

Evidentiality across age and gender: a corpus-based study of variation in spoken British English

Erika Berglind Söderqvist
Uppsala University / Sweden

Abstract – This paper investigates the effects of age and gender on the frequency of the evidentiality categories ‘sensory’, ‘hearsay’ and ‘inferential’ in spontaneous spoken British English. The findings from the main data sample from the British National Corpus (BNC) are also compared to patterns in a smaller data set from the Diachronic Corpus of Present-day Spoken English (DCPSE) in order to estimate the relative effects of age-grading versus historical change. The results confirm statistically significant differences between men and women in their use of evidentiality, but show no significant effect of age or the interaction of age and gender. The comparison of the findings from the BNC data and the DCPSE data suggests that age-related patterns in evidentiality use are more diachronically stable than gender-related patterns.

Keywords – evidentiality, corpus linguistics, pragmatics, sociolinguistics

1. INTRODUCTION

The present study takes a rare approach to evidentiality in English in using corpus linguistics methods to investigate the sociolinguistic patterning of a large number of evidentiality markers in casual speech. At a glance, evidentiality – the linguistic expression of information source – may seem a relatively straightforward concept. Put simply, it communicates to the addressee what kind of evidence the speaker or writer has for the truth of a given proposition.

- (1) Of course *I heard* a car. And then I thought oh that’s Clayton coming in as well and then I thought no it isn’t, he’s going out and *I heard* him go out. (KBE 6342–6343)
- (2) Oh Wayne *told me* it was twenty five pence off a pint. (KCA 1419)
- (3) But you *must have* met him, he’s been up here a couple of times. (KC7 536)

Examples (1) through (3) (from the British National Corpus; BNC) illustrate what are typically considered three categories of evidentiality: sensory, hearsay and inferential (cf. e.g. Willett 1988). The speaker in (1) refers to her own sense of hearing as the source of the information that Clayton left the house. The speaker of (2) tells the addressee not only that there was a discount on pints of something, but refers to someone else’s report as her information source. Finally, the speaker of (3) says that she believes the addressee has met ‘him’, the word *must* communicating that she does not consider this an absolute truth, but rather a logical conclusion based on the evidence that is available to her. She even specifies this evidence: the fact that ‘he’s been up here a couple of times’.

Communicating information source is not, however, the only effect of the use of evidentiality – and this is where it becomes clear that evidentiality is rather complex. Choosing to express evidentiality in a certain manner relays information regarding the speaker’s confidence in the information source; it is likely that different types of

evidentiality lead to different degrees of reliability. For example, the fact that the speaker in (1) specifies that she bases her information on sensory experiences, rather than allow the addressee to assume that she is only guessing, strengthens her claim. The reliability of (2) is dependent on the perceived reliability of Wayne, but the wording suggests that the speaker refuses to take full responsibility for the information. In (3), the evidential expression implies that the speaker considers the information highly reliable, yet conveys that this is not perceived as an absolute truth but only a conclusion drawn by the speaker.

Traditionally, evidentiality has been researched as a ‘closed grammatical system whose use is obligatory’ (Aikhenvald 2007: 209) in languages where it is generally morphologically marked. However, the conceptual category of evidentiality exists in languages where this definition does not hold as well, such as English, where it is non-obligatory and generally marked by lexical or pragmatic means. Since evidentiality is optional in English, using evidentiality may be considered a pragmatic choice; for example, Alonso-Almeida and González-Cruz (2012: 340) suggest that frequent use of evidentiality in English indicates a need to show authority.

The fact that expressing evidentiality is optional, and that using it conveys information on how the speaker positions her/himself in relation to her/his utterance and to the addressee(s), suggests that extra-linguistic variables, such as factors regarding the speaker’s identity, background and/or the type of speech situation might affect the way evidentiality is used in English. Targeting such extra-linguistic variables in order to investigate the nature of the interactional functions and effects of evidentiality would thus add to the existing body of knowledge not only of evidentiality, but also of the reasons for certain pragmatic choices in interaction. Previous studies have found some gender variation in evidentiality use (Alonso-Almeida and González-Cruz 2012; Berglind Söderqvist 2017), but since it is generally argued (e.g. Canary and Hause 1993; Eckert and McConnell-Ginet 2013) that gender should not be viewed as “acting” in isolation, it is likely that any gender variation found in interaction is caused not only by gender, but by the intersection of gender with other factors.

Since it has previously been found (Paunonen 1994; Barbieri 2007) that gender impacts the kind of age-related changes to a person’s language, the present study will investigate whether age is a factor in any gender variation found in evidentiality use. Three different age groups will be targeted: 20–29, 40–49 and 70+. Any gender variation found will be compared across the age groups to test whether age is a factor in determining the degree and nature of gender-based variation in this linguistic feature. The data will be collected from the spoken-language section of the BNC. In order to retain some control of whether any differences found between the age groups are likely to be due to age-grading or to historical change, the patterns found in the main data set – recorded in the early 1990s – will be compared to patterns found in a smaller control group from the Diachronic Corpus of Present-day Spoken English (DCPSE), recorded in the 1960s and 1970s. Even though the BNC material is around 25 years old, the results of this study will not only provide a synchronic snapshot of the sociolinguistic patterning of evidentiality, but also an approximation to the diachronic stability of this patterning in spoken Present-day British English.

2. BACKGROUND

2.1. Evidentiality

Concerning the scope of evidentiality, the present study has been informed by Carretero and Zamorano-Mansilla (2015: 147), who claim that since evidentiality is not fully grammaticalized in English, “the evidential interpretation often emerges from the pragmatic interpretation of a wide range of expressions”. In other words, the operationalization of evidentiality in this paper relies on function rather than form; while certain lexical forms are often used to express evidentiality, very few lexical items in the English language can be said to *always* carry evidential meaning. Examples (4) through (6) demonstrate the necessity of studying the context of a word in order to determine whether or not it marks evidentiality.

- (4) (a) It’s not at home as far as I can *see*. (KD5 6201)
 (b) Did you *see* that Stars in Their Eyes last night? (KD8 3440)
- (5) (a) He *said* it’s bec-- all because the tipping’s been privatized or something. (KD8 3304)
 (b) It can hear everything you *say* or do. (KBA 212)
- (6) (a) I, I wasn’t watching while [unclear] at the time, but erm [unclear] basically it *seemed* like it was six of one and half a dozen of the other. (KDA 7488)
 (b) *Seemed* a bit daft to me though [pause] you’d think they, mind you they’d, they’d supposed to of had erm [unclear] er sub contractors down there [...]. (KDA 6303)

Examples (4a), (5a) and (6a) show *see*, *say* and *seem* used as evidential markers; they modify the utterances by adding information regarding the kind of evidence that the speakers have for their claims. Their counterparts (4b), (5b) and (6b) show the very same lexical forms, but in contexts where they do not carry evidential

meaning. Examples (4b) and (5b) are relatively straightforward; the speaker/writer is not the agent of the *seeing* in (4b) nor the recipient of any report in (5b) (cf. Section 4.2.1). Example (6b), on the other hand, represents a case of *opinion*. It is integral to the definition of evidentiality that it concerns evidence for something that the speaker perceives as true, or likely to be true, and that only verifiable statements can be said to contain evidentiality (Carretero and Zamorano-Mansilla 2014). This rule of verifiability necessitates the exclusion of statements such as the one in (6b). The present study's operationalization of Carretero and Zamorano-Mansilla's verifiability criterion entails that a state of affairs can be said to be modified by an evidential expression if it refers to an event or activity that is either true or not true. In (4a), *it* is either at home or it is not; in (5a), the reason is that the tipping has been privatized or it is not; and in (6a), it is either true that there were six of one and half a dozen of the other, or it is not. In the case of (6b), however, it would not be as straightforward to determine whether or not something is *a bit daft*, as whether or not something is daft at all is a question of opinion.

Cornillie (2009) argues that it is essential to the degree of reliability of an evidential expression whether it refers to shared or non-shared information, and it has often been suggested in the literature that sensory, hearsay and inferential evidentiality communicate different levels of certainty; information is likely perceived as more reliable coming from a person who witnessed the proposed state of affairs first-hand (expressing sensory evidentiality) than from a person who cites logical likelihood (i.e. inferential evidentiality) as source of information. These implications from the theoretical discussions of evidentiality are echoed in empirical investigations, which have found evidence of variation pertaining to types of evidentiality. For example, Alonso-Almeida and González-Cruz (2012) find that women use inferential expressions more often than men, and Berglind Söderqvist's (2017) results suggest that men are more likely to use intersubjective evidential expressions – citing shared information – whereas women are more likely to use subjective evidential expressions, citing non-shared information. In line with such findings, the present study will focus on patterns regarding *category* of evidentiality (sensory, hearsay or inferential) rather than patterns regarding the individual lexical items targeted in order to find instances of evidentiality.

The nature of the preference under investigation should also be pointed out here. Assuming that people speak the truth most of the time, the choice is not between using sensory, hearsay or inferential evidentiality. Any person who decides to communicate a proposition will have a given basis of evidence to support that statement; if the person decides to express evidentiality, the nature of that existing basis of evidence will determine which type of evidentiality is used. In other words, the choice in question lies in whether or not to express evidentiality at all in any given situation. Hence, the present study will investigate gender- and age-related patterns in speakers' tendencies to use evidentiality more or less often as an effect of what kind of evidence is available to them in a particular situation.

2.2. Gender

The research on language and sex or gender has moved from assuming that the societal inequities relating to gender must be reflected by a men's style and a women's style of using language (Lakoff 1973) to adopting a more nuanced view. Eckert and McConnell-Ginet (2013: 47) liken gender to "a set of constraints that one embraces or simply accepts, that one struggles within, or that one struggles against. But these constraints are not set for all time, and it is people's day-to-day actions that make them change". In modern language and gender research, gender is generally seen as a fluid concept that has more to do with learned practices and internalized attitudes than with chromosomes and reproductive organs.

Although it is becoming increasingly recognized that there is reason to problematize the traditional view of gender as a binary, categorical concept (e.g. MacDonald 1998; Ansara and Hegarty 2014), the decision was made, for the purpose of this study, to operationalize gender as a variable with the variants *male* and *female*. The reasoning behind this is twofold: first, the present study is a quantitative one focusing on informal spoken language, and to my knowledge, no large spoken corpora exist with information about speakers' age and gender that include more than two gender categories. Second, the norms telling us what a woman should be and what a man should be are still something that we have no choice but to relate to. Returning to Eckert and McConnell-Ginet's "set of constraints", it is likely that the constraints that a person who is labeled *female* embraces or struggles against are not entirely the same constraints that a person labeled *male* embraces or struggles against. For this reason, it is still relevant to study behavioral patterns based on binary gender categories.

A number of studies that looked at gender differences in various linguistic strategies for expressing evaluation and speaker attitude (which can be said to be effects of evidentiality) have found that there is some variation. For example, in a study of spoken British and New Zealand English data (Holmes 1990), it has been found that men hedge their utterances more often than women do. Moreover, Grob, Meyers and Schuh (1997) have shown in a study of 64 American university students performing a discussion task that men also use more disclaimers, which include phrases like *I guess* and *I suppose*. Mulac et al. (1988) analyzed the speech of 96 American university students performing discussion tasks in pairs and found that women justify their statements more often

than men do. As hedging, disclaiming and justifying are possible “side effects” of the markers that are used to express evidentiality in English, these findings support the relevance of the present study. Alonso-Almeida and González-Cruz (2012) showed that women use more evidentiality than men; Precht (2008), in an investigation of evidentiality as a subcategory of stance, found small but statistically significant differences between men’s and women’s use of evidentiality, and Berglund Söderqvist (2017) proved that men and women have different preferences when it comes to inferential evidentiality markers.

The fact that most of the studies cited above do not focus exclusively or explicitly on evidentiality indicates that there is a gap to be filled by the present study in the body of gender and language research. However, all of the studies referenced in the previous paragraph also found similarities between the genders, and the authors generally recommend that future research take into consideration additional extra-linguistic variables in order to tease out why and in what way gender affects language use. This recommendation is in agreement with Canary and Hause’s (1993: 141) advice that gender should be researched as *happening* in relation to other variables, or as Eckert and McConnell-Ginet put it: “Gender must be recognized in its full glory – in its inseparability from the rest of life experience” (2013: 61). For the purpose of the present paper, it was decided to investigate whether the variable *age* has an effect on any gender variation found in evidentiality use.

2.3. Age

While the language of children and adolescents has been extensively researched as an acquisition process, and the language of the elderly has often been researched in terms of how various pathological states cause language loss, the language of adults in between adolescence and old age has simply been considered a stable norm (Eckert 1997; Murphy 2010) and as such somewhat neglected. This is a problem because, as Eckert (1997: 152) notes, aging “involves changes in family status, gender relations, employment status, social networks, place of residence, community participation, institutional participation, engagement in the marketplace – all of which have implications for patterns of variation”. The present study will focus specifically on adulthood, testing whether there is evidence of age-grading in the degree of gender variation in evidentiality use targeting speakers at various stages in their lives. Following Murphy (2010), age groups have been defined based on the life stages they are likely to correspond to (see further Section 3.1). The reason why age has been selected rather than some other variable that might conceivably be a factor in gender variation (such as formality level, socioeconomic class of interlocutors or single-/mixed-gender interaction) is that it has previously been found that people change the way they speak in various ways throughout their lives. The findings of this study will add to the existing knowledge of in what ways language changes through adulthood.

Previous studies that have combined the variables age and gender have tended to find that the variables do interact to influence language use. For example, Paunonen (1994) showed that gender as well as age affect language change on the community level as well as individual level. Barbieri (2007) found that men and women change their use of quotatives in differing directions as they age, which could be relevant to evidentiality, considering the hearsay type which often involves quotatives. Among studies that have looked solely at the variable of age, it has been shown that speakers of American English change their use of stance markers qualitatively as they age (Barbieri 2008), and that there is age-related variation at the levels of discourse, grammar and lexis in casual conversation among women (Murphy 2010). These studies are relevant to the present one in that evidentiality can be argued to overlap with stance in certain respects, and in that some of the linguistic features investigated by Murphy (2010), such as hedges and boosters, include some evidential expressions. Hence, it is likely that the results of the present investigation will show that age has an effect on evidentiality use.

As Eckert (1997: 151) points out, investigations on age-related language variation “can reflect change in the speech of the community as it moves through time (*historical change*), and change in the speech of the individual as he or she moves through life (*age grading*)”. It can be difficult to know which of these effects are being demonstrated by any given study, and as shown by Paunonen (1994), it is likely that historical change and age-grading interact to form the way the language of a person, or a group of people, changes over time. The present study will compare the main data set of synchronically-captured material from speakers of different age groups to data from similarly sampled age groups recorded a couple of decades earlier. The purpose of doing this is to try to obtain a hint of the nature and/or extent of the effects of historical language change, on the one hand, and age-grading, on the other.

3. MATERIAL

3.1. The BNC data

The main data set that was used for this study was extracted from the BNC. The BNC was accessed via the BNCweb interface, which is an online concordancing tool that also allows the user to define and save subcorpora based on all of the categories that are included in the metadata of the BNC.

First, it was decided to control carefully for the genre of language use, to make sure that any variation found is not due to some hidden skewed genre distribution. The present study focuses on spoken language because of its interactive nature; the features of evidentiality that are interesting to the present investigation are specifically relevant to interaction. Further, following Murphy (2010: 37), it was decided to target the life stages *young adults*, *middle-aged professional adults* and *retired elderly adults* in order to cover as much as possible of the adult life span and yet get a picture of changes that happen in between the stages. Like Murphy (2010), the present study seeks to operationalize these life stages by collecting data from speakers of the following age groups: 20s, 40s and 70s/80s.

The data sample was restricted to informal conversation, because this genre constraint generated the most even distribution across genders and ages. In the demographically sampled part of the BNC, 124 volunteers, selected to represent an even distribution of ages, genders, accents and social classes, were equipped with tape recorders and instructed to record all of their conversations over a range of days.¹ This portion of the BNC consists mainly of casual conversation in private settings and was recorded in the early 1990s. Via BNCweb, the demographically sampled part is only searchable by metatextual information regarding the volunteers holding the recording equipment, even though their recordings tend to include a large number of additional interlocutors, many of them with metatextual information of their own. Hence, it was quite a challenge to define subcorpora that would include individual interlocutors as well.

BNC64 (Brezina 2013), which is a corpus of 32 male speakers and 32 female speakers with exact age information from the demographically-sampled part of the BNC, was used as a starting point for finding speakers that had enough metatextual information to be useful for my purposes. The remaining speakers from the demographic section fitting my age/gender groups were then manually harvested by looking through the information of all of the relevant text files. The category of male speakers in their 70s/80s comprised the smallest group, of only 17 speakers. One of the speakers, who had a word count of four, was removed, leaving 16 speakers. In the other five groups, speaker IDs were removed one by one, based on their word count, with the goal of having six subcorpora with as-equal-as-possible word counts. Table 1 shows the results of this endeavor, with word counts in parentheses.

	20s	40s	70/80s	Total
Women	16 (139,533)	15 (160,504)	16 (136,494)	47 (436,531)
Men	14 (125,689)	12 (168,343)	16 (124,599)	42 (418,631)
Total	30 (265,222)	27 (328,847)	32 (261,093)	89 (855,162)

Table 1: Speakers from the BNC

Thus, the main data set used for the present investigation consists of 89 speakers and 855,162 words. The word counts are not exactly equal, but these differences were compensated for during the statistical analysis.

3.2. The DCPSE data

In order to retain some control over whether the age-related variation can be argued to be mainly due to age-grading or historical change, a smaller sample of data was collected from the Diachronic Corpus of Spoken Present-day English (DCPSE). The DCPSE is an 800,000-word corpus of spoken British English in which half of the material is taken from the London-Lund Corpus (LLC), recorded mainly in the 1960s and 1970s, and the other half is sampled from the spoken section of the International Corpus of English: Great Britain (ICE-GB), recorded in the early 1990s (Nelson, Wallis and Aarts 2006).

The DCPSE was accessed through the concordancing program ICECUP, which was specifically written for ICE-GB and the DCPSE. As with BNCweb, it is possible to customize subcorpora in ICECUP. It was decided that only two age groups would be sampled from the DCPSE, as it was not possible to find enough speakers in their 70s/80s, drawn from the LLC portion and from a genre that would make it comparable to the BNC data.

The DCPSE and the ICECUP are differently organized from the BNC and BNCweb, and so the collecting process for this data sample had to be differently approached. A search was carried out for files from the LLC portion of the DCPSE, marked as face-to-face interaction (as opposed to telephone conversations) and containing

¹ <http://www.natcorp.ox.ac.uk/corpus/creating.xml>

speakers of the age spans 18–25 and 46–65. Nine female speakers and nine male speakers were selected from each of the two age groups, including text files that were labeled informal (to correspond to the casual conversations of the BNC material), with the goal of arriving at approximately equal word counts across all groups. The result of the DCPSE data collection is outlined in Table 2.

	20s	50s	Total
Women	9 (15,353)	9 (15,654)	18 (31,007)
Men	9 (15,413)	9 (17,727)	18 (33,140)
Total	18 (30,766)	18 (33,381)	36 (64,147)

Table 2: Speakers from the DCPSE

The data set of 36 speakers and 64,147 words shown in Table 2 was recorded between 1969 and 1977. The age group called *20s* is what it sounds like, but the other age group consists of speakers aged 48–65. This somewhat liberal approach to the time and age spans had to be accepted in order to accumulate enough speakers from the relevant genre in this rather small corpus.

4. METHOD

4.1 Searching for evidentiality

While the view held in this paper is that evidentiality in English is primarily a functional category, it can still be argued that evidentiality is often communicated by means of certain lexical items with semantic scopes that overlap with the functional scope of evidentiality. By searching for a pre-defined list of lexical items that potentially communicate evidentiality, it should thus be possible to find many of the occurrences. This study makes no claim to total recall; still, in order to harvest as many occurrences of evidentiality as possible, it was decided to adopt a combined bottom-up and top-down approach. First, a thorough review of the literature supplied a number of lexical items that have been argued to carry evidential meaning (e.g. de Haan 2001a; Bednarek 2006; Whitt 2009; Berglind Söderqvist 2017). Second, a manual analysis of a subset of the BNC material was carried out as described below.

The goal was to use a representative, randomized subset of the BNC material and scan it manually for any and all occurrences of evidentiality in order to find forms that may not have been brought up previously in the literature. In order to achieve this, I carried out a simple search for the definite article *the* in each of the subcorpora, which generated large numbers of hits from all of the speakers, where each token of *the* was found within its original s-unit (approximately equivalent to a sentence). The order of the hits was randomized, and the first 100 s-units from each subcorpus were saved, yielding 600 s-units. These 600 were then analyzed manually for expressions of evidentiality, and the “core” word or phrase of the evidential expression (i.e. what I would search for in order to find other similar cases) was noted down. The original idea was to add more s-units to this subset until I could find no more new tokens, but this measure did not prove necessary. Aside from markers that were already represented in the literature on evidentiality, the manual scanning process yielded mainly phrasal verbs such as *go on about* and *make (something) out*.

The literature review and the manual analysis yielded a list of 26 verbs, adverbs, phrases and modals, for which I conducted lemma searches in the BNC material. In order to enable statistical testing of variation across individual markers,² only markers that occur more than 100 times per million words were included in the study, since it was thought that any items below this threshold would be rendered too infrequent after the non-evidential occurrences were weeded out. The 18 lexical items that were left after this procedure are listed in the Appendix of the present paper, along with the eight markers that were discarded. It should be noted that the fact that only spontaneous spoken data was used is likely to have affected which markers were found during the manual scan as well as the frequencies of all the markers.

4.2. Coding for evidentiality

The particular challenge of conducting a corpus study of a functional category is the risk of allowing the coding procedure to become too subjective for the investigation to be replicable. In order to reduce this risk to the greatest extent possible, much time and effort was put into defining a set of criteria that would enable a consistent coding process.

² Unfortunately, it was found that many of the lexemes still so infrequently mark evidentiality that the individual markers could not be statistically tested for gender/age interaction.

The coding was conducted in two steps: first, all tokens that were so obviously non-evidential that there was no need to look at any co-text beyond the s-unit of the token were excluded. Second, the less straightforward tokens were analyzed and coded, often by consulting context beyond the s-unit in which the token was found. This procedure allowed the formulation of categories based on which factors promoted or blocked evidential reading of the markers. In the following sections, all of the subsets of criteria and constraints will be dealt with under the title of the group to which they were applied.

4.2.1. General criteria

The criteria and constraints described in the present section were adhered to consistently throughout the coding regardless of the marker. First, a corpus example was coded as evidentiality when *the targeted item* was used to add information to the proposition regarding the kind or source of evidence for or against the truth of the proposition (cf. Carretero and Zamorano-Mansilla 2014: 319). Second, only examples in which the speaker was the deictic center, i.e. the experiencer of sensory perception, the recipient of report or the agent of inferential processes, were coded as evidentiality. The following cases were always coded as non-evidential:

- Opinion; non-verifiable statements (cf. Section 2.1); typical statements of this category are those containing evaluative/emotive adjectives (Biber et al 1999: 509), such as *bad*, *nice* or *best*.
- False starts, repetitions (i.e., if a speaker said *I think I think the dog was there*, only the first *think* was included).
- Unclear cases; when the transcription had so many [*unclear*] tags, overlaps or the like that it was impossible to be certain that it was evidentiality.
- Cases where the evidential expression was part of reported speech, i.e. when the speaker reported someone else's evidential expression or their own from a different situation.

Further, all corpus examples that were coded as evidentiality were also coded for *sensory*, *hearsay*, *inferential* or *unspecified* evidentiality. Items were coded as *sensory* or *hearsay* evidentiality when the marker itself referred to the sensory perception or the act or reception of the (verbal or written) report that constituted the evidence. Cases when the marker referred to a logical cognitive process, or only indirectly to the evidence, were marked as *inferential*. *Unspecified* was used for examples that were clearly evidentials when it was not possible to determine the category of evidentiality.

4.2.2. Adverbs

The adverbs in the present study include *obviously* and *apparently*. In an exploration of the epistemic and evidential properties of some English adverbials, Carretero and Zamorano-Mansilla (2014: 351) found that *obviously* and *apparently* are “suitable for automatic annotation”, i.e. can be said always to carry evidential meaning. Hence, they were always coded as inferential evidentiality (since they do not explicitly refer to a report or to sensory perception) in the present study, except for the non-evidential contexts described in Section 4.2.1.

4.2.3. Cognitive verbs

The cognitive verbs include *believe*, *know*, *suppose*, *think* and *remember*. This is a particularly complicated group, due to their wide range of uses, and my treatment of it has benefitted greatly from the work of Cappelli (2007). Her comprehensive exploration of the epistemic and evidential uses of cognitive verbs in English resulted, among other things, in detailed descriptions of the conditions under which such words carry evidential meaning. Cappelli argues that a *qualificational, non-descriptive* construal is a condition for these verbs to carry evidential meaning, meaning that, for example, *know* serves as a qualificational measure as regards the proposition it refers to, rather than to describe the state of being aware of something (which is a different, non-evidential function of *know*). She further notes that the prototypical case of a qualificational non-descriptive reading occurs in the first person singular and in the simple present tense (Cappelli 2007: 140), and lists the following factors, which tend to trigger either a non-qualificational or a descriptive reading and thus block evidentiality:

- *-ing* form.
- Imperative mood.
- Co-occurrence with *of*, *about*, *in*.
- Reduced occurrences (*Yes, I think/believe so/know*) and *I know, but...*

The existence of these constraints were seen as preventing evidentiality, except in cases where *I think so/know* was directly followed by a clause explicating the speaker's reason for thinking or knowing so; i.e., the source of evidence.

Of the cognitive verbs, *think*, *believe* and *suppose* behave similarly. These are always inferential, and their negated forms do not act as constraints against evidentiality. In the past tense, however, they tend to express that

the speaker believed something which, at the time of the utterance, has turned out not to be true, which means the construction is non-evidential. It should be pointed out here that I have *not* followed Cappelli in the case of *think*. She argues that *think* is not an evidential verb, but states that it lexicalizes the conceptual dimension “computational process over available evidence” (Cappelli 2007: 185). The description of this dimension can be argued to be a paraphrase of Carretero and Zamorano-Mansilla’s (2014: 320) description of the “evidential semantic feature that points to the important role of the evidence in the assessment of the reliability of information”. Since the working definition of evidentiality of the present study is modeled on Carretero and Zamorano-Mansilla’s definition (cf. Section 4.2.1), it was decided to adhere to their scope of evidentiality and include *think* in the study.

Know and *remember* constitute another functional sub-group of the cognitive verbs. Cappelli (2007: 156) argues that “by using *know*, the evaluator wishes to point out that he/she has indeed reliable evidence to say that *p* is a fact with a high degree of confidence on his/her part”, and shows that *know* often co-occurs with lexicalized references to the relevant evidence. In the present study *know*, as well as *remember*, which exhibits similar behavior, were treated as “pointers” to evidence, following Cappelli’s discussion, and were coded as the type of evidence they pointed to. In cases when they fulfilled the criteria for a qualificational non-descriptive reading but did not explicitly refer to evidence, they were coded as evidential but unspecified. For this subgroup, negations block evidentiality, as they tend to denote an absence of knowledge and evidence.

4.2.4. Speaker-subject verbs

Speaker-subject verbs include *hear*, *see*, *watch*, *find out* and constructions where *look* occurs with a first-person subject (*I looked at it*). As the title of this section suggests, the feature they have in common is that the speaker is almost always the explicit subject; the exception to the explicit speaker-subject rule of this category is shown in (7).

- (7) so they’re trying to corner it and it's munching on grass quite happily, you could *see* it was watching a minute [...] (KB9 90)

In (7), the speaker is describing an event which she witnessed to people who were not there, using a second person pronoun even though she is referring to her own observation. Such uses of *see* were coded as sensory evidentiality.

4.2.5. Non-speaker subject verbs

The category non-speaker subject verbs includes constructions of the type PRONOUN/NP + LOOK/SOUND/SEEM + ADJ/PP (*it seems like they are gone / the door looks open*). *Seem* is always inferential, and the other two are sensory or inferential. In particular, *look* and *sound* are sometimes used for figurative language, i.e. describing something by comparing it to something else, in which case it is closer to opinion than evidentiality. Examples (8) and (9) demonstrate a non-evidential and an evidential occurrence of this category, respectively.

- (8) She *looks* like somebody got her bloody in a trance. (KBE 5574)
 (9) It doesn't *look* cut off to me, it *looks* as if it's all there (KE2 1320)

When determining whether or not a speaker is making a statement regarding something they *believe to be true*, context is often crucial. The statement in (8) is best described as figurative language. From the context of this utterance, it is clear that the speaker does not actually believe that *she* has been put in a trance; rather, the speaker believes the person’s appearance is caused by medication but opts to describe it by likening it to being in a trance. In other words, (8) is not an example of evidentiality. In (9), the context reveals that the speaker is referring to the appearance of a person’s hair in a photograph, talking about what the visual evidence leads him to conclude regarding this person’s hair.

4.2.6. Speaker-recipient verbs

The category of speaker-recipient verbs, comprising *say* and *tell*, is the one where the speaker has the role of the recipient of the report. In this category, deontic contexts (*he told me to do it*) were excluded, as well as instances where the speaker was quoting a conversation rather than citing someone else’s report as source of evidence. In order to distinguish between quoting a conversation and using hearsay evidentiality, reported direct speech was categorized as non-evidential, while reported indirect speech was quoted as evidential, as in (10) and (11). This practice was decided upon because indirect speech is often used in the material to refer to pieces of information in order to strengthen or inform a claim, whereas direct speech tends to be used to retell longer stretches of conversation seemingly without any intention to modify a proposition with evidentiality.

(10) I know Pauline was *saying* she's had to buy some coal (KB2 1054)

(11) [...] she *said* oh I'd like them, but Sue *said* she went sort of a bit earlier and see how they got on and eh, she *said* oh it wasn't the same at five, they didn't respond and. (KCD 4977)

As shown in these examples, the difference between reported indirect speech as in (10) and reported direct speech as in (11) can often be seen from the use of first person pronouns and discourse markers such as *oh* in the latter kind.

4.2.7. Modals

The category modals is the smallest one, only including *must*. When found to mark evidentiality, it was always coded as inferential, and deontic contexts (*they mustn't do that*) were excluded. While it has also been argued, e.g. by Alonso-Almeida and Cruz-García (2011), that the modal *may* sometimes falls within the evidential scope, it was not included in the present study. As Alonso-Almeida and Cruz-García find in their material of academic writing, *may* is epistemic rather than evidential in the majority of cases. Possibly because of the nature of the data used for this study, *may* is less frequent than *must* overall and was not found to mark evidentiality during the manual scan of the material.

4.3. Statistical method

In order to test whether and to what extent gender, age and the interaction of the two have an effect on the use of evidentiality markers and categories in English, a two-way Multivariate Analysis of Variance (MANOVA) was carried out on the BNC data. The frequencies were normalized per 100,000 words for the purpose of the MANOVA, as the word counts of the individual speakers differed so much that raw frequencies would have rendered invalid results. The category *Unspecified* was left out of the category analysis; it would be difficult to discuss such results in any meaningful way, since the tokens in this category only have in common that it was impossible to place them in any other category.

As the DCPSE material is so small, no statistical method was applied to it. The frequencies were simply normalized per 100,000 words, so that the findings could be compared to those in the BNC material and any patterns could be detected.

5. RESULTS AND DISCUSSION

The null hypotheses that were examined by the two-way MANOVA include:

- 1) Speakers use the three evidentiality categories equally often regardless of gender.
- 2) Speakers use the three evidentiality categories equally often regardless of age.
- 3) There is no interaction effect for age and gender pertaining to the frequency by which the three evidentiality categories are used.

These three null hypotheses will serve as starting points for the presentation of the results of the statistical analysis of the variation across age and gender as regards the evidentiality categories *sensory*, *hearsay* and *inferential*. For each of the three null hypotheses, the BNC results will be compared to findings from the DCPSE material. Raw and normalized frequencies for the three evidentiality categories in all speaker groups can be found in the Appendix.

First, Figure 1, based on frequencies per 100,000 words, represents the differences between the female and male speakers in the material.

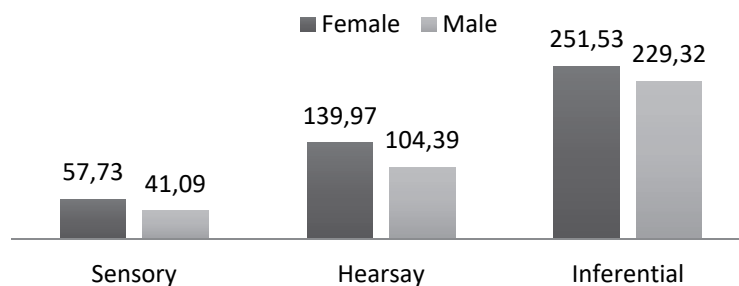


Figure 1: Gender variation in evidentiality categories in the BNC material

According to the results of the MANOVA for the entirety of the material, null hypothesis (1) can be rejected; the p -value 0.0022 shows statistically significant variation between men and women. As can be seen in Figure 1, the female speakers use all three evidentiality categories more frequently than the male speakers do. This finding is in line with the results of Alonso-Almeida and González-Cruz (2012); additionally, as their study is conducted on material consisting of nineteenth-century travel chronicles, the finding of the present study suggests that gender variation is not necessarily restricted to that particular genre or period in time.

Each of the three evidentiality categories were also tested for statistical significance, which shows that the gender variation is statistically significant in *sensory* and *hearsay* at p -values 0.0021 and 0.0303, respectively. The difference between men and women in the frequency of *inferential* evidentiality is, however, not quite statistically significant at $p = 0.0768$. This finding contradicts the results of Berglind Söderqvist (2017), whose study focuses only on inferential evidentiality and shows that the gender variation in the use of inferential evidentiality markers is statistically significant. It could be the case that this divergence is an effect of differing speech registers. While in Berglind Söderqvist's (2017) study, material consisted mainly of conversations taking place in professional contexts was used, the present study uses exclusively conversations taking place in private settings, usually among interlocutors who know each other well. Potential differences between men's and women's tendencies to use evidentiality depending on the kind of context they are in is something that is yet to be explored.

Figure 2, based on frequencies per 100,000 words, shows the corresponding gender differences in the DCPSE data.

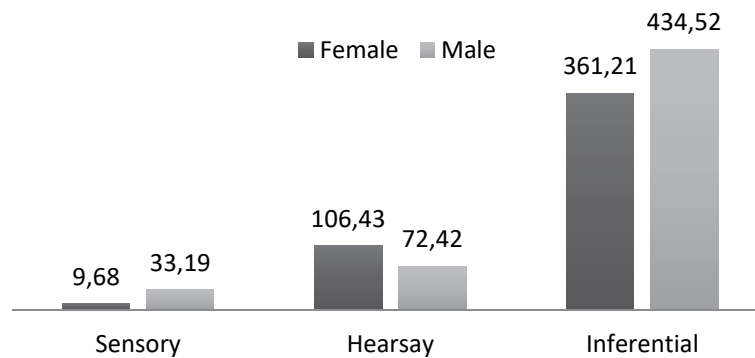


Figure 2: Gender variation in evidentiality categories in the DCPSE material

Just as in the BNC material, there appear to be differences between the male and the female speakers in the DCPSE material regarding their tendencies to use the different evidentiality categories. Unlike the BNC patterns, however, here male speakers display the higher frequencies in two out of three categories. An additional dissimilarity between the findings from the two data sets is the appearance of a stronger preference for *inferential* over the other two categories in the DCPSE material; in particular on the part of male speakers. What these findings tell us, then, is that while gender differences in evidentiality use are confirmed, it seems that the nature of this difference may be changing over time. Further, the evidential category that appears to stand out is *inferential*; it displays the most marked change over time in terms of gender variation, since it goes from being used the most frequently by men to being (non-significantly so, unlike the other categories) used the most frequently by women.

So, what might be the pragmatic implications of statistically significant gender variation for *sensory* and *hearsay* and of the seeming instability of *inferential*? As argued in Section 1 above, it is likely that *sensory* evidentiality is often perceived as more reliable than other evidentiality categories, since it refers to first-hand information. De Haan (2001b: 197) argues that there is an evidential hierarchy, where *sensory* is first, *inferential* is second and *hearsay* is third. He finds that if a speaker has access to more than one source of evidence for the same proposition, the “highest” type of evidentiality available will be preferred. This suggests that *sensory* evidentiality can be thought of as a tool that makes a proposition be perceived as reliable *and* is likely used by speakers to strengthen the force of a proposition rather than softening it. In other words, the frequent use of sensory evidentiality by the female speakers of the BNC material might indicate that they are more likely than the male speakers to use evidentiality to boost their utterances.

Hearsay evidentiality is a complex category. In de Haan's hierarchy, it comes last, as it is neither coming directly from the speaker nor a result of sensory perception. To Cornillie (2009), however, it would probably be considered to communicate high reliability, as it signals shared information (cf. Section 2.1). It is likely that

hearsay evidentiality can be employed to add force to propositions as well as to resist responsibility for the reliability of a proposition depending on the context, as illustrated by (12) and (13) from the BNC material.

(12) Anyway it *says* in the letter that he's got over her now. (KB7 7843)

(13) Is only Paul who can do gravy, so he *tells* us (KP1 2370)

In (12), the speaker uses *hearsay* evidentiality to signal that she knows this to be true because she read it in the words of the person who is in possession of first-hand information; it serves to strengthen the reliability of her utterance. In (13), the speaker, after announcing to the other interlocutors that Paul will make the gravy, tells them that Paul is the only one who can, with the (seemingly) jocular addition that this information is coming from Paul. In other words, it appears that he is using *hearsay* to show that he cannot be held accountable for the reliability of the information. While a quantitative study of the pragmatic functions of evidentiality categories is beyond the scope of the present paper, the difficulty with which (13) was found suggests that the function represented by (12) is more frequently occurring in the material. In other words, a high frequency of *hearsay* evidentiality indicates an inclination to use evidentiality to signal high reliability.

Inferential is the most frequent and probably the most pragmatically diverse category. For example, the most common marker of inferential evidentiality – *think* – which has been extensively researched, has been found to have as many as five different functions: emphatic, evaluative, tentative, mitigating and discursive (Zhang 2014). Further, Plungian (2001: 354) distinguishes between three different kinds of inferential evidentiality, namely synchronous inference (based on observing signs of something currently happening), retrospective inference (based on observing traces of something that happened in the past) and reasoning (based on general knowledge or experience relating to the resulting inference). Although pragmatic functions and semantic values refer to distinct features, this indicates that *inferential* is a heterogeneous linguistic tool. Examples (14) through (17) illustrate manifestations of *inferential* in the material of this paper.

(14) Wind in the Willows, it's pretty *obviously* it's not gonna be that one (KCU 1000)

(15) Maybe he just didn't think when they showed him it, that it wasn't the right one *cos I suppose* if you've seen them every day you don't, it doesn't register [...] (KB9 726)

(16) It's all a load of bull shit basically, it *seems*, it *seems* that the para's were certainly under fire [...] (KDA 7491)

(17) I know it *must* be you because it wasn't me. (KCH 452)

These extracts exemplify some of the categories used to outline the coding criteria in Section 4.2; (14) has an *adverb*, (15) shows a *cognitive verb*, (16) features a *non-speaker subject verb* and (17) contains a *modal*. What they have in common is that they signal evidentiality without explicitly naming any direct evidence. They appear to express different degrees of certainty, however, and to exemplify somewhat pragmatically different strategies; for example, (15) appears less assertive than the others, and the marker in (16) seems to signal a disinclination to take full responsibility for the truth of the utterance rather than a lack of certainty. It is possible that the diversity of inferential evidentiality is the reason for its diachronic instability and the fact that results differ across studies.

Next, the findings regarding the variation between age groups will be explored. Figure 3 illustrates the differences between the usage frequencies of the three evidentiality categories across the age groups in the BNC material, normalized per 100,000 words.

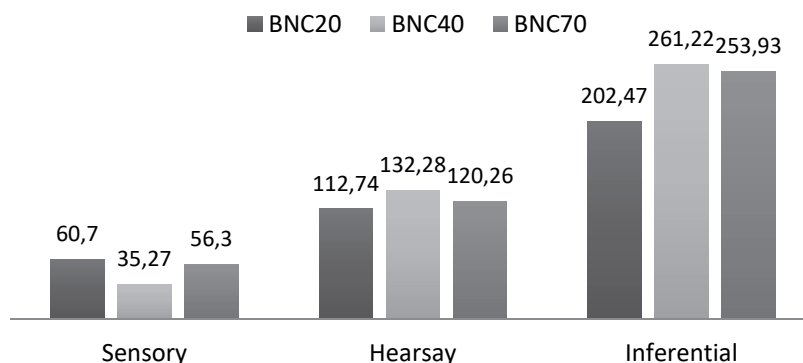


Figure 3: Age group variation in evidentiality categories in the BNC material

According to the MANOVA results, there is no statistically significant effect of age,³ neither in the material as a whole nor in any of the three evidentiality categories. For inferential evidentiality the p -value of age group variation does approach statistical significance at 0.0639; Figure 3 suggests that the rather marked difference between BNC20 and BNC40 in the category *inferential* could be the reason for this. Figure 4 below shows the age variation in the DCPSE material.

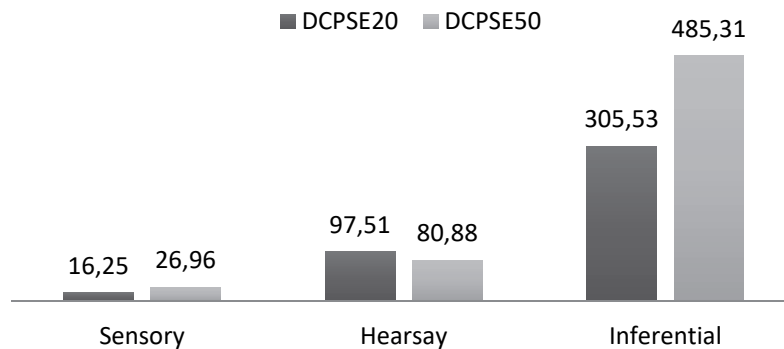


Figure 4: Age group variation in evidentiality categories in the DCPSE material

As the speakers in DCPSE50 do not appear to be retired and are no more than 65 years old, they should be thought of as corresponding to BNC40 rather than BNC70, even though they are slightly older than most of the speakers in BNC40. In other words, the findings shown in Figure 4 should be compared to the bars representing BNC20 and BNC40 in Figure 3. It should also be noted that DCPSE20, who were in their 20s in the 1970s, belong to the same generation as BNC40, who were in their 40s in the 1990s. Likewise, many speakers in DCPSE50 would have been over 70 years old by the 1990s. In other words, this study could be said to ‘follow’ two generations across a span of approximately 20 years. While there are differences in *sensory* and *hearsay* in Figure 4, they are rather slight and unlikely to have been statistically significant, if tested. The difference between DCPSE20 and DCPSE50 in *inferential*, however, is more marked and mirrors the difference between BNC20 and BNC40.

These findings suggest that while there is no statistically significant age group variation, the age-related patterns of evidentiality use appear to be diachronically stable in Present-day spoken English. As in the case of gender variation, *inferential* stands out, but here in showing very nearly statistically significant age group variation, unlike the other two categories. Again, it could be the diversity and versatility of this category that appears to give it a special standing.

Moving on to the interaction of the variables gender and age, Figures 5, 6 and 7 show the gender variation within each of the three age groups in the BNC material, based on frequencies per 100,000 words.

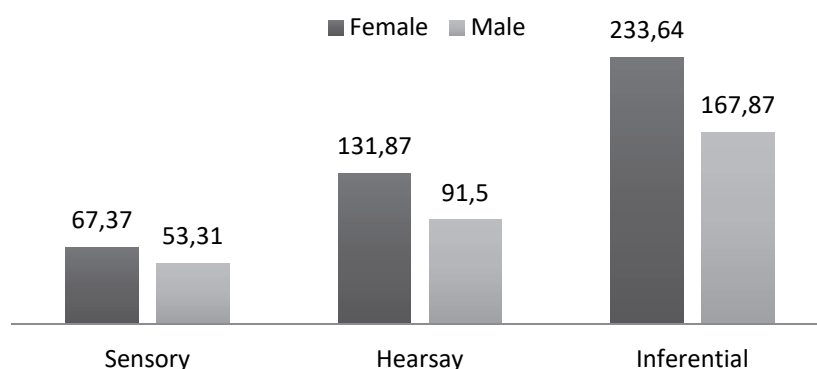


Figure 5: Gender differences in evidentiality categories within BNC20

³ See Appendix for all p -values, including the non-significant ones.

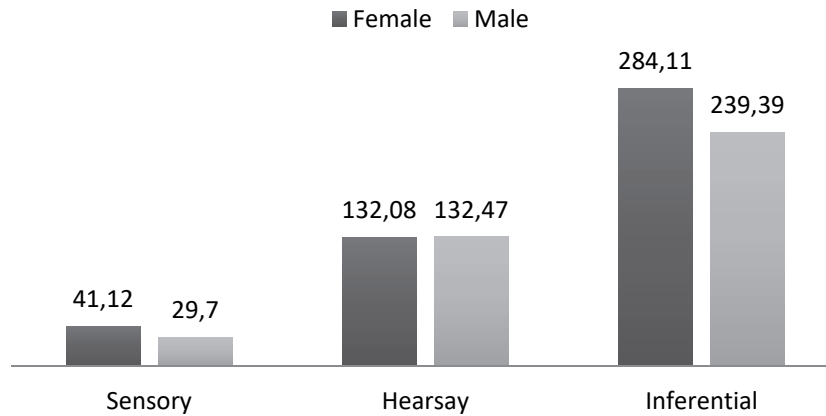


Figure 6: Gender differences in evidentiality categories within BNC40

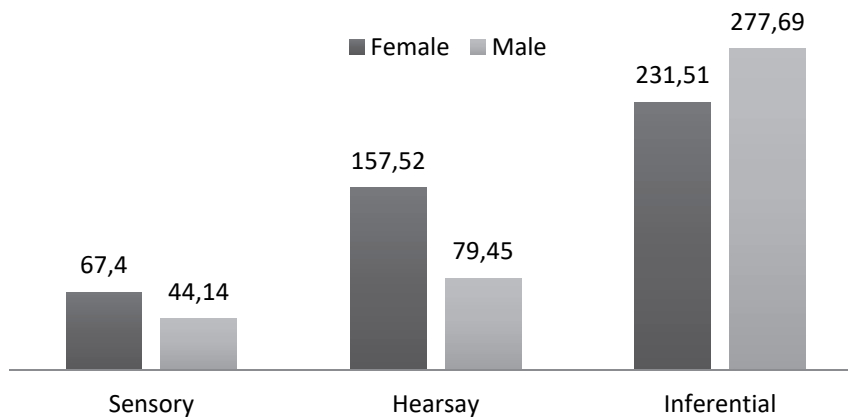


Figure 7: Gender differences in evidentiality categories within BNC70

The results of the MANOVA show that there is no statistically significant effect of the interaction of the variables age and gender in terms of the variation in the use of evidentiality categories, in the material as a whole or in any of the three evidentiality categories. The visual impression of Figures 5 through 7 suggests some variation, the most striking being that the gender variation appears to be the least pronounced in BNC40. While this finding should be viewed with some caution, as it is not statistically significant, it would be interesting to see what results might emerge from a considerably larger material; after all, the combination of the BNC and the DCPSE samples yielded only 3,937 tokens of evidentiality, divided across 10 groups of speakers (see Appendix).

Figures 8 and 9 represent the findings of gender variation within age groups in the material sampled from the DCPSE.

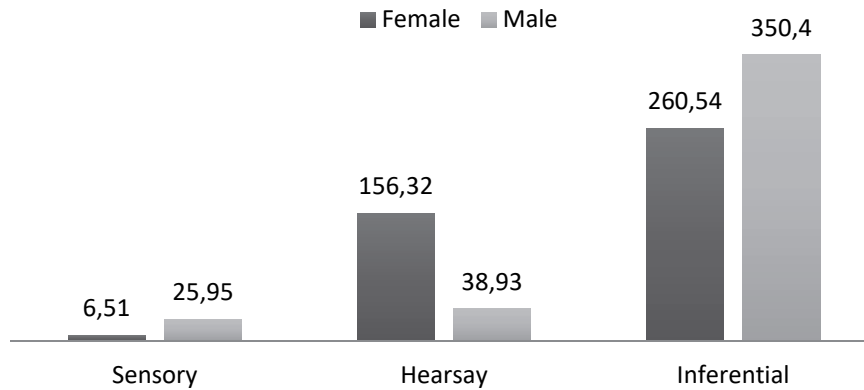


Figure 8: Gender differences in evidentiality categories within DCPSE20

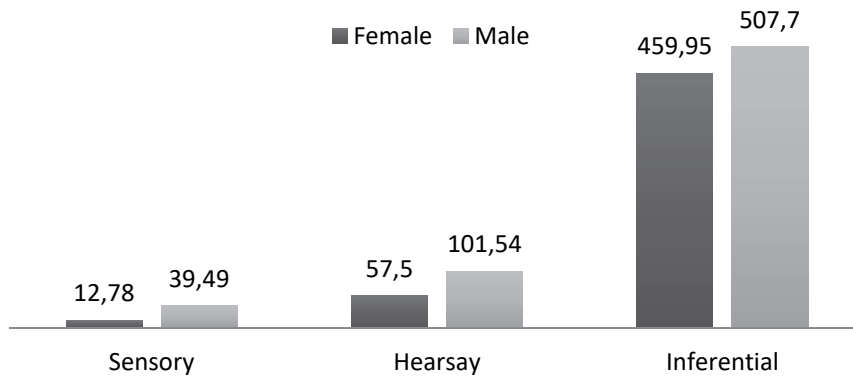


Figure 9: Gender differences in evidentiality categories within DCPSE50

Figures 8 and 9 indicate, similarly to Figure 2, that the DCPSE material differs from the BNC material in that the female speakers do not use evidentiality more frequently throughout. The two data samples show similar patterns, however, in that the gender differences appear less pronounced among the middle-aged speakers than among the young speakers.

The implications of the findings regarding gender-age interaction in variation in evidentiality use should certainly be viewed with caution, as they are not statistically significant; the pattern that gender differences decrease from the young group to the middle-aged group in both data samples is, however, interesting. Following Eckert's (1997) suggestion that age-grading in language should be viewed as happening as an effect of the social changes that happen as a person moves through various life stages, one has to wonder what happens at the onset of middle-age – in the 1970s *and* in the 1990s – that might cause gender differences in evidentiality use to decrease. Lindsey (2016: 301) suggests that the events that tend to be associated with what is commonly referred to as midlife – such as children growing up – could cause men and women to become more equal in their approach to life, as they move into “each other's” domains. Referring to research by Kasen et al. (2006) and Mann (2007), she argues that while women tend to view this “empty-nest” stage as an opportunity to pursue personal development in the form of increased independence and competence, men treat it as a relief from the stereotypical role of provider, which allows them to become more nurturing and *interdependent*. This shifting of roles might be what extends into people's language use, causing gender variation to become less pronounced.

6. CONCLUDING REMARKS AND SUGGESTIONS FOR FUTURE RESEARCH

Using corpus linguistics methods, the present study has explored the sociolinguistic variation of evidentiality in spoken British English across groups of men and women and three different stages of adulthood. While it was confirmed that women use evidentiality statistically significantly more often than men, the statistical testing

shows that age has no significant effect on this variation. Further, the comparison of the BNC material to a smaller subset from the DCPSE, recorded approximately 20 years earlier, suggests that while gender variation in evidentiality use appears to have changed in nature, age-related patterns seem diachronically stable.

The category of inferential evidentiality repeatedly stands out from the other two in being the most frequently used, diachronically unstable in terms of gender variation and showing contradictory results when the findings of this study are compared to those of Berglind Söderqvist (2017). An in-depth qualitative study of the pragmatic functions of various markers of inferential evidentiality in English might help shed light on this diverse category, and whether and how it might be divided into subcategories. Further, there are some indications that factors which have to do with the context of interaction might have an impact on the use of evidentiality, which indicates that there is reason for future research into this variable and its potential effect on evidentiality use.

REFERENCES

- Aikhenvald, Alexandra Y. 2007. Information source and evidentiality: what can we conclude? *Rivista di Linguistica* 19/1: 209–227.
- Alonso-Almeida, Francisco and Laura Cruz-García. 2011. The value of *may* as an evidential and epistemic marker in English medical abstracts. *Studia Anglica Posnaniensia* 46/3: 59–73.
- Alonso-Almeida, Francisco and M^a Isabel González-Cruz. 2012. Exploring male and female voices through epistemic modality and evidentiality in some Modern English travel texts on the Canaries. *Research in Language* 10/3: 323–343.
- Ansara, Y. Gavriel and Peter Hegarty. 2014. Methodologies of misgendering: recommendations for reducing cisgenderism in psychological research. *Feminism and Psychology* 24/2: 259–270.
- Barbieri, Federica. 2007. Older men and younger women: a corpus-based study of quotative use in American English. *English World-Wide* 28/1: 23–45.
- Barbieri, Federica. 2008. Patterns of age-based linguistic variation in American English. *Journal of Sociolinguistics* 12/1: 58–88.
- Bednarek, Monika. 2006. Epistemological positioning and evidentiality in English news discourse: a text-driven approach. *Text and Talk* 26/6: 635–660.
- Berglind Söderqvist, Erika. 2017. Gender differences and similarities in the use of inferential evidentiality markers in spoken British English: a corpus-based study. In Juana I. Marín-Arrese, Julia Lavid-López, Marta Carretero, Elena Domínguez Romero, M^a Victoria Martín de la Rosa and María Pérez Blanco eds. *Evidentiality and modality in European languages: discourse-pragmatic perspectives*. Bern: Peter Lang, 371–399.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.
- Brezina, Vaclav. 2013. *BNC64*. <http://corpora.lancs.ac.uk/bnc64/>. Data used in the BNC64 Search & Compare tool have been extracted from the British National Corpus, distributed by Oxford University Computing Services on behalf of the BNC Consortium. Used with permission. (3 November 2016)
- The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. <http://www.natcorp.ox.ac.uk>.
- Canary, Daniel J. and Kimberley S. Hause 1993. Is there any reason to research sex differences in communication? *Communication Quarterly* 41/2: 129–144.
- Cappelli, Gloria. 2007. *'I reckon I know how Leonardo da Vinci must have felt...': Epistemicity, evidentiality and English verbs of cognitive attitude*. Paris: Paris Publishing.
- Carretero, Marta and Juan Rafael Zamorano-Mansilla. 2014. Annotating English adverbials for the categories of epistemic modality and evidentiality. In Juana I. Marín-Arrese, Marta Carretero and Jorge Arús Hita eds. *English modality: core, periphery and evidentiality*. Berlin: De Gruyter Mouton, 317–355.
- Carretero, Marta and Juan Rafael Zamorano-Mansilla. 2015. Evidentiality as conversational implicature: implications for corpus annotation. *Procedia – Social and Behavioral Sciences* 212: 146–150.
- Cornillie, Bert. 2009. Evidentiality and epistemic modality: on the close relationship between two different categories. *Functions of Language* 16/1: 44–62.
- de Haan, Ferdinand. 2001a. The relation between modality and evidentiality. *Linguistische Berichte* 9: 201–216.
- de Haan, Ferdinand. 2001b. The place of inference within the evidential system. *International Journal of American Linguistics* 67/2: 193–219.
- The Diachronic Corpus of Present-Day Spoken English (DCPSE). Distributed by the Survey of English Usage, University College London. <http://www.ucl.ac.uk/english-usage/projects/dcpse>.
- Eckert, Penelope. 1997. Age as a sociolinguistic variable. In Florian Coulmas ed. *The handbook of sociolinguistics*. Oxford: Blackwell, 151–167.

- Eckert, Penelope and Sally McConnell-Ginet. 2013. *Language and gender*, 2nd ed. Cambridge: Cambridge University Press.
- Grob, Lindsey M., Renee A. Meyers and Renee Schuh. 1997. Powerful/powerless language use in group interactions: sex differences or similarities? *Communication Quarterly* 45/3: 282–303.
- Holmes, Janet. 1990. Hedges and boosters in women's and men's speech. *Language and Communication* 10/3: 185–205.
- Kasen, Stephanie, Henian Chen, Joel Sneed, Thomas Crawford and Patricia Cohen. 2006. Social role and birth cohort influences on gender-linked personality traits in women: a 20-year longitudinal analysis. *Journal of Personality and Social Psychology* 91/5: 944–958.
- Lakoff, Robin. 1973. Language and woman's place. *Language in Society* 2/1: 45–80.
- Lindsey, Linda L. 2016. *Gender roles: a sociological perspective*, 6th ed. London: Routledge.
- MacDonald, Eleanor. 1998. Critical identities: rethinking feminism through transgender politics. *Atlantis* 23/1: 3–12.
- Mann, Robin. 2007. Out of the shadows?: grandfatherhood, age, and masculinities. *Journal of Aging Studies* 21/4: 281–291.
- Mulac, Anthony, John M. Wiemann, Sally J. Widenmann and Toni W. Gibson. 1988. Male/female language differences and effects in same-sex and mixed-sex dyads: the gender-linked language effect. *Communication Monographs* 55: 315–335.
- Murphy, Bróna. 2010. *Corpus and sociolinguistics: investigating age and gender in female talk*. Amsterdam: John Benjamins.
- Nelson, Gerald, Sean Wallis and Bas Aarts. 2006. *The International Corpus of English: the British Component (Release 2): The Diachronic Corpus of Present-Day Spoken English. Getting Started with ICECUP 3.1*. London: Survey of English Usage, University College London.
- Paunonen, Heikki. 1994. Language change in apparent time and in real time. Paper delivered at the 23rd Annual Conference on New Ways of Analyzing Variation, Stanford CA.
- Plungian, Vladimir A. 2001. The place of evidentiality within the universal grammatical space. *Journal of Pragmatics* 33: 349–357.
- Precht, Kristen. 2008. Sex similarities and differences in stance in informal American conversation. *Journal of Sociolinguistics* 12/1: 89–111.
- Whitt, Richard J. 2009. Auditory evidentiality in English and German: the case of perception verbs. *Lingua* 119/7: 1083–1095.
- Willett, Thomas. 1988. A cross-linguistic survey of the grammaticization of evidentiality. *Studies in Language* 12/1: 51–97.
- Zhang, Grace. 2014. The elasticity of *I think*: stretching its pragmatic functions. *Intercultural Pragmatics* 11/2: 225–257.

Corresponding author

Engelska parken, Thunbergsvägen 3 L
Box 527
751 20 Uppsala
e-mail: erika.soderqvist@engelska.uu.se

received: April 2017
accepted: December 2017

Appendix A:

Table A1 shows the number (in the BNC material) of types and tokens analyzed and coded in this study, as well as which types were not included because of their low frequency and how all types were found – in the literature or through the manual scanning process. Please note that only the types that were *only* found during the manual scan are marked *search* here; several of those marked *literature* were found during the manual search as well. Table A2 shows the raw and normalized frequencies for the three evidentiality categories in all speaker groups. Table A3 presents the *p*-values of the MANOVA results for the effects of the independent variables in the three evidentiality categories as well as in the entire BNC material.

Lexical item	Normalized freqs	Raw freqs	Origin
<i>according to</i>	excluded	excluded	literature
<i>apparently</i>	107.58	92	literature
<i>appear</i>	excluded	excluded	literature
<i>believe</i>	153.19	131	literature
<i>clearly</i>	excluded	excluded	literature
<i>evidently</i>	excluded	excluded	literature
<i>find out</i>	107.58	92	search
<i>go on about</i>	excluded	excluded	search
<i>hear</i>	478.27	409	literature
<i>I BE in</i>	excluded	excluded	search
<i>know</i>	7,458.24	6,378	literature
<i>look</i>	2,530.51	2,164	literature
<i>make out</i>	excluded	excluded	search
<i>must</i>	765.94	655	literature
<i>obviously</i>	167.22	143	literature
<i>remember</i>	612.75	524	literature
<i>reportedly</i>	excluded	excluded	literature
<i>say</i>	8,951.52	7,655	literature
<i>see</i>	3,929.08	3,360	literature
<i>seem</i>	323.92	277	literature
<i>sound</i>	156.7	134	literature
<i>suppose</i>	424.48	363	literature
<i>tell</i>	1,499.13	1,282	literature
<i>think</i>	5,624.67	4,810	literature
<i>watch</i>	360.17	308	search
Total	33,650.93	28,777	

Table A1. Lexical items – frequencies in the BNC data prior to coding

Groups	Sensory	Hearsay	Inferential	Unspecified	Total
F20	94 (67.37)	184 (131.87)	326 (233.64)	12 (8.6)	616 (441.47)
M20	67 (53.31)	115 (91.5)	211 (167.87)	19 (15.12)	412 (327.79)
F40	66 (41.12)	212 (132.08)	456 (284.11)	6 (3.74)	740 (461.05)
M40	50 (29.7)	223 (132.47)	403 (239.39)	15 (8.91)	691 (410.47)
F70	92 (67.4)	215 (157.52)	316 (231.51)	9 (6.59)	632 (463.02)
M70	55 (44.14)	99 (79.45)	346 (277.69)	12 (9.63)	512 (410.92)
DCPSE-F20	1 (6.51)	24 (156.32)	40 (260.54)	2 (13.03)	67 (436.4)
DCPSE-M20	4 (25.95)	6 (38.93)	54 (350.4)	2 (12.98)	66 (428.21)
DCPSE-F50	2 (12.78)	9 (57.5)	72 (459.95)	1 (6.39)	84 (536.6)
DCPSE-M50	7 (39.49)	18 (101.54)	90 (507.7)	2 (11.29)	117 (660.01)
Total	438	1,105	2,314	80	3,937

Table A2. Raw frequencies (and freqs/100,000 words) of evidentiality categories in all speaker groups

Category	Gender	Age	Gender + age
All categories	0.0022	0.1654	0.2541
Sensory	0.0021	0.2915	0.1449
Hearsay	0.0303	0.8921	0.1019
Inferential	0.0768	0.0639	0.8154

Table A3. Results of the MANOVA: *p*-values across evidentiality categories and independent variables