconds.
    Men denkt dat een tsunami zo voorbij is.

(19) These bubbles suggest to geologists that the Krakatoa was a stratovulcano.
    Hierdoor denken geologen dat de Krakatau een stratovulkaan was.

The change of type as a translation strategy is represented by example (16), here re-numbered as (19), which was described as adding uncertainty to the inferential type. The other changes (from reported to uncertainty, direct to inference etc.) given in the previous section, are not useful for this formal analysis. The structure of these subtitles does not
differ from the source structure, except when the translations use *blijken* and *volgens* which will be discussed below as separate categories.

When analysing the examples above, which do express a formal discrepancy, one notices the following. First of all, example (17) shows that the divalent verb *to notice* has a human subject (*he*) and a *that*-clause as object (even though *that* is not expressed). In Dutch the *that*-clause becomes the main clause, because the evidential marker (*he noticed*) is deleted. As a result, the internal structure of the *that*-clause is maintained, and the subject (*a great gorge*) of the *that*-clause becomes the subject (*een grote kloof*) of the Dutch main clause. The other examples (cf. section 4.2.1.1) with a deletion of the evidential meaning are translated according to the same principle. Secondly, in the example of addition, the opposite occurs. The main English statement becomes a *dat*-clause in Dutch. Thirdly, in example (19) the change of type causes a change of arguments. In the original sentence the subject denotes a nonhuman thing (*these bubbles*), and the object is a *that*-clause. In the translation, by contrast, the subject is human but the object remains a *dat*-clause. What is striking is that *geologists* is a prepositional object with a benefactive function in the source text while in the Dutch translation *geologen* functions as the subject. So, even though the two verbs are divalent, and using a *that*-clause (*I suggest that, Ik denk dat*), the arguments are changed.

Concerning the formal encoding of the **data without change in type**, four different Dutch formulations are distinguished. First one finds a small amount of the preposition *volgens* (*according to*) (11 occurrences), followed by 17 occurrences of the verb *blijken* (*to become clear*), 24 of *men* (*one/they*) and finally, 26 examples of various other structures.

First of all, the smallest group is typified by the preposition *volgens*. Next to the examples in section 4.2.1.1 (cf. examples 13, 14 and 15) which express certainty, seven tokens of which the subtitles are characterised by *volgens* express reported hearsay evidence (cf. example 20) and one signals inferential evidence (cf. example 21). These examples prove that the preposition *volgens* is used to indicate either that the statement is derived from something a person said (reported evidence) or that it is derived from actual facts (inferential evidence), although the latter occurs less frequently in the corpus.

(20) *Experts believe* that in the US New York is the third most vulnerable city after Miami and New Orleans, for a hurricane disaster.

*Volgens experts* is *New York de op twee na kwetsbaarste Amerikaanse stad na Miami en New Orleans, voor een orkaanramp.*
(21) The study estimates that a major earthquake in the LA metro area would cost 2000 dead, 50,000 injuries and 200 billion dollars of damage.

Volgens een studie uit 2008 zou een zware beving in LA 2000 doden, 50.000 gewonden en 200 miljard dollar schade eisen.

Most of the subtitles with volgens change the NP + verb structure or the PN + verb group. The subtitler turns the English human subject X into a part of the prepositional phrase volgens X. As can be seen in example (20), the subject which acts in the English sentence (experts) is preceded by volgens in the Dutch subtitles. Again, the object, which is a that-clause, becomes the main clause. Consequently, the subject of the that-clause (New York) becomes the subject of the main clause. This causes the deletion of the original verb in favour of the verb of the that-clause. Semantically speaking, the agent of the original sentence (experts) becomes a source in the subtitle. This is also the case with the sentences expressed above, which express certainty.

Nevertheless, in spite of the frequent use of the preposition volgens, 31 English expressions with a human noun or proper noun as subject are subtitled literally, of which example (22) is an illustration:

(22) Geologists believe that the trench formed the volcanoes by a process called subduction.

Geologen denken dat de vulkanen ontstonden bij een proces dat subductie heet.

Aside from the literal translation and the structure volgens X, the reported evidence is also introduced by the pronoun men, which occurs in 24 subtitles. Men is used with expressions of reported evidence (10 occurrences), (un)certainty (10 occurrences) and inference (4 occurrences).

First, when looking at the reported evidence type and the (un)certainty type, the generic subjects scientists, geologists, etc. and the personal pronoun they are replaced by the more general term men. As deducible from example (23), scientists now believe is translated more generally as men denkt, which justifies the classification of the generic nouns as subjects of hearsay reported evidence instead of quotative. Furthermore, the replacement of the subject by men does not affect the internal structure, given that men is also subject in the Dutch sentences. In addition, its use reinforces the statement made in the first study, which said that for the target group of the science documentary the fact that the information being communicated is the result of scientific research is more important than
knowing exactly who carried out this research. Additionally, men is the default translation of the passive voice (cf. example 24).

(23) *Many scientists now believe* that the damage was greater than expected because the 1906 had traveled a super shear speed.

\[*Men denkt nu dat de schade groter was dan verwacht doordat de aardbeving uit 1906 een supershear snelheid had.*\]

(24) *It is thought* that eruptions began in this area over 36 million years ago.

\[*Men denkt dat de uitbarstingen hier 36 miljoen jaar geleden begonnen.*\]

Secondly, four cases of inference are translated by *men*. At first sight, this is surprising because most of the inferential subjects are nonhuman (*research, clues, dating, etc.*), but two of these four cases are assumptions, in which humans are implied. The two other expressions follow sentence (25), in which a human subject is introduced by *men*:

(25) *Ice cores From Antarctica suggest that Krakatau may have exploded before.*

\[*Toch leest men uit ijskernen af dat hij al eens uitgebarsten kan zijn.*\]

A third group is characterised by the verb *blijken*. 12 of these 17 sentences express inference, followed by five examples of direct evidence with the verb *to show* which are mentioned in the previous section because of their change of type. Example (26) and (27) exemplify the two possible structures with *blijken*. In example (26) the subject of the inferential sentence, that expresses the crucial evidence \(Y\) from which the statement is inferred, is changed into the pattern *uit/na \(Y\) blijkt + statement*. Similar to the *volgens* group, the subject is changed into an adverbial.

(26) *Radiometric dating proves this rock is about 450 million years old.*

\[*Uit radiometrische datering blijkt dat de steen 450 miljoen jaar oud is.*\]

(27) *The map revealed* that the Mariana trench is just a tiny fraction of a network of enormous underwater canyons, stretching right around the planet

\[*De Marianentrog bleek te behoren tot een netwerk van enorme zeekloven dat de hele planeet omvatte.*\]

Example (27), by contrast, is an instance of subject to subject raising (Auwera & Noël 2011). Instead of saying *Het bleek dat de Marianentrog behoort tot een netwerk van enorme zeekloven dat de hele planeet omvatte*, the subject of the *dat*-clause is raised to the subject of the main clause, followed by the verb *blijken* and a *te*-infinitive. In a similar way, when comparing the
English and Dutch sentences (cf. example 17,20,21,26,27), the Dutch translation offers a kind of 'subject to subject raising' as well. The subject of the that-clause in English is raised to the subject of the main clause in the subtitles.

In addition, the use of *blijken* causes the omission of the expert in direct evidence examples, such as

(28) Recreating how the land had moved showed Hudnut that the two parts of the creek had traveled more than three hundred feet apart.

*Uit reconstructie van de beweging van het land *blijkt* dat de twee delen van beek ruim 90m uit elkaar geschoven zijn.*

In example (28) one notices a decrease in valency because of the use of the verb *blijken*. The subject (*recreating how the land had moved*) is followed by the predicate (*show*), the indirect object, and the object (*that-clause*). The Dutch translation, on the other hand, consists of two arguments, one being a prepositional phrase as counterpart of the English subject, the other being the *dat*-clause, which is the translation of the *that*-clause. The human beneficiary is deleted because of *blijken*.

The deletion of an argument happens also with the verb *schijnen*. In example (29) the reporters are deleted, due to the use of the specific Dutch verb. *Schijnen* designates that the utterance is based on what is said. As a consequence, *schijnen* encompasses perfectly the hearsay reported evidence which is expressed by *people say*.

(29) *People say* that more than half a million people died from the drought in India.

*Daarbij *schijnen* meer dan een half miljoen mensen te zijn omgekomen.*

The last group comprises *various forms*. First of all, the Dutch subtitles tend to verbalise the English expressions. For instance, *to be proof* becomes *bewijzen*, *to be clues* becomes *vertellen*, *it is a clue* becomes *er is een aanwijzing zijn*, *to be evidence* becomes *bewijzen*, *to come to the conclusion* becomes *concluderen*. What is more, the expression *to be proof* is never translated literally in the complete corpus, while *to indicate* becomes *een aanwijzing zijn*, *evidence suggests* becomes *er is bewijs dat* and *to prove* becomes *bewijs hebben* in only three occurrences. Furthermore, in spite of the frequent use of *volgens*, the English counterpart *according to* is never used.

---

*Translation of definition found in Van Dale dictionary, consulted online http://www.vandale.be/opzoeken?pattern=schijnen&lang=nn#.UZONncoyc80. 05/05/2013*
As with the *blijken* group, some subtitles reduce the evidential markers to one element. The Dutch subtitles of example (30) express inference only by the conjunction *dus*, while in example (31) the additional reported evidence *leading investigators to conclude* disappears and the inferential meaning remains:

(30)  *Finding this told* them the Mid-Atlantic ridge was highly volcanic.
*De Mid-Atlantische Rug was dus zeer vulkanisch.*

(31)  *There are hundreds of layers of salt, leading investigators to conclude* the sea must have dried up and refilled hundreds of times.
*Er zijn honderden zoutlagen. De zee moet dus honderden malen drooggevallen en weer volgelopen zijn.*

Other distinct formulations are the use of an adverb (*vermoedelijk*) to replace a NP + Verb structure (*scientists suggest*), a lexical phrase (*naar we denken*) instead of a hearsay translation (*it is thought*), the transfer from direct to indirect speech, and the deletion of a conjunction like *so*.

To sum up, the Dutch language uses some particular structures to express evidentiality. First, the pronoun *men* illustrates clearly the unspecified nature of the hearsay subjects, but does not change the internal structure drastically. Second, the preposition *volgens* is used to express reported evidence. When *volgens* is used, the English subject, which is the agent, turns into an adverbial which expresses the source. This change from subject to adverbial also happens with the first possibility of the verb *blijken*, which is *uit/na Y blijkt + statement*, where *Y* denotes the inferential evidence. The second possibility with *blijken* is subject to subject raising, which means that the subject of the *that*-clause becomes the subject of the main clause. This raising can be applied to the subtitles in general. Most of the examples of the *volgens* group and the *blijken* group are translated as follows. The subject of the English *that*-clause becomes the subject of the Dutch main clause, which could be considered as the subject of the *that*-clause that 'raises' to the subject position in the Dutch translation.

### 4.2.2 Influence of subtitles

The radical differences in formulation, as expressed above, are often caused by the subtitling restrictions, led by reading speed and synchronicity. This section focuses on the
shortened subtitle sentences in order to find out whether (part of) the evidentiality meaning is deleted because of the advantage the shortening offers for the viewer.

In the literature review, I introduced the following subtitling techniques: (1) condensing the text, (2) omission or paraphrase, (3) muddled speech, (4) ellipsis, (5) merging short dialogues, (6) simplifying the syntax, (7) simple vocabulary, (8) subtitle breaks, (9) borrowing time. I will not focus on (3) muddled speech and (5) merging short dialogues. The use of dialogues in a television documentary is minimal, or in this case, nonexistent because the majority of the text is spoken by means of a voice-over. The other parts consist of monologues by professors, geologists or other scientists. They speak as clearly as possible to explain their theories or ideas. Consequently, in the current corpus, neither muddled speech nor dialogues occur.

The remaining strategies can be divided into three categories, being deletion, time or simplifying strategies. As seen previously, some subtitles delete (evidential) parts of the original English sentence. The condensation of the sentences is due to the limited time and/or space to display the subtitles. Moreover, the restrictions on time and space are led by the reading speed of the viewers and the need of synchronicity between subtitle and sound and subtitle and image. According to Ivarsson and Carroll (1998) the reading speed is much slower than for instance when reading a book, because of the images and sound the viewer has to absorb while reading the subtitles. For a subtitle to be completely read and understood, one line should be kept on the screen for about three seconds, while one and a half lines need at least four seconds (Ivarsson & Carroll 1998). In addition, when dealing with reading speed, it is important to bear in mind the target group. For this documentary, which is originally broadcasted on History Channel, one has to take into account all kinds of audiences; youngsters who are accustomed to subtitles and the English language, as well as elders who experience more difficulties with the multimodal genre. The mix of both possible target groups automatically slows down the reading speed.

Next to reading speed, another important notion to remember is the synchronicity between the subtitle and both the image and the sound. The most important thing for the subtitler is the beginning and end of the subtitles which have to coincide with the speech of the speaker. In addition, it is also crucial that the subject of which the narrator is talking, is displayed on the screen. (Ivarsson & Carroll 1998).

The following table, which is based on de Linde and Kay's (1999) analysis, sheds light on the reading speed and synchronicity of some data.
Table 4 illustrates how long the speaker needs to utter the sentence (5 seconds), how long the subtitle appears on the screen (5 seconds) and how long the image remains the same (5 seconds). First of all, according to the optimal time in seconds proposed by Ivarsson and Carroll, the subtitle which consists of two lines appears long enough on the screen to be read and understood by the viewer. Secondly, the 'on-set' and 'off-set' times of the speech coincide with those of the subtitle. In the corpus in general, a subtitle never appears longer on the screen than the voice of the speaker, even though two seconds without speech follow. Thirdly, the image does not change during the speech and the appearance of the subtitle. As a result, the criteria of synchronicity and reading speed are fulfilled. This would not have been the case when the English expression would have been translated literally: *Men is akkoord dat zo'n krater alleen door een enorme kracht kon zijn veroorzaakt.* The non-abbreviated sentence needs more than two lines, and therefore requires more seconds to be read and understood, which would challenge the synchronicity between the speech and the subtitle. Also the synchronicity between the subtitle and image would change, since a new line commonly coincides with a new image.

When a subtitle translation does consist of three lines, a good application of **subtitle breaks** is needed. In general, the subtitle breaks (//) are marked in accordance with the intonation of the speaker or with the image that follows. The division of the example below, which, albeit the deletion of the evidential meaning, consists of three lines, is done as follows:
### TABLE 5. Time analysis of a 'deletion' example

<table>
<thead>
<tr>
<th></th>
<th>On-set</th>
<th>Off-set</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time in seconds</strong></td>
<td>0 seconds (17:30)</td>
<td>7 seconds (17:37)</td>
</tr>
<tr>
<td><strong>Speech</strong></td>
<td>The map revealed that the Mariana trench is just a tiny fraction of a network of enormous underwater canyons // stretching right around the planet</td>
<td></td>
</tr>
<tr>
<td><strong>Subtitle</strong></td>
<td>De Marianentrog bleek te behoren tot een netwerk van enorme zeekloven // dat de hele planeet omvatte.</td>
<td></td>
</tr>
<tr>
<td><strong>Image</strong></td>
<td>Image of the earth from above</td>
<td>Close-up of the earth</td>
</tr>
</tbody>
</table>

In this case, the speaker pauses after the words *underwater canyons*, which means the long first part (cf. image below) has to be read and understood during seven seconds. In addition, the part that follows the *underwater canyons* coincides with a new image, which means the first part has to be read before the image changes. This is important because "when a shot change occurs in the middle of a subtitle, viewers return to the beginning of a partially read subtitle and start re-reading" (Baker 1982 qtd. in de Linde & Kay 1999: 48). Accordingly, the subtitler deletes the first words, in order to be sure the first long part of the sentence is read and understood before the image changes. Additionally, the sentence has to be shortened because the first part already occupies the available space:
The synchronicity between subtitle and image has also an influence on the reading speed. In the example below, Dr. Cheryl Jaworowski of the Yellowstone national Park utters the first sentence in five seconds, so the first image, which portrays Dr. Cheryl herself, and the concomitant subtitle appears for five seconds. She talks fast and the second sentence follows immediately after the first. What is more, the second sentence is accompanied by a new image, which displays the hills and trees she was talking about. The viewer needs a split second to absorb the new image and to start reading again. So, in order to take into account the time the viewer will need to read the two sentences and to synchronise with the speech and image, the first words of the second sentence are left out.

(32) Today, when you're looking out across the landscape, you see here hills and trees. //

*Nu zie je hier alleen glooiende heuvels en bomen. //

So, obviously, there's not a big hole left in the ground from the eruption of the Yellowstone volcano.

*Er is geen gat van de uitbarsting van de Yellowstone-vulkaan.*

These three examples show that, sometimes, the subtitler has to condense the source sentence to meet up with the reading speed and synchronicity, so that the viewer is able to read and understand the subtitles properly. Nevertheless, this does not mean the viewer misses crucial information. The motivation to delete particular parts is discussed in the following section.

Finally, the last technique is the *simplification of the syntax and vocabulary*. As stated in the literature review, using simple vocabulary is only done when it does not affect the style and spirit of the genre (Ivarsson & Carroll 1998). No data were found in which geographical or scientific difficult terms were simplified in the subtitles. What is more, the complicated terms used by the speaker are explained in the documentary itself. Examples include:

He proposed that Hawaii was created by something called a 'hotspot', an exceptionally hot region beneath the earth's crust, that was concentrated under Hawaii's Big Island.
Concerning the syntax, the most frequent distinctions in sentence structure were found with the structures *volgens X* and *uit/na Y blijkt + statement*. Even though they may cause a change in arguments, they are not considered as simplifying the syntax, given that the original syntax is not difficult at all.

### 4.2.3 Functional motivations of the subtitler

As mentioned in the literature review, when the original sentence has to be condensed the subtitler has to decide what is the essential content. This decision is affected by the aim of the genre, which consists of providing a vital explanation to people outside of the scientific community. Consequently, the essential information in a documentary is the statement which contains the information and/or explanation needed to understand the message. Therefore, when condensation is required, the need to express the fundamental information excludes the accessory matters, which are in the 'deletion' examples, the sources of information.

Following Martin and White’s forms of engagement (denial, countering, concurrence, pronouncement, endorsement, entertain, acknowledgment and distancing) examples (33) to (35) represent acknowledgement, endorsement and concurrence respectively. The speaker of example (33) is presenting the statement (cf. underlined part) as a neutral observer. Nevertheless, by saying *people agree* he acknowledges that several persons agreed upon the statement, which increases the degree of certainty. Example (34) represents endorsement, which means the speaker sees the statement as correct, or undeniable. The example of concurrence (cf. example 35) expresses, via the certainty adverb, an agreement with the addressee. Thus, when looking at examples (33) to (35), it is noted that the evidentiality expressions imply a high degree of certainty. This certainty is also expressed by the bare assertion in the translation. So, from an epistemic point of view, the English expressions and their translations express the same meaning. Both languages show that the statements are facts. The evidence for the fact (*a map*) or who reported them (*people*) becomes accessory, since the epistemic meaning is considered as more essential than the evidential one.

(33) People agree that only a massive force could have created such a huge crater.

*Zo’n krater kon alleen door een enorme kracht zijn veroorzaakt.*
The map revealed that the Mariana trench is just a tiny fraction of a network of enormous underwater canyons stretching right around the planet.

De Marianentrog bleek te behoren tot een netwerk van enorme zeekloven dat de hele planeet omvatte.

Today, you see here hills and trees. So, obviously, there's not a big hole left in the ground from the eruption of the Yellowstone volcano.

Nu zie je hier alleen glooiende heuvels en bomen. Er is geen gat van de uitbarsting van de Yellowstone-vulkaan.

In addition, the examples above illustrate that not only the epistemic meaning remains the same, but also the motivations to use a particular structure are the same for the subtitler as for the English speaker. If possible, the subtitler aims to maintain the functional motivation that moved the speaker to use an evidential marker in the source language. The endorsement examples (36) and (37) show how the literal translation expresses the same certainty and validity added by the speaker as the non-literal translation where the fact that there is clear evidence is implied in the verb blijken (to become clear). In example (38) as well, in spite of the different structure, the specific meaning of schijnen (according to what they say) expresses the neutral reporting meant by the English speaker.

Contemporary reports show that most of Krakatau victims were not killed by these deadly avalanches.

Recent onderzoek wijst uit dat de meeste slachtoffers niet door gloedwolken zijn omgekomen.

This research shows that although the air in the desert is around 10% humidity, on rare occasions it rises as high as 75%.

Het bleek dat de luchtvochtigheid die hier gemiddeld 10% is bij zeldzame gelegenheden kon oplopen tot 75%.

People say that more than half a million people died from the drought in India.

Daarbij schijnen meer dan een half miljoen mensen te zijn omgekomen.

The most significant example, which illustrates that the subtitler aims at expressing the same meaning as the source sentence, is the unique example of denial.
(39) A tsunami isn't over in just a few seconds.

Men denkt dat een tsunami zo voorbij is.

As seen in section 4.1.3, the English denial is used to anticipate beliefs by others and to refute them. This particular meaning is expressed explicitly in Dutch by the use of the evidential markers men denkt ('people think'). So, the case of denial proves that, albeit the different structure, the subtitler and the speaker share the same functional motivation, which is in this case denying the general beliefs.
5 Discussion

The present chapter discusses the findings of the corpus analysis. In particular, it zooms in on three aspects treated in the analysis:

1. The evidentiality types
2. The formal encoding of evidentiality
3. The functional motivations for evidentiality

All three topics examined the English expressions and their accompanying subtitles. The results of these two studies will be combined and summarised in the sections that follow.

When interpreting the results, it is important to bear in mind that this is a small-scale corpus, limited to the specific language of a science documentary. This popularised genre of science and its particular aim has a considerable influence on the results.

5.1 Types of evidentiality in popularised science

In the first study four different types to express English evidentiality were distinguished. These distinctions were based on the findings presented in the literature review and consist of the indication of direct evidence, reported evidence, and inferential evidence and the expression of (un)certainty. The same types were examined when dealing with the Dutch subtitles. This comparative approach helped to gain more insight into the features of the four evidentiality types found in the popularised genre of science, i.e. the television documentary.

The results concerning the direct evidence type suggest that, albeit the assumptions in the literature review, in this particular corpus the pure perception verbs are not the main indicators of direct evidence. The majority of the data indicates direct evidence with the causative perception verbs to show and to reveal. Nonetheless, these verbs are not exclusively used for the direct evidence type. The subtitle analysis revealed that the two verbs may also indicate inference, which shows the evidential multifunctionality of the English verbs.

In a similar way, with regard to the reported evidence, the occurrences of to believe were systematically treated as reported evidence, but the translation warned that its meaning should be analysed separately and based on the context. The translation study also helped distinguishing between hearsay and quotative evidentials. The subjects of the
reported evidence, i.e. generic nouns like scientists, geologists etc., are considered hearsay subjects. This decision is backed up by the translation of these generic subjects, which is in most cases the general Dutch term men. The frequent occurrence of generic hearsay subjects resulted in a high incidence of hearsay expressions, which is due to the specific aim of the popularised scientific discourse. For laypeople, it is more important to know that the claims are based on actual research than to hear the names of the (for them unknown) scholars behind it. The slightly lower number of quotative expressions may be explained by the fact that the narrator only names the speaker who will feature in the documentary. Thirdly, the low number of quotation marks is due to the specific multimodal genre. Because the speaker is shown on the screen and his/her name is displayed, quotation marks are not needed. Also, in a scientific narrative it is easier to use indirect speech.

Another consequence of the genre-specific study is the adjustment of the inferential evidence type. The inclusion of sentences like example (1) as inferentials was justified by tailoring the analytical toolkit to the genre under investigation, which is centred around searching for and providing evidence.

(1) The crystals are clear evidence that under the springs is a volcano.

Firstly, this reconsideration has as result that the subjects of the inferential evidence expressions are all nonhuman. However, personified subjects do also occur. The investigation proves that the examples of personification have to be classified according to the level of humanity (reported evidence) or investigation (inferential evidence) implied in their source. Second, the results show that the verbs on which the recognition of evidential expressions in the science documentary was based (cf. Figure 2), did not appear in this genre. Ifantidou's verbs were derived from spontaneous conversations, where the direct speaker-hearer interaction is central, but this interaction does not take place between viewer and narrator. The different communicative needs of the two genres result in different evidential verbs and prove that a genre-specific approach to evidentiality is needed.

Even though Ifantidou's verbs are not represented, her inclusion of the expression of (un)certainty as a part of evidentiality, is. Again, the (structure of the) documentary itself is responsible for the occurrence of the (un)certainty expressions. As explained in the data collection (section 3.1), the documentary starts from questions and doubts and ends
with evidence and conclusions, which creates expressions of both certainty and uncertainty. The findings reinforced the statement that in non-grammatical evidentiality languages the expression of evidentiality frequently coincides with epistemic expressions. Moreover, in the translation as well, the addition of an epistemic marker beside the original English evidentiality type was noticed. Lastly, also the use of modal must as an evidential marker is proved by the data.

5.2 The formal encoding of evidentiality in popularised science

One observes in the results that 213 of the 255 English tokens mark evidentiality by means of the following formations: noun phrase (NP) + verb, proper noun (PN) + verb and pronoun (Pron) + verb. The high frequency (83.53%) of the structure subject + verb + that-clause contrasts with the variety of forms which exist to express evidentiality lexically and which were proposed in the literature review (modal verbs, adverbs, adjectives, nouns, conjunctions and lexical phrases). This observation is again explainable by the specific genre from which the data are taken. Most of the sentences are spoken by the narrator. His role is to narrate a prepared text accurately and clearly in order to elucidate things to laypeople. The simple structure above serves this goal perfectly. In addition, the subject + verb + that-clause structure is able to express all evidentiality types, which may also be an explanation for the high incidence.

In Dutch, by contrast, the simple English syntactic structure is often disturbed in the subtitles. The subtitles that use volgens or the non-raising use of blijken are translated according to the same pattern. First, the subject of the original main clause becomes an adverbial in the subtitle (volgens X, uit/na Y). Second, the subject of the English that-clause becomes the subject of the Dutch main clause and the evidential verb is deleted in favour of the verb of the that-clause. This may be seen as the subject that 'raises' from the subject in the English that-clause to the subject in the Dutch main clause. The term 'raising' is borrowed from the second structure of blijken, which is also marked by subject to subject raising.

All these small differences in formulation between the English and the Dutch sentences are due to the specific evidentiality nature of each language. To express reported evidence Dutch tends to use the preposition volgens and the raising verb schijnen. In Dutch, the expression of hearsay evidence is perfectly captured in just one verb: schijnen. Similarly, the specific verb blijken is used for inference, because its meaning 'to become clear' implies
that the speaker has evidence, which causes that something has 'become clear'. The fact that the structural distinctions are forced by the evidential preferences of Dutch, is proved by the finding that the English counterpart of volgens, according to, is never used in the complete corpus. What is more, the default translation of schijnen (seem) is also never translated with schijnen. Furthermore, albeit these specific words to express Dutch evidentiality, the word order and the essential message is maintained, as will be emphasised in the following section.

Next to restructuring the arguments, some subtitlers delete the arguments. The (partial) deletion of the evidential markers in the subtitles is justified by the subtitle restrictions, led by the reading speed and synchronicity.

5.3 The functional motivations for evidentiality in popularised science

The functional motivations proposed in the literature review (acknowledging responsibility, denying responsibility and assigning authority) were classified as insufficient when dealing with the actual data. A broad view on evidentiality needs a broad view of possible motivations to use the evidentiality forms. The engagement types of Martin and White (2005) (denial, countering, concurrence, pronouncement, endorsement, entertain, acknowledgement and distancing) were introduced because it was possible to represent them all by an evidentiality example of the corpus. Nevertheless, not all engagement types affect the scientific text equally. When starting from the specific genre, which main character is the narrator, and its accompanying aim, which is convincing and informing the unknowing audience, it is noted that the documentary makers must be moved by endorsement and acknowledgement. These motivations, which are expressed by the inferential verbs to show, to prove and reporting verbs to say, to believe, explain the high incidence of the reported evidence type and the inference type.

In spite of the change in type or in structure, the subtitler works according to the same communicative needs as the speaker. That is why, when the source sentence has to be condensed, which is frequently the case, the subtitler maintains the most essential information. In accordance with the aim, this essential information is the statement or the explanation of a certain scientific finding, which means that the evidential information becomes accessory. In this view, when dealing with the ten subtitles which delete the evidential markers completely, one sees that the evidential meaning is not crucial given
that the epistemic meaning is not changed. The maintenance of the essential message is led by the maintenance of the functional motivation. This goal is made perfectly clear by the exception, which expresses evidentiality in Dutch and not in English. The English corrective type of denial is used to refute beliefs the audience may have and this is made explicit in the translation by adding the evidentials *men denkt*.

Additionally, also the explanation of the vocabulary and the simple syntax in English as well as in Dutch reinforce the importance of the aim.
6 Conclusion

The goal of this dissertation was twofold. In the introduction, two questions were raised: Does evidentiality occur in English and if it does, does the delineation of the concept change according to the genre under investigation? The first question was answered in the literature review. By comparing the different opinions, definitions and taxonomies, it was decided that English evidentiality was possible but had to be approached in its broadest sense. This meant that the speaker's epistemic attitude, the evidential must and pragmatic inference were included in the concept. Secondly, evidentiality was defined in a functional light. The lexical means to express evidentiality in English were summarised in Figure 2. This overview was useful as a starting point to compose the corpus and to analyse the data, taken from the American science documentary How the Earth Was Made. The corpus analysis consisted of the analysis of the types, the formulations and the functional motivations. The investigation of these categories applied to the Dutch subtitles served as a meaningful completion. The results showed that the characteristics outlined in the literature review were not represented in the corpus. The main outcomes of the corpus analysis and consequently the main differences with the default definitions and features presented in the literature review, were explained by two factors: the genre and its aim.

First, the genre is marked by a particular structure on the one hand, and a specific main character on the other hand. Due to the structure of searching and providing evidence, the inferential evidence type had to be adjusted to the genre under investigation. This adjustment was responsible for the dehumanisation of the inferential subjects. On its turn, these nonhuman subjects are the reason for the exclusion of verbs with a first person singular. The particular scientific structure was also responsible for the emphasis on the (un)certainty evidentiality expressions, which proved that these often combine with the expression of information source. The high incidence of reported and inferential evidentiality expressions on the other hand, was due to the role of the narrator, who was moved by the features of endorsement and acknowledgement.

Second, the aim was defined by Meyers (2003: 273) as "convey[ing] scientific knowledge to a wider audience." This definition implies that the audience consists of laypeople who want to learn more about a specific scientific subject. The aim explains, among others, the almost equal division between referring to the specific source (quotative) and not specifying the source (hearsay), which implies the translation by men, and the frequent use of the simple structure: subject+ verb+ that-clause.
Additionally, the corpus analysis revealed some other important findings. For instance, some differences between the source and target language could not be explained by the genre or the special features of evidentials. As a result, it was discovered that small differences in formulation (*blijken, schijnen, volgens*) were due to the specific devices used to express evidentiality in Dutch. The drastic differences (deletion of evidential meaning) were explained by the subtitle time restrictions. Finally, the study also gained an insight into the work of the subtitler, who is also affected by the aim and tries to convey the message as accurately as possible.

In conclusion, evidentiality is possible in English, but it is advisable to adopt a broad approach and to tailor the analytical toolkit according to the genre under investigation. This is confirmed by the effect the aim and the genre had on the outcomes of the studies. Consequently, the evidential types, formulations and the functional motivations will be different when investigating a different genre. For instance, in spontaneous conversations the speakers will be more motivated by the 'concurrence' or 'entertain' category. Finally, the analysis also provided additional information about evidentiality in Dutch, which can be expressed by *schijnen, volgens* and *blijken*. These findings may serve as a starting point for further investigation about evidentiality in Dutch.
7 References


