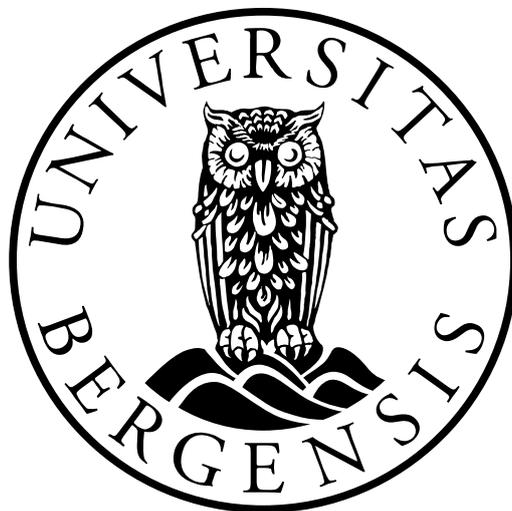


*Say that she rail; why then I'll tell her plain
She sings as sweetly as a nightingale*

A corpus-based study of the development of *say* as a conditional subordinator in English

Kari Johnsen Skjelvik



Master's Thesis in English Linguistics

Department of Foreign Languages

University of Bergen

May 2018

Abstract in Norwegian

Denne masteroppgaven tar for seg hvordan *say* ‘si’ kan brukes som en betingelseskonjunksjon på engelsk. Dette er i utgangspunktet et verb, og dermed er bruken som konjunksjon et eksempel på grammatikalisering. Grammatikalisering er en fellesbetegnelse for språkfenomen hvor en enhet, enten et ord eller en frase, utvikler seg fra å være innholdsord til å bli funksjonsord. Med innholdsord menes ordklasser som substantiv, verb og adjektiv. Funksjonsord er for eksempel pronomen, preposisjoner eller konjunksjoner. Dette er en gradvis prosess hvor enheten blir assosiert med en grammatisk funksjon, og til slutt blir omkategorisert som en funksjonsenhet. Bruken av *say* som betingelseskonjunksjon har fått lite oppmerksomhet i tidligere forskning, og dette er så vidt det er meg bekjent den første studien dedikert til utviklingen og bruken av betingelseskonjunksjonen *say*.

For å finne ut hvordan *say* har utviklet seg til en konjunksjon, har jeg samlet data fra forskjellige korpora. Kombinert dekker disse perioden 1500–2015, og kan dermed fortelle mye om utviklingen av konjunksjonen *say* fra opprinnelsen rundt 1600 til dagens bruk i engelsk. Oppgaven er først og fremst diakronisk, og utviklingen av *say* blir vektlagt fremfor en grundig analyse av dagens bruk. Betingelsessetninger med *say* blir sammenlignet med andre, vanligere betingelsessetninger, spesielt de introdusert av *if*. Jeg introduserer begrepet ‘prototypiske betingelsessetninger’, og dette danner grunnlaget for en stor del av analysen.

Funnene fra oppgaven viser at bruken av *say* som en betingelseskonjunksjon har steget over tid. Flere trekk ved utviklingen er typiske for grammatikalisering, og forsterker dermed inntrykket av at *say* har blitt grammatikalisert. Betingelsessetninger med *say* har derimot ikke utviklet seg mot den prototypiske betingelsessetningen, men ser ut til å være i ferd med å utvikle sin egen ‘norm’. Spesielt viser funnene at konstruksjoner der betingelsessetningen er syntaktisk uavhengig fra den tilhørende hovedsetningen er vanlig.

Acknowledgements

I would like to thank my supervisor, Dr Jerzy N. Nykiel, for all his help and guidance throughout this process. I have truly felt that he has taken a personal interest in my project, and his comments and nudges in the right direction have been invaluable.

I would also like to thank the crowd at the reading hall. We have shared lots of laughs and nervous speculations, and I have more than once been tempted into taking a much longer lunch break than I should. I would especially like to mention Erika, who has gone through every step of the process with me, and whose friendship I value immensely.

Finally, I would like to thank my family. Without their help this past year, my dog would certainly have been the unhappiest puppy in the world.

Thank you!

Table of contents

Abstract in Norwegian	iii
Acknowledgements	iv
List of tables	vii
List of figures	viii
Abbreviations and conventions	ix
1. Introduction	1
1.1 Aim and scope	1
1.2 Research questions and hypotheses	2
1.3 Outline of thesis	3
2. Theoretical background	4
2.1 Grammaticalisation	4
2.1.1 Central mechanisms	5
2.1.2 Pragmatic factors	7
2.1.3 Subordination and clause combining	10
2.1.4 Stages and contexts	13
2.2 Conditionals	18
2.2.1 The link between protasis and apodosis/conditional relation	19
2.2.2 Formal characteristics	25
2.2.3 Prototypical and non-prototypical conditionals	30
2.3 About <i>say</i>	32
3. Methods and results	38
3.1 Method	39
3.1.1 The corpora	39
3.1.2 Searches in COCA and COHA	40
3.1.3 Searches in corpora covering earlier time periods	41
3.2 Results in COCA (1990–2015) and COHA (1810–1980)	41
3.2.1 Overall results	41
3.2.2 Genre	44
3.2.3 <i>That</i> -/zero clauses	45

3.2.4 Individual searches in COCA	46
3.2.5 Individual searches in COHA	48
3.2.6 Formal structure	50
3.2.7 Mental spaces	56
3.2.8 Isolating and critical contexts	59
3.3 Results before 1810.....	61
4. Discussion	67
4.1 Findings in relation to the prototypical conditional	67
4.2 Findings in relation to grammaticalisation theory	71
5. Conclusion	77
5.1 Shortcomings	78
5.2 Further research	79
References.....	80

List of tables

Table 3.1: Overall results from COCA and COHA.....	42
Table 3.2: Distribution of <i>say</i> -conditional across genres in COCA.....	45
Table 3.3: Distribution of <i>say</i> -conditional across genres without ‘Spoken’ in COCA.....	45
Table 3.4: Distribution of <i>say</i> -conditional across genres in COHA.....	45
Table 3.5: Distribution of <i>that</i> -clauses and zero-clauses following conditional <i>say</i> in COCA and COHA.....	46
Table 3.6: Distribution of <i>that</i> -clauses and zero-clauses following imperative <i>say</i> in COCA and COHA.....	46
Table 3.7: Distribution of <i>say</i> -conditional with <i>that</i> -clauses by individual searches in COCA.....	47
Table 3.8: Distribution of <i>say</i> -conditional with zero-clauses by individual searches in COCA.....	48
Table 3.9: Distribution of <i>say</i> -conditional with <i>that</i> -clauses by individual searches in COHA.....	49
Table 3.10: Distribution of <i>say</i> -conditional with zero-clauses by individual searches in COHA.....	50
Table 3.11: Distribution of <i>say</i> -conditional across formal variants in COCA.....	54
Table 3.12: Distribution of <i>say</i> -conditional across formal variants in COHA.....	55
Table 3.13: Distribution of <i>say</i> -conditional across mental spaces in COCA.....	56
Table 3.14: Distribution of <i>say</i> -conditional across mental spaces in COHA.....	58
Table 3.15: Distribution of <i>say</i> -conditional across mental spaces and formal variants in COCA.....	58
Table 3.16: Distribution of <i>say</i> -conditional across mental spaces and formal variants in COHA.....	59
Table 3.17: Top three results for . <i>say</i> <i>PRN</i> <i>V</i> in COCA.....	61
Table 3.18: Results before 1810, by corpus.....	62
Table 3.19: Results from Shakespeare’s plays, by genre.....	65
Table 3.20: Distribution of <i>say</i> -conditional across mental spaces in earlier corpora.....	65

List of figures

Figure 2.1: The development of conditional <i>say</i> , adapted from Brinton (2008: 89).....	36
Figure 3.1: Variables for searches in COCA and COHA.....	41
Figure 3.2: Normalised frequency of <i>say</i> -conditional per million words in COHA.....	43
Figure 3.3: Normalised frequency of <i>say</i> -conditional per million words by decade from 1810–2009 in COHA and COCA.....	44
Figure 3.4: Distribution of <i>say</i> -conditional by formal variants across time periods in COHA.....	55

Abbreviations and conventions

OED	–	Oxford English Dictionary
CCCs	–	Complex conditional connectives
PIE	–	Proto-Indo-European
BT	–	Bosworth-Toller Anglo-Saxon Dictionary
MED	–	Middle English Dictionary
COCA	–	Corpus of Contemporary American English
COHA	–	Corpus of Historical American English
PPCEME	–	Penn-Helsinki Corpus of Early Modern English
ARCHER	–	A Representative Corpus of Historical English Registers
Lampeter	–	The Lampeter Corpus of Early Modern English Tracts
Shakespeare	–	<i>The Oxford Shakespeare</i>

> – developed into

When discussing the role of the speaker and hearer in a specific speech event, I follow the example of Declerck and Reed (2001) in referring to the speaker as ‘she’ and the hearer as ‘he’.

I will invariably use the terms ‘speaker’ and ‘hearer’, even when interlocutor or some other term might be more appropriate.

OE	Old English (c. 600–1125)
ME	Middle English (c. 1125–1500)
eModE	Early Modern English (c. 1500–1750)
ModE	Modern English (c. 1750–1950)
PDE	Present-Day English (c. 1950–)

1. Introduction

1.1 Aim and scope

The aim of the present study is to trace the development of *say*-conditional from its origins to its use in Present Day English (PDE). Although *say*-conditional is both mentioned in the Oxford English Dictionary (OED) and mentioned by authors such as Poutsma (1926) and Brinton (2008), it seems to have received little attention in its own right. Indeed, the OED lists the construction as a subentry of ‘verb and interjection’, and Poutsma’s mention is limited to a single sentence.

The earliest attestation of *say*-conditional in the OED is given in [1]. This dates from 1596. The corpora I have used for data collection cover the period 1500–2015, and as such should be sufficient to provide a reliable overview of the development of *say*-conditional. Although a contrastive view between British and American English would be preferable, the data after 1810 is taken from American English. This is due to both time constraints and to the ease of access to large corpora covering this period in American English.

[1] Well, **say** tis read, what is your further meaning in the matter.

Sir Thomas More (c1596). i. i. 159 [OED]¹

[2] **Say** you want to hunt buffalo or passenger pigeons. We beam you back to, **say**, the 1700s.

Outdoor Life (2000) [COCA]

This study is concerned solely with the conditional use of *say*. Contrasting the two uses of *say* in [2] provides a provisional illustration. The first use is a conditional subordinator and introduces a hypothetical situation. The second use has the meaning ‘about, approximately’, and has no subordinating function.

The study is situated within the framework of grammaticalisation theory, and relies especially on the works of Hopper and Traugott (2003) and Diewald (2002) to

¹ Throughout this study, the emphases in all examples are mine unless otherwise stated.

account for the changes *say* goes through. The analysis draws a great deal upon Dancygier (1999) and Dancygier and Sweetser (2005) for the categorisation of conditionals.

1.2 Research questions and hypotheses

The research questions and hypotheses listed below are based on grammaticalisation theory in general, as well as both surveys of conditionals and previous studies on the grammaticalisation of other subordinators. As the question of when conditional *say* originated has already been tentatively answered, this is not one of my research questions. However, while the present study does not collect data prior to 1500, I still have hopes of finding indications of which contexts conditional *say* originated in, and seeing whether this knowledge can indicate anything about the motivation for the change. As the main focus of this study is the development of *say*-conditional, the diachronic perspective is favoured in the research questions.

Research questions:

1. Has *say*-conditional become more frequent over time?
2. Are there any changes in which clauses follow conditional *say*, and in their distribution?
3. Are there any changes in which registers *say*-conditional is used?
4. How does *say*-conditional relate to descriptions of more established conditionals, especially *if*-conditionals? Are there any changes in this?
5. Can the observed developments of *say*-conditional indicate something about its future development?

Hypotheses:

1. As *say*-conditional has become more accepted, the frequency of the construction has increased accordingly.
2. There are both *that*-clauses and zero-clauses following conditional *say*. Since *that* is a complementiser on its own, there will be a decrease of *that*-clauses over time.
3. As *say*-conditional becomes more established, it is also used in a wider range of registers.

4. As *say*-conditional becomes more established, it develops to more closely resemble the ‘norm’ of *if*-conditional.
5. Based on the use of *say*-conditional in PDE, and on its recent development, it is possible to indicate how it will be used in the future.

1.3 Outline of thesis

This study is divided into six chapters, including references. Relevant theoretical background will be presented in Chapter 2. Chapter 3 is divided into two main parts, so that it both details the methods used in my data collection, as well as presents the results thereof. Chapter 4 is dedicated to the comparison and discussion of my results. The conclusion is presented in Chapter 5.

2. Theoretical background

In this chapter I introduce aspects of theory relevant for the present study. The chapter is divided into three main parts. Section 2.1 concerns grammaticalisation theory, and looks at different aspects of how lexical items can develop into grammatical items such as conditional subordinators. Section 2.2 details different views of conditionals, and establishes the framework I will be using in my analysis. Section 2.3 is a chronological overview of the development of *say*, with focus on its various uses.

2.1 Grammaticalisation

Grammaticalisation is the study of “how lexical items come in certain linguistic contexts to serve grammatical functions and how grammatical items develop new grammatical function.” (Hopper and Traugott 2003: 1) The term is twofold, as it refers both to the field of study and to the language phenomena themselves. The term was first coined by Antoine Meillet in 1912, but the interest for the origins of grammatical forms had been present for some time already.

We can classify words into two major categories: lexical words and grammatical words. To lexical words belong word classes such as nouns, verbs and adjectives, while grammatical words are prepositions, connectives, pronouns and demonstratives, to name a few (Hopper and Traugott 2003: 4). The function words serve, among other things, to indicate relationships between content words and to tie together parts of discourse. The main idea of the grammaticalisation framework is that grammatical words have their origins in lexical words which have developed over time to acquire grammatical functions. This change does not happen as a sudden shift from one category to another, but rather as a movement along a continuum. This gradual change is often illustrated metaphorically in terms of a cline of grammaticality (Hopper and Traugott 2003: 6):

content item > grammatical word > clitic > inflectional affix

A cline illustrates the set of changes an item typically goes through. The item in question might move all the way from a lexical item (or content item) to an inflectional affix – or even become reduced to zero – or it can stop at any point of the cline. The idea of a continuum also implies that the item can stop developing at a point between two of the

categories listed in the cline, as the development from one to another is gliding, not sudden. Further, the hypothesis of unidirectionality states that there is a strong tendency for development towards the grammatical ‘end’ of the cline, while developments from grammatical to more lexical are sporadic (Hopper and Traugott 2003: 99).

Something that is important to note from the onset is that grammaticalisation happens in a specific context. In Diewald’s (2002: 103) words: “a new grammatical function does not arise homogeneously in all uses of the item concerned, but in its origin is bound to specific linguistic contexts or constructions”. Thus it is not the lexical item in itself that is grammaticalised, but rather the entire structure it appears in (Bybee, Perkins, and Pagliuca 1994: 11).

In the remainder of this section I will take a closer look at the central mechanisms in grammaticalisation, as well as some pragmatic factors which may motivate these kinds of changes. I will then introduce Diewald’s three successive stages of the grammaticalisation process, and explore how these can be relevant to the present study.

2.1.1 Central mechanisms

The two central mechanisms of grammaticalisation – and indeed of all language change – are reanalysis and analogy (Hopper and Traugott 2003: 69). Through these, new items are created and spread to become part of the established language. Reanalysis is defined by Hopper and Traugott (2003: 50) as the hearer understanding “a form to have a structure and a meaning that are different from those of the speaker”. A good example is the development of the futurity marker *be going to*, which is illustrated in [1–2]. [1] shows the structure before reanalysis has taken place. Here, the lexical verb of motion has an auxiliary *am* and a non-finite complement *to marry Bill*, which explains the purpose of the movement. The meaning is ‘I am moving in order to marry Bill’. In [2], on the other hand, the utterance is restructured so that *am going to* is perceived as a unit. The meaning is now ‘I will marry Bill (in the immediate future)’, and there is no movement indicated. Hopper and Traugott (2003: 3) write that this reanalysis is possible because the inference of futurity has become salient, i.e. the movement in order to marry Bill implies that the action will take place in the future. I will return to inferences in Section 2.1.2.

[1] I am going [to marry Bill].

[2] I [am going to] marry Bill.

Hopper and Traugott (2003: 3)

It is important to note that at this point, there is no outward evidence that a change in interpretation, i.e. reanalysis has taken place. The interlocutors can go on for some time, the speaker intending [1] and the hearer interpreting [2], without them realising they see the utterance differently. This is reflected in Langacker's definition (1977: 58; in Hopper and Traugott 2003: 51), where he calls reanalysis a "change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation". We could say that reanalysis is covert (Hopper and Traugott 2003: 63).

Analogy, on the other hand, is overt. Hopper and Traugott (2003: 63–4) define analogy as "the attraction of extant forms to already existing constructions". There is no creation of new forms through analogy, but rather a spreading of an already existing form into contexts where it was not used previously. In terms of the development of *be going to*, after reanalysis has taken place, this new futurity marker can be used in contexts which were incompatible with the previous 'moving in order to' meaning (Hopper and Traugott 2003: 3).

[3] "I am going to hate it – hate being a civilian," he says.

Denver Post (2004) [COCA]

[4] "I am going to go clean up," I said.

Analog Science Fiction & Fact (2006) [COCA]

This is clear in [3], where the speaker is unlikely to mean 'I am moving in order to hate being a civilian', as stative verbs such as *hate* and *like* are not generally thought of as something one does on purpose. Another example of a previously unavailable context is seen in [4], where the reanalysed meaning of *be going to* is combined with the earlier meaning. To interpret *going* here as a verb of motion is difficult for two reasons. First, *am going to* is followed by another verb of motion, i.e. *go*. With a non-reanalysed interpretation of *am going to*, this would give the meaning 'I am moving in order to move', which makes little sense. Second, the phrase *clean up* is an explicit statement of purpose. As the pre-reanalysis meaning of *be going to* is a "purposive directional

construction” (Hopper and Traugott 2003: 2), this would make the meaning of [4] ‘I am moving in order to move in order to clean up’.

Because reanalysis is covert, analogy is often the only way we can realise that a change in interpretation has taken place. In the case of *be going to* above, the two language users can, as mentioned, go on for quite some time with different interpretations of the phrase without realising it. It is first when the hearer from the setting above employs analogy to create other kinds of phrases, such as those in [3–4], that the change in reanalysis has taken place.

2.1.2 Pragmatic factors

While reanalysis and analogy are very important to grammaticalisation, so are the potential enabling or motivating factors behind language change. Paraphrasing Dressler (1985), Hopper and Traugott (2003: 72) write that “the development of language involves conflict and problem solving.” A speaker normally seeks to be understood, and will shape her utterance accordingly. Likewise, a hearer will “seek the most unambiguous interpretation” and make active use of reasoning strategies to do so (Hopper and Traugott 2003: 72). Although it is only implicitly assumed in this section, we might also add that this view on language change necessarily involves that change occurs first in the spoken language, where speakers and hearers make use of various strategies in order to understand one another. To this Mair (2011: 245) adds that some few grammaticalisation phenomena seem to have originated in the written language.

In the speaker’s desire to be understood by the hearer, Grice (1975: 45) imagines that she is guided by the Cooperative Principle: “make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.” Under this main principle, he formulates a series of maxims, which can be summed up in the four categories of Quantity, Quality, Relation and Manner. The following two maxims serve as an expression of the category of Quantity:

1. Make your contribution as informative as is required (for the current purposes of the exchange).²
2. Do not make your contribution more informative than is required.

² All the maxims are direct quotes from Grice (1975: 45–6)

To the category of Quality belong two maxims:

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

These could be summed up as the supermaxim “Try to make your contribution one that is true.” There is a single maxim under the category of Relation:

1. Be relevant.

Finally, the category of Manner contains the supermaxim “Be perspicuous” and at least four more specific maxims:

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief (avoid unnecessary prolixity).
4. Be orderly.

Through these maxims a speaker should be able to make cooperative and understandable contributions to a conversation. Hopper and Traugott (2003: 79) note that there is much debate concerning whether these maxims are the correct ones, and whether others are equally important. They also note that grammaticalisation draws mostly on the second maxim of Quantity, although in a slightly altered form, namely “Say no more than you must and mean more thereby.” Additionally, Hopper and Traugott (2003: 79) feel that *maxim* is a misleading term, as it implies principles and obligations. They follow Levinson (2000) in deeming *heuristic* a more appropriate term, as it “evokes strategic choice in the communicative situation.”³

These maxims – or heuristics – are linked to conversational implicatures. These are “inferences that are made in linguistic contexts from one clause or constituent to another, or even from one utterance to another” (Hopper and Traugott 2003: 78). By relying on the linguistic context, the speaker is able to imply more meaning than is ‘lexically’ present in her utterance. Before concluding that a conversational implicature is made, three criteria must be satisfied. If we assume a situation where a speaker says *x*, and where there is the potential implicature *y*, we must be able to assume that (1) the speaker is observing the conversational maxims; (2) the speaker is aware that implying *y* is necessary in order to observe the maxims; and (3) both the speaker and the hearer

³ I will use the two terms interchangeably, so that when referring to Grice, I use ‘maxim’, and when referring to Hopper and Traugott, I use ‘heuristic’.

believe that the hearer is capable of understanding that the requirement in (2) is necessary (Grice 1975: 49–50).

- [5] A: I am out of petrol.
B: There is a garage around the corner.

Grice (1975: 51)

We can look at these criteria in relation to [5]. B's utterance has the actual meaning as stated above, but the implicature 'your problem can be solved by visiting the garage around the corner, which sells petrol.' If B is lying, or is stating that there is a garage nearby while knowing it is closed, he has not followed the maxims of Quality and Relation. Thus, criterion (1) is not fulfilled, and there is no conversational implicature. Likewise, B must be aware that his speaking of the presence of a garage is only relevant if it supplies petrol, thus solving A's problem. If he is not, criterion (2) is unfulfilled. Finally, B must believe A capable of understanding that B would not mention this garage if he did not believe it was relevant to A in his present situation, i.e. A must attempt to interpret B's utterance on the basis of the conversational maxims. If B does not believe that A is able to do this leap of thought, there is no purpose in attempting an implicature, and criterion (3) remains unfulfilled. In the case of [5], B would have no reason to doubt A's ability to understand the connection between A's problem and B's utterance. In a slightly different scenario, however, where B knows that there is a flower shop nearby which also happens to be selling petrol, B might not believe that the connection is clear, and might resolve not to use an implicature at all. This touches upon the point that the speaker must consider the knowledge of the hearer when deciding whether to imply or explicitly state her meaning.

Thus far we have looked mainly at the role of the speaker. As mentioned, the hearer is also an active participant in the conversation, and seeks to understand the speaker's utterance by making use of various strategies. Andersen (1973: 774–6) explains the role of the three modes of inferencing, i.e. deduction, induction and abduction, and emphasises the role of abduction in language change. Abductive inferencing "proceeds from an observed result, invokes a law, and infers that something may be the case" (Andersen 1973: 775). In terms of language change, language learners observe the utterances of others, i.e. the results, and must relate these to the various laws they are

aware of in order to explain the underlying structure of them. Andersen (1973: 776–7) notes that this is a “goal-oriented process”, but that the aim of the language learner is not a ‘true’ or ‘optimal’ grammar, but rather any grammar which conforms to the observed data.

We might attempt to apply these principles to the case of *be going to*. In terms of Grice’s model, the utterance in [1–2], repeated here as [6], can be seen as having the literal meaning x and the following implicature y . Because the speaker is observing the conversational maxims, and especially the maxim of Quantity rephrased as “Say no more than you must and mean more thereby” (Hopper and Traugott 2003: 79), she will not explicitly say y , but will leave it to the hearer to understand the implicature.

- [6] I am going to marry Bill
 x : ‘I am moving in order to marry Bill’
 y : ‘When I reach my destination, I will commence marrying Bill’

The hearer will then attempt to interpret the utterance, and, understanding the implied meaning y , will conclude that the speaker will marry Bill in the near future. This is where both abduction and reanalysis, as mentioned in Section 2.1.1, come in. Abductive reasoning allows the hearer to come to the faulty conclusion that *am going to* is indicating a future action, and not describing a present one. Because the ‘movement meaning’ is so general, the ‘future action meaning’ might become salient, and the hearer will reanalyse the structure of the utterance as detailed in Section 2.1.1. The implicature y thus becomes the literal meaning of the utterance in the hearer’s grammar. Because reanalysis is covert, the hearer might then produce similar utterances which seem acceptable, but are not structured according to the same principles as the original utterance. The hearer’s grammar will thus be different from that of the speaker, while still conforming to the observed data.

2.1.3 Subordination and clause combining

Since the present study focuses on how *say* has developed from a *verbum dicendi* to a conditional subordinator, we should take a look at how some clauses come to be subordinated, and how markers of subordination are ‘chosen’. Before continuing, we

should note that the notion of subordination is by no means unproblematic. For instance, Haiman (1985: 216–228) argues that the criteria used to prove that the clauses in [7] have a relationship of subordination, whereas the clauses in [8] are coordinated, are to a great extent not satisfactory.

[7] *Leaving her children*, she fled for safety

[8] *She left her children* and fled for safety

Haiman (1985: 196; emphasis original)

Likewise, Cristofaro (2014) illustrates how four common ‘tests’ to determine embeddedness, i.e. clause internal word order; clause extraposition; backwards pronominal anaphora; and the Coordinate Structure Constraint, fail to provide definite conclusions as to which clause types are embedded and which are not. In other words, a certain clause type might be seen as embedded according to one test, while another test will categorise the same clause as non-embedded. This, then, indicates that the criteria for what we consider embedded clauses are unclear. While keeping this in mind, I will follow the framework of Hopper and Traugott (2003) in this section.

We can divide clauses into two main types, namely nuclei and margins (Hopper and Traugott 2003: 176–7). A nucleus is a clause which can stand on its own, while a margin is dependent on a nucleus. However, how strong this dependence is varies with the type of margin in question. With regards to semantics, we can divide margins into three subtypes: complements, relative clauses and adverbial clauses. Of these, adverbial clauses are of particular interest, as these are clauses which modify either a verb phrase or, more importantly for us, an entire proposition. This is, of course, what a conditional clause does. From an evolutionary perspective, complex clauses probably originated as the result of two nuclei being seen as mutually relevant, and the consequent combining of the two (Hopper and Traugott 2003: 177).

Hopper and Traugott (2003: 177) offer a “cline of clause combining” where we can see a movement from loose to tighter relationships between the clauses:

parataxis > hypotaxis > subordination

Parataxis is the loosest relationship of the three, where two or more nuclei are juxtaposed, and where the relationship between them is left to inference. We can see this in both [9] and [10]. The difference here is that while in [10] the utterance has a “single intonation

contour” (Hopper and Traugott 2003: 180), the utterance in [9] does not. As such, [10] is counted as a single, complex sentence, while [9] is not.

[9] Fort Sumter has been fired on. My regiment leaves at dawn.

[10] You keep smoking those cigarettes, you’re gonna start coughing again.

[11] If you keep smoking those cigarettes, you’re going to start coughing again.

Hopper and Traugott (2003: 179–83)

We can also compare examples [10] and [11]. [10] is a paratactic construction, and the relationship between the two clauses must be inferenced. The two clauses are both nuclei, and each can appear on their own. [11] on the other hand, is hypotactic. A hypotactic construction contains a nucleus and one or more margins, which cannot stand by themselves. These are thus relatively dependent, but are “typically not wholly included within any constituent of the nucleus” (Hopper and Traugott 2003: 177). In [11], the conditional marker *if* has been added to the first clause, thus making it a margin rather than a nucleus. Hopper and Traugott classify conditionals as hypotactic constructions rather than subordinated, as is common in traditional grammars. They explain this by pointing to the research by Matthiesen and Thompson (1988), which shows that in conditionals, “the relationship of dependency is different from that of the prototypical cases of embedding” (Hopper and Traugott 2003: 183). Ohori (2011: 636) notes adverbial clauses are less homogenous than complement and relative clauses, and states that “the border between adverbial subordination and coordination may be viewed as a continuum.” This view seems to be shared by Quirk et al. (2005) who compare the central coordinators *and* and *or* with other linking items like *but* and *for*, subordinators like *if* and *because*, and conjuncts like *yet*, *so* and *nor*. Using six qualities of *and* and *or* as criteria, they find that the differences between coordinators, conjuncts and subordinators are gradual, not discrete.

Hopper and Traugott (2003: 185) theorise that the development of clause linkage markers was probably “motivated by speakers’ desire to be clear and informative, particularly to give directions to hearers for interpreting clauses in terms of the linguistic environment.” In the case of conditional markers, Traugott (1985: 290-92) lists five main “lexico-grammatical domains” from which they have derived, that is (1) words for modality; (2) copula constructions; (3) interrogatives; (4) words that mark something as

known or given; and (5) words that are temporal in origin, especially words that express the temporal notion of ‘for a time’. This last category seems to be the most common structure. For the present study, (1) is of particular interest. These words for modality often have epistemic modality, expressing possibility and doubt. Traugott (1985: 291) gives *suppose* as an example from English, noting that this is often the only conditional marker in Creoles. As for the motivation behind the change from a ‘modal word’ to a conditional marker, Traugott notes that conditionals deal with possibility along the whole range from ‘real’ to ‘counterfactual’. This “may directly motivate the use of epistemic modals in the choice of marker to signal a conditional clause” (Traugott 1985: 293).

Heine and Kuteva (2002: 329) identify four sources of conditional markers, namely (1) copulas⁴; (2) s-questions, i.e. markers of polar (yes-no) questions; (3) *say*; and (3) temporals. Some of these are similar to Traugott’s categories. S-questions, for instance, match with Traugott’s interrogatives, while both works list temporals and copulas as possible sources of conditionals. Interestingly, Heine and Kuteva recognise *say* as a conditional marker. I will return to this in Section 2.3.

2.1.4 Stages and contexts

Although we often tend to focus on the end result of grammaticalisation and contrast it to the earlier, more lexical meaning, we should also be aware of the stages an item goes through before the grammaticalisation process is complete. Diewald (2002) and Heine (2002) present slightly different models of this process. While they both focus on the role of contexts in grammaticalisation, and indeed define the stages in terms of their corresponding contexts, the stages – and contexts – they propose overlap without matching completely. In my analysis, I will make use of Diewald’s model, and as such, it will be given more attention in this section. I will illustrate the stages with examples from Brems and Davidse (2010), who discuss the grammaticalisation of the complex subordinator of purpose *in the hope that* – and variations of it – in terms of Diewald’s model.

⁴ With the term ‘copula’, Heine and Kuteva (2002: 19; footnote) mean “a range of different predicative notions, including identification, classification, specification, and characterization (see Hengeveld 1992). Excluded are existential copulas [...] and locative copulas”

Heine (2002) proposes three types of context of the grammaticalisation process: bridging contexts, switch contexts and conventionalisation. If we add to this a pre-stage in which only the “source meaning” – the earlier, more lexical meaning – is possible, we have four separate stages. We can note that Heine (2002: 86), too, emphasises that the development happens along a continuum, and not by sudden shifts from one discrete stage to another.

As mentioned, Heine’s model has four stages. In the first, the pre-stage, only the source meaning is possible. In the second stage, we find the bridging contexts. In this context, inference makes an interpretation other than the source meaning possible. This new meaning – the “target meaning” – becomes the preferred reading in the context, but can still be cancelled. In other words, the source meaning still cannot be ruled out completely.

The ‘bridging stage’ is followed by the stage characterised by the switch context. Switch contexts are incompatible with “some salient property of the source meaning” (Heine 2002: 85). Because the context is incompatible with some important aspect of the source meaning, that means that the source meaning is inaccessible, and that the target meaning is the only possible interpretation. Unlike the conventional meanings, to which we turn next, the target meaning at the ‘switch stage’ still needs to appear in that specific context in order to be understood correctly.

Conventionalisation, then, entails that the item can be used in contexts other than the one the meaning arose in. This involves analogy, in that the already existing item has spread to new contexts. These new contexts can “violate or contradict the source semantics” (Heine 2002: 85), and it is possible for the source meaning and the target meaning to appear in the same clause.

Diewald (2002) also operates with three successive stages of grammaticalisation, each associated with a different kind of context. These three contexts are the “untypical context”, the “critical context” and the “isolating context” (Diewald 2002: 103-4). For the sake of comparison with Heine, we could also assume a similar pre-stage where only the earlier, more lexical meaning is possible. Nevertheless, I will follow Diewald in referring to stage I, II and III, without amending the numbering of stages to accommodate this potential pre-stage. In [12], Brems and Davidse (2010: 103) illustrate *in the hope that* before the grammaticalisation process has begun. This utterance “lacks any notion, or even

inference, of causality between the SoA [state of affairs] depicted in the matrix clause and that in the clause complementing *that*.”

[12] We look out on bags of balloons in jelly bean colors ready to drop on Democrats who are gathering **in the hope that** this time they can retire the Republicans and win a presidential election for a change.

Brems and Davidse (2010: 103; emphasis original)

In Stage I, the lexical unit expands into contexts in which it has not been used earlier. These new contexts are called “untypical contexts”. Here, it is possible to interpret the new, soon-to-be grammaticalised meaning as a conversational implicature. These untypical contexts do not “display any definite clues as to their preferred reading” (Diewald 2002: 106), and can thus be interpreted in several ways with equal plausibility. This can be seen in [13], where the stockpiling of waste cannot be said to cause people to talk common sense, but where there nevertheless seems to be a weak inference of purpose. It is important to note that Stage I is not actually part of the grammaticalisation process. Rather, this stage sets the preconditions which are necessary for grammaticalisation to begin.

[13] Mr Corbet ... said his mill had stock piled waste **in the hope that** “people might start talking some common sense”.

Brems and Davidse (2010: 104; emphasis original)

Thus, the actual grammaticalisation process begins in Stage II. This stage is associated with the “critical context”, which is “characterized by multiple structural and semantic ambiguities” (Diewald 2002: 103). Unlike Stage I, where there were several untypical contexts into which the lexical unit had spread, Stage II is centred on a single, critical context. Illustrating with the development of German modals, Diewald (2002: 112) writes that the critical context is characterised by “the coincidence – one could even say: the clash – of two verbal forms which both are morphologically or morphosyntactically ambiguous, and therefore cannot mutually disambiguate each other.” In a more general sense we can probably assume that the critical context contains several items, either lexical or grammatical, which are all ambiguous, and thus cannot help clarify each other.

Several interpretations are possible, and there are no clues as to which of these is the most plausible. Faced with this ambiguity, the hearer must make use of inferencing and conversational implicatures to guess at the most plausible interpretation. One of these interpretations would be the new grammatical meaning.

[14] We know that aging skin is dry, for example, so we try to replace moisture **in the hope that** it will look younger.

Brems and Davidse (2010: 104; emphasis original)

[14] illustrates the critical context of *in the hope that*, where two interpretations are possible (Brems and Davidse 2010: 104). In the first interpretation, the cosmetic designers are hoping that replacing moisture will make the skin look younger. In the second, the cosmetic designers are replacing moisture, the purpose being to prevent the skin from aging. In this interpretation, *in the hope that* is a subordinator of purpose.

When we get to the third and last stage of the grammaticalisation process, the new grammatical meaning has become the preferred interpretation in that context. It is now seen as “a semiotic unit independent of the older, more lexical meaning” (Diewald 2002: 104). To this stage belongs the isolating context, in which one meaning is preferred to such an extent that the other meaning is excluded. There would be two such contexts: one for the older, more lexical meaning and one for the new grammatical meaning. Diewald (2002: 106) notes that these isolating contexts are important not only in a diachronic perspective, but also to account for synchronic variation.

[15] This year’s project is digging out the dewponds in the woods **in hopes of** attracting migrant ducks.

Brems and Davidse (2010: 105; emphasis original)

In [15], an isolating context for *in hopes of* is given. Brems and Davidse (2010: 105) note that because the complex subordinators they are studying are in the early stages of grammaticalisation, there are few isolating contexts. This might account for why the subordinator in [15], i.e. *in hopes of*, is different from the structures in the preceding examples. Nevertheless, the example illustrates how the backgrounding of conscious agents makes the interpretation of ‘people hoping’ unlikely, thus favouring the new,

grammaticalised purpose meaning. Although the different form of [15] compared to the illustrations of the earlier stages might be seen as a weakness, it also illustrates an important point about reanalysis and analogy, as discussed in Section 2.1.1. Although reanalysis has taken place, and *in the hope that/in hopes of* in many cases is perceived as a subordinator of purpose, it might take considerable time before this is made clear through the analogous spread to new contexts.

To sum up the development of *in the hope that*, we can compare the structure of the pre-grammaticalisation utterance in [12] with the grammaticalised utterance in [15]. The more lexical meaning has the structure [matrix clause [*in* [*the hope* [*that*-clause]]]], while the more grammatical meaning found in [15] would give the structure [matrix clause [[*in hopes of*] non-finite clause]]. We see that the structure of *in the hope that* has been reanalysed from a prepositional phrase where the noun *hope* is followed by a complement clause, to a prepositional purpose subordinator introducing a clause. With this stage, the grammaticalisation process is completed. It is not possible to reverse the process to an earlier stage, as the two meanings are now seen as separate items.

As the approaches of Heine and Diewald are presented in the same volume, the authors have made a point of comparing their own model to that of the other. Heine (2002: 84–5) notes that his bridging context is roughly equivalent to Diewald’s critical context, and that his switch context is similar to her isolating context. Diewald (2002: 117) does not wholly agree. Her opinion is that Heine’s bridging context would belong partly in both her untypical and critical contexts. She does agree that the switch context is similar to the isolating context, but emphasises that there are in fact two isolating contexts in any given case – one for the more grammatical meaning and one for the more lexical meaning. She concludes that while their models are “principally compatible, both models focus on different aspects of grammaticalization processes” (Diewald 2002: 117). Since both authors agree that the switch and isolating contexts are similar, it follows that Heine traces the grammaticalisation process one step further than Diewald with his conventionalisation stage. Whereas Diewald considers grammaticalisation to be complete when the item in question can appear in contexts where the earlier, more lexical meaning cannot, Heine sees a final stage where the item is ‘released’ from the context it arose in.

Both of these models offer insights into the grammaticalisation process, and could easily serve the purposes of this study. Heine’s first two contexts, bridging and switch,

provide an illustration of the grammaticalisation process which ties nicely in with the mechanisms of reanalysis and analogy as mentioned in 2.1.1. However, as conditional *say* is probably still developing, it is uncertain whether it will have reached Heine's conventionalisation stage. That would potentially leave me with only two stages to operate with. On the other hand, finding Diewald's isolating contexts in my own data analysis would allow me to show that *say* has gone through a grammaticalisation process which left conditional *say* as a separate item from *say* as a *verbum dicendi*. Therefore, as previously mentioned, I will make use of Diewald's context model in the present study. In the present study, therefore, I will make use of Diewald's context model.

2.2 Conditionals

A knowledge of conditionals in general will be necessary in order to see how *say*-conditionals relate to more established conditional constructions. In this section, therefore, I will briefly introduce some essential terms and principles of conditional theory, before moving on to subsections with more specific focus.

A conditional is made up of two parts, the protasis and the apodosis (Comrie 1986, Bybee, Perkins, and Pagliuca 1994, Greenbaum 1996, Dancygier 1999, Huddleston and Pullum 2002). In other approaches, these parts have different names, such as 'conditional clause' and 'matrix clause' (Quirk et al. 1985, Greenbaum and Quirk 1990), 'P-clause' and 'Q-clause' (Declerck and Reed 2001) or '*If*-clause' and '*Then*-clause' (Thompson and Longacre 1985). These two parts each contain a proposition or assumption which are related to each other in a certain way. The relation is often one of causality, as we see in [16], where it is the putting down of the baby that causes the baby to scream. Other relations are also possible, however, as we will see later on.

[16] If you put the baby down, she'll scream.

Quirk et al. (1985: 1088)

Some initial notes on terminology are in order. I will be using the terms protasis *p* and apodosis *q* to refer to the two parts of the conditional. When paraphrasing other works, I will use the terms of the authors, if necessary combined with the term preferred by myself. The contents of the protasis and apodosis are usually called 'propositions' or

‘assumptions’. I use these interchangeably, but favour ‘proposition’. For practical reasons I often refer to ‘the proposition in the protasis’ merely as ‘the protasis’, and likewise for the proposition in the apodosis.

In the remainder of this section, I will first take a look at which kinds of relationships are possible between the protasis and the apodosis, before looking at some formal characteristics of conditionals. After that I will look at two different takes on what constitutes a prototypical conditionals. I will not go into the truth value of conditionals, as this will not be reflected in my analysis. Further, Dancygier and Sweetser (2005: 14) note that a conditional implies more than a coincidence of truth values, and thus more than standard definitions of implicature. As a result of this, I will not mention counterfactuals in the following.

2.2.1 The link between protasis and apodosis

As mentioned above, several relations between the protasis and the apodosis are possible. To account for this, Quirk et al. (1985: 1088–9) introduce the distinction between direct and indirect conditions. The direct conditions are the most central use of conditionals. Here the situation in the matrix clause is directly contingent on that of the conditional clause, or in other words, the truth of the proposition of the matrix clause is a consequence of the fulfilment of the condition in the conditional clause. [16] above is an example of a direct conditional, where the prediction that the baby will scream depends upon the listener putting the baby down. In practice, the opposite prediction is also inferred, so that [17] is the natural implication of [16]. However, this implication can be cancelled by hedging, as we see in [18].

[17] If you don’t put the baby down, she won’t scream.

[18] If you put the baby down, she’ll scream. But she may scream anyway.

Quirk et al. (1985: 1089)

More peripheral uses of conditionals, Quirk et al. (1985) continue, express an indirect condition in which the condition is not related to the situation in the matrix clause. We can see examples of this in [19] and [20], where it is not the situation in the apodosis so

much as the speech act of asserting that is dependent on the listener's permission or understanding.

[19] She's far too considerate, if I may say so.

[20] She and I are just good friends, if you understand me.

Quirk et al. (1985: 1089)

Comrie (1986) takes a slightly different approach. He begins by defining conditionals in logic as the relationship between two propositions, the protasis p and the apodosis q , such that either p and q are both true, or p is false and q is true, or p is false and q is false. Excluded is the possibility of p being true while q is false (Comrie 1986: 78). In [21], for example, if it is Sunday (p is true), the only possibility is for the priest to be in church (q is true). However, if it is any other day (p is not true), it is possible for the priest to either be in church or not be in church without affecting the truth of the conditional as a whole.

[21] If today is Sunday, the priest will be in church.

Comrie (1986: 78)

This view of conditional relationships seems to conflict with that of Quirk et al. above, where a conditional is understood to also imply the opposite prediction. This is because Comrie distinguishes strictly between the meaning of a construction and its interpretation, where the opposite prediction is part of the interpretation, not the meaning. He notes that when interpreting actual utterances in actual contexts, the interpretation of a conditional may be more restrictive than in logic. In [22], for example, the statement is intended as a prohibition and should be interpreted so that the listener will escape punishment if he does not do the action referred to. If the listener believes that the speaker will hit him anyway, the statement loses its prohibitive purpose. The case is similar in [23], where the normal interpretation would be that bringing the umbrella will ensure that the listener does not get wet. However, this is part of the interpretation, not the meaning; if the speaker knows that the umbrella has holes in it and will not protect the listener from becoming wet, she still has not made a false statement, only a very misleading one (Comrie 1986: 78–9). These two examples show that the double interpretation of conditionals is very common. In [24], however, we can see that it is possible to have a conditional where the opposite

interpretation is explicitly made possible as well (= even if you don't buy the stocks, you'll probably lose your money).

[22] If you do that, I'll hit you.

[23] If you go out without the umbrella, you'll get wet.

[24] If you buy those stocks, then you'll lose your money, but of course you'll probably lose your money anyway.

Comrie (1986: 77–8)

Comrie does not mention a distinction between direct and indirect conditions. He does, however, attempt to describe a prototypical conditional. This prototypical conditional has three characteristics (Comrie 1986: 82). First, it has the logical relation between the two propositions, as described above. Second, it must have an identifiable syntactic form. I will return to this in Section 2.2.2 below. Third, the link between the protasis and the apodosis must be causal. Causal is here intended in a broad sense (Comrie 1986: 80–2). The causal relation is from the protasis as cause to the apodosis as effect, but it is not necessarily the literal content of the propositions that is considered, but also the speaker's motivation for making the claim that includes a proposition. Thus in [25], the protasis does not provide the reason for ten not being a prime number, but rather for the speaker's asserting this.

[25] If you want to know, ten isn't a prime number.

Comrie (1986: 81)

A third approach (Dancygier 1999, Dancygier and Sweetser 2005) sees the protasis and apodosis as being related in a mental domain. This approach is situated within the framework of Mental Spaces Theory (Fauconnier 1994). As this approach will be important for my analysis, a quick introduction of the framework is in order. The main idea of Mental Spaces Theory is that in any form of thought, which is then again mediated by language, domains – or mental spaces – are set up, structured and connected (Fauconnier 1994: xxxvii). Mental spaces are “constructs distinct from linguistic structures but built up in any discourse according to guidelines provided by the linguistic expressions” (Fauconnier 1994: 16). The speaker guides the listener into understanding

that the proposition is not to be understood as asserted in Reality,⁵ but in some other space.

[26] In that play, Othello is jealous. He believes that Desdemona is unfaithful.

[27] If Max has gone to the meeting, Max's children are alone.

Fauconnier (1994: 18, 91)

This allows for several possibilities. We can, for instance set up several spaces. In [26], the first mental space M_1 is within the play. Then we see that a second mental space M_2 is set up, where Othello's (mistaken) belief is presented. In this example, the proposition "Desdemona is unfaithful" would then be true in M_2 while false in M_1 . We can also assume properties of a mental space M through Space Optimisation (Fauconnier 1994: 91), which states that unless otherwise is stated, the mental space M inherits the properties of its parent space R . We can see this in [27], where the background assumption "Max has children" is inherited from R to M . This, of course, assumes a situation where the speaker knows that Max has children.

As part of the "guidelines provided by linguistic expressions" we have space-builders: expressions which may establish new mental spaces or refer back to one already introduced in the discourse (Fauconnier 1994: 16–8). Space-builders vary in form, as we see from [26–7] above, where the phrases "in that play", "he believes" and "if" are all space-builders. In Fauconnier's (1994: 89) terms, *If S* sets up a space in which *S* is satisfied.

This brings us back to conditionals. Dancygier and Sweetser (2005) focus on the function of the conditional, and write that the Mental Spaces Theory allows for talking about several kinds or classes of mental spaces. The relations linking the protasis and the apodosis are construed differently depending on the domain in which the conditional relationship applies (Dancygier 1999). Dancygier (1999: 13) argues against the idea of an infinite number of possible relations, which can only be understood against the context of the speaker's or hearer's belief, and instead notes that the relations she has observed

⁵ 'Reality' in the following should not be understood as referring to an objective reality, which could be argued to be unavailable to us. We could, for instance, have utterances like "John believes X, but in reality it's Y", but where the speaker herself is mistaken. Instead, 'reality' here refers to "the speaker's mental representation of reality" (Fauconnier 1994: 15).

fall into several classes based both on the cognitive domain and the context. In their latest work (Dancygier and Sweetser 2005) there are five such domains or spaces: the content space, the speech-act space, the epistemic space, the metalinguistic space and the meta-metaphorical space. These can be illustrated with examples [28–32]:

[28] If I tie my handkerchief around it it'll stick.

Dancygier and Sweetser (2005: 16)

[29] If you need any help, my name is Ann.

Dancygier and Sweetser (2005: 110)

[30] If John went to that party, (then) he was trying to infuriate Miriam.

Dancygier (1999: 7)

[31] That's what we're in business to do, get this cocksucker nailed, if you'll excuse my Greek.

Dancygier and Sweetser (2005: 128)

[32] If the beautiful Golden Gate is the thoroughbred of bridges, the Bay Bridge is the workhorse.

Dancygier and Sweetser (2005: 132)

We can construct spaces which are either representations of content being talked about, or representations of aspects of the speech-act structure itself. Thus we can talk about a main division between the content space and the others. The utterance in [28], where a boy is discussing bandaging a cut with his younger sister, is an example of a content conditional. The speaker sets up a space of mental content, which means that the space is about a possible state of affairs in his world. In this space (which is further filled out by Space Optimisation), the speaker makes the prediction of the handkerchief sticking to the cut. We can note here that he only makes this prediction in the case where he ties the handkerchief around the cut (within the space he set up), and not in any other case (Reality). This kind of conditionals is, thus, predictive. It also follows that the link between the protasis and the apodosis is causal, as the only way of making a prediction is by setting up the preconditions which will lead to the actualisation of the prediction. Content conditionals do not necessarily need to involve direct causation, however. Examples like [33], where it is the enablement of the speaker that is in focus, also concern a possible state of affairs.

[33] If I get funding, I'll go to the conference

Dancygier and Sweetser (2005: 34)

Unlike the content space, conditionals connected in the speech-act space are not concerned with making a prediction. In [29], for instance, the speaker does not imply a causal relation between the hearer needing help and her own name. Rather, the speaker uses the protasis to set up a discourse context – a speech-act space – in which the utterance in the apodosis is appropriate or valid. The content of the apodosis – the name – is not as important as the offering of it, and as such we can say that the focus is on the speech act itself.

In an epistemic conditional the structure follows the speaker's reasoning process. Unlike content conditionals, which must move from cause to effect, epistemic conditionals can move from a known cause to a likely effect, or from a known effect to a possible cause (Dancygier and Sweetser 2005: 17). In [30] we follow the speaker's process from the known effect of John going to the party to what the speaker assumes must have caused him to go – trying to infuriate Miriam. We could also imagine a similar utterance, where the speaker does not know whether or not John went to the party, but does know that he wanted to infuriate Miriam:

[34] If John was trying to infuriate Miriam, he went to that party.

This utterance looks very similar to a content conditional. The difference would be its function, namely that if the speaker is predicting John going to the party, the conditional belongs in the content space. If, as here, the speaker shows a reasoning process where she is “conditionally obliged to conclude” (Dancygier and Sweetser 2005: 17; footnote) with the apodosis, it should be seen as an epistemic conditional.

In utterances such as [31], the focus is on metalinguistic negotiation. The speaker is using the protasis to comment on or question the linguistic choices made in the apodosis. In [31], the speaker does not intend to say that his profession depends upon the hearer excusing his rude language. Rather, he is excusing his language use in the apodosis.

The last class would then be the meta-metaphorical conditionals. These kinds of conditionals express a relationship between metaphorical mappings. In [32], the speaker establishes mappings between horses and bridges, which then allows for a comparison

between the two bridges in those terms. In a way, meta-metaphorical conditionals are a variety of metalinguistic conditionals, in that the hearer must accept the language use (the metaphorical mappings) before being able to make sense of the utterance (Dancygier and Sweetser 2005: 132). It should also be noted that meta-metaphorical conditionals can be combined with an epistemic function. In [35] we see that the protasis sets up metaphorical mappings between cities and organisms, while the apodosis continues with an inference drawn within that mapping.

[35] If public transit is the lifeblood of a dynamic city, Vancouver's in a coma.

Dancygier and Sweetser (2005: 134)

In this section we have looked at three different approaches to the classification of conditionals. There seems to be a great deal of correspondence between the three. The direct conditionals of Quirk et al., for instance, seem equivalent to Dancygier and Sweetser's content conditionals. Likewise, the indirect conditionals correspond to the other categories, i.e. speech-act, epistemic, metalinguistic and meta-metaphorical. There is some overlap between Comrie's approach and that of Dancygier and Sweetser. Comrie's prototypical conditionals corresponds to both content conditionals and speech-act conditionals, as Comrie sees a protasis which explains the speaker's motivation for asserting the apodosis as being causal. The other categories of Dancygier and Sweetser's approach, i.e. epistemic, metalinguistic and meta-metaphorical, on the other hand, would be deemed non-prototypical while following Comrie's approach.

2.2.2 Formal characteristics

As mentioned in Section 2.2.1, Comrie (1986: 82) writes that a conditional needs an identifiable syntactic form. This form should have the encoding of conditionals as its basic function, or at least one of its basic functions. In English, *if p, (then) q* is widely recognised as the standard conditional structure (Quirk et al. 1985, Comrie 1986, Ferguson et al. 1986, Dancygier 1999, Declerck and Reed 2001, Dancygier and Sweetser 2005). Comrie notes that other structures in English can have conditional meaning, even though this is not their basic meaning. He compares the utterances in [36–7] to illustrate,

and shows that the use of *whenever* in temporal clauses such as in [36] can result in an interpretation similar to a conditional.

[36] Whenever he came late, he was scolded.

[37] If he came late, he was scolded.

Comrie (1986: 83)

[38] You so much as touch alcohol and your boss will fire you.

Dancygier and Sweetser (2005: 243)

Dancygier (1999: 1) defines conditionals as complex sentences, composed of a main clause *q* and a subordinate clause *p*. The subordinate clause is introduced by a conjunction, where the least marked of English conditional conjunctions is *if*. For the purposes of this study, we should note that this definition allows for other conjunctions than *if*. Dancygier also mentions some other structures which may have conditional meaning, among them the imperative *and*-conditional illustrated in [38]. This possibility is also mentioned by Comrie (1986: 19–20), who acknowledges the conditional meaning of these structures, while noting that encoding conditional meaning is not their basic function.

[39] There are biscuits on the sideboard if you want them.

[40] #There are biscuits on the sideboard provided that/so long as you want them.

Montolío (2000: 157)⁶

In her study of complex conditional connectives (CCCs) – phrases such as *so long as*, *on the condition that* and *except if* – Montolío (2000)⁷ notes that conditional structures other than *if*-clauses have received little attention from linguists. This, she writes, might be due to the general impression that CCCs are equivalent to each other and completely parallel to *if*-clauses. She shows that this is not the case, for example in [39–40], where she illustrates how CCCs are difficult to use in speech-act conditionals, especially with a postposed protasis. We see this in [40], where the use of a CCC leads us to interpret the

⁶ Montolío uses # to indicate that “the utterance under consideration is grammatically correct but pragmatically inadequate.” (2000: 144)

⁷ Montolío’s data is mostly from Spanish, but she believes that, to a large extent, her findings are valid for both Romance and Germanic languages.

relationship between the protasis and the apodosis as causal. Where the protasis in [39] explains the relevance of the information in the apodosis, the protasis in [40] is seen as the requirement for the actualisation of the apodosis.

Declerck and Reed (2001) list several possible conditional connective devices, including *unless*, *in case*, *lest*, *even if*, *whether...or*, *supposing*, and *assuming*. The last two are of particular interest to the present study. They belong to the same category, where participles⁸ such as *supposing* and *assuming*, exhortatives such as *let's suppose/assume/say* and imperatives *suppose* and *imagine* are included. Imperative *say* is not listed, but would probably belong in this category. The authors note that these protases use the same verb-forms as in *if*-clauses, and have roughly the same meaning (Declerck and Reed 2001: 27). A particularity of conditionals with exhortatives and imperatives as subordinators is that these can have a protasis that is syntactically independent from the apodosis, as illustrated in [41] (Declerck and Reed 2001: 27). It seems clear, then, that it is possible to have conditionals of the pattern *if p, (then) q*, but with a conjunction other than *if*. Even more interesting for my purposes is Poutsma (1926), who notes that *say* in sentences such as [42] functions as a conjunction.

[41] Imagine you were the murderer. Where would you have hidden the knife?

Declerck and Reed (2001: 26)

[42] Say I should succeed at the Bar, is that fame which would satisfy my longings?

Poutsma (1926: 202)

So far I have only mentioned the clause order where the protasis precedes the apodosis, and not without reason. In his Universal 14, Greenberg (1963: 66) states that in conditional statements, the conditional clause, i.e. the protasis, precedes the conclusion, i.e. the apodosis as the normal clause order in all languages. This is also reflected in the statement by Quirk et al. (1985: 1088) that conditional, concessive and contrastive clauses tend to assume initial position, preceding their matrix clauses. Dancygier (1999) also views preposed protasis as the norm, but notes two other structures, *q if p* and *q, if p*. These can be seen in [44–5] below, with *if p, q* in [43] for comparison.

⁸ *Supposing* and *assuming* might more correctly be called ‘former participles’, as they are decategorialised. However, I will follow the terminology of the authors.

- [43] If it stops raining, I'll take you to the park tomorrow.
[44] I'll take you to the park tomorrow if it stops raining.
[45] I'll take you to the park tomorrow, if it stops raining.

Dancygier (1999: 146)

The examples here are predictive (content conditionals), but Dancygier also tested the clause orders for epistemic and speech-act conditionals. Her finding is that the structure without comma is the most restrictive one, as it requires a causal link between *p* and *q*. Thus it is unavailable to epistemic and speech-act conditionals. Regarding metalinguistic conditionals Dancygier notes that a sentence-medial position of the protasis is possible, as the protasis 'wants' to be as close to the language use being commented on as possible. We can see an example of this in [46]. We can also see that a preposed or postposed position in this case would make for a strange utterance.

- [46] My husband, if I can still call him that, hates onion soup.

Dancygier (1999: 152)

Finally, Montolío (2000) notes that the normal order for CCCs seems to be a postposed protasis. The protases in such structures seem to function more as afterthoughts, evaluations or justifications.

One last thing to note regarding the formal structure of conditionals, is the possibility of having covert apodoses. Declerck and Reed (2001: 383-9) list nine different situations where excluding the apodosis is possible. An example from each situation is given in [47–56]. In [47], the apodosis is excluded in order to avoid repetition, and can be retrieved from the preceding context. [48] is an example of a protasis that is purely case-specifying – it specifies the case(s) in which the situation in the apodosis actualises, or in which the apodosis is true (Declerck and Reed 2001: 304). If the apodosis is dropped in these kinds of conditionals, they are interpreted similarly to conditionals starting with *what if*. The speaker does not supply the apodosis to the hearer, but allows him to fill it in himself. [49] expresses a wish of the speaker's, which is either unlikely to be fulfilled or counterfactual. [50] signals that the speaker has doubts about the presupposition expressed in the preceding statement to which the conditional is an answer, in this case,

the supposition that ‘she’ will come. In [51], we see that the conditional is used to express an indignation.

- [47] [I only wanted to help. And I’m not at all sure if I could recognise her.] Perhaps if she dolled herself up and... took that ring out of her nose.
- [48] [I’m sure Lillian saw you take the snuff-box.] Suppose she went to the police?
- [49] If only she would listen to me!
- [50] [“I will be happy if she comes.”] – “IF she comes.”
- [51] [I don’t know what they paid you. You’ve never told me.” – “Exackly! So you know fuck-all about it, right? Look at you! You go out for a bloody teacher and what’s ‘er rates, eh? Four quid an hour? Less?] Christ, if you add up what *she* gets an hour – [all those ‘olidays and everything].”
- [52] Suppose we left in half an hour.
- [53] [“She’ll be here by ten.”] – “If you say so.”
- [54] If you have a lump sum to invest? [Please tick here opposite.]
- [55] Well, if it isn’t the manager himself!
- [56] If you are shot at, this bullet-proof jacket will protect you.

All examples from Declerck and Reed (2001: 384–7)⁹

Conditionals such as in [52] are used for weak manipulation. They express polite directives such as requests, offers or invitations, or a suggestion, as in the example. [53] expresses the speaker’s acquiescence to the listener’s previous utterance, but usually also implies that she has reservations about its contents. [54] is classified by the authors as the protasis being an independent question, although they also note that the following clause could be seen as the missing apodosis. In sentences such as [55] the conditional expresses surprise, and is often preceded by either *why* or *well*. Lastly, [56] is described by the authors as a “pseudo-Q conditional” (Declerck and Reed 2001: 387), meaning that the apparent ‘apodosis’ is not really so; it is in fact another, covert clause which is the true apodosis of the construction (Declerck and Reed 2001: 56). As such, [56] could be interpreted as *if you are shot at you won’t get hurt, as this bullet-proof jacket will protect you*. These kinds of conditionals, however, seem very similar to what Dancygier and Sweetser (2005) call speech-act conditionals, as explained in Section 2.2.1 above.

Some of the kinds of conditionals mentioned above seem to be idiomatic, and are likely to be tied to *if*-conditionals. However, the deletion of the apodosis in especially [47], [48] and [52] seem motivated by pragmatic factors, and could possibly occur in my

⁹ Declerck and Reed use square bracket to separate the conditional from its context.

data. As for [56], I have opted to follow the framework of Dancygier and Sweetser (2005) and will thus classify these as having an overt apodosis, though linked in the speech-act mental space.

2.2.3 Prototypical and non-prototypical conditionals

I have already mentioned the idea of a prototypical conditional in relation to Comrie (1986). To recall, his three requirements of a prototypical conditional were that there must be a logical relationship between p and q (one of material implication), that the conditional must have a identifiable syntactic form, and that the link between p and q must be causal. Comrie uses the term ‘causal’ in a very broad sense here, to also include the speaker’s motivation. Thus, utterances such as [25], here repeated as [57], would be counted as prototypical by Comrie. This classification conflicts with Quirk et al.’s (1985), who would call this an indirect conditional, and as such more peripheral than the central direct conditionals.

[57] If you want to know, ten isn’t a prime number.

Comrie (1986: 81)

Dancygier (1999) also makes use of the idea of a prototypical conditional. To her, it is the predictive, causal conditionals that are prototypical, while other uses such as epistemic and speech-act are seen as derived from that prototypical use. The prototypical conditional is characterised by “three aspects of the construction: ‘sequence of events’ iconicity, non-assertiveness of the protasis, and a causal relation between the protasis and the apodosis” (Dancygier 1999: 187). This reflects much that has already been said. The fact that *if p, (then) q* is more common than other possible clause orders can in part be explained by iconicity, that is, the cause in the protasis precedes the consequence in the apodosis both textually and in reality. Non-assertiveness is linked to the use of *if*, which marks the assumption in its scope as “not assertable in the usual way” (Dancygier 1999: 110). This is usually linked to the speaker’s state of knowledge, with the notable exception of metalinguistic conditionals. A causal relationship between the protasis and the apodosis indicates that only content conditionals can be prototypical.

In addition to the three characteristics of prototypical conditionals, Dancygier (1999: 39) elsewhere adds that a predictive conditional is distinguished from non-predictive ones by the use of systematic backshifting. The definition of backshifting in this case is defined as “every case of language use such that the time marked in the verb phrase is earlier than the time actually referred to” (Dancygier 1999: 37), where ‘earlier’ also includes the use of present tense to refer to the future. Examples [58–60] illustrate the systematic backshifting of predictive conditionals. All three examples could be seen as referring to the future, although the speaker’s belief in the likelihood of the actualisation of the protasis decreases with each example. In [59] and [60], the verb phrases of both the protasis and the apodosis are marked as earlier than the actual time reference. The situation is different in [58], where the apodosis has a verb phrase which matches the future time reference. However, the present tense verb phrase in the protasis also marks a time earlier than the actual time reference, and is thus also an example of backshifting.

[58] If it rains, the match will be canceled.

[59] If it rained, the match would be canceled.

[60] If it had rained, the match would have been canceled.

Dancygier (1999: 25)

Non-predictive conditionals might also make use of backshifting, but these cases are not systematic, and are connected to some other kind of distancing, such as politeness (Dancygier 1999: 38).

There are similarities and differences between Comrie’s and Dancygier’s prototypes. The formal description seems similar, in that both Comrie and Dancygier see *if p, (then) q* as the most common structure. The greatest difference is perhaps the interpretation of ‘causal’. Comrie has interpreted this broadly, while Dancygier restricts the definition and separates utterances such as [29], repeated here as [61], from the causal ones by calling them speech-act conditionals.

[61] If you need any help, my name is Ann.

In the remainder of this study, ‘prototypical’ and ‘non-prototypical’ will refer to Dancygier’s approach. In addition to the three characteristics explicitly stated by Dancygier, we might argue that two others are implicitly present, namely that prototypical conditionals show systematic backshifting, and that they follow the pattern *if p, (then) q*, where *if* refers to any conditional subordinator.

2.3 About *say*

In this section I will attempt to give an overview of the development of *say* using the dictionary entries of its various diachronic forms. This will hopefully offer insights into some tendencies of its development, as well as how it has been able to grammaticalise into a conditional subordinator. I will move from proto-Indo-European (PIE) through Old English (OE) and Middle English (ME) to PDE. While looking at the use of *say* in PDE, I will also make use of Brinton’s (2008) work on *say* in comment clauses.

In Watkins (2011) the PIE root of the modern English *say* is listed as *sek* ‘to say, utter’. From this we might presume that *say* originated as a *verbum dicendi*, i.e. a verb of speaking, and that it was directly linked to a speech act. This immediate link to a specific speech act is an important assumption, as it shows some tendencies in the following developments.

Moving on to OE, PIE *sek* has developed into *secgan* ‘to say’. Here more uses are listed, most of which are still *verba dicendi* (BT *s. v. secgan*). However, it is now possible to use *secgan* in impersonal and absolute ways, as in [62] and [63] respectively. Both of these expansions are very interesting, as they move away from the immediate setting of a speech act in different ways. The impersonal use allows *secgan* to be used with subjects which are incapable of speech, such as the book in [62]. This could suggest that the use of *say* in these cases is intended to focus more on the information conveyed, rather than the actual speech act. The absolute use also focuses less on the speech act, although in a different way. Here, the focus seems to be on a longer discourse, rather than any particular speech act. These are both tendencies we will see more of in PDE.

[62] Hit **segþ** on bōcum, ðæt...
 It says on books that
 ‘It is said about books that ...’

Wulfstan 146, 16 [BT]

[63] Tō ðǣm gesǣlþum, ðe wē **secgaf** ymb, [...]
 For those fortunate, that we say about
 ‘As for those fortunate, who we speak about’

The metres of Alfred, 21, 4 [BT]

In ME there are 20 subentries to *seien* ‘to say’, the development of OE *secgan* (MED *s. v. seien*) . Most of these can be described as verba dicendi, but there is also an increase in uses with impersonal subjects. In addition, two new uses are introduced in this period. The first, which is illustrated in [64], is *say* used as an interjection. The other innovation is for *say* to be used in situations where it does not refer to a specific speech act, nor any conveyance of information by writing or other means. This is illustrated in [65–6]. In [65] the speaker presents the statement as common knowledge by using the phrase *þey sey*. The case is similar in [66], where the speaker presents the statement as the opinion of an authority figure, thus reinforcing its truth-value. We also find proverbial sayings in this period, as illustrated in [67]. Finally, by presenting the statement as hearsay, as in [68], the speaker can safeguard herself from being seen as a liar.

[64] He cryde to hym yn despyte, ‘**Say**, þou felaw yn whyt, Tell me what art þou’
 ‘He cried to him in spite, “Say, you fellow in white, tell me who you are”’

Libeaus Desconus (a1500). 1280 [MED]

[65] Þe whiche oile, þey **sey**..is good for al-manere schabbe.
 ‘The oil of which, they say, is goof for all kinds of scabs.’

Treatise on Horses (a1500). 109/324 [MED]

[66] Of courtaysye, as **saytz** Saynt Poule, Al arn we membrez of Jesu Kryst.
 ‘Of course, as says Saint Paul, we are all members of Jesus Christ’

Pearl (c1400). 457 [MED]

[67] The prouerbe **seith** that for to do synne is mannyssh.
 ‘The proverb says that to sin is human.’

Canterbury Tales (c1390). B.2454 [MED]

[68] I herde **sey** he brake hys neke ab [read: as] he rode in Fraunce.
 ‘I heard it said that he broke his neck as he was riding in France.’

Mankind, The Macro Plays (c1475). 597 [MED]

In the Oxford English Dictionary (OED), the verb *say* is listed with 22 different subentries, some of which are obsolete today (OED *s. v. say v.1 and int.*). Looking again at the degree to which *say* is related to a specific speech act, we find that the majority of these entries either refer to an actual speech act or to the conveyance of information through some other means like writing, body language or symbolism.

[69] Online discourse, **it is said**, is characterised by personal insult, childish mudslinging, pranksterish vandalism and empty threats.

The Guardian 14 Apr. 2007 25/2 [OED]

In addition to these, and to the continuations of the uses mentioned in OE and ME, we find for instance [69], which resembles [65] in modifying the overall statement with *it is said*. This could be done to insert some doubt into the statement, i.e. as a hedging or face-saving technique – or to strengthen the argument by presenting it as a widely known fact. This ties in with Heine and Kuteva's (2002: 265) observation that *say* is the source of evidential markers in English and Taiwanese, among others.

The dictionary entries indicate that *say* has over time developed uses which are more removed from the immediate speech act context. The impersonal use is a removal, in that it involves no oral speech act. Interjections and evidential markers are even further removed, in that they do not refer to any specific situation at all, and the focus is on the nature of the statement rather than the act of conveying it.

[70] **Say** there were actual vultures on his tail.

[71] Keep the reconstructed stuff down to, **say**, 5% of the whole.

[72] I need some fruit, **say**, apples.

Brinton (2008: 92 (footnote); my emphasis)

[73] “**Say**, isn't that – ” Lance started, but Buck answered before the question was even asked.

[74] “**I say**, hospital life certainly suits you!”

[75] Jump, **I say**, and be done with it.

[76] these appeared to be grounded on scientific or quasi-scientific notions. **I say** quasi-scientific, because...

All other examples from Brinton (2008: 74-5; emphases original)

In her work concerning comment clauses, Brinton (2008) identifies six uses of *say* other than as a *verbum dicendi*. Although her focus is on *say* used in comment clauses, she provides an overview of the six uses, and dates their first attestations in English.

The first of these uses, which Brinton terms *say*¹, has the meaning of ‘suppose, assume’. It occurs clause-initially and alternates with *let’s say* and sometimes *what say*. An example can be seen in [70], where the *say*-clause, i.e. the protasis, occurs on its own. As mentioned in Section 2.2.2, this is possible in certain contexts. As this is the focus of the present study, I might add here that this use is an example of the attested cross-linguistic tendency for verbs of saying to become conditional markers (Heine and Kuteva 2002: 265). This is attested in several languages, among them in Ga, a Kwa language of the Niger-Congo family, where the verb *kɛɛ* ‘say’ is the source of conditional subordinator *kɛ*. This is illustrated in [77].

[77] máha o níyení kɛ oba.
 give:1:SG:FUT you food (say) you:come
 ‘I’ll give you some food if/when you come.’

Heine and Kuteva (2002: 265)

Returning to Brinton’s six uses of *say*, *say*² and *say*³ are illustrated in [71–2]. These uses are similar to one another, but have slightly different meanings. Where *say*² signifies ‘about, approximately’, *say*³ means ‘for example, suppose, let’s imagine’. In [71], the speaker does not necessarily mean *exactly* 5%, but rather something around that number. Likewise, the speaker in [72] intends *apples* to be an example of the mentioned set, i.e. ‘fruit’. This does not imply that apples are absolutely necessary; if the hearer were to respond by giving the speaker bananas and grapes, that would be a perfectly valid interpretation. *Say*² and *say*³ usually occur in clause-medial or final position.

The last three uses are less relevant for this study, as their meaning is very different from that of *say*¹. *Say*⁴, as seen in [73], introduces a question. It usually occurs initially or following a particle. *Say*⁵ is typically sentence-initial. It is an interjection, functioning either to express the speaker’s strong emotion, or to evoke the hearer’s attention. An example of the former can be seen in [74], where the speaker signals her surprise with *I say*. Lastly, there is *say*⁶, which is found in the phrase *I say* and is either an emphatic marker, as in [75], or is used to recapitulate a word or phrase, as in [76].

In her survey of PDE corpora, Brinton (2008: 79) found that *say*³ is by far the most common form, with 59% of the occurrences. *Say*¹, on the other hand, is relatively infrequent with 8%. The six uses have developed from two types of sources. *Say*¹⁻⁴ “originate as second-person imperative verbs taking a clausal complement”, whereas *say*⁵⁻⁶ have developed “from matrix *I say* (shortened to *say*) with a nominal or clausal complement” (Brinton 2008: 88).

The development of *say*¹⁻⁴ can be seen in Figure 2.1. Here we see that *say*¹ has developed directly from imperative *say* as a matrix, i.e. main clause with a following clausal complement. It is “fossilized in form and reduced syntactically from a matrix clause to a subordinating conjunction (rather like *if*)” (Brinton 2008: 88). The other three uses have gone through an additional stage where *say* first changed from a matrix clause to a parenthetical, after which the meanings *say*²⁻⁴ developed. This indicates that these three are later developments. The change in *say*¹ appears to have happened in the sixteenth century. Brinton acknowledges the OED’s first occurrence in 1596, and notes that her earliest findings are from approximately the same time period. Moreover, she notes that there are no occurrences of *let’s say* or *let us say* in her corpora of that period, indicating that this is a later development. This is interesting, as it makes it unlikely that conditional *say* has developed from an earlier *let’s say*, which Brinton (2008: 82) notes has been proposed as a possible path of development.

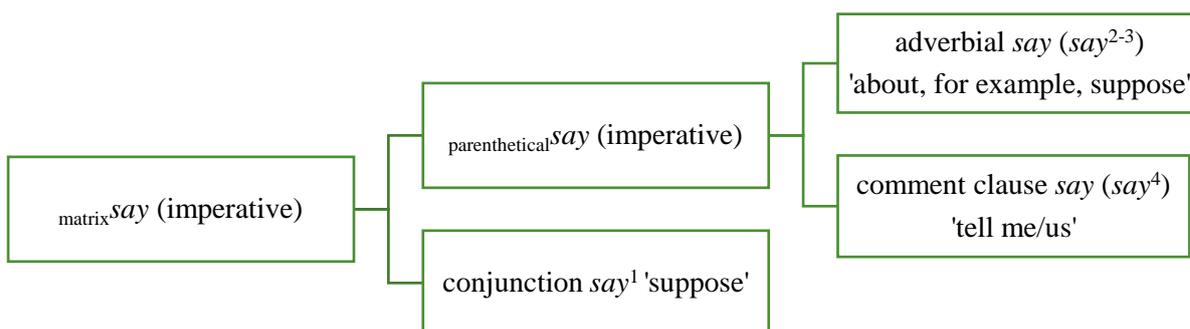


Figure 2.1: The development of conditional *say*, adapted from Brinton (2008: 89)

Finally, Brinton notes that the development of all the six forms of *say* “exhibit most of the hallmarks of grammaticalization” (Brinton 2008: 90). Regarding *say*¹ (Brinton 2008: 91–2), she notes that *say* is fixed in the imperative form, and has lost verbal behaviours

such as the ability to take complements or be modified by an adverbial. We also see the shift from a major to a minor words class, i.e. from a verb to a conjunction. There is also evidence of semantic weakening, in that the *verbum dicendi* meaning of *say*, i.e. ‘speak, utter words’, is weakened. Brinton notes, however, that this weakening is accompanied by the strengthening of other, more abstract meanings. There is also evidence of “pragmatic strengthening” (Brinton 2008: 91). The meaning of *say* as a *verbum dicendi* invites inferences which can be illustrated as ‘speak’ > ‘suppose’. This was possible because “in an evidential sense, something that is said cannot be taken as fact, but must be assumed or supposed to be” (Brinton 2008: 92). These implicatures have become strengthened and conventionalised, so that *say*¹ came to signify ‘suppose’ rather than ‘speak’.

To sum up this section, we have seen that *say* has gone through a development which allows it to be used in a wider range of contexts. We might loosely relate this to the “unspecific expansion of the distribution of the lexical unit in question to contexts in which it had not been used before” associated with Diwald’s (2002: 103) Stage I, which precedes the onset of the grammaticalisation process. Brinton (2008) lists conditional *say* as one of six uses in PDE, and suggests the reasoning ‘say’ > ‘suppose’ as motivation for the change. Cross-linguistically, Heine and Kuteva (2002: 325) note that equivalents of *say* are the source of both conditional subordinators and evidential markers, as well as six other uses.

3. Methods and results

In the following chapter I present the corpora and the methods I have used in order to extract relevant tokens, before looking at my results. Since my knowledge of *say*-conditional is for the most part based on literature on similar structures such as *let's say* and *suppose*, the natural place to start is with a survey of the use of *say*-conditional in PDE. Collecting and analysing data in a reversed chronological makes sense for two reasons. First, knowledge about how *say*-conditional is used in PDE will provide a foundation from which tracing the construction back in time will be easier. Second, the corpora from before 1810 have provided very few relevant results. As such, it is more practical to relate the few results from the earlier period to the more substantial results from after 1810, than to attempt to establish categories based on the earlier findings and use that as a point of departure.

The results section is structured so that the results from COCA and COHA¹⁰ are presented in a contrastive manner in Section 3.2. There are two main reasons for doing so. COCA and COHA are created by the same person, are part-of-speech tagged using the same criteria, and are accessed using the same interface. Thus, the results from these corpora are entirely comparable. Additionally, COCA and COHA are the largest corpora, and therefore the ones where quantitative analysis of the results is possible. Results from earlier time periods are presented in Section 3.3.

When presenting my results, I will wherever possible do so without the use of decimals. However, since *say*-conditional is a relatively infrequent construction, it will sometimes be necessary to include decimals.

¹⁰ The corpora are presented in full in Section 3.1.1.

3.1 Methods

3.1.1 The corpora

In this section I introduce the corpora used to collect data, together with relevant information such as size and which time period they cover. I also include in square brackets the shorter name I will refer to them by throughout this study.

The Corpus of Contemporary American English [COCA] (Davies 2008– .) consists of 520 million words, with 20 million from each year covering the period 1990–2015.¹¹ The genres represented are spoken, fiction, popular magazines, newspapers and academic texts. The corpus has been tagged, allowing for part-of-speech searches.

The Corpus of Historical American English [COHA] (Davies 2010– .) consists of more than 400 million words from the period 1810–2009. The genres represented are fiction, popular magazines, newspapers (from the 1860s onwards) and non-fiction books. Within each decade, the corpus has an equal number of words in each genre. However, the total number of words increases with each decade, so that there are 1.2 million words from the 1810s, and 29.4 million words from the 2000s. Thus the corpus is balanced by genre decade by decade, but the results from, for example, ‘Popular magazines’ in the 1850s are not directly comparable to the results from ‘Popular magazines’ in the 1960s. To avoid overlap with COCA, I have limited my searches to the period 1810–1980. That lowers the total number of words to approximately 350 million. As with COCA, COHA is also tagged and allows for part-of-speech searches.

ARCHER: A Representative Corpus of Historical English Registers [ARCHER] (Biber and Finegan 2013) contains over 3 million words from the years 1600–1999. I have only made use of the data from 1600–1799, which consists of approximately 1.1 million words. The corpus is part-of-speech tagged, but this proved unnecessary for my purposes.

The Penn-Helsinki Parsed Corpus of Early Modern English [PPCEME] (Kroch, Santorini, and Delfs 2004) consists of over 1.7 million words. These are collected from 448 text samples from the years 1500–1720. The corpus is both parsed and part-of-speech tagged.

¹¹ COCA was updated in December 2017 with 20 million words from both 2016 and 2017. This was after my data collection.

The Lampeter Corpus of Early Modern English Tracts [Lampeter] (Siemund and Schmied 2001) consists of 1.1 million words from the years 1640–1740. The creators have aimed to create a balanced corpus, selecting tracts and pamphlets from six domains: religion, politics, economy and trade, science, law, and miscellaneous.

In addition to these, I have made use of the electronically available *The Oxford Shakespeare* [Shakespeare] (1916), which contains all of Shakespeare's plays. The dating of the plays is uncertain, and there are many editions. Because the language used in the plays can have changed from the original version to the First Folio, on which this collection seems to have been based, I have followed the dating provided in *The Oxford Shakespeare*. This corpus amounts to approximately 836 000 words.

3.1.2 Searches in COCA and COHA

As mentioned in the introduction to this chapter, COCA and COHA are accessed using the same interface. In order to be able to compare the results from these two corpora in detail, I have performed the same searches in each corpus.

The first step to collecting data was to identify effective ways of searching for *say* as a conditional subordinator. As a subordinator, it will be followed by a subordinate clause. As it marks the beginning of a new clause, it is also logical to assume that it may be preceded by a punctuation mark or a coordinating conjunction.

With these criteria the question arises how a clause is structured. Naturally, there are a multitude of different clause structures, but due to time restrictions I have focused on a few of the more basic structures, namely a subject followed by a verb, alternatively with an adverb in between. The subject can be realised as either a pronoun, a noun or a noun with a preceding article. These variables create six possible clause structures. In addition to the clause structure itself, I also wanted to see whether it was more common for a *that*-clause or a zero-clause to follow conditional *say*. Combined with *say* being preceded by either a coordinating conjunction or a punctuation mark, this leaves 24 searches. All the variables are presented in Figure 3.1. It should be noted that it is impossible to search for *say* only in the imperative as then all matching forms will be given, including the infinitive and the present tense form of *say*. Thus the data collection involves not only identifying conditional uses of *say*, but also separating the imperatives from the non-imperatives.

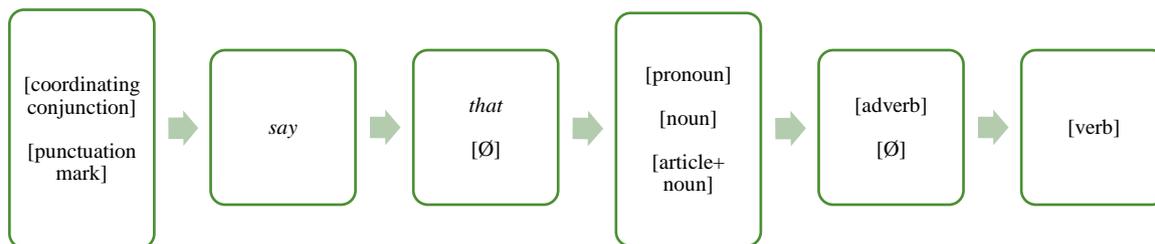


Figure 3.1: Variables for searches in COCA and COHA

Were it not for the time constraints, I could of course have searched for a greater variety of clause structures. All cases where the subject is modified by an adjective, for example, will be irretrievable by these searches. In addition, elements such as adverbs and interjections could alter the basic structure of the clause, and would thus cause it to fall outside the limits of my searches. However, as the objective was not an exhaustive quantitative study of conditional *say* but rather a study of a representative sample of instances, this procedure has yielded sufficient results.

3.1.3 Searches in corpora covering earlier time periods

The corpora covering periods earlier than those of COCA and COHA were of considerably smaller size. Because of this, and because some of the corpora were not part-of-speech tagged, it was unnecessary and impractical to use elaborate search methods. In addition, I would risk missing tokens which were conditional, but did not follow the clause structures searched for in COCA and COHA. For the ARCHER, PPCEME, Lampeter and Shakespeare corpora, therefore, I have simply searched for *say* and gone through the search results manually. To ease this process, I have made use of the corpus analysis tool AntConc while searching some of the corpora.

3.2 Results in COCA (1990–2015) and COHA (1810–1980)

3.2.1 Overall results

The overall results from both COCA and COHA can be seen in Table 3.1. Looking at the results from COCA in isolation first, we see that the searches yielded 5593 instances of *say*. As mentioned in Section 3.1.2, it is not possible to search for *say* exclusively in the imperative. After excluding those instances which were non-imperative, there were 777

instances of *say* to be analysed. I read through these in order to identify the ones which were conditional and which were imperative, and found 504 conditional tokens and 273 imperative tokens. The conditional constitutes 65% of the relevant tokens and, in terms of normalised frequency, 1 occurrence per million words. For a more accurate comparison with the results from COHA, I have also calculated the normalised frequency without ‘Spoken’. This comes to 0.9 occurrences per million words.

Table 3.1: Overall results from COCA and COHA

	COCA	COHA
Hits	5593	2366
Relevant tokens	777	879
Imperatives	273 (35%)	732 (83%)
Conditionals	504 (65%)	147 (17%)

Note: the percentages of *say*-conditional are calculated according to the number of relevant tokens, not the total number of hits.

In COHA, the total number of hits was 2366, of which 879 were either imperative or conditional. We can note here that while the total number of hits is considerably higher in COCA, the number of imperatives in COHA is much higher, perhaps indicating that imperatives were more common in the time period covered by COHA. Despite this, only 147, or 17%, of the tokens were conditional subordinators. The normalised frequency for this corpus is at 0.4 occurrences per million words. This is considerably lower than that of COCA, even without ‘Spoken’.

As COHA extends over a period of 180 years, a more nuanced look at the development of conditional *say* might be useful. Figure 3.2 shows the development of conditional *say* over periods of 45 years, namely 1810–54, 1855–99, 1900–44 and 1945–89. By using normalised frequency, it is possible to see how frequent *say*-conditionals are in relation to the total number of words in each time period. The total for each time period is 46 million, 86 million, 107 million and 109 million words, from the earliest period to the latest. Figure 3.2 shows that the frequency of *say*-conditionals within COHA fluctuates, and that there is an increase in the last time period, 1945–89, so that the frequency here is nearly equal to that of COCA without ‘Spoken’, i.e. 0.7 occurrences per

million words. Interestingly, we also see that *say*-conditionals were slightly more common in the first period than in the two following.

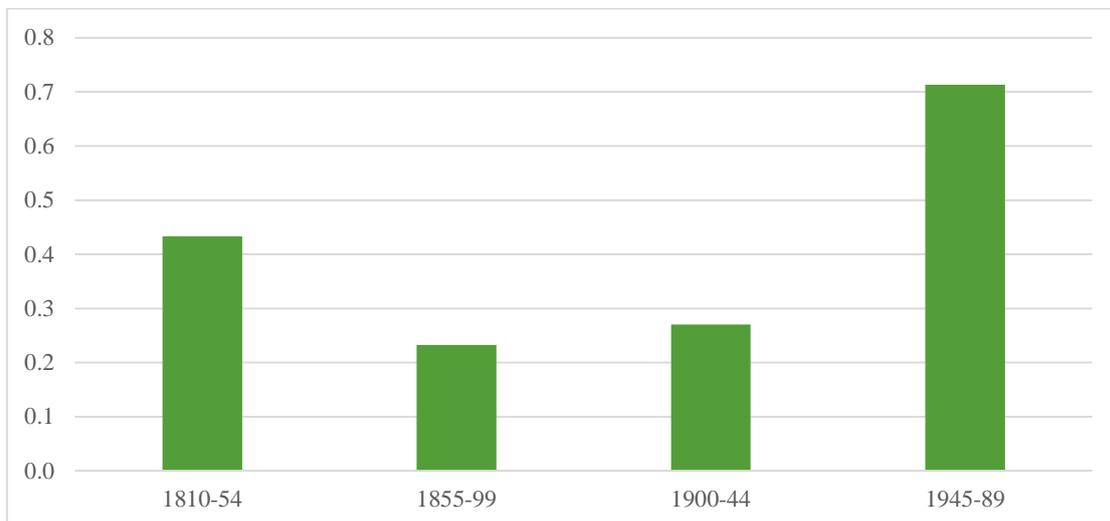


Figure 3.2: Normalised frequency of *say*-conditional per million words in COHA

While looking at the periods of 45 years is useful when it comes to looking for overall trends, especially with such low-frequency structures as *say*-conditional, it might be useful to look at the development per decade. As we see from Figure 3.3, the frequency of *say*-conditionals in fact fluctuates throughout the period 1810–1939 before rising from the 1940s and into the 2000s. The clear exception to this pattern is the first decade, 1810–19, which has a remarkably high frequency. This is due to the same author using two *say*-conditionals in immediate proximity to one another, as shown in [1]. As the total number of words for the decade 1810–9 is only 1.2 million words, this causes a higher frequency than in any other decade.

- [1] And what if it should conclude that there is something inconsistent? Is it or is it not the province and duty of the executive to enforce the same laws? **Say** it is, you destroy the right to judge. **Say** it is not, you cancel the duty to preserve peace, till war is declared.

The Federalist, on the new Constitution (1817) [COHA]

Disregarding the decade 1810–19, the highest frequency of *say*-conditionals is in 1970–9 with 0.9 occurrences per million words. *Say*-conditionals are, then, not very frequent in

the English language. There is an increase in frequency from 1820 to 1859, after which there is a sudden decrease. There are low frequencies throughout 1860–1939, before rising into the 1940s. After this there are some fluctuations, but the frequency is generally higher than in the earlier decades – except 1810–9.

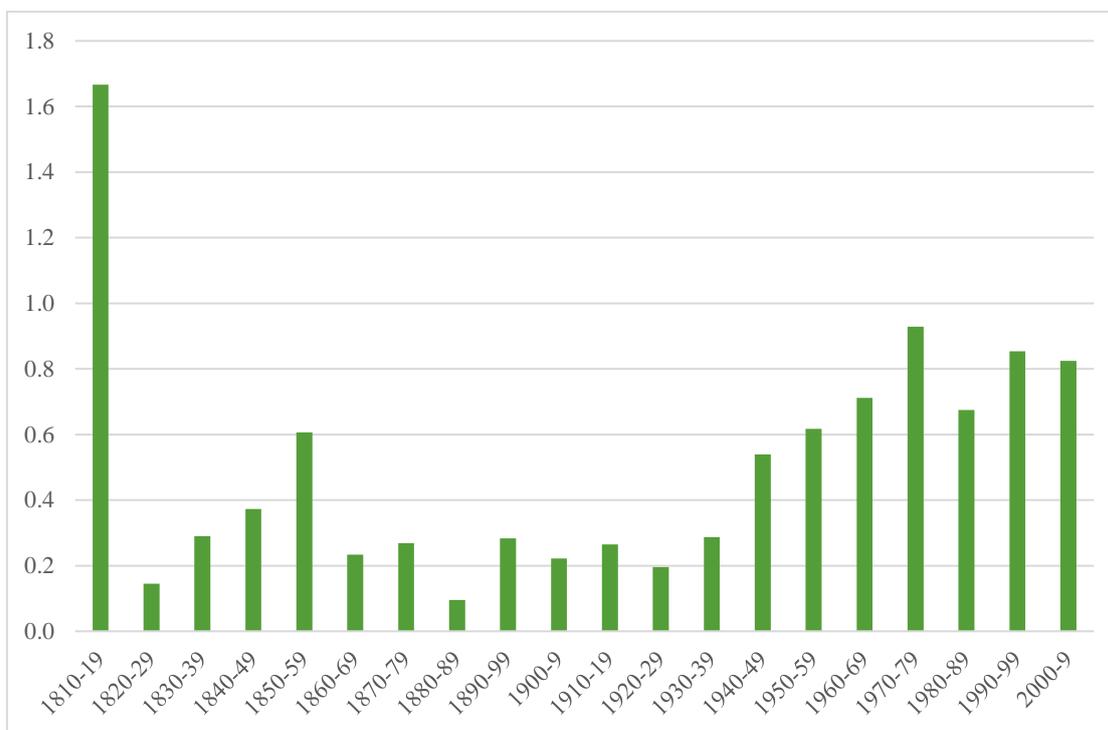


Figure 3.3: Normalised frequency of say-conditional per million words by decade from 1810–2009 in COHA and COCA

Note: The frequencies from COCA (1990–99 and 2000–9) are calculated without ‘Spoken’.

3.2.2 Genre

The distribution of tokens across genres is presented in the tables below. The genres in COCA and COHA are, as mentioned, largely comparable, with the exception of ‘Spoken’, which is only represented in COCA. Thus, Tables 3.2 and 3.3 show the distribution in COCA with and without ‘Spoken’, respectively. ‘Spoken’ accounts for 24% of the tokens. ‘Popular magazines’ is the most frequent genre with 38% of the tokens, while ‘Academic journals’ accounts for only 4%. Without ‘Spoken’, on the other hand, ‘Popular magazines’ is by far the most common genre with half the tokens. What this is indicative of is that *say*-conditionals are associated with a less formal register. This is in keeping

with language change being initiated in everyday language, and spreading to more formal registers once established, as mentioned in Section 2.1.2.

Table 3.2: *Distribution of say-conditional across genres in COCA*

	N	%
Spoken	123	24
Academic journals	22	4
Newspapers	58	12
Popular magazines	190	38
Fiction	111	22

Table 3.3: *Distribution of say-conditional across genres without ‘Spoken’ in COCA*

	N	%
Academic journals	22	6
Newspapers	58	15
Popular magazines	190	50
Fiction	111	29

As we see from Table 3.4, there is a clear tendency for *say*-conditionals to occur in ‘Fiction’ in COHA. This genre accounts for 71% of the tokens, while ‘Non-fiction’ and ‘Popular magazines’ have 12% and 17%, respectively. ‘Newspapers’, on the other hand, is by far the least common genre, with only 1% of the tokens.

Table 3.4: *Distribution of say-conditional across genres in COHA*

	N	%
Non-fiction	17	12
Newspapers	1	1
Popular magazines	25	17
Fiction	104	71

3.2.3 *That*-/zero clauses

Table 3.5 shows the distribution of *that*-clauses and zero-clauses following conditional *say* in COCA and COHA. We see that the majority of tokens in COCA, i.e. 98%, have

zero-clauses. Only 8 tokens, i.e. 2%, have a following *that*-clause. The situation is more varied in COHA, where 14% of the tokens have *that*-clauses. Zero-clauses are still clearly preferred, however, with 86% of the tokens.

Table 3.5: *Distribution of that-clauses and zero-clauses following conditional say in COCA and COHA*

	<i>That</i> -clauses	Zero-clauses	Total
COCA	8 (2%)	496 (98%)	504
COHA	20 (14%)	127 (86%)	147

I have also presented the distribution of *that*-clauses and zero-clauses following imperative *say*, so as to compare the distribution in *say*-conditional to the structure from which it originates. These findings are presented in Table 3.6.

Table 3.6: *Distribution of that-clauses and zero-clauses following imperative say in COCA and COHA*

	<i>That</i> -clauses	Zero-clauses	Total
COCA	29 (11%)	244 (89%)	273
COHA	243 (33%)	489 (67%)	732

The general tendencies shown in this table mirrors those of Table 3.5. There is a decrease in the use of *that*-clauses following imperative *say*, and a corresponding increase in the use of zero-clauses. In fact, the decrease of *that*-clauses is sharper following imperatives and following conditionals, with 22 versus 12 percentage points.

3.2.4 Individual searches in COCA

The results from the individual searches in COCA can be seen below. For ease of reading, I have divided these into two tables, so that Table 3.7 shows the results for *that*-clauses, while Table 3.8 shows the results for zero-clauses. As such, the results show the same overall tendency as discussed in Section 3.2.3, namely that *that*-clauses are much less frequently combined with conditional *say*. Thus, Table 3.7 contains very few tokens. The 8 tokens which have *that*-clauses represent three different structures, namely , *say that*

PRN V; , *say that N V* and , *say that ART N V*.¹² In two of these structures, namely , *say that ART N V* and , *say that N V*, all occurrences are conditional, leading to a frequency of 100% in these categories. However, considering the low number of tokens, this might not mean much. All of these three structures, we might note, are structures where *say*-conditional is preceded by a punctuation mark. There are no cases where *say*-conditional is preceded by a coordinating conjunction such as *and*. There are also several structures where neither imperative nor conditional *say* occurs at all, which could mean either that these structures are uncommon for imperatives, or that they are uncommon in a general sense.

Table 3.7: Distribution of say-conditional with that-clauses by individual searches in COCA

	Imperatives	Conditionals	Total
, say that PRN ADV V	-	-	-
, say that PRN V	25 (89%)	3 (11%)	28
, say that N ADV V	-	-	-
, say that N V	-	2 (100%)	2
, say that ART N ADV V	-	-	-
, say that ART N V	-	3 (100%)	3
and say that PRN ADV V	1 (100%)	-	1
and say that PRN V	3 (100%)	-	3
and say that N ADV V	-	-	-
and say that N V	-	-	-
and say that ART N ADV V	-	-	-
and say that ART N V	-	-	-

Note: , = any punctuation mark; PRN = pronoun; ADV = adverb; V = verb; N = noun; ART = article; and = any coordinating conjunction¹³

Considerably more *say*-conditionals are combined with zero-clauses, as Table 3.8 shows. Discounting structures with less than 5 instances in total, *say*-conditional is by far the most frequent in , *say ART N V* with 96% of the tokens. Three other structures are also

¹² The symbols are explained in the notes beneath the tables below.

¹³ I have opted to change the symbols used in COCA and COHA into symbols which are easier to follow. For instance, the corpus search symbol for punctuation would be `_y*`, but is here represented by a comma.

very common, namely , *say PRN V* with 67%, **and say PRN V** with 49% and , *say N V* with 42%.

Table 3.8: Distribution of *say*-conditional with zero-clauses by individual searches in COCA

	Imperative	Conditional	Total
, <i>say PRN ADV V</i>	-	3 (100%)	3
, <i>say PRN V</i>	193 (33%)	398 (67%)	591
, <i>say N ADV V</i>	-	-	-
, <i>say N V</i>	7 (58%)	5 (42%)	12
, <i>say ART N ADV V</i>	-	2 (100%)	2
, <i>say ART N V</i>	2 (4%)	46 (96%)	48
and <i>say PRN ADV V</i>	2 (67%)	1 (33%)	3
and <i>say PRN V</i>	39 (51%)	38 (49%)	77
and <i>say N ADV V</i>	-	-	-
and <i>say N V</i>	-	1 (100%)	1
and <i>say ART N ADV V</i>	-	-	-
and <i>say ART N V</i>	1 (33%)	2 (67%)	3

Note: , = any punctuation mark; PRN = pronoun; ADV = adverb; V = verb; N = noun; ART = article; and = any coordinating conjunction

We can also note that here as well, three structures have no tokens whatsoever, namely , *say N ADV V*, **and say N ADV V** and **and say ART N ADV V**. The equivalent structures with *that*-clauses have also, as is clear from Table 3.7, no tokens. All of these structures contain an adverb, which makes them part of the least ‘basic’ clause structures in this study. As such, they might simply be too infrequent to yield many results.

3.2.5 Individual searches in COHA

As with COCA, the results from COHA are presented in two tables, so that Table 3.9 shows the results of the *say*-conditionals with *that*-clauses, while Table 3.10 shows the results for zero-clauses. As noted in Section 3.2.3, the distribution of *that*/zero is more even in COHA than in COCA, although there is still a clear preference for zero-clauses. Table 3.9 shows that the two most frequent structures for *that*-clauses are , *say that ART N V* and **and say that ART N V**, with 46% and 33% respectively. However, both these

structures are fairly uncommon, and so one cannot read much into these results. Other than this, Table 3.9 shows low values, as did the equivalent for COCA, i.e. Table 3.7.

Table 3.9: Distribution of say-conditional with that-clauses by individual searches in COHA

	Imperative	Conditional	Total
, say that PRN ADV V	-	-	-
, say that PRN V	154 (94%)	10 (6%)	164
, say that N ADV V	-	-	-
, say that N V	4 (80%)	1 (20%)	5
, say that ART N ADV V	-	-	-
, say that ART N V	7 (54%)	6 (46%)	13
and say that PRN ADV V	1 (100%)	-	1
and say that PRN V	74 (97%)	2 (3%)	76
and say that N ADV V	-	-	-
and say that N V	1 (100%)	-	1
and say that ART N ADV V	-	-	-
and say that ART N V	2 (67%)	1 (33%)	3

Note: , = any punctuation mark; PRN = pronoun; ADV = adverb; V = verb; N = noun; ART = article; and = any coordinating conjunction

More substantial findings are, once again, to be found with the zero-clauses. These are presented in Table 3.10. The most common structure here is , **say ART N V** with 63%. The structures , **say PRN ADV V** and , **say PRN V** are nearly equally frequent with 25% and 22%, respectively. It should be noted, however, that the former's percentage is based on a considerably fewer tokens.

Overall, what we see in Table 3.10 is lower frequencies for all the structures, as compared to the equivalent results from COCA presented in Table 3.8. All the structures which have tokens in COHA, have increased in COCA. In addition, four structures, i.e. , **say N V**; , **say ART N ADV V**; **and say PRN ADV V**; and **and say N V**, which are all without conditional tokens in Table 3.10, i.e. COHA, have tokens in Table 3.8, i.e. COCA. This increase in less common structures might indicate analogy to various clause types following the reanalysis of *say*.

Table 3.10: Distribution of say-conditional with zero-clauses by individual searches in COHA

	Imperative	Conditional	Total
, say PRN ADV V	6 (75%)	2 (25%)	8
, say PRN V	322 (78%)	92 (22%)	414
, say N ADV V	-	-	-
, say N V	3 (100%)	-	3
, say ART N ADV V	-	-	-
, say ART N V	7 (37%)	12 (63%)	19
and say PRN ADV V	-	-	-
and say PRN V	136 (87%)	20 (13%)	156
and say N ADV V	-	-	-
and say N V	3 (100%)	-	3
and say ART N ADV V	-	-	-
and say ART N V	12 (92%)	1 (8%)	13

Note: , = any punctuation mark; PRN = pronoun; ADV = adverb; V = verb; N = noun; ART = article; and = any coordinating conjunction

3.2.6 Formal structure

As stated in Section 2.2.2, *if p, (then) q* is widely recognised as the standard structure of conditionals. Other structures are also recognised, but do not have conditionality as their basic meaning. This structure, then, will be my point of departure, as it will be interesting to see to what extent *say*-conditionals behave like the prototypical *if*-conditionals. As mentioned in Section 2.2.2, Declerck and Reed (2001) remarks on how conditionals with markers such as *let's suppose/assume/say* and *suppose* and *imagine* can have protases that are syntactically independent from the apodosis, and indeed, that the apodosis can be merely implied and have no overt presence in the sentence.

Based on the prototypical *if p, (then) q* and my observations while collecting data, I have distinguished four variants of the formal structure of *say*-conditionals as follows: *say p, (then) q*; *say p. (then) Q*; *say p.*; *say o. if p, (then) q*. The first of these, *say p, (then) q*, is identical to the prototypical *if*-conditional, except that the subordinator *if* is replaced with *say*, as exemplified in [2]. In order for these results to be entirely comparable to *if*-conditionals, no deviations from the pattern were allowed in this formal variant. This meant I had to exclude 3 tokens where the protasis and the apodosis were

separated by a comma, but where the apodoses were introduced by *and*. We can see an example of this in [3]. These are listed as ‘Other’ in Table 3.11. Likewise, there was one token which had to be excluded for similar reasons in COHA. I have, however, allowed for punctuation other than a comma, so long as it is not sentence-final. That means that tokens such as [4] are also included in this category.

[2] But **say** he did die, what would happen?

New England Review (2006) [COCA]

[3] Let’s start off with a pretty simple question here. **Say** they do the weapons inspections, *and then what?*

CNN Iraq (2002) [COCA]

[4] But **say** you did want to sell it - it would be sold as a working forest.

Associated Press (2005) [COCA]

The second variant, *say p. (then) Q*, is similar to the first, but has a protasis which is syntactically independent from the apodosis. We can see an example of this in [5]. I have allowed for some variation of the pattern here, as long as the protasis is syntactically independent from the following apodosis. First of all, I have allowed for the protasis to extend over several sentences, as in [6]. I have found that some of these can be quite long, building up an entire scenario before concluding with the apodosis. I have also allowed for the apodosis to be introduced by other words or phrases than *then*, such as *so*, *therefore*, *in that case*, *and then* and *so then*. This seems like a natural effect of dividing the protasis and the apodosis into separate sentences, as the speaker might feel the need to recapitulate and enhance the connection between the two parts. An example of this can be seen in [7]. As a side note, I have also observed restructurings of the apodosis as a response to having an independent protasis. In [8] we see that the apodosis has a preposed adverbial *because you’re afraid of spiders*.

[5] **Say** you’re lucky and it’s one of the few filoviruses we have antiserum for. Then we can treat you and you’ll probably live.

Outbreak (1995) [COCA]

[6] But **say** you're unlucky. And you get one of those filoviruses we don't have an antiserum for, which is most of them. There's no medicine, no cure, nothing we can do to help you.

Outbreak (1995) [COCA]

[7] **Say** he was working a con and had a sucker on the hook [,] that might keep him here until he made his score. *In that case*, would he keep on using Lucas Zeller's ID?

Betrayers: a Nameless Detective novel (2010) [COCA]

[8] **Say** you have arachnophobia [sic] and you spot a tiny spider. Because you're terrified of spiders, you may "see" it as gigantic.

Cosmopolitan (2006) [COCA]

The third variant is *say p.*, in which the protasis stands alone. In order to keep the focus of this study solely on conditionals, I have excluded many tokens. There are, for example, many instances where *say* seems to introduce a hypothetical scenario, but where there is no conditional link between the *say*-clause and what follows, and no covert apodosis. An example of this can be seen in [9].

[9] **Say** you want to move your finger: The thought probably originates in the prefrontal cortex. It then spreads through the neural networks to the back part of the frontal lobe, which then coordinates the activity through the cerebellum and the basal ganglia. All this eventually leads down through the spinal cord and results in the movement of your finger.

Esquire (2010) [COCA]

Erring on the side of caution, I have included only the tokens in which there is a clearly implied apodosis. These seem to correspond particularly to two of the cases listed in Section 2.2.2 as being able to exclude the apodosis: purely case-specifying protases and conditionals used for weak manipulation. Of these, the case-specifying type is the most common, and are exemplified in [10–2] below. In each of these, *say* can be interpreted as equal to *what if*, and indeed elicit responses from the respective hearers which indicate that it has been interpreted as such. Additionally, it could also be that [11] deletes the apodosis because it is retrievable from the preceding context. Examples [13–4] show instances of weak manipulation, where a *say*-protasis is used as a suggestion. This is perhaps clearest in [13], where the listener responds to the suggestion with an approval (*All right; do that*).

[10] Mr. DONALDSON: But **say** the president makes a decision to commit the United States to war. Sen. MITCHELL: Then he must come to Congress and ask for a formal declaration.

ABC_Brinkley (1990) [COCA]

[11] “OK. Let’s change the scenario. **Say** you could name her.” [...] “I don’t know.” She brought her hand to her mouth, tenderly bit her thumb. “Heather?”

The Hudson Review (2002) [COCA]

[12] “All right. And this pilot, Charles Pettikin. **Say** he won’t take us?” “Make him. One way or another.

Whirlwind (1986) [COHA]

[13] The sergeant’s grin widened. “High-pressure fire hose, one at the head of each escalator, and a couple more that can be dragged over from other outlets. **Say** we put two men on each hose, lying down at the head of the escalators. And we got plenty of firearms; we can arm some of these clerks, up here –” “All right; do that.

Null-ABC (1953) [COHA]

[14] “Now, I’m going to propose something that may surprise you. It may seem bizarre, but give it a chance, hear me out, Gleason, because I think it will work. Are you ready?” Gleason nodded. “**Say** you kidnap my wife.”

New Yorker (2006) [COCA]

The last variant is *say o. if p, (then) q*, where the *say*-clause does not have its own apodosis, but precedes another, full conditional structure. I have allowed for the following conditional to be introduced by either *if, say, let’s assume/suppose/say* – as in [15–7], respectively – or even to be an imperative *and*-conditional, as in [18]. I have chosen to do so because it is not the form of the following conditional that is important, but rather that the preceding *say*-clause should be associated with its conditional meaning. I have also allowed, as with *say p. (then) Q*, for the ‘*say*-protasis’ to extend over several sentences, as we see is the case in [17].

[15] **Say** you got a milk bottle. Now, **if** I asked you to fill that milk bottle with cotton, you could do that in ten minutes.

New Yorker (1973) [COHA]

[16] **Say** you’re a teen-aged athlete feeling pressure to perform. You hear about a substance you could take that would help. And **say** this dietary supplement was perfectly legal, hadn’t been banned by any sport and was taken by a number of professional athletes. Would you take it?

NPR_ATCW (2001) [COCA]

[17] **Say** you’re in your attic one day sorting through old copies of Life magazine and dusting off your brother’s high school yearbook. Then in a particularly neglected

corner you come across a bound book and discover a diary, a stranger's, maybe, someone who lived in the house a century before, a thrifty housewife or a glamorous socialite, who managed to record all the mundane, scandalous, intimate and now historical details of her life; or maybe a relative, someone you barely knew or thought you knew, at least until the secrets began spill out. So **let's assume** that you went ahead and read it. Then what?

NPR_TalkNation (2002) [COCA]

[18] **Say** you are in the country; in some high land of lakes. **Take** almost any path you please, **and** ten to one it carries you down in a dale, and leaves you there by a pool in the stream.

Moby Dick, or, the whale (1851) [COHA]

These, then, are the four formal variants I have operated with. Tables 3.11 and 3.12 show the results obtained from COCA and COHA, respectively.

Table 3.11: Distribution of say-conditional across formal variants in COCA

	N	%
<i>Say p, (then) q</i>	55	11
<i>Say p. (then) Q</i>	333	66
<i>Say p.</i>	25	5
<i>Say o. If p, (then) q</i>	86	17
Preposed apodosis	2	< 1
Other	3	1

Looking at COCA in Table 3.11 first, we see that *say p. (then) Q* is by far the most common structure, with 66% of the tokens. *say p, (then) q* and *say o. if p, (then) q* are roughly of the same frequency with 11% and 17%, respectively. It is surprising to see a relatively low frequency in *say p, (then) q*, as this reveals that *say*-conditionals largely do not follow the prototypical pattern we know from *if*-conditionals. Least common is *say p.* with only 5% of the tokens.

Table 3.12 shows that the distribution in COHA is slightly more even. *say p. (then) Q* is still the most frequent variant, although with a lower percentage: 52%. Here *say p, (then) q* is the second most frequent with 27%, while *say p.* and *say o. If p, (then) q* each accounts for 10% of the tokens. The decrease in *say p, (then) q* is quite unexpected, as it indicates a movement away from the prototypical form, rather than towards it.

Table 3.12: Distribution of say-conditional across formal variants in COHA

	N	%
<i>Say p, (then) q</i>	39	27
<i>Say p. (then) Q</i>	77	52
<i>Say p.</i>	15	10
<i>Say o. If p, (then) q</i>	14	10
Preposed apodosis	1	1
Other	1	1

It should also be mentioned that in both COCA and COHA, a preposed apodosis is very uncommon. In both corpora, these account for less than 1%. As stated in Section 2.2.2, it is the norm for the protasis to precede the apodosis, and so these findings are not surprising.

Figure 3.4 shows the development of the formal variants across the four time periods of 45 years. As in Section 3.2.1 I have calculated the normalised frequency of each variant, so as to compensate for the increasing number of words for each decade in COHA.

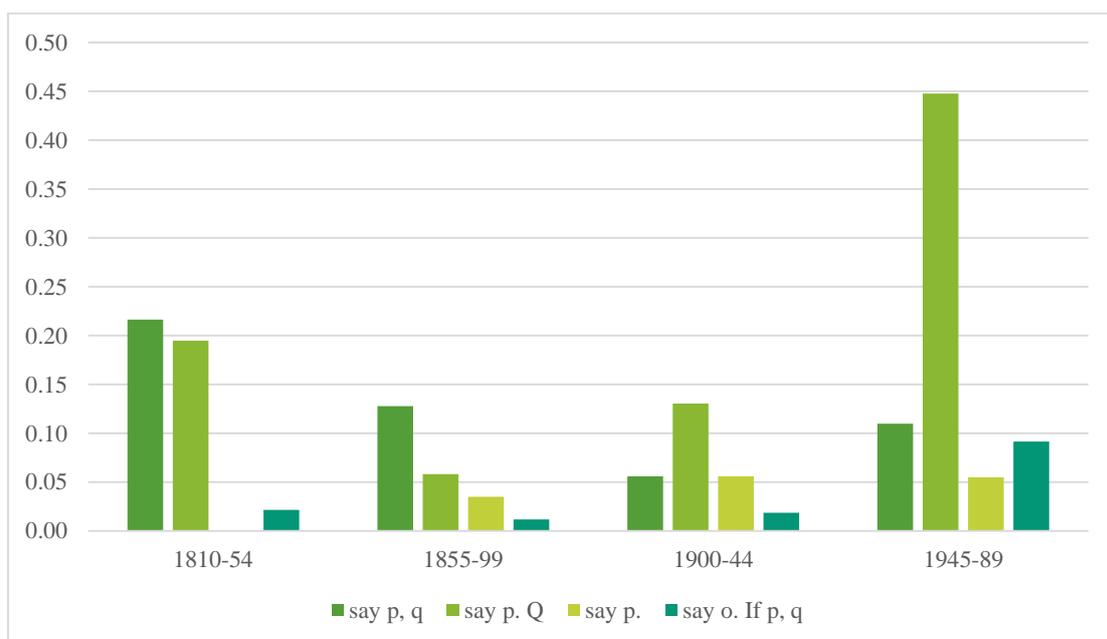


Figure 3.4: Distribution of say-conditional by formal variants across time periods in COHA

The figure shows the same general development as we have seen earlier in Figure 3.1; there is a small decrease from the first time period and into the next two, before an increase in the last time period. What we can see here, however, is that these changes in frequency are also accompanied by shifts in which formal variant is the most common. Of all the variants, *say o. if p, (then) q* seems to be the only with a steady, although modest, growth. *Say p. (then) Q* goes through the most dramatic changes, ranging from being very frequent in the first time period, decreasing in frequency in the two middle time periods, and then increasing rapidly after 1945. *Say p, (then) q* shows the same general trend, but with a much smaller increase in the last time period. Comparing only the first and last time periods, we see that this variant was twice as common in 1810–54 than in 1945–89. This adds in with the development we saw while comparing the results from COHA and COCA, and indicates that the prototypical formal variant is on the decline. Lastly, *say p.* develops from having no presence in the first time period into having a modest, although even frequency in the following three time periods. The clear trend here then, is that *say p. (then) Q* is becoming the dominant formal variant for *say*-conditionals.

3.2.7 Mental spaces

As explained in Section 2.2.1, there are five possible mental spaces in which the protasis and the apodosis of a conditional may be linked: the content space, the speech-act space, the epistemic space, the metalinguistic space and the meta-metaphorical space. Of these, the content space, which implies a causal, predictive relationship, is seen as the ‘original’, from which the other meanings are derived.

Table 3.13: *Distribution of say-conditional across mental spaces in COCA*

	N	%
Content	187	48
Speech-act	151	38
Epistemic	53	14
Meta-metaphorical	2	1

Table 3.13 shows the distribution of *say*-conditionals across mental spaces. It is clear the content space is the most common, with 48% of the tokens. The speech-act space is also frequent, with 38%. The epistemic conditionals are less frequent, but still have 14% of the tokens. Meta-metaphorical conditionals are present, but are very rare, with only 1%. An example of each mental space can be seen in [19–22].

[19] **Say** you come out winners, what will you do with the prize?

The Prince of India — Volume 02 (1893) [COHA]

[20] **Say** you're exhausted and you have to sit and rest for a while, but you desperately need to stay awake. Just put a razor in your mouth. You start to doze, sure, but part of you remains conscious. You never slip too far under.

Literary Review (1999) [COCA]

[21] **Say** the guy got hit in his arm or got his arm chopped off and he's spitting up blood -- what happens is you create a huge amount of pressure inside your chest, kind of like a retching cough, which can cause microtears in your mucosal lining, which can actually bleed.

Esquire (2009) [COCA]

[22] You and I aren't exact opposites. Like, **say** I'm Death and Darkness. Then you should be Immortality and Light instead of just some poor fool scrambling to stay afloat.

Fantasy & Science Fiction (1997) [COCA]

In [19], the apodosis is framed as a question. This does not mean that it cannot be a content conditional. Rather than actually making a prediction, the apodosis here is requesting a prediction. In [20], which is an example of a speech-act conditional, the protasis sets up a situation in which the advice in the apodosis is relevant or useful – although it is difficult to imagine a situation in which putting a razor in one's mouth would be advisable. The epistemic conditional in [21] follows a line of reasoning from a known effect (a man missing an arm and spitting blood) to its likely cause (microtears in his mucosal lining). As the last example, [22] shows a meta-metaphorical conditional. As mentioned in Section 2.2.1, it is possible for these conditionals to combine with epistemic ones. We can see this here, where the metaphorical mappings from the speaker to 'Death' and 'Darkness' enables her to conclude that the listener should be equally mapped to 'Immortality' and 'Light'.

Table 3.14: Distribution of say-conditional across mental spaces in COHA

	N	%
Content	68	58
Speech-act	33	28
Epistemic	17	14

Table 3.14 shows the distribution of mental spaces in COHA. In comparison to Table 3.13, there is a category missing; there are no instances of meta-metaphorical conditionals in COHA. Once again, the content conditionals are the most numerous, with 58%. Speech-act conditionals make up 28% of the tokens, while epistemic conditionals are the least frequent with 14%. In comparison with COCA, we see that content conditionals have declined over time from 58% to 48%. Speech-act conditionals have increased from 28% to 38%, while epistemic conditionals have remained stable at 14%. The increase of speech-act conditionals and meta-metaphorical conditionals at the expense of content conditionals is in keeping with Dancygier’s theory that the causal, predictive meaning (i.e. the one represented in content conditionals) is the most basic, and from which other meanings derive. Assuming that *say*-conditionals originated with a causal meaning, what we now see is a spread into new meaning relations.

Table 3.15: Distribution of say-conditional across mental spaces and formal variants in COCA

	<i>Say p, (then) q</i>		<i>Say p. (then) Q</i>	
	N	%	N	%
Content	34	62	150	45
Speech act	11	20	138	41
Epistemic	10	18	43	13
Meta-metaphorical	0	-	2	1

Table 3.15 shows the distribution of *say*-conditionals across both mental spaces and formal variants in COCA. As the formal variants *say p.* and *say o. if p, (then) q* are both without apodoses, these are not included. I have also excluded the smaller variants ‘Other’ and ‘Preposed apodosis’, as the results from these are unlikely to provide much insight. By looking for potential ‘preferences’ of a formal tendency towards a mental space, we

can see whether there are any correlations between form and meaning. In *say p, (then) q*, the clear preference is for content conditionals, with 62%. The other two other kinds of conditionals represented, i.e. speech-act conditionals and epistemic conditionals, account for approximately 20% each. In *say p. (then) Q*, on the other hand, the distribution is much more even between content conditionals and speech-act conditionals, with 45% and 41%, respectively. Epistemic conditionals are slightly less frequent in this formal variant than in *say p, (then) q*, with 13%. In addition to these, it also here we find the two meta-metaphorical conditionals.

Table 3.16: Distribution of say-conditional across mental spaces and formal variants in COHA

	Say p, (then) q		Say p. (then) Q	
	N	%	N	%
Content	27	69	40	52
Speech-act	12	31	20	26
Epistemic	0	-	17	22

In Table 3.16, we see the equivalent results from COHA. In this corpus, *say p, (then) q* does not have any instances of epistemic conditionals. As a result, the frequencies of content conditionals and speech-act conditionals are higher, respectively 69% and 31%. When it comes to *say p. (then) Q*, the preference for content conditionals is much clearer than in COCA, with 52%. The other two types, i.e. speech-act conditionals and epistemic conditionals, are evenly distributed with respectively 26% and 22%.

3.2.8 Isolating and critical contexts

In order to show that *say* has grammaticalised into a conditional subordinator, I should be able to find isolating contexts for both the conditional and the ‘lexical’ use of *say*. The tables in 3.2.1 give a good indication of where to start. We have seen that when *say* is followed by *that*, it is very unlikely to be a conditional subordinator. Unfortunately, most of the structures do not seem very common for *say* as an imperative verbum dicendi either. The only structure with more than 10 relevant tokens is , *say that PRN V*, which has 25 instances of imperative *say*. However, there are also 3 conditional tokens in this

category, and as such, the context does not favour the *verbum dicendi* to the exclusion of the conditional.

Another possible isolating context presented itself during my data collection. It appears that the phrase *you will* or its contracted form *you'll* cannot combine with *say* to become a conditional. Searching for these phrases in COCA, I found that all instances of *say you'll*, *say you will*, *say that you'll* and *say that you will* are *verba dicendi*. There were 70 such instances. Examples of all these phrases can be seen in [23–6]. This might very well be an indication of an isolating context for *say* as a *verbum dicendi*.

[23] “Don’t say ‘okay’ to me. **Say you’ll** stop smoking.” “I don’t smoke,” I told her. We were whispering. “Say it.” “I’ll stop smoking.” “Good. Thank you.”

Atlantic (2013) [COCA]

[24] Oh, stop it. **Say you will** come. **Say you will** have dinner at the Adamses’ with me. Tell me you’ll come to dinner next Saturday night. It will be nice. I promise!

Being with him (2008) [COCA]

[25] If I should go, forget me never. Please **say that you’ll** remember me.

ABC_PrimeTime (1993) [COCA]

[26] So don’t hide behind this charlatan, David Keen. Just **say that you will** support all of law enforcement, all the cops, and policemen and women who are afraid of being shot by these kinds of terrible bullets that don’t kill deer and aren’t used for target practice, but simply are aimed at killing policemen.

ABC_Nightline (1994) [COCA]

Finding potential isolating contexts for conditional *say* is more challenging. With a normalised frequency of 1 occurrence per million words, *say*-conditionals are not very common to begin with. Even though some of the searches presented in Section 3.2.4 show high frequencies for *say*-conditional, most of these have less than 5 tokens. This is the case for all searches which are exclusively conditional. The structure , **say ART N V** is notable for having 96% conditional tokens. This shows a near exclusion of *say* as a *verbum dicendi*, although not yet entirely so.

[27] However he runs things, it's gotta be quiet. Let him hide upstairs in the office. **Say** he's the janitor, I don't give a shit. But, please, whatever job he takes, make sure it's something quiet.

Casino (1995) [COCA]

Another potential isolating context is the form *say p, (then) q*. The primary function of this form is to encode conditional meaning. Therefore, one might expect that this contexts favours the conditional meaning to the exclusion of the *verbum dicendi* meaning. However, it is still possible for *say* to have *verbum dicendi* meaning in such contexts, as is illustrated in [27]. Even though the form of this construction is *say p, q*, and even though the relationship between the clauses seem conditional-like, *say* is a *verbum dicendi*.

As it is the normal order for the *say*-conditional, as with most other conditionals, for the protasis to precede the apodosis, we might assume that *say* is often sentence-initial. By refining one of the most common search structures, *, say PRN V*, to include only those cases where *say* is preceded by a period, we find the results given in Table 3.17.

Table 3.17: Top three results for *. say PRN V* in COCA

	Imperative	Conditional	Total
<i>. say you're</i>	11 (21%)	42 (79%)	53
<i>. say you have</i>	2 (10%)	18 (90%)	20
<i>. say you want</i>	2 (10%)	19 (90%)	21

Here, the top three results are given, with the number of conditionals and imperatives for each structure. The percentages of conditionals are, as we see, higher than those where other kinds of punctuation are included, with 79% for the first phrase and 90% for the two others. This shows that these phrases clearly favour the conditional meaning. However, an isolating context should prefer one meaning to the exclusion of all others, and that is not the case here. Further, a problem with excluding all other punctuation marks is that the *say*-clause is often preceded by an interjection such as *well*, which would make *say* non-initial in the sentence, while the protasis still precedes the apodosis. We might very well have to conclude that there are no completely isolating contexts for conditional *say*, and that the grammaticalisation process is still ongoing.

3.3 Results before 1810

When moving on to the period before 1810, the corpora become considerable smaller. Combined, the corpora ARCHER, PPCEME, Lampeter and Shakespeare make up approximately 4.8 million words. Comparing with COHA, which consists of 350 million

words and in which there were 147 tokens, this is a small data sample. As a result, the number of tokens is considerably lower than in the later corpora, and no quantitative analysis will be attempted. The results from the four earlier corpora are listed in Table 3.18.

Table 3.18: Results before 1810, by corpus

	Corpus size	Years	Tokens
PPCEME	1.7 mill.	1500–1720	1
Shakespeare	836 000	1594–1623	25
ARCHER	1.2 mill.	1600–1799	1
Lampeter	1.1 mill.	1640–1740	1

As we see, the corpora PPCEME, ARCHER, and Lampeter provided moderate findings. All three instances are shown below. [28], which is from 1630, seems like a concessive, and could perhaps be interpreted similarly to *even if*. In [29] we see an example of a conditional with a covert apodosis. This seems very similar to [33] in Section 2.2.2, repeated here as [30], where the apodosis is deleted to avoid repetition of the preceding discourse, and in which the speaker signals her doubts about the presupposition expressed in the preceding discourse. In [29], the speaker apologises for having thought the recipient unkind, but at the same time doubts that it was only his perception that was faulty. He did not *think* her unkind; she *was* unkind. [31] is another example of a conditional with an apodosis framed like a question. We might note that both conditionals with an overt apodosis, i.e. [28] and [31], are of the formal variant *say p, (then) q*. Additionally, I would like to point out before moving on that even with these three tokens only, the normalised frequency of PPCEME, ARCHER and Lampeter combined exceeds that of COHA overall, with 0.6 to 0.4 occurrences per million words.

[28] Well grant all this,
say now his deeds are blacke,
 Pray what serues marriage, but to call him backe,

Middleton, Thomas. *A chaste maid in Cheapside*. 1630. [PPCEME]

[29] If I have been more Sensibly Grievd at what I thot Unkind in my Sophi (**Say** it was Only that I thot So), if I Took Fire more Than another Would have done, it Was because I Loved you More Than Ever any Ever any lovd, or will or can love you (he that has you Excepted).

Daniel Defoe, *To MRS. SOPHIA DEFOE BAKER*, 1729 [ARCHER]

[30] ["I will be happy if she comes."] – "IF she comes."

Declerck and Reed (2001: 385)

[31] Nay, if some sorts of Stone could bee but found out in some other parts, what might it arise unto? Nay, **say** that either Marl, Chalk, or Lime, or some other fat Earth could be found in some other parts where they are wanting, how much would it enrich those parts?

Walter Blith. *The English improver improved or the Survey of husbandry surveyed* [...]. 1653. [Lampeter]

The overwhelming majority of tokens can be found in the plays of Shakespeare. There are 26 instances of *say*-conditionals here, some of which can be seen in the examples below. In [32] we see two instances of conditional *say*, both of which can be interpreted as *what if*. The first of these is a complete conditional, with the apodosis *what then*. The second has a covert apodosis, which is deleted in order to avoid repetition, or perhaps because the protasis is of the purely case-specifying type, where the speaker leaves it to the hearer to 'fill in' the missing apodosis. Another of these can be seen in [33], which is interpreted as a *what if*-question and answered accordingly.

[32] Well; and **say** that Marcius
Return me, as Cominius is return'd,
Unheard; what then?
But as a discontented friend, grief-shot
With his unkindness? **say**'t be so?

Coriolanus (1623) [Shakespeare]

[33] K. RICHARD
Say, that I did all this for love of her?

Q. ELIZABETH
Nay, then indeed, she cannot choose but hate thee,
Having bought love with such a bloody spoil.

The Tragedy of King Richard the Third (1597) [Shakespeare]

The excerpt in [34] is notable for having three consecutive *say*-conditionals followed by two *if*-conditionals. The change of subordinator does not seem to elicit any change in the

structure or tone of the following clause, and might indicate that *say*-conditionals and *if*-conditionals are seen as interchangeable, at least in these types of context. We can also note that in the apodosis of the second *say*-conditional (*Say that she frown...*), *say* also appears as a *verbum dicendi* (*I'll say she looks as clear...*), although with a preceding auxiliary. It might be worth remembering the characteristic of Heine's conventionalisation stage, as mentioned in Section 2.1.4. In this stage, which is the final of the grammaticalisation process in Heine's model, the source meaning and target meaning can occur in the same clause. This is not applicable here, as the *verbum dicendi* *say* is in the infinitive and not the imperative. In fact, it is difficult to imagine an imperative and a conditional *say* occurring in the same clause at all, as they both tend to occupy the first position. Even so, the occurrence of the two in close proximity to one another suggests that the speaker (or rather, author) is not afraid of any confusion between the two: one is clearly a subordinator while the other is clearly a verb.

[34] I pray you do; I will attend her here,
 [...]
 And woo her with some spirit when she comes.
Say that she rail; why then I'll tell her plain
 She sings as sweetly as a nightingale:
Say that she frown; I'll say she looks as clear
 As morning roses newly wash'd with dew:
Say she be mute and will not speak a word;
 Then I'll commend her volubility,
 And say she uttereth piercing eloquence:
If she do bid me pack; I'll give her thanks,
 As though she bid me stay by her a week:
If she deny to wed; I'll crave the day
 When I shall ask the banns, and when be married.
 But here she comes; and now, Petruchio, speak.

The Taming of the Shrew (1623) [Shakespeare]

There are also a few instances where the phrase *say if* is used, as in [35–6]. In these, the role of *say* is different. In [34], *say* could be seen as introducing the question in the apodosis of the conditional (*Who takes offence at that would make me glad?*) in the same way as described in Section 2.3. In [35], on the other hand, *say* seems more of an interjection, emphasising the following statement.

[35] PERICLES

Then, as you are as virtuous as fair,
 Resolve your angry father, if my tongue
 Did e'er solicit, or my hand subscribe
 To any syllable that made love to you.

THAISA

Why, sir, **say if** you had,
 Who takes offence at that would make me glad?

Pericles Prince of Tyre (1609) [Shakespeare]

[36] DAUPHIN

Say, if my father render fair return,
 It is against my will; for I desire
 Nothing but odds with England: to that end,
 As matching to his youth and vanity,
 I did present him with the Paris balls.

The Life of King Henry the Fifth (1600) [Shakespeare]

Table 3.19: Results from Shakespeare's plays, by genre

	Number of plays	Tokens
Comedies	17	14
Tragedies	10	4
Historical plays	10	7

The plays of Shakespeare are divided into three groups, or genres: the comedies, the tragedies and the historical plays. Table 3.19 shows the distribution of tokens across the genres. As we can see, the comedies account for over half the tokens with 14, while the tragedies have 4, and the historical plays have 7.

Table 3.20: Distribution of say-conditional across mental spaces in earlier corpora

	N	%
Content	15	60
Speech act	8	32
Epistemic	1	4
Meta-metaphorical	1	4

Table 3.20 shows the distribution of *say*-conditionals across mental spaces. 28 tokens are perhaps not enough to read much into these results, but they do seem to confirm the

tendencies we have seen in later developments. Content conditionals have 60% of the tokens, which is higher than in both COHA and COCA, where the percentages are at 48% and 58%, respectively. Speech-act conditionals account for 32% of the tokens, whereas epistemic and meta-metaphorical conditionals account for 4% each. The presence of a meta-metaphorical conditional here seems to indicate that the assumed innovation in COCA is not so innovative after all. As we can see in [36], this conditional seems to be combining the meta-metaphorical and the speech-act mental spaces.

[37] Ah! what a shame! ah, what a fault were this.
Say, Warwick was our anchor; what of that?
And Montague our top-mast; what of him?
Our slaughter'd friends the tackles; what of these?
Why, is not Oxford here another anchor?
And Somerset, another goodly mast?
The friends of France our shrouds and tacklings?
And, though unskilful, why not Ned and I
For once allow'd the skilful pilot's charge?

The Third Part of King Henry the Sixth (1594) [Shakespeare]

I have not looked at the earlier results in relation to formal variants in a systematic sense. As most tokens are from Shakespeare, and therefore in verse, the findings might not be comparable to findings in COCA and COHA. However, we could note that the 4 tokens in which there is no apodosis, makes up a higher percentage, i.e. 14%, than the formal variant *say p.* has in either COCA or COHA.

4. Discussion

In this chapter, I will compare and discuss the findings presented in Chapter 3. The discussion is divided into two sections. In Section 4.1 I will relate my findings to the characteristics of a prototypical conditional, as described by Dancygier (1999: 187). In Section 4.2, I look at how my findings relate to various points of grammaticalisation theory.

4.1 Findings in relation to the prototypical conditional

As mentioned in Section 2.2.3, the prototypical conditional, in Dancygier's sense, is characterised by “‘sequence of events’ iconicity, non-assertiveness of the protasis, and a causal relation between the protasis and the apodosis” (Dancygier 1999: 187) To this I added that the conditional should be of the form *if p, (then) q*, where *if* symbolises any conditional subordinator.

As for the first of these characteristics, almost all instances of *say*-conditional have a protasis which precedes the apodosis. Tables 3.11 and 3.12 show that only two tokens from COCA and 1 token from COHA have postposed protases, and thus do not show ‘sequence of events’ iconicity. So far, then, *say*-conditional fits with the description of a prototypical conditional. However, this characteristic is not very restrictive. As stated in Section 2.2.2, the clause order where the protasis precedes the apodosis is seen as the universal norm. Although there are exceptions to this, such as the ‘complex conditional connectives’ described by Montolío (2000), it not surprising that the majority of *say*-conditionals should adhere to this norm.

The non-assertiveness of the protasis is linked both to the use of a conditional subordinator and to the use of systematic backshifting. It seems clear, both from accounts such as Brinton's (2008) and Poutsma's (1926) and from my own data collection, that *say* can and is used as a conditional subordinator. In [1], this use of *say* is emphasised by the speaker introducing the *say*-conditional with *I'll give you an example* and following it with a clause introduced by *imagine*, which both suggest that what is said should not be interpreted as asserted in ‘reality’.

- [1] Yeah. I'll give you an **example**. **Say** you have a live show on TV and this very famous chef is making a meal on the show. With our technology and our device, you'll be able to smell the meal that he's doing through your TV live. Or **imagine** you go out to a holiday, to India, and you have a video camera. You're taking a picture of the market in Bombay and you're going to catch the smell, too. You'll come back home, you'll stick your video in your VCR and you will see, hear and smell your experience on your trip.

NPR_Morning (2000) [COCA]

Although I have not provided any definite numbers relating to backshifting, I might add that my general observation while analysing data is that backshifting as described in Section 2.2.3 is rare in *say*-conditional. For the potentially prototypical conditionals, namely content conditionals with a syntactically dependent protasis, for example, less than half of the tokens show systematic backshifting.

The last of Dancygier's characteristics concerns the relation between the protasis and the apodosis. In order for the conditional to be prototypical, this relation should be causal, or in other words, they should be connected in the mental space of content. We can compare Tables 3.13, 3.14 and 3.20 to see the development in mental spaces. While keeping in mind that Table 3.20 is based on a much smaller sample than the other two, the tendencies are usually clear. Content conditionals show a decrease throughout, from 60% in the corpora before 1810 to 58% in COHA and 48% in COCA. The development of speech-act conditionals show more fluctuation, with 32% in the earliest corpora, 28% in COHA and 38% in COCA. As I will return to later in this section, however, the proportion of speech-act conditionals is higher in certain formal variants. Epistemic conditionals have increased in frequency from the earliest corpora, but the growth seems to have stagnated, so that the percentage for both COHA and COCA is 14%.

Finally, meta-metaphorical conditionals have occurrences in both the earliest and the latest corpora, but not in the middle, i.e. COHA. This is puzzling, as the presence of meta-metaphorical conditionals in the earliest corpora refutes my initial belief that this use of *say*-conditional was an illustration of how the use has spread into new mental spaces. The conditionals in question, namely [22] and [37] of the previous chapter, are repeated here as [2] and [3].

- [2] You and I aren't exact opposites. Like, **say** I'm Death and Darkness. Then you should be Immortality and Light instead of just some poor fool scrambling to stay afloat.

Fantasy & Science Fiction (1997) [COCA]

- [3] Ah! what a shame! ah, what a fault were this.
Say, Warwick was our anchor; what of that?
And Montague our top-mast; what of him?
Our slaughter'd friends the tackles; what of these?
Why, is not Oxford here another anchor?
And Somerset, another goodly mast?
The friends of France our shrouds and tacklings?
And, though unskilful, why not Ned and I
For once allow'd the skilful pilot's charge?

The Third Part of King Henry the Sixth (1594) [Shakespeare]

The use of metaphor in the two examples is slightly different. The modern conditional, i.e. [2], shows a combination of the meta-metaphorical and the epistemic mental spaces. Although the protasis is syntactically independent, the reasoning process is concise, and the metaphor is carried into the apodosis to be concluded there. In [3], on the other hand, the conditional *Say, Warwick was our anchor; what of that?* is part of a larger passage based on the same metaphor, i.e. allies portrayed as parts of a ship. We might also note that there is another 'conclusion' later in the passage, namely *Why, is not Oxford here another anchor?*, which answers the question *what of that?*. We might thus argue that [3] is not a meta-metaphorical conditional per se, but rather a speech-act conditional situated within the larger context of a metaphor. However, the absence of meta-metaphorical conditionals in COHA might just be the result of this being a minor use of an already infrequent construction.

The absence of metalinguistic conditionals should also be commented on. I noted in Section 2.2.1 that Dancygier and Sweetser (2005) consider there to be five mental spaces in which the protasis and the apodosis of a conditional can be connected. It appears that of these five, only the content, speech-act and epistemic spaces were part of Sweetser's original study (Sweetser (1990); in Dancygier 1999: 12-13), to which Dancygier (1999: 13) suggests the addition of a metalinguistic space. In Dancygier and Sweetser (2005), meta-metaphorical conditionals have been added to the other four. From this we might gather that metalinguistic conditionals are less frequent than content, speech-act and epistemic conditionals, but probably not less frequent than meta-

metaphorical ones. The absence of metalinguistic *say*-conditionals, as the absence of meta-metaphorical conditionals in COHA, might be due to the low frequency and chance. On the other hand, it could be possible that *say*-conditionals do not occur with metalinguistic content because of the origin of *say* as a *verbum dicendi*. If so, this would indicate persistence, as I will return to in the next section.

Overall, the variation in mental spaces shows a development away from content conditionals, and as such also from the prototype. This might indicate that *say*-conditional originated as a content conditional and has since spread to be used in with other meaning relations, particularly the speech-act mental space. This ties in with Dancygier's (1999: 185) theory of "content level, i.e., causal predictive conditionality as a center from which semantic extension occurs to epistemic, speech act, and metatextual domains." While it seems clear that such a semantic extension has taken place, the 48% of content conditionals in COCA indicates there are many potentially prototypical instances of *say*-conditional.

Turning now to the form of *say*-conditional, Tables 3.11 and 3.12 show that the formal variant most closely tied to prototypical conditionals, i.e. *say p, (then) q*, has decreased from 27% in COHA to 11% in COCA. Moreover, Table 3.15 shows that in this formal variant, only 34, or 7% of the total 504 tokens in COCA, are content conditionals. As such, there are very few tokens which fulfil all the requirements for Dancygier's prototypical conditional.

Meanwhile, the variant with a syntactically independent protasis, i.e. *say p. (then) Q*, is increasingly becoming the preferred variant for *say*-conditional. It has increased from 52% in COHA to 66% in COCA. Further, the more detailed development shown in Figure 3.4 reveals fluctuation in frequency 1810–1944, followed by a jump in frequency after 1945. Table 3.15 shows that in COCA, *say p. (then) Q* has a near equal preference for content conditionals and speech-act conditionals, with respectively 45% and 41%. This is opposed to the distribution in COHA, shown in Table 3.16. Here, content conditionals are clearly preferred with 52%, whereas speech-act conditionals account for 26%. This means that in this formal variant, speech-act conditionals have increased with 15 percentage points from COHA to COCA. This is interesting, as it suggests that the increase in frequency of *say p. (then) Q* after 1945 might be connected to speech-act

conditionals becoming ‘normalised’ and thus being used more. If the increase in speech-act continues in this formal variant, this might become the new norm for *say*-conditional.

4.2 Findings in relation to grammaticalisation theory

First of all, the overall increase in the use of *say*-conditional shows both that it has become more frequent in recent years. As Figure 3.3 shows, the frequency of *say*-conditional fluctuates after 1940 as well as before, although with higher values after 1940. It is unclear whether *say*-conditional has stabilised in terms of frequency, or whether it will increase more. As we have seen in Section 4.1, there are still developments in the meaning and form of the construction. This, combined with the potential for further increase in frequency, gives the impression that the ‘norm’ for *say*-conditional is not yet firmly set.

The development in genre distribution reflects the principle that language change initiates in everyday language, and spreads to more formal registers once the new use is established. Tables 3.2 and 3.4 show the development in genre between COHA and COCA. The most striking development is in ‘Fiction’, where the frequency has decreased from 71% in COHA to 22% in COCA. We can tentatively add to this that 26, i.e. 93% of the 28 tokens from earlier corpora are from fiction. All these are from plays, which are, of course, based on dialogue. Moreover, there are more tokens from comedies than from tragedies and historical plays, perhaps indicating a preference for less formal registers.

The increase in ‘Popular magazines’ is also sizeable, from 17% in COHA to 38% in COCA. At 12%, the usage in ‘Non-fiction books’ is more common in COHA than in its (near) equivalent ‘Academic journals’ in COCA, where the percentage is 4%. This might be accounted for by the difference in the two genres. However, the material for both ‘Non-fiction books’ and ‘Academic journals’ is selected according to the Library of Congress classification system, which indicates that the content of each genre is fairly similar. ‘Newspapers’ shows an increase of 11 percentage points, from 1% in COHA to 12% in COCA.

The increase in both ‘Newspapers’ and ‘Popular magazines’ might be accounted for by hypothesising that *say*-conditionals are becoming acceptable in more formal registers. On the other hand, the low frequency in ‘Academic journals’ does not encourage this interpretation. It is perhaps more likely that the register used in newspapers and magazines has become more informal with time, and thus allowing for more

innovative language use. This hypothesis is supported by Mair's (2006: 183-93) observations about the colloquialisation of English which happened throughout the twentieth century. The term 'colloquialisation' refers to a stylistic shift in the written norm, in which there is a movement away from a formal, elaborated style intended to be as removed from spoken language as possible, and towards a style more resembling of speech. The colloquialisation of English happened throughout the twentieth century, but the process has accelerated in the second half of the century. Mair also notes that the colloquialisation of English is unevenly distributed by genre: press writing shows a great tendency towards a more informal style, while the shift in academic writing is more conservative. This tendency, of course, perfectly reflects the development of *say*-conditional.

The decrease in *that*-clauses following conditional *say* from COHA to COCA, as shown in Table 3.5, emphasises that *say* has become reanalysed as a subordinator. In COCA, zero-clauses are preferred to such an extent that *that*-clauses are nearly excluded, with 98% and 2% respectively. The distribution in COHA is more even, although the preference for zero-clauses is still strong, with 86%. A similar development is observed by Molencki (2012) while looking at the grammaticalisation of the conditional subordinators *providing* and *provided (that)*. He writes that the grammaticalisation path of these subordinators seems straightforward:

The legal provision might have started as a full finite clause with the passive form of provide complemented by a *that*-clause. The next stages might have been the reduction to a nonfinite clause (*being provided that*), then the loss of the copula, thus leaving only the participle, later reinterpreted as a conjunction. [...] The final stage of the grammaticalization down the participle-to-conjunction cline was the elimination of *that*, which took place in Shakespearian English (Molencki 2012: 205, 207)

Deletion of *that* suggests that the clause following is no longer seen as a complement to the verb *say*, but rather as an adverbial clause. In terms of the grammaticalisation cline presented in Section 2.1, *say* has developed from a content item to a grammatical word. As in Section 2.1.1, the reanalysis of *say* becomes clear when we look at how the clause structure has changed., which is shown in [4].

[4] *say* [clause complement *that* S V] > [adverbial clause *say* S V]

Looking at the results from earlier corpora, the preference for zero is smaller, with 61% to 39% *that*-clauses. From examples such as [34] in Chapter 3 and [6] and [7] below in comparison to one another, it seems that the use of *that* is optional – or perhaps a stylistic choice.

[5] **Say** I do speak with her, my lord, what then?

Twelfth Night (1623) [Shakespeare]

[6] Sooth, but you must.

Say that some lady, as perhaps, there is,
Hath for your love as great a pang of heart
As you have for Olivia: you cannot love her;
You tell her so; must she not then be answer'd?

Twelfth Night (1623) [Shakespeare]

Before moving on, we might add here as a word of caution that Table 3.6 shows the same decrease in *that*-clauses following imperative *say*. Although this is an interesting development in its own right, it lies outside the scope of this study. Further, it does not remove the fact that the decrease in *that*-clauses following conditional *say* can be accounted for by reanalysis, as explained above.

When it comes to Diewald's (2002) context types, it is clear that both the untypical and critical contexts must precede the time period covered in the present study. The first example attested in the OED is, as stated in Section 2.3, from 1596, and both my own and Brinton's (2008) first findings are approximately from that time. Only one corpus, namely PPCEME, covers the period 1500–1600, and there are no relevant tokens from this period. Nevertheless, the change from one stage to another is gradual, and examples of a critical context could be found among the collected data. A potential candidate is presented in [8], which is situated in the larger context of King Richard asking Queen Elizabeth for help in wooing her daughter. In this context, the *say*-clause can be interpreted either as a conditional with a covert apodosis *what then*, or as asking Elizabeth to repeat his words to the daughter, i.e. as an imperative.

[7] K. RICHARD
Say, that I did all this for love of her?

Q. ELIZABETH
Nay, then indeed, she cannot choose but hate thee,
Having bought love with such a bloody spoil.

The Tragedy of King Richard the Third (1597) [Shakespeare]

If this is the critical context in which *say*-conditional originated, it implies that *say*-conditional originated without a covert apodosis. A *say*-clause such as in [8] might have been intended as an imperative, but because of the ambiguous context, the hearer interpreted it as having a covert apodosis, i.e. an implicature, which led to the reanalysis of *say* as a conditional subordinator.

As for untypical contexts, we might loosely relate this to the semantic widening of *say* witnessed in Section 2.3. As *say* expanded its range to include more uses, the connection to an immediate speech-act was weakened in various ways. Brinton's (2008: 92) comment, also in Section 2.3, that something which is said cannot be taken as fact, but must be supposed, seems a likely explanation of the motivation for the onset of grammaticalisation.

The isolating contexts, although within the scope of this study, have proved difficult to identify. As the general observation seems to be that *say*-conditional is still evolving, this is perhaps to be expected. As discussed in Section 3.2.8, the only potential isolating context is *say you'll* (with variations), in which all of the 70 tokens in COCA were *verba dicendi*. It seems that this particular phrase is connected with eliciting a promise from the hearer. As such, the focus on the speech act might have prevented it from being interpreted as a conditional. We also have the case of the 'missing' metalinguistic *say*-conditionals, as mentioned in Section 4.1, which may be explained along the same line. This, as well as *say you'll*, might be an instance of 'persistence', the phenomenon where a form undergoing grammaticalisation retains "some traces of its original lexical meanings" and where "details of its lexical history may be reflected in constraints on its grammatical distribution" (Hopper and Traugott 2003: 96). As conditional *say* originates from a *verbum dicendi*, it could follow that there are restrictions on its use with words of that kind, such as *say* (as a *verbum dicendi*), *call*, *tell*, *whisper* and so on. Indeed, of the 504 instances of *say*-conditional found in COCA, only the one

given in [9] was combined with any of those verbs. This might explain why there are no occurrences of metalinguistic *say*-conditionals. As metalinguistic conditionals are concerned with commenting on a speech act, this might be seen as incompatible with conditional *say*.

Table 3.17 shows that when the phrases *say you're*, *say you have* and *say you want* occur sentence-initially, i.e. following a period, they are very likely to be conditional. These phrases favour conditional meaning, but not to the exclusion of the imperative meaning. As such, we cannot speak of a genuine isolating context.

[8] Now, **say I tell** those guys down there if they shoot or make a move, I'm killing Wayne Gale. And they shoot or make a move anyway. Now say by some freak accident, you don't die, you live through it. What would you do?

Natural Born Killers (1994) [COCA]

Although the form *say p, (then) q* nearly always involves *say* being a conditional subordinator, this is not always the case. This was also illustrated in Section 3.2.8. The absence of a clear isolating context for conditional *say* indicates that the process of grammaticalisation is still ongoing. *Say*-conditional might be approaching Stage III of Diewald's model, but does not seem to have reached it as of yet.

Finally, we might note that Table 3.1 shows a higher number of imperatives in COHA than in COCA. Given the differences in size of the two corpora – COCA has 520 million words while COHA has 350 – this effectively means that the use of imperatives in this context has decreased. In terms of normalised frequency, the frequency of imperatives in COCA is 0.5 per million words, while in COHA the frequency is at 2.5. Although this is pure speculation, it could be that this context is increasingly being associated with conditionality, and as such is avoided by language users wishing to convey imperative meaning. If that is the case, it would mean that *say*-conditional is becoming increasingly established. However, as observed earlier in this section, imperative *say* in this context seems to be going through changes of its own. Thus, this decrease in *say*-imperatives could be caused by something entirely unrelated.

To sum up this discussion, it seems that *say p. (then) Q* is becoming the dominant formal variant. This form favours both content and speech-act conditionals, but the recent growth of speech-act conditionals in this variant suggests that these are likely to become

normative. The stagnation in the development of epistemic conditionals makes it unlikely that these will become more common in the future.

The development of *say*-conditional also relates to several points of grammaticalisation theory. The earliest attestations are primarily from fiction, which often seeks to imitate spoken language. The spread from fiction, into ‘new’, more formal genres is in keeping with language change happening in everyday language, as mentioned in Section 2.1.2. The decrease in *that*-clauses and the presence of at least one isolating context also touch upon characteristics on grammaticalisation.

5. Conclusion

In Chapter 4 I discussed the findings presented in Chapter 3 in relation to the theoretical background presented in Chapter 2. In this chapter I aim to sum up my findings and relate them to the research questions presented in Section 1.2. I will also look at shortcomings of this study, and indicate possible further research.

Below are presented a summary of my findings based on the research questions in Section 1.2:

1. The overall frequency of *say*-conditional has more than doubled from the period covered by COHA, i.e. 1810–1989 to the period covered by COCA, i.e. 1990–2015. There is especially a jump in usage after 1940.
2. The frequency of *that*-clauses following conditional *say* has decreased over time. In COCA, *say*-conditional shows a preference for zero-clauses almost to the exclusion of *that*-clauses. This development is similar to that of other related cases of grammaticalisation, such as *providing/provided (that)*, and reinforces the observation that *say* has become reanalysed as a conditional subordinator.
3. My findings from COCA show that *say*-conditional is found in all genres, although with a low frequency in academic writing. Findings from COHA are predominantly from fiction, as are the findings from earlier corpora. As fiction, and most notably plays, often imitate everyday language, these findings indicate a spread into more formal registers which is still ongoing.
4. The development of *say*-conditional does not indicate a movement towards the prototypical *if*-conditional. Rather, recent development shows a preference for a syntactically independent protasis, and the proportion of prototypical *say*-conditionals has decreased over time rather than increased.

5. A *say*-conditional with a syntactically independent protasis is preferred over a dependent one. Further, the recent development within this formal variant shows that speech-act conditionals have increased rapidly, and may become the new norm for *say*-conditional.

These points show that all the hypotheses except Hypothesis 4 is confirmed by my findings. The confirmation of Hypotheses 1–3 is expected, as these are all related to commonly observed tendencies in grammaticalisation.

5.1 Shortcomings

There will always be things one would wish one had done differently, and the present study is no exception. Here I present some aspects of the study which, for time considerations or other reasons, have left something wanting.

My intention in this study was to both trace the development *say*-conditional, and to see whether I could discover some motivating factors for the reanalysis of *say* into a conditional subordinator. This last part has largely been left untouched. This is due both to time constraints and to the fact that large corpora covering the relevant period, i.e. approximately the sixteenth century, are difficult to come by. This point is also made by Mair (2011: 247), who notes that smaller corpora might be sufficient to study phenomena of higher frequency, but are “insufficient when it comes to rarer types of clausal subordination.” *Say*-conditional is an infrequent construction, and as such, a large corpus is necessary in order to find enough relevant data.

A concern which manifested during my data analysis was that the formal variant *say p. (then) q* might be too large a category. Although all the tokens in this category have in common a syntactically independent protasis and a following, overt apodosis, it varies how soon the apodosis follows. Had there been time for further analysis, I would have liked to have attempted potential subgroupings of this formal variant.

Another concern relating to formal variants is the potential problem caused by the written presentation of spoken language. My formal variants rely on orthographic signs, i.e. commas designate one formal variant while periods designate another. The distinction between the two might not always be clear while transcribing spoken language. A similar problem presented itself in fictional writings. It is a common practice in narratives to

interrupt a direct discourse with phrases such as *he said, said she*, etc. and then continue the direct discourse as though there had been no interruption. This becomes a problem when a conditional is presented as [‘protasis’ *she said* ‘apodosis’], as it could be placed in either *say p, (then) q* or *say p. (then) Q*. I have consistently opted for the latter.

Finally I would like to add that my analysis, particularly that part which relates to the semantic relationship between the protasis and the apodosis, might include the occasional error of judgement. As the conditionals in my data show great variation in both shape and meaning, it has sometimes been difficult to know where to draw the line. However, I am confident that the overall tendencies which I have presented are valid despite any potential lapses.

5.2 Further research

The present study is, to my knowledge, the first study dedicated to *say* as a conditional subordinator. As such, there is much left to learn.

As also mentioned as a shortcoming, it would be interesting to look for evidence of a critical context in the period prior to the reanalysis of *say*, i.e. the sixteenth century. This would offer more insight into the motivations for the change.

Similarly, a closer look at the period around 1945 could provide interesting findings. In this period the syntactically independent protasis became the dominant formal variant of *say*-conditional, and this is probably the cause of the increase in frequency after 1945. As such, a more qualitatively focused study of this period could give new insights.

It could also be worthwhile to look for distributional differences between British and American English. This could easily be done by use of the British National Corpus, which is available using the same interface as COCA and COHA. This would provide results which are entirely comparable to those of this study.

Finally, I believe that a cross-linguistic study between English and Norwegian could be worthwhile. Since beginning this project, I have observed *si*, i.e. the Norwegian equivalent of *say*, being used similarly in Norwegian. Whether this is an independent development or whether there has been influence between the languages one way or another is unknown to me, but it would certainly be interesting to discover.

References

Primary sources

- Biber, Douglas, and Edward Finegan. 2013. *ARCHER: A Representative Corpus of Historical English Registers*. <http://www.projects.alc.manchester.ac.uk/archer/> (accessed January 2018).
- Davies, Mark. 2008–. *The Corpus of Contemporary American English (COCA): 560 million words, 1990–present*. <https://corpus.byu.edu/coca/> (accessed October 2017).
- Davies, Mark. 2010–. *The Corpus of Historical American English (COHA): 400 million words, 1810–2009*. <https://corpus.byu.edu/coha/> (accessed October 2017).
- Kroch, Anthony, Beatrice Santorini, and Lauren Delfs. 2004. *The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME)*. CD-ROM, first edition, release 3.
- Shakespeare, William. 1916. *The Complete Works of William Shakespeare (The Oxford Shakespeare)*. ed. William James Craig. Oxford: Oxford University Press. <http://oll.libertyfund.org/titles/1612> (accessed April 2018).
- Siemund, Rainer, and Josef Schmied. 2001. *The Lampeter Corpus of Early Modern English Tracts*. <https://www.tu-chemnitz.de/phil/english/sections/linguist/real/independent/lampeter/lamphome.htm> (accessed February 2018).

Dictionaries

- An Anglo-Saxon Dictionary Online*, s. v. 'secgan'. Joseph Bosworth and Thomas Northcote Toller. www.bosworthtoller.com/027305 (accessed November 2017).
- The American Heritage dictionary of Indo-European roots*, s. v. 'sek'. Calvert Watkins. (accessed November 2017).
- Middle English Dictionary*, s. v. 'seien'. Frances McSparran, Paul Schaffner, John Latta, Alan Pagliere, Christina Powell and Matt Stoeffler. <http://quod.lib.umich.edu/cgi/m/mec/med-idx?type=id&id=MED39209> (accessed November 2017).
- Oxford English Dictionary*, s. v. 'say, v.1 and int.'. OED Online. <http://www.oed.com/view/Entry/171590?rskey=WDZHkk&result=5> (accessed 18.09.17).

Software

- Anthony, Laurence. 2018. AntConc (Version 3.5.7) [Computer Software]. Tokyo: Waseda University.

Secondary references

- Andersen, Henning. 1973. Abductive and Deductive Change. *Language* 49 (4):765–793. doi: 10.2307/412063.
- Brems, Lieselotte, and Kristin Davidse. 2010. Complex subordinators derived from noun complement clauses: grammaticalization and paradigmaticity. *Acta Linguistica Hafniensia* 42 (1):101–116. doi: 10.1080/03740463.2010.486920.
- Brinton, Laurel J. 2008. Comment clauses with *say*. In *The comment clause in English: syntactic origins and pragmatic development*, 73–110. Cambridge: Cambridge University Press.
- Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The evolution of grammar: tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Comrie, Bernard. 1986. Conditionals: a typology. In *On Conditionals*, edited by Elizabeth Closs Traugott, Alice ter Meulen, Judy Snitzer Reilly and Charles A. Ferguson, 77–99. Cambridge: Cambridge University Press.
- Cristofaro, Sonia. 2014. Is there really a syntactic category of subordination? In *Contexts of Subordination: Cognitive, typological and discourse perspectives*, edited by Laura Visapää, Jyrki Kalliokoski and Helena Sorva, 73–91. Amsterdam: John Benjamins Publishing Company. <http://ebookcentral.proquest.com/lib/bergen-ebooks/detail.action?docID=1772295> (accessed April 2018).
- Dancygier, Barbara. 1999. *Conditionals and Prediction: Time, Knowledge and Causation in Conditional Constructions*. Cambridge: Cambridge University Press. <http://ebookcentral.proquest.com/lib/bergen-ebooks/detail.action?docID=144737> (accessed January 2018).
- Dancygier, Barbara, and Eve Sweetser. 2005. *Mental spaces in grammar: conditional constructions*. Vol. 108, Cambridge studies in linguistics. Cambridge: Cambridge University Press.
- Declerck, Renaat, and Susan Reed. 2001. *Conditionals: A Comprehensive Empirical Analysis*. Vol. 37, Topics in English linguistics. Berlin: Mouton de Gruyter.
- Diewald, Gabriele. 2002. A Model for Relevant Types of Contexts in Grammaticalization. In *New Reflections on Grammaticalization*, edited by Ilse Wischer and Gabriele Diewald, 103–120. Philadelphia: John Benjamins Publishing Company. <https://ebookcentral-proquest-com.pva.uib.no/lib/bergen-ebooks/detail.action?docID=622912>. (accessed December 2017).
- Dressler, Wolfgang U. 1985. *Morphology: the Dynamics of Derivation*. Ann Arbor: Karoma.
- Fauconnier, Gilles. 1994. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge: Cambridge University Press.
- Ferguson, Charles A., Judy Snitzer Reilly, Alice ter Meulen, and Elizabeth Closs Traugott. 1986. Overview. In *On Conditionals*, edited by Elizabeth Closs et al. Traugott, 3–20. Cambridge: Cambridge University Press.
- Greenbaum, Sidney. 1996. *The Oxford English grammar*. Oxford: Oxford University Press.
- Greenbaum, Sidney, and Randolph Quirk. 1990. *A student's grammar of the English language*. Harlow: Longman.
- Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In *Universals of language: report of a conference held at Dobbs Ferry, New York, April 13–15, 1961*, edited by Joseph H. Greenberg, 58–90. Cambridge, Mass: MIT Press.

- Grice, Herbert Paul. 1975. Logic and conversation. In *Speech Acts*, edited by Peter Cole and Jerry L. Morgan, 41–58. New York: Academic Press.
- Haiman, John. 1985. *Natural Syntax: Iconicity and erosion*. Vol. 44, Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.
- Heine, Bernd. 2002. On the role of context in grammaticalization. In *New Reflections on Grammaticalization*, edited by Ilse Wischer and Gabriele Diewald, 83–102. Philadelphia: John Benjamins Publishing Company. <http://ebookcentral.proquest.com/lib/bergen-ebooks/detail.action?docID=622912> (accessed March 2018).
- Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press. <http://ebookcentral.proquest.com/lib/bergen-ebooks/detail.action?docID=202017> (accessed November 2017).
- Hopper, Paul J., and Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd ed. ed, Cambridge textbooks in linguistics. Cambridge: Cambridge University Press.
- Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Langacker, Ronald W. 1977. Syntactic reanalysis. In *Mechanisms of Syntactic Change*, edited by Charles N. Li, 57-139. Austin: University of Texas Press.
- Levinson, Stephen. 2000. *Presumptive Meanings: the Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press, Bradford Book.
- Mair, Christian. 2006. *Twentieth-Century English*, Studies in English Language. Cambridge: Cambridge University Press.
- Mair, Christian. 2011. Grammaticalization and corpus linguistics. In *The Oxford Handbook of Grammaticalization*, edited by Bernd Heine and Heiko Narrog, 239–50. Oxford: Oxford University Press. <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199586783.001.0001/oxfordhb-9780199586783-e-19> (accessed May 2018).
- Matthiesen, Christian, and Sandra A. Thompson. 1988. The structure of discourse and "subordination". In *Clause Combining in Grammar and Discourse*, edited by John Haiman and Sandra A. Thompson. Amsterdam: Benjamins.
- Molencik, Rafał. 2012. "Providing/provided that": Grammaticalization or loan translation? In *English Historical Linguistics 2010: Selected Papers from the Sixteenth International Conference on English Historical Linguistics (ICEHL 16), Pecs, 23-27 August 2010*, edited by Irén Hegedus and Alexandra Fodor, 197–214. Amsterdam: John Benjamins Publishing Company. <http://ebookcentral.proquest.com/lib/bergen-ebooks/detail.action?docID=1053085> (accessed March 2018).
- Montolío, Estrella. 2000. On Affirmative and Negative Complex Conditional Connectives. In *Cause-Condition-Concession-Contrast: Cognitive and Discourse Perspectives*, edited by Elizabeth Couper-Kuhlen and Bernd Kortmann, 143–71. Berlin, Germany: Berlin, Germany: Mouton de Gruyter.
- Ohori, Toshio. 2011. The Grammaticalization of subordination. In *The Oxford Handbook of Grammaticalization*, edited by Bernd Heine and Heiko Narrog, 636–45. Oxford: Oxford University Press. <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199586783.001.0001/oxfordhb-9780199586783-e-52> (accessed April 2018).

- Poutsma, H. 1926. *A grammar of late modern English: Pt. 2 Section 2: The parts of speech The verb and the particles*. Vol. Pt. 2 Section 2. Groningen: Noordhoff.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive grammar of the English language*. London: Longman.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 2005. The Coordination-Subordination Gradient. In *Fuzzy Grammar: a reader*, edited by Bas Aarts, David Denison, Evelien Keizer and Gergana Popova, 423–430. Oxford: Oxford University Press.
- Thompson, Sandra A., and Robert E. Longacre. 1985. Adverbial clauses. In *Language typology and syntactic description: Vol. 2: Complex constructions*, edited by Timothy Shopen, 171–205. Cambridge: Cambridge University Press.
- Traugott, Elizabeth Closs. 1985. Conditional markers. In *Iconicity in Syntax: Proceedings of a Symposium on Iconicity in Syntax, Stanford, June 24–6, 1983*, edited by John Haiman, 289–307. Amsterdam: John Benjamins Publishing Company.